VOLUME LXXII NUMBERS 137 AND 138

PROCEEDINGS

OF THE

Casualty Actuarial Society

ORGANIZED 1914

1985

VOLUME LXXII

Number 137 — May 1985

Number 138 — November 1985

Library of Congress Catalog No. HG9956.C3

Printed for the Society by
Recording and Statistical Corporation
Boston, Massachusetts

FOREWORD

The Casualty Actuarial Society was organized in 1914 as the Casualty Actuarial and Statistical Society of America, with 97 charter members of the grade of Fellow; the Society adopted its present name on May 14, 1921.

Actuarial science originated in England in 1792, in the early days of life insurance. Due to the technical nature of the business, the first actuaries were mathematicians; eventually their numerical growth resulted in the formation of the Institute of Actuaries in England in 1848. The Faculty of Actuaries was founded in Scotland in 1856, followed in the United States by the Actuarial Society of America in 1889 and the American Institute of Actuaries in 1909. In 1949 the two American organizations were merged into the Society of Actuaries.

In the beginning of the twentieth century in the United States, problems requiring actuarial treatment were emerging in sickness, disability, and casualty insurance—particularly in workers' compensation—which was introduced in 1911. The differences between the new problems and those of traditional life insurance led to the organization of the Society. Dr. I. M. Rubinow, who was responsible for the Society's formation, became its first president. The object of the Society was, and is, the promotion of actuarial and statistical science as applied to insurance other than life insurance. Such promotion is accomplished by communication with those affected by insurance, presentation and discussion of papers, attendance at seminars and workshops, collection of a library, research, and other means.

Since the problems of workers' compensation were the most urgent, many of the Society's original members played a leading part in developing the scientific basis for that line of insurance. From the beginning, however, the Society has grown constantly, not only in membership, but also in range of interest and in scientific and related contributions to all lines of insurance other than life, including automobile, liability other than automobile, fire, homeowners and commercial multiple peril, and others. These contributions are found principally in original papers prepared by members of the Society and published in the annual *Proceedings*. The presidential addresses, also published in the *Proceedings*, have called attention to the most pressing actuarial problems, some of them still unsolved, that have faced the insurance industry over the years.

The membership of the Society includes actuaries employed by insurance companies, ratemaking organizations, national brokers, accounting firms, educational institutions, state insurance departments, and the federal government; it also includes independent consultants. The Society has two classes of members, Fellows and Associates. Both classes are achieved by successful completion of examinations, which are held in May and November in various cities of the United States and Canada.

The publications of the Society and their respective prices are listed in the *Yearbook* which is published annually. The *Syllabus of Examinations* outlines the course of study recommended for the examinations. Both the *Yearbook*, at a $10 charge, and the *Syllabus of Examinations*, without charge, may be obtained upon request to the Casualty Actuarial Society, One Penn Plaza, 250 West 34th Street, New York, New York 10119.

JANUARY 1, 1986

*EXECUTIVE COUNCIL

PHILLIP N. BEN-ZVI .... *President*
MICHAEL A. WALTERS................................ *President-Elect*
RICHARD H. SNADER .... *Vice President-Administration*
DAVID G. HARTMAN...................... *Vice President-Development*
WAYNE H. FISHER .... *Vice President-Membership*
MICHAEL FUSCO.......................... *Vice President-Programs*

THE BOARD OF DIRECTORS

**Officers:*

PHILLIP N. BEN-ZVI .... *President*
MICHAEL A. WALTERS................................ *President-Elect*

†Immediate Past President:

C. K. KHURY..1986

†Elected Directors:

ROBERT A. BAILEY 1986
E. FREDERICK FOSSA...1986
ROBERT B. FOSTER..1986
GEORGE D. MORISON 1986
LINDA L. BELL...1987
HERBERT J. PHILLIPS (a) 1987
KEVIN M. RYAN 1987
MICHAEL L. TOOTHMAN..1987
JAMES R. BERQUIST...1988
CHARLES F. COOK 1988
DAVID P. FLYNN..1988
MAVIS A. WALTERS...1988

---

* Term expires at 1986 Annual Meeting. All members of the Executive Council are officers. The Vice President-Administration also serves as Secretary and Treasurer.

† Term expires at Annual Meeting of year given.

(a) Appointed by the Board to fill unexpired term of Michael Fusco.

1985 PROCEEDINGS
CONTENTS OF VOLUME LXXII

1985 PROCEEDINGS
CONTENTS OF VOLUME LXXII—CONTINUED

Volume LXXII, Part 1 No. 137

PROCEEDINGS

May 8, 9, 10, 11, 1985

ON STEIN ESTIMATORS: "INADMISSIBILITY" OF ADMISSIBILITY AS A CRITERION FOR SELECTING ESTIMATORS

JAMES E. BUCK

Abstract

Stein estimators are an alternative (non-Bayesian) explanation for credibility. Until this year, the syllabus for Part 4 of the Society's examinations contained an article discussing Stein estimators, or James-Stein estimators, as part of the credibility readings for the exam [2]. The article focuses on some examples where Stein estimators are applied to baseball players' batting averages, among other things. In the examples, Stein estimators seem much like Bayesian credibility estimators and, in fact, credibility estimators derived from Stein's theory have been used by the Insurance Services Office for products liability classification rate-making.

Alike as Stein estimators and Bayesian credibility estimators are in practice, the theory behind Stein estimators is very much different and does not make much sense from the author's point of view. This paper consists of a discussion of the theory that underlies Stein estimators, including an example which illustrates the flaw in logic behind this alternative explanation of credibility.

INTRODUCTION

The literature of the Casualty Actuarial Society has been replete for years with papers on the theory of credibility (for instance, [3], [7], [8]). Practice, at least for most direct lines of business, has lagged far behind. In 1980, the Insurance Services Office (ISO) Credibility Subcommittee [5] produced a comprehensive report on credibility which recommended adoption of an empirical Bayes credibility procedure for products liability classification ratemaking. Normally, one would rejoice at this attempt of life to imitate art. However, the method chosen for use was adapted from the method of Morris and Van Slyke [9], which in turn is based on Stein estimation. Stein estimation is derived from the work of Charles Stein [10] (also, James and Stein [6]), and herein lies the reason for the author's less-than-jubilant reaction to the method of estimation chosen: the theory underlying Stein estimators does not make sense.

From a practical point of view, the adapted Morris-Van Slyke procedure worked better than the Buhlmann-Straub empirical Bayesian procedure in the testing done by the ISO. This is not all that surprising, given that the Morris-Van Slyke procedure is biased upwards and the testing included groups where the expected class loss ratios trended up or down over time. One of the assumptions underlying the Buhlmann-Straub credibility procedure is that the expected loss ratio of a class remains fixed over time. If the expected loss ratio changes, then the credibility to be applied to the most recent experience should be higher, since this recent experience is more related to the expected future experience of the class than the rate based on past class data.

While the Morris-Van Slyke procedure seems to work well in the simulations performed by the ISO, its theoretical flaws make the application of the technique to other problems dangerous. For example, the degree of upward bias in the class credibilities is directly related to the number of classes in the group: in its report, the ISO Subcommittee notes (p. I-19), "An interesting observation is that this process [the adapted Morris-Van Slyke method] effectively produces a minimum credibility of 3/k [where k is the number of classes] for each class in the group." Interesting, indeed. The ISO testing procedure involved groups with between 9 and 24 classes, and so the minimum credibility for each class was between 1/3 and 1/8. However, if the Morris-Van Slyke credibility procedure were applied without adjustment to private passenger auto territorial ratemaking for Rhode Island, the experience of each of the three territories would be given full credibility, regardless of the amount of experience! In this instance, it is clear that the Morris-Van Slyke procedure would not work well. What we have, then, is a procedure which works well in some instances, and yet produces poor results in other instances. Why? In the author's view, it is because the Morris-Van Slyke procedure used by the ISO is based, among other things, on Stein estimators, and Stein estimators are theoretically unsound. To understand the flaw in the theory, it is necessary to review the underlying statistical assumptions that form the basis of the development of Stein estimators.

THE THEORETICAL BASIS FOR STEIN ESTIMATORS

The focus of Bayesian estimation and Bayesian credibility is on modifying an estimate based on additional data. That is, the Bayes approach assumes that we already know something about the parameter to be estimated (the prior distribution). Bayes theorem and Bayesian credibility give us a way to combine that prior knowledge with additional information to produce a revised estimator of the parameter.

Stein estimators are based on a different (sometimes called frequentist or classical) view of estimation. According to this view, it is meaningless to discuss prior distributions of parameters; the parameters of a distribution are fixed values, even though the values may be unknown. Frequentists study the distribution of estimators about parameters in order to make inferences about the quality of different estimators. One of the properties of estimators used for comparison is expected squared error. To use a more specific example, let's take the normal distribution of mean μ and variance 1, or $N(\mu,1)$. If we select a sample point x from the distribution and use it as an estimator of μ, we know from the definition of variance that the estimator has an expected squared error of 1. Are there better estimators of μ? That is a very tough question to answer directly, if you believe that talking about the distribution of μ is meaningless. Since μ is fixed but unknown, there may well be better estimators, depending on the particular value of μ. For instance, if μ happens to be between 1 and 3, the fixed estimator $f(x)=2$ has smaller expected squared error than 1, the expected squared error of the estimator x.

Thus, we need an additional requirement besides low expected squared error if we are to choose among estimators in the frequentist framework. One such requirement is that an estimator be unbiased. An estimator is said to be unbiased if its expected value is always equal to the parameter to be estimated. In terms of the example, an estimator is unbiased if the expected value of the estimator is equal to μ for all values of μ. The sample point, x, is an unbiased estimator of μ and has been shown to be the unbiased estimator of minimum expected squared error (see, for example, [4], pp.362–365).

The requirement that an estimator be unbiased is one way to help define what is meant by best estimate, but in some cases it is felt to be too stringent. After all, an estimator that is biased but with low expected squared error may well be more desirable than an unbiased estimator of high expected squared error. This led to the alternate standard of admissibility for estimators. An estimator is said to be admissible with respect to a loss function (e.g. expected squared error) for a class of distributions if there is no other estimator which has expected squared error less than or equal to the expected squared error of the estimator for all distributions in the class, with the strict inequality holding for at least one distribution. Admissibility certainly sounds like an admirable quality for an estimator to have, but using it produces some disturbing results. In fact, the theoretical basis for Stein estimates is a proof by Stein [10] that the sample mean is not an admissible estimator of the mean of the n-variate normal distribution, $n \geq 3$. (This result is sometimes referred to as Stein's paradox.)

In order to discuss Stein's results, let's review briefly the multivariate normal distribution and its notation. Conceptually an n-variate normal distribution can be thought of as a collection of n separate variables, each normally distributed. Using vector notation, any particular multivariate normal distribution can be specified as $N(\vec{\mu},\Sigma)$, where $\vec{\mu}$ is a mean vector $\vec{\mu}(\mu_1,\ldots,\mu_n)$, with μ_i representing the mean of the i-th variable, and Σ is a symmetrical n-by-n covariance matrix, with each element of the matrix, σ_{ij}^2, representing the covariance between the i-th and j-th variables. If the n-variate distribution is independent, then the covariances between variables are equal to zero, and Σ is a diagonal matrix.

Stein considers the task of estimating $\vec{\mu}(\mu_1,\ldots,\mu_n)$ given a single sample point $\vec{x}(x_1,\ldots,x_n)$ selected from the multivariate independent normal distribution of variance 1, i.e., $\vec{x} \sim N(\mu, I)$, where I is the identity matrix. The usual estimator, $\vec{x}$, has expected squared error of n, the number of parameters to be estimated. James and Stein [6] developed an estimator with smaller squared error. The development of the estimator is based on the following property of the multivariate normal distribution: for any point $\vec{p}$,

$$P(|\vec{x} - \vec{p}| > |\vec{\mu} - \vec{p}|) > .50 .$$

In words, there is always a better than even chance that a point chosen at random from the multivariate normal distribution is farther away from $\vec{p}$ than $\vec{\mu}$, the mean of the distribution, is from $\vec{p}$, no matter what $\vec{p}$ is chosen to be. Stein estimators which shrink $\vec{x}$ to an arbitrary $\vec{p}$ by a factor of

$$\frac{n - 2}{|\vec{x} - \vec{p}|^2}$$

have smaller expected squared error than $\vec{x}$ for all $\vec{\mu}$. That is,

$$E|\hat{\mu} - \vec{\mu}|^2 < n,$$

$$\text{for } \hat{\mu} = \left[1 - \frac{n-2}{|\vec{x} - \vec{p}|^2}\right]\vec{x} + \frac{n-2}{|\vec{x} - \vec{p}|^2}\vec{p},\ n \geq 3$$

When Stein estimators are applied to problems, $\vec{p}$ is usually chosen to be the average result for the group—in the notation above, the average of the x_i—and the resulting formula looks a lot like a Bayesian credibility estimate.

It's *important* to note, however, that there is *no* requirement that $\vec{p}$ be chosen as the average of the group in the theoretical work by Stein. And this flexibility with regard to $\vec{p}$ produces unusual results, particularly if we change the frame of reference. For instance, consider the three-dimensional case, where we select $\vec{x}(x_1,x_2,x_3)$ from a multivariate normal distribution of mean $\vec{\mu}(\mu_1,\mu_2,\mu_3)$ and covariance matrix I, the identity matrix. To make the presentation simpler, let $\vec{x} = (0,0,0)$, the origin. According to Stein, $\vec{x}$ can be combined with any arbitrary $\vec{p}$ (shrunk toward $\vec{p}$) to produce a better estimate of $\vec{\mu}$. For example, if we select $\vec{p} = (1,0,0)$, the Stein estimate combining $\vec{p}$ and $\vec{x}$ is $0\vec{x} + 1\vec{p}$, or $\vec{p}$ itself. In fact, for any point chosen from the sphere of radius 1 centered at origin, the estimate is the point itself. Thus, every point on the sphere of radius 1 centered at the origin is a "better" estimate of $\vec{\mu}$ than $\vec{x}$, the origin.

If that were not unusual enough, we can go further and show that any point $\vec{a}$ is a Stein estimate of $\vec{\mu}$, if we select an appropriate $\vec{p}$. The $\vec{p}$ to choose, for any given $\vec{a}$, is determined from the formula $\vec{a}/|\vec{a}|^2$. So, to show that $\vec{a} = (100,0,0)$ is a Stein estimate, we need only choose $\vec{p} = (.01,0,0)$. Therefore, based on the theory underlying Stein estimators, even a point as far away as $(100,0,0)$ is a better estimate of $\vec{\mu}$ than $\vec{x} = (0,0,0)$, the sample point!

THE CIRCLE DISTRIBUTION

To understand what's wrong with Stein estimators, it helps to go through the development of a Stein-like estimator for a simpler distribution. The chosen distribution is the one-dimensional distribution defined on a plane by the function

$$f(x_1,x_2) = \frac{1}{2\pi},\ x_1^2 + x_2^2 = 1$$
$$= 0, \text{ elsewhere.}$$

This distribution represents the chance of randomly picking a point on the circle of radius 1 centered at the origin. The mean of the distribution is also the origin.

The circle distribution was chosen because from any point $\vec{p}$ on the plane, there is a better than 50% chance that the distance between a randomly selected point on the circle and $\vec{p}$ will be greater than the distance between the origin and $\vec{p}$. Geometrically, we can see this by noting that, for any point $\vec{p}$, the arc around $\vec{p}$ through the origin contains less than half the circle. Because the circle distribution shares this property with the multivariate normal distribution, we should be able to shrink the values of the circle distribution to an arbitrary $\vec{p}$ and get an estimate that is, on average, closer to the mean. Indeed, if we notice that, for any $\vec{p} = (p_1, p_2)$, the average squared distance between the circle distribution and $\vec{p}$ is

$$\frac{1}{2\pi}\int_0^{2\pi}(p_1 - \sin\theta)^2 + (p_2 - \cos\theta)^2\, d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} p_1^2 - 2p_1\sin\theta + \sin^2\theta + p_2^2 - 2p_2\cos\theta + \cos^2\theta\, d\theta$$

$$= p_1^2 + p_2^2 + 1,$$

we might consider estimators of the form

$$\hat{\mu} = \frac{1}{|\vec{x} - \vec{p}|^2 + c}\,\vec{p} + \frac{|\vec{x} - \vec{p}|^2 + c - 1}{|\vec{x} - \vec{p}|^2 + c}\,\vec{x}.$$

And, in fact, Appendix I shows that if c is greater than $(|\vec{x} - \vec{p}| + 1)^2$, the expected squared error of this estimator, $\hat{\mu}$, is always less than 1, the expected squared error of the usual estimator, $\vec{x}$.

If one were to take the classical viewpoint, and the viewpoint that underlies the standard of admissibility of estimators, we should use this form of estimator in determining $\vec{\mu}$, given a particular $\vec{x}$. The fallacy in this approach can be seen by taking a Bayesian point of view. Let's again use the circle distribution of radius 1 and choose at random a point $\vec{x}$ from a circle of radius 1 with an unknown center. Without loss of generality, we can set $\vec{x} = (0,0)$. Now, we want to estimate the center of the circle, given that $\vec{x}$ is a point on the circle. If we consider all possible circles of radius 1 equally likely, then a good candidate for the distribution of $f(\vec{\mu}|\vec{x} = (0,0))$ would be

$$f(\vec{\mu}) = \frac{1}{2\pi},\ \mu_1^2 + \mu_2^2 = 1$$

$$= 0,\ \text{elsewhere}.$$

In fact, if we represent equally likely (or no prior knowledge) as the prior distribution

$$h(\vec{\mu}) \lim_{n\to\infty} g_n(\mu_1,\mu_2) = \frac{1}{4n^2}, \; -n \le \mu_1 \le n, \; -n \le \mu_2 \le n$$

$$= 0, \text{ elsewhere},$$

among others, then the candidate distribution shown as $f(\vec{\mu})$ above can be derived through the use of Bayes Theorem for continuous functions (see Appendix II).

Now, from a Bayesian point of view, we have determined the distribution of $f(\vec{\mu}|\vec{x})$. The next step is to determine the best point etimate of the $\vec{\mu}$ distribution (uniform distribution on the unit circle centered at the origin). The squared error function between the $\vec{\mu}$ distribution and any estimate $\vec{e} = (e_1,e_2)$ is given by

$$\frac{1}{2\pi}\int_0^{2\pi} (\sin\theta - e_1)^2 + (\cos\theta - e_2)^2 \, d\theta$$

$$= 1 + e_1^2 + e_2^2$$

which obviously is at a minimum at (0,0), or $\vec{x}$.

Stein estimation takes another approach. Stein's argument in this case would be, let us select an arbitrary point $\vec{p}$, say $\vec{p} = (2,0)$. It was previously shown that, if we shrink the $\vec{\mu}$'s to $\vec{p}$ by a factor of

$$1 - \frac{1}{|\vec{\mu} - \vec{p}|^2 + c}, \; c \ge (|\vec{\mu} - \vec{p}| + 1)^2,$$

the transposed $\vec{\mu}$'s are closer to $\vec{x}$. Based on this, it is therefore appropriate to shift $\vec{x}$ towards $\vec{p}$ by a factor of

$$1 - \frac{1}{|\vec{x} - \vec{p}|^2 + c}, \; c = (|\vec{x} - \vec{p}| + 1)^2$$

(or, equivalently, choose an estimate of (2/13, 0)) to give a better estimate of $\vec{\mu}$.

A geometric analogy may be of some help in understanding this point. Figure 1 shows the problem in graphical terms, from a Bayesian standpoint. Imagine that $\vec{\mu}$ represents the rim of a dartboard attached to the back of a door, and $\vec{p}$ represents the doorknob. The problem is to place a dart on the wall that is closest, on average, to the points on the rim of the dartboard ($f(\vec{\mu}|\vec{x})$). From the calculations above, and from common sense, we can see that the dart should be placed at the center of the dartboard ($\vec{x}$).

Figures 2 and 3 represent the Stein estimator approach. Figure 2 shows that if one squishes the rim of the dartboard a bit towards the doorknob (shifts the $\vec{\mu}$'s), there is a smaller average distance between the rim of the dartboard and the center of the dartboard ($\vec{x}$). This is then used to justify aiming the dart at a point closer to the doorknob, even though the problem is to get as close as possible, on average, to the rim of the original (unshifted) dartboard (Figure 3).

APPLICABILITY TO MULTIVARIATE NORMAL DISTRIBUTION

While it is easier to see the fallacy of admissibility and Stein estimators with respect to the circle distribution, Stein estimators are equally invalid for the multivariate normal distribution. Let's again take the problem of estimating $\vec{\mu}$ given $\vec{x}$, $\vec{x} \sim N(\vec{\mu}, I)$.

Using a variety of "flat" prior distributions, including $N(0,\infty)$ and the rectangular distribution used above, we can derive $\vec{\mu} \sim N(\vec{x}, I)$. Here, also, the standard of admissibility asks the wrong question from the Bayesian viewpoint. The proper question to ask is not what function $f(\vec{\mu})$ minimizes $E|f(\vec{\mu}) - \vec{x}|^2$, but rather, what value $\vec{p}$ minimizes $E|\vec{\mu} - \vec{p}|^2$. Because the multivariate normal distribution is independent and can be expressed as the product of one-dimensional normal distributions, the minimum of $\vec{p}$ at $\vec{x}$ follows from the fact that the squared distance function is minimized at the mean in the one-dimensional case.

From a theoretical point of view, it would seem that the major accomplishment of Stein estimators is to show that admissibility as applied by Stein isn't a very good criterion for choosing estimators and that the Bayesian theory of estimation, when properly applied, gives consistent and reasonable results. In fact, Stein's paradox is not a paradox at all when viewed from a Bayesian standpoint. From a practical point of view, biased estimators are still appropriate to use in many cases, but not those derived from this particular theory of estimation.

ACKNOWLEDGEMENTS

The author would like to thank Messrs. Chris Svendsgaard and Steven Yaset of the ISO for their review and comments on earlier drafts of this paper. The Committee on Review of Papers was also very helpful in this regard.

Figure 1

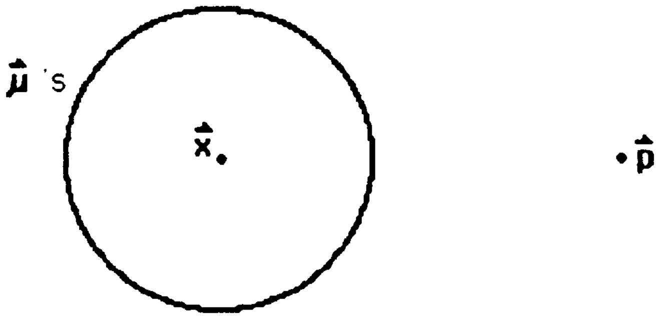


Figure 2

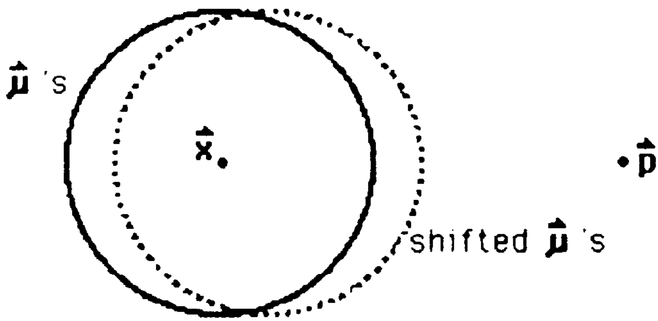
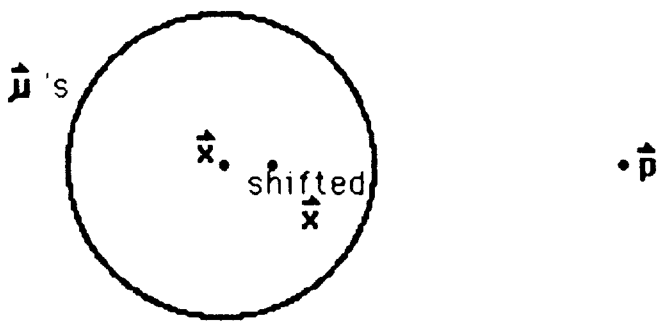


Figure 3



REFERENCES

[1] B. Efron, "Controversies in the Foundations of Statistics", *American Mathematical Monthly,* April, 1978, pp. 231–245.

[2] B. Efron and C. Morris, "Stein's Paradox in Statistics", *Scientific American,* May, 1977, pp. 119–127.

[3] C. C. Hewitt, "Credibility for Severity", *PCAS* LVII, 1970.

[4] P. G. Hoel, *Introduction to Mathematical Statistics,* 4th Edition, New York: John Wiley and Sons, Inc., 1971.

[5] Insurance Services Office, *Report of the Credibility Subcommittee: Development and Testing of Empirical Bayes Credibility Procedures for Classification Ratemaking,* New York, Insurance Services Office, 1980.

[6] W. James and C. Stein, "Estimation with Quadratic Loss", *Proceedings of the Fourth Berkeley Symposium,* Vol. 1, Berkeley, University of California Press, 1961, pp. 361–379.

[7] A. L. Mayerson, "A Bayesian View of Credibility", *PCAS* LI, 1964.

[8] A. L. Mayerson, D. A. Jones and N. L. Bowers, Jr., "On the Credibility of the Pure Premium", *PCAS* LV, 1968.

[9] C. Morris, and O. E. Van Slyke, "Empirical Bayes Methods for Pricing Insurance Classes", *Proceedings of the Business and Economics Statistics Section,* American Statistical Association, 1978.

[10] C. Stein, "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution", *Proceedings of the Third Berkeley Symposium,* Vol. 1, Berkeley, University of California Press, 1955, pp. 197–206.

APPENDIX I

DEMONSTRATION THAT THE MEAN IS INADMISSIBLE AS AN ESTIMATOR OF THE CIRCLE DISTRIBUTION

The following shows that there is a function that combines a data point, $\vec{x}$, with any $\vec{p}$ to produce an estimate of $\vec{\mu} = (0,0)$ that has expected squared error less than 1, the squared error of $\vec{x}$. This treatment is consistent with the frame of reference discussed in the text. However, this is equivalent to showing that for a circle distribution centered at $\vec{p}$, there is a function which combines $\vec{x}$, a randomly selected point on the circle, and the origin to produce an estimate of $\vec{p}$, the mean of the circle distribution, with expected squared error of less than 1. We consider estimators of the form

$$\frac{1}{|\vec{x} - \vec{p}|^2 + c}\vec{p} + \frac{|\vec{x} - \vec{p}|^2 + c - 1}{|\vec{x} - \vec{p}|^2 + c}\vec{x}, \; c \geq 1$$

$$= \vec{x} + \frac{\vec{p} - \vec{x}}{|\vec{x} - \vec{p}|^2 + c}.$$

For $\vec{x} = (\sin\theta, \cos\theta)$ and $\vec{p} = (p_1, p_2)$, the expected squared error is given by

$$\frac{1}{2\pi}\int_0^{2\pi} \left[\sin\theta + \frac{p_1 - \sin\theta}{(p_1 - \sin\theta)^2 + (p_2 - \cos\theta)^2 + c}\right]^2$$

$$+ \left[\cos\theta + \frac{p_2 - \cos\theta}{(p_1 - \sin\theta)^2 + (p_2 - \cos\theta)^2 + c}\right]^2 d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \sin^2\theta + \frac{2(p_1 - \sin\theta)\sin\theta}{(p_1 - \sin\theta)^2 + (p_2 - \cos\theta)^2 + c}$$

$$+ \frac{(p_1 - \sin\theta)^2}{[(p_1 - \sin\theta)^2 + (p_2 - \cos\theta)^2 + c]^2} + \cos^2\theta$$

$$+ \frac{2(p_2 - \cos\theta)\cos\theta}{(p_1 - \sin\theta)^2 + (p_2 - \cos\theta)^2 + c}$$

$$+ \frac{(p_2 - \cos\theta)^2}{[(p_1 - \sin\theta)^2 + (p_2 - \cos\theta)^2 + c]^2} d\theta$$

$$< \frac{1}{2\pi}\int_0^{2\pi} 1 + \frac{2(p_1 - \sin\theta)\sin\theta + 2(p_2 - \cos\theta)\cos\theta + 1}{(p_1 - \sin\theta)^2 + (p_2 - \cos\theta)^2 + c}\, d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} 1 + \frac{2p_1\sin\theta + 2p_2\cos\theta - 1}{p_1^2 - 2p_1\sin\theta + p_2^2 - 2p_2\cos\theta + 1 + c}\, d\theta$$

$$= \frac{1}{2\pi}\int_0^{2\pi} 1 + \frac{2p_1\sin\theta + 2p_2\cos\theta - 1}{2p_1\sin\theta + 2p_2\cos\theta - 1 - (p_1^2 + p_2^2 + c)}\, d\theta$$

Using the relation $\frac{a}{a-b} = 1 + \frac{b}{a-b}$,

$$\frac{1}{2\pi}\int_0^{2\pi} 1 - \left[1 + \frac{p_1^2 + p_2^2 + c}{2p_1\sin\theta + 2p_2\cos\theta - 1 - (p_1^2 + p_2^2 + c)}\right] d\theta$$

$$= \frac{-(p_1^2 + p_2^2 + c)}{2\pi}\int_0^{2\pi} \frac{d\theta}{2p_1\sin\theta + 2p_2\cos\theta - (p_1^2 + p_2^2 + c + 1)}$$

$$= \frac{-(p_1^2 + p_2^2 + c)}{2\pi}\int_{-\pi}^{\pi} \frac{d\theta}{\left[\begin{array}{c} 2p_1\sin(\pi + \theta) + 2p_2\cos(\pi + \theta) \\ -(p_1^2 + p_2^2 + c + 1) \end{array}\right]}$$

$$= \frac{p_1^2 + p_2^2 + c}{2\pi}\int_{-\pi}^{\pi} \frac{d\theta}{2p_1\sin\theta + 2p_2\cos\theta + p_1^2 + p_2^2 + c + 1}$$

Using integral tables, we find that the above is equivalent to

$$\frac{p_1^2 + p_2^2 + c}{2\pi} \cdot \frac{2}{\sqrt{(p_1^2 + p_2^2 + c + 1)^2 - (2p_1)^2 - (2p_2)^2}}$$

$$\tan^{-1} \frac{2p_1 + (p_1^2 + p_2^2 + c + 1 - 2p_2)\tan(\theta/2)}{\sqrt{(p_1^2 + p_2^2 + c - 1)^2 - (2p_1)^2 - (2p_2)^2}} \Bigg|_{-\pi}^{\pi}$$

$$= \frac{p_1^2 + p_2^2 + c}{2\pi} \cdot \frac{2\pi}{\sqrt{(p_1^2 - p_2^2 + c + 1)^2 - (2p_1)^2 - (2p_2)^2}}$$

$$= \sqrt{\frac{(p_1^2 + p_2^2 + c)^2}{(p_1^2 - p_2^2 + c)^2 + 2(p_1^2 + p_2^2 + c) + 1 - 4p_1^2 - 4p_2^2}}$$

$$= \sqrt{\frac{(p_1^2 + p_2^2 + c)^2}{(p_1^2 + p_2^2 + c)^2 + 2c + 1 - 2p_1^2 - 2p_2^2}}$$

So, if $c > p_1^2 + p_2^2 - 1/2$, the squared error is less than 1. In particular, since $(|\vec{p} - \vec{x}| + 1)^2 > \mathrm{p}_1^2 + \mathrm{p}_2^2 - 1/2$, if we choose $c \geq (|\vec{p} - \vec{x}| + 1)^2$, the estimator has expected squared error of less than 1.

APPENDIX II

DERIVATION OF THE POSTERIOR DISTRIBUTION USING THE CIRCLE DISTRIBUTION AND A "FLAT" PRIOR DISTRIBUTION

The purpose of this appendix is to determine $f(\vec{\mu}|\vec{x})$ for

$$f(\vec{x}|\vec{\mu}) = \frac{1}{2\pi}, \ |\vec{x} - \vec{\mu}|^2 = 1$$

$$= 0, \text{ elsewhere and}$$

$$h(\vec{\mu}) = \lim_{n\to\infty} g_n(\mu_1,\mu_2) = \frac{1}{4n^2}, \ -n \le \mu_1 \le n, \ -n \le \mu_2 \le n,$$

$$= 0, \text{ elsewhere.}$$

For $\vec{x} = (0,0)$ and any particular $n \geq 1$, the joint distribution is given by

$$f(\vec{x}|\vec{\mu})g_n(\vec{\mu}) = \frac{1}{8\pi n^2}, \ |\vec{x} - \vec{\mu}|^2 = 1, \ -n \le \mu_1 \le n, \ -n \le \mu_2 \le n,$$

$$= 0, \text{ elsewhere.}$$

$$= \frac{1}{8\pi n^2}, \ |\vec{\mu}|^2 = 1,$$

$$= 0, \text{ elsewhere, and}$$

$$\int f(\vec{x}|\vec{\mu})g_n(\vec{\mu})\, d\vec{\mu} = \int_0^{2\pi} \frac{d\theta}{8\pi n^2}$$

$$= \frac{1}{4n^2}$$

From Bayes Theorem for continuous functions, we have, for all $n \geq 1$,

$$f_n(\vec{\mu}|\vec{x}) = \frac{f(\vec{x}|\vec{\mu})g_n(\vec{\mu})}{\int f(\vec{x}|\vec{\mu})g_n(\vec{\mu})\, d\vec{\mu}}$$

$$= 4n^2 \cdot \frac{1}{8\pi n^2} = \frac{1}{2\pi}, \ |\vec{\mu}|^2 = 1,$$

$$= 0, \text{ elsewhere.}$$

and thus the distribution of $f(\vec{\mu}|\vec{x})$ is given by

$$f(\vec{\mu}|\vec{x}) = \frac{f(\vec{x}|\vec{\mu})\ h(\vec{\mu})}{f(\vec{x}|\vec{\mu})\ h(\vec{\mu})\ \mathrm{d}\vec{\mu}}$$

$$= \lim_{n\to\infty} f_n(\vec{\mu}|\vec{x})$$

$$= \frac{1}{2\pi} \quad , \quad |\vec{\mu}|^2 = 1,$$

$$= 0, \text{ elsewhere.}$$

DISCUSSION BY CHRISTIAN SVENDSGAARD AND PAUL BRAITHWAITE

INTRODUCTION

Having worked on Empirical Bayes credibility for a combined total of over ten years, we share Mr. Buck's frustration at the slow acceptance in practice of Empirical Bayes techniques. Part of the reason, we believe, is the inherent conservatism of the insurance business. Considering the sums at stake, practicing actuaries are reluctant to adopt new methods until they have been thoroughly researched and tested. It is gladdening, then, to see further discussion of credibility in the *Proceedings*. Only after undergoing thorough scrutiny can new methods hope to be adopted in practice.

Mr. Buck has written a paper that considers one aspect of credibility—Stein estimation—from a theoretical point of view. Lay actuaries hoping to see a comparison based on real data of an Empirical Bayes credibility procedure and (say) the square root rule must look elsewhere. But Mr. Buck's paper could still have relevance to lay actuaries. If a method can be shown to be theoretically incorrect, there is no reason to test it on real data.

While we applaud further exchange of ideas on credibility, we find parts of Mr. Buck's fundamental approach, and several of his conclusions, problematic. Our thoughts regarding his approach can be summarized as follows:

1. There are three schools of statistical thought: classical, Bayesian, and Empirical Bayesian. A case may be made for one school or another on philosophical grounds, or possibly on practical grounds. But from a mathematical viewpoint, arguing against one school based on the assumptions underlying another begs the question. This, Mr. Buck has done, treating Stein (classical) estimation from a Bayesian viewpoint.

2. The circle distribution example is interesting but not really relevant to actuarial problems.

More specific criticisms are:

1. In the normal case, the mean is the Bayes estimator *only* for a flat prior. In that case, the Stein estimator approaches the mean with probability one. (The concept of a "flat prior" is an attempt to extend the concept of a uniform prior distribution to an infinitely large parameter space. This is done by examining a sequence of uniform prior distributions, each covering a larger area of the parameter space. Loosely speaking, a flat prior gives the Bayes estimate if there is no prior belief.)

2. The circle distribution example, besides suffering from irrelevance, has the same problems as the normal case.

3. The adapted Morris-Van Slyke procedure, while akin to the Stein estimator, is based on Empirical Bayes rather than classical ideas. While Mr. Buck is correct in pointing out the bias in the procedure, the bias is due to the logical constraint that estimates of variances should not be negative. Because the bias is non-linear, it cannot be corrected by a linear transformation of the estimate. By making distributional assumptions, it might be possible to construct an unbiased estimator. However, the procedure would be valid only in situations where the posited distribution held and would lose its generality. In practice, Empirical Bayes credibility procedures have been applied to loss ratios. The distributional properties of loss ratios are complicated and it seems unlikely that an unbiased Empirical Bayes estimator could be constructed based on a realistic loss ratio distribution.

4. Tests on simulated data show that the $3/k$ adjustment factor that Mr. Buck criticizes should be used whether or not individual classes are trending at different rates from one another.

We explain and elaborate on these comments below.

COMMENTS ON THE APPROACH

Currently, there exist at least three schools of statistics: classical, Bayesian, and Empirical Bayesian. Each school makes different assumptions. Bayesians assume a prior distribution; classicists do not. Empirical Bayesians assume the parameters of the prior distribution are unknown; Bayesians do not.

It is easy to "prove" that one school is wrong by examining it from the viewpoint of another school. However, this makes no more sense than "proving" non-Euclidean geometry is wrong by making Euclidean assumptions. ("Assuming the parallel postulate holds, then any geometry where it does not hold has a contradiction. Therefore the parallel postulate holds.")

One set of assumptions may be more useful than another, because it fits reality better. Prior to general relativity, non-Euclidean geometry was an interesting curiosity. Afterwards, non-Euclidean geometry became *the* geometry. It is conceivable that *in practice* (say) Bayesian estimators will always perform best. The data could tell us which school of statistics is right. But it cannot be decided *a priori*.

The approach we advocate for selecting estimators in practice is:

(1) Selection of reasonable models;
(2) Testing of the model assumptions using the data;
(3) Derivation of estimators based on the models; and
(4) Testing of the estimators using the data.

The different schools of statistical thought might select different models, and testing of the model assumptions might not eliminate any of the models. This seems especially true of (pure) Bayesian models, which incorporate prior belief. But the various estimators derived will yield different results when tested on the data. Given enough data, one estimator will prove most attractive.

One of the major themes of Mr. Buck's paper is an argument that, from a certain Bayesian viewpoint, Stein estimators (i.e., classical estimators) do not make sense. He presents no empirical data. In our view, this argument is no more convincing than the argument against non-Euclidean geometry from a Euclidean point of view.

While you cannot "prove" one school is wrong from the point of view of another, it may be that the assumptions underlying one school are self-contradictory. Mr. Buck hints that the disturbing property of the Stein estimator, that it is not translation-invariant (i.e., that for a given data point, the Stein estimator could be anywhere, depending on the location of the origin you are shrinking toward) is such a contradiction.

While the non translation-invariance of the Stein estimator is disturbing, Mr. Buck has not shown that it is paradoxical. As a footnote, the Morris-Van Slyke and Bühlmann-Straub Empirical Bayes credibility procedures *are* translation-invariant. This is accomplished by shrinking towards the group mean, rather than the origin.

Mr. Buck attempts to illustrate the failings of Stein estimation by means of a similar estimator derived for the circle distribution. Reasoning by analogy is, of course, inappropriate in a mathematical context. The success or failure of the illustration must therefore be judged on its effectiveness as a pedagogical device. In our case, at least, we were not convinced by the illustration.

SPECIFIC COMMENTS

Inadmissible Estimator is Bayes Only For Flat Prior

We argue above that it is incorrect to criticize a classical estimator by making Bayesian assumptions. You will not find us making the converse mistake here. However, we do wish to show that in the normal case the mean is not a Bayes estimator except in the case of a flat prior.

We believe this is an important fact because, strictly speaking, a flat prior is not a prior distribution at all. The concept of a flat prior is based on a sequence of ever-flatter distributions. For any distribution in the sequence, the mean is not the Bayes estimate; this follows from Stein's result, as we will show below. Thus, while it is true that in the limit the mean is the Bayes estimate, it is not true for any intermediate point.

Moreover, in the limit, the Stein estimator approaches the mean with probability one. This means that under a flat prior the mean is not better than the Stein estimator: it is essentially equal to it.

Proof

We are attempting to estimate an (at least) 3-dimensional vector of means, θ, given a vector of observations, X, distributed normally around θ with covariance matrix, I, the identity:

$$X \sim N(\theta,I).$$

For an estimator $\hat{\theta}$ of θ, the squared error loss is

$$\text{Loss} = L(\hat{\theta},\theta) = |\hat{\theta} - \theta|^2$$

The risk is the expectation given θ:

$$\text{Risk} = R_\theta(\hat{\theta}) = E(L(\hat{\theta},\theta)|\theta)$$

The Bayes risk is

$$B(\theta) = E_\theta(R_\theta(\hat{\theta})) = \int_\Omega R_\theta(\hat{\theta})\, \mathrm{d}F(\theta)$$

where E_θ denotes "expectation with respect to θ," the integral is over all possible values of θ (i.e., the sample space is Ω), and $\int \ldots \mathrm{d}F(\theta)$ denotes Riemann-Stieltjes integral where $F(\theta)$ is the CDF of the prior of θ.

When Stein proved that the mean is not admissible, he proved it by showing that

$$R_\theta(\hat{\theta}_{\text{Stein}}) < R_\theta(\hat{\theta}_{\text{Mean}})$$

for all θ's, where $\hat{\theta}_{\text{Stein}}$ denotes the Stein estimator, and $\hat{\theta}_{\text{Mean}}$ denotes the mean (i.e., X). (See "Estimation with Quadratic Loss," p. 363 [2].)

The size of the difference

$$R_\theta(\hat{\theta}_{\text{Mean}}) - R_\theta(\hat{\theta}_{\text{Stein}})$$

depends on the value of θ. It is greatest at the origin (which makes sense, since the Stein estimator shrinks the estimate towards the origin). It decreases as θ moves away from the origin, but it is always positive.

How does the Bayes risk of $\hat{\theta}_{\text{Mean}}$ compare to the Bayes risk of $\hat{\theta}_{\text{Stein}}$?

$$B(\hat{\theta}_{\text{Mean}}) - B(\hat{\theta}_{\text{Stein}}) =$$

$$\int_\Omega R_\theta(\hat{\theta}_{\text{Mean}})\, dF(\theta) - \int_\Omega R_\theta(\hat{\theta}_{\text{Stein}})\, dF(\theta) =$$

$$\int_\Omega [R_\theta(\hat{\theta}_{\text{Mean}}) - R_\theta(\hat{\theta}_{\text{Stein}})]\, dF(\theta).$$

The expression inside the integral

$$R_\theta(\hat{\theta}_{\text{Mean}}) - R_\theta(\hat{\theta}_{\text{Stein}})$$

is greatest at the origin and decreases as θ moves away from the origin. But it is always positive. This is what Stein proved.

The value of the integral will depend on $F(\theta)$, the prior distribution. The more weight given to θ's away from the origin, the smaller the integral will be. *But it will always be positive.*

Since the Bayes estimator minimizes Bayes risk, the mean cannot be the Bayes estimator for any prior. Only by taking the limit of distributions throwing more and more weight away from the origin can the mean be made to approach the Bayes estimate.

Note, however, that in the limit the difference in Bayes risk is *zero*. The mean and the Stein estimator are equal in the limit. In the expression

$$\left(1 - \frac{n-2}{|\vec{x} - \vec{p}|^2}\right)\vec{x} + \left(\frac{n-2}{|\vec{x} - \vec{p}|^2}\right)\vec{p},$$

$|x - p|^2$ is greater than 100 (1,000,000, 10^{26}, ...) with probability .99 (.999, $1 - 10^{-500}$, ...) in the limit: the Stein estimator reduces to the mean $\bar{x}$.

A careful re-reading of the above proof should convince the reader that it can be made entirely general. An inadmissible estimator cannot be a Bayes estimate. In other words, Bayes estimators are admissible. (See [3].)

Generality of the Circle Distribution Example

To repeat our main concern, this is not relevant. If normal-distribution Stein estimation has faults, they cannot be discerned by examining circle distribution Stein-like estimators.

We showed earlier that Bayes estimators are admissible. Mr. Buck claims to "prove" that his Stein-like estimator dominates the mean (in Appendix I). He then derives the mean as a Bayes estimator. This is a contradiction due to the use of a flat prior. Note that in Appendix I, he claims to show that the Stein-like estimator has risk less than one. Then he shows that the risk of the mean is one. In the limit, the Stein-like estimator is the mean.

There is also a mistake in the derivation of the Stein-like estimator in Appendix I. The quantity c is treated as a constant in all the integrals—but at the end "... we chose $c \geq (|\vec{p} - \bar{x}| + 1)^2$ ...," i.e., c depends on $\bar{x}$.

ISO's Empirical Bayes Credibility Procedure

Mr. Buck says that the Morris-Van Slyke Empirical Bayes credibility procedure is "based on 'Stein estimation'." This is not entirely accurate. While the Morris-Van Slyke procedure is similar to, and to an extent suggested by, the Stein procedure, it is developed in an Empirical Bayes framework. In fact, Efron and Morris, in [4], show that the Stein estimator itself can be developed as an Empirical Bayes estimator.

Mr. Buck states that the adapted Morris-Van Slyke procedure with the $3/k$ factor is biased upwards and that, because the tests included groups where the expected class loss ratios trended up or down over time, the results were slanted in favor of the $3/k$ factor. While the testing included the "residual trend" case, the original testing was done on the no residual trend case. For instance, in [5], page 79 ff., among other things, the adapted Morris-Van Slyke procedure is tested on simulated data against the same procedure *without* the $3/k$ factor where no residual trend is in effect. The with-$3/k$ procedure does better than the without-$3/k$ procedure in 86 out of 110 cases.

For instance, Table 1 reproduces the results given in [5] of simulated consecutive reviews for six different groups of simulation parameters. The error ("premium weighted test statistic," which is defined as the premium-weighted sum over all classes of the squared difference between the class loss ratio after the rate change and the expected loss ratio, see [6] p. II-15) is shown for the first through fifth reviews after the implementation of the new credibility procedure. Each entry is the average of 21 independent simulations.

As a footnote, group 4 was constructed with a very low original between-variance. This is why the procedure without the $3/k$ correction did better—lower credibilities were called for.

Mr. Buck says that Empirical Bayes credibility procedures using the $3/k$ correction factor are biased. The adapted Morris-Van Slyke procedure is biased, but not due to the $3/k$ factor. The bias is caused by logical constraints imposed on the variance estimators and there are good Empirical Bayesian reasons for these constraints.

The credibility formula depends on using an estimate of between-variance (parameter variance) in the denominator. Even though the between-variance estimator is unbiased, the credibility is not unbiased, because the credibility is not a *linear* function of the between-variance. To correct this, the indicated credibility is adjusted as follows:

$$Z^{\text{Adjusted}} = \frac{k-3}{k} Z^{\text{Indicated}} + \frac{3}{k}$$

where k is the number of classes. The derivation of this bias correction is given in [6].

At this point, we have an unbiased estimate of the credibility. Technically, the procedure is only unbiased for a highly restrictive set of assumptions. But, even without these assumptions, the bias correction is in the right direction.

Unfortunately, this unbiasedness depends on allowing the estimate of the between-variance to be negative. While the explanation of this [6] is complicated, the reader can see this intuitively by looking at the above equation and trying to imagine what value of $Z^{\text{Indicated}}$ is necessary for a class whose "true" credibility is less than $3/k$.

A priori, negative between-variances are impossible. Restricting the estimate of the between-variance to be non-negative leads to the minimum credibility of $3/k$ that Mr. Buck mentions. Unfortunately, restricting the estimate of the

TABLE 1
COMPARISON OF MORRIS-VAN SLYKE PROCEDURE WITH AND WITHOUT THE $3/k$ FACTOR
(ERRORS $\times 10^{-3}$)

Group	Review	Error With $3/k$	Error Without $3/k$	% Reduction in Error With $3/k$
1	1	2.30	2.43	5.3
	2	2.11	2.39	11.7
	3	1.84	2.25	18.2
	4	1.66	1.97	15.7
	5	1.47	1.69	13.0
2	1	1.49	1.58	5.7
	2	1.43	1.53	6.5
	3	1.38	1.52	9.2
	4	1.25	1.49	16.1
	5	1.14	1.23	7.3
3	1	2.61	2.67	2.2
	2	2.40	2.67	10.1
	3	2.18	2.62	16.8
	4	1.96	2.53	22.5
	5	1.66	2.11	21.3
4	1	0.72	0.66	−9.1
	2	0.83	0.65	−27.7
	3	1.03	0.74	−39.2
	4	1.16	0.73	−58.9
	5	1.19	0.72	−65.3
5	1	2.59	2.55	−1.6
	2	2.26	2.51	10.0
	3	2.09	2.37	11.8
	4	1.89	2.23	15.2
	5	1.77	2.21	19.9
6	1	3.10	3.40	8.8
	2	3.04	3.43	11.4
	3	3.03	3.32	8.7
	4	3.10	3.21	3.4
	5	2.62	2.99	12.4

between-variance also biases the procedure. Because the bias is a non-linear function of the between-variance, it cannot be corrected for all values of the between-variance. The bias is less in cases where there are many classes (k is large) and the underlying between-variance is high (so the probability of a negative between-variance estimate is low). These two conditions tend to hold in ISO's Products review.

As a practical matter, we do not recommend applying Empirical Bayes procedures if k is less than six. At least for the current generation of Empirical Bayes procedures, there simply aren't enough degrees of freedom to get a good estimate of the between-variance when there are five or fewer classes.

We do not regard current Empirical Bayes credibility procedures as the final word. Empirical Bayes procedures are, in general, the best credibility procedures we've seen thoroughly tested. However, other families of techniques offer hope for greater accuracy, particularly when data are classified by more than one rating variable (e.g., class and territory). As estimates of loss costs improve, insurers, insured, and society as a whole receive benefits. By focusing interest on improving credibility procedures, Jim Buck has performed a valuable service.

REFERENCES

[1] James E. Buck, "On Stein Estimators: 'Inadmissibility' of Admissibility as a Criterion for Selecting Estimators," *PCAS* LXXII, 1985.

[2] W. James and C. Stein, "Estimation with Quadratic Loss," *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability,* Vol. 1, University of California Press, Berkeley, 1961, pp. 361–379.

[3] Thomas S. Ferguson, *Mathematical Statistics: A Decision-Theoretic Approach,* New York, Academic Press, 1967.

[4] Bradley Efron and Carl Morris, "Stein's Estimation Rule and Its Competitors—An Empirical Bayes Approach," *JASA*, March 1973, Volume 68, Number 341.

[5] Minutes, ISO Credibility Subcommittee Meeting of December 13, 1979.

[6] Insurance Services Office, Report of the Credibility Subcommittee: Development and Testing of Empirical Bayes Credibility Procedures for Classification Ratemaking, September, 1980.

AN ESTIMATE OF STATISTICAL VARIATION IN DEVELOPMENT FACTOR METHODS

ROGER M. HAYNE

Abstract

This paper explores some properties of the lognormal distribution. It is possible that these properties can provide information not only regarding the variability of age-to-age development factors but also regarding that of age-to-ultimate factors, if the actuary is willing to assume that these factors are lognormally distributed. Considered are problems of parameter estimation and uncertainty under the assumption of independence of the age-to-age factors. Some results are generalized by weakening the independence hypothesis, and a method of parameter estimation with missing observations is presented. This paper is intended as a starting point, indicating useful results if factors are assumed to be lognormally distributed. Still an open question is the specific situations where such is the case.

1. INTRODUCTION

Development factor techniques have long been in the casualty actuary's repertoire of projection methods and are used extensively in both ratemaking and in the estimation of loss reserves. There are, however, numerous sources of variability in such results. For this reason an actuary often applies several techniques to obtain several estimates of ultimate losses. The actuary then selects a "best estimate" which reflects his or her best judgment of the amount of those losses.

The complex interactions in the data and the influence of non-random events (such as changes in claims practices) add to the variability inherent in the results of any projection technique. This makes it difficult to assess whether random fluctuations alone can be responsible for a range of estimates provided by various techniques or whether the various methods detect different patterns actually present in the data, or whether some combination of the two exists.

This paper will not present a loss projection technique but rather will explore properties of the lognormal distribution which will allow for some estimation of the statistical variability inherent in development factor projections if certain specific assumptions are satisfied. This can be useful in judging the differences among several projection techniques. For example, if wide fluctuations can be expected in projections (evidenced by wide confidence intervals), then variations in projections using different methods can be expected. If, on the other hand, there is little evidence of statistical variability and if the results of two methods are "far" apart, there is reason for further investigation to determine the cause of such differences.

Since the objective here is to study variability, the results depend on the underlying probability distributions and not on the particular age-to-age factors selected in practice. Thus, the resulting estimates of variability can be used as a measure of a "range of reasonableness" for various development factor selections and projections of other methods.

There are several useful properties of the lognormal probability distribution which motivated its choice as the statistical model here. First, the lognormal is defined for positive values of the random variable (development factors, except for extreme situations, are positive). Next, the distribution is skewed to the right but retains positive probabilities for large factors. This also has intuitive appeal for development factors which can be very large and experience anomolous swings. A third, and most useful, property of the lognormal distribution which suggested its consideration is its reproductive property. As will be stated in more precise form and greater generality below, the product of two lognormal random variables is, under certain assumptions, itself lognormal. In addition, the parameters of the product distribution can be determined easily from those of the two distributions. In terms of development factors this allows inferences regarding the age-to-ultimate factors to be made based on assumptions on the age-to-age factors.

The purpose of this paper is to explore the consequences of the assumption that the development factors are lognormally distributed. As stated above, the lognormal was chosen due to its useful properties and not on the basis of empirical data. Just as the normal distribution has been used to derive approximations in other areas of statistical work, it is possible that the lognormal model here may provide useful approximations in practice.

2. NOTATION AND BASIC CONCEPTS

This paper will deal with development factor techniques (see, for example, [1]). For this purpose, let $L_{i,j}$ denote data for exposure period i valued j valuation periods from the start of the exposure period i. Exhibit 1 gives a few of the possible choices for each of the parameters. For simplicity, $L_{i,j}$ will generally be referred to here as incurred losses for accident year i valued j years from the beginning of year i. This is by no means an attempt to limit the results shown here.

Let D_j denote the factor to develop losses valued at j years into losses valued at $j + 1$ years (commonly called the age-to-age development factor). Thus, if the loss data strictly followed this model, the following relation would hold:

$$L_{i,j+1} = D_j L_{i,j} \tag{2.1}$$

Let D_j^* denote the factor to develop losses from age j years to their ultimate value (commonly called the age-to-ultimate development factor). From the above definition of D_j the following formula holds:

$$D_j^* = \prod_{k=j}^{\infty} D_k \tag{2.2}$$

Extensive use will be made here of the lognormal probability distribution which depends on two parameters, denoted here by μ and σ^2. As used in this paper the probability density function is defined as:

$$f(x) = \exp\{-[(\ln(x) - \mu)^2]/2\sigma^2\}/x\sigma\sqrt{2\pi} \tag{2.3}$$

Here $\ln(x)$ denotes the natural logarithm and $\exp(x)$ its inverse. This distribution has been used rather extensively in actuarial work especially in the modeling of size-of-loss distributions (see, for example, [2], [3] and [4]) and has many useful properties.

In particular, if the random variable X is lognormally distributed with parameters μ and σ^2, then the random variable $Y = \ln(X)$ is normally distributed with mean μ and variance σ^2. This fact, and reference to tables of normal probabilities, allows for easy calculation of confidence intervals for this distribution.

Probably of greatest use here is the following theorem (see p. 11 of [5]):

<u>THEOREM 2.1</u>

If $\{X_j\}$ is a sequence of independent variables, where X_j is lognormally distributed with parameters μ_j and σ_j^2, $\{b_j\}$ a sequence of constants and $c = \exp(a)$ a positive constant, then provided $\Sigma b_j\mu_j$ and $\Sigma\sigma_j^2 b_j^2$ both converge, the product

$$c \prod X_j^{b_j} \tag{2.4}$$

is lognormally distributed with parameters $a + \Sigma b_j\mu_j$ and $\Sigma b_j^2\sigma_j^2$.

This theorem thus gives rise to the primary result used in this paper. In particular:

COROLLARY

If: (1) each age-to-age development factor D_j is lognormally distributed with parameters μ_j and σ_j^2 ($j = 1, 2, 3, \ldots$),

(2) all age-to-age development factors are independent, and

(3) $\sum_{j=1}^{\infty} \mu_j$ and $\sum_{j=1}^{\infty} \sigma_j^2$ both converge;

then each age-to-ultimate development factor D_j^* is lognormally distributed with parameters

$$\sum_{k=j}^{\infty} \mu_k \text{ and } \sum_{k=j}^{\infty} \sigma_k^2 \tag{2.5}$$

In most applications the third assumption above is fulfilled. Usually it is assumed that loss development stops after some finite point in time so that $\mu_j = \sigma_j^2 = 0$ for j sufficiently large.

3. SIMPLIFIED EXAMPLE

As an example of an application of these results assume that there is no development after four years (i.e. $\mu_j = \sigma_j^2 = 0$ for $j > 3$), that the age-to-age development factors D_1, D_2, and D_3 are each lognormally distributed with known hypothetical parameters given in the top half of Exhibit 2, and that all D_1, D_2, and D_3 are independent. Then the age-to-ultimate factors D_1^*, D_2^*, and D_3^* are all lognormally distributed with parameters as shown in the bottom part of Exhibit 2.

These parameters then allow for the calculation of various percentiles for the distributions of the age-to-age and age-to-ultimate development factors. To this end, let t denote the p ($0 < p < 1$) quantile of a standard normal random variable Z. That is, t satisfies the equation

$$\mathrm{P}(Z < t) = p \tag{3.1}$$

Since D_j is assumed to be lognormally distributed with parameters μ_j and σ_j^2 the following formula holds:

$$P(D_j < \exp\{\mu_j + t\sigma_j\}) = p. \tag{3.2}$$

Formula (3.2) then allows the computation of percentiles for the age-to-age factors using μ_j, σ_j and various percentiles of the normal distribution. Some examples are presented in the top part of Exhibit 3. For example, $\exp(.175 + .674 \times \sqrt{.075}) = 1.433$, and so forth. Similar examples for the age-to-ultimate factors are shown on the bottom of that exhibit.

4. REFINEMENTS

Under the assumption of lognormality, this procedure provides a means to estimate the *statistical* uncertainty inherent in the development factor method. This method assumes the parameters μ_j and σ_j^2 are known. Yet to be addressed, however, is the uncertainty surrounding μ_j and σ_j^2. In actual practice μ_j and σ_j^2 are not known for certain. Most often the only source of knowledge lies in the historical development factors themselves.

Assume here that there are n_j accident years of incurred loss data available valued at year j and year $j + 1$ and that there are k such periods of development under consideration. Thus $L_{i,j}$ and $L_{i,j+1}$ are defined for $i = 1, 2, 3, \ldots, n_j$ and $j = 1, 2, 3, \ldots, k$. Assume further that the historical development factors at age j, defined by

$$d_{i,j} = L_{i,j+1}/L_{i,j} \tag{4.1}$$

form a random sample from a lognormal distribution of unknown parameters μ_j and σ_j^2. Moreover, define the statistics:

$$Y_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \ln(d_{i,j}) \tag{4.2}$$

$$S_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} (\ln(d_{i,j}) - Y_j)^2 \tag{4.3}$$

It follows that Y_j and S_j^2 are the maximum likelihood estimators of μ_j and σ_j^2 respectively (see [5], p. 39) but, as in the normal analog, S_j^2 is biased. However, the statistic

$$V_j^2 = \frac{n_j}{n_j - 1} S_j^2 \tag{4.4}$$

is an unbiased and minimum variance estimator for σ_j^2, though it is no longer a maximum likelihood estimator (see [6], p. 165). Using these statistics, confidence intervals for μ_j and σ_j^2 can be obtained. This follows from the fact that, under these assumptions of lognormality and independence, for each j, $\ln(d_{i,j})$ form a sample from a normal distribution and thus

$$\frac{Y_j - \mu_j}{V_j/\sqrt{n_j}} \text{ has a } t \text{ distribution with } n_j - 1 \text{ degrees of freedom, and} \tag{4.5}$$

$$\frac{(n_j - 1)V_j^2}{\sigma_j^2} \text{ has a chi-squared distribution with } n_j - 1 \text{ degrees of freedom.} \tag{4.6}$$

Exhibit 4 shows a hypothetical development factor triangle which is generated using lognormally distributed random numbers. Since it is assumed that each column represents a random sample from a lognormal distribution, Y_j and V_j^2 provide estimators for μ_j and σ_j^2, respectively, for each value of j. In addition, the above observations regarding the distributions of Y_j and V_j^2 lead to the confidence intervals for μ_j and σ_j^2 given on the bottom of that exhibit. This information is helpful in estimating the degree of parameter uncertainty contained in the various age-to-age estimates. It cannot, however, be easily extended in general to the age-to-ultimate factors without additional assumptions, usually made about σ_j^2.

Since $\ln(d_{i,j})$ are normally distributed with mean μ_j and variance σ_j^2, the Y_j values are normally distributed with mean μ_j and variance σ_j^2/n_j. Moreover, any sum, such as $Y_1 + Y_2 + \ldots + Y_k$, is also normally distributed, in this case with mean $\mu_1 + \mu_2 + \ldots + \mu_k$ and variance $\sigma_1^2/n_1 + \sigma_2^2/n_2 + \ldots + \sigma_k^2/n_k$. If σ_1, $\sigma_2, \ldots, \sigma_k$ are all known then the normal distribution can be used to obtain a confidence interval for $\mu_1 + \mu_2 + \ldots + \mu_k$ of the form

$$Y_1 + \ldots + Y_k \pm t\sqrt{\sigma_1^2/n_1 + \ldots + \sigma_k^2/n_k} \tag{4.7}$$

where t is the selected percentile for the standard normal distribution.

Normal distribution theory also provides results in the case where $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \ldots = \sigma_k^2 = \sigma^2$ but all are unknown. The obvious generalizations can be made in this case; however, it is quite unlikely that this would occur in development factor applications. The author is not aware of further statistics which do not need a restrictive assumption such as those on σ_1^2 through σ_k^2 above and which can be used to obtain estimates of parameter variability.

Exhibit 5 provides an example of an application of the first of these assumptions, that the values of $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$ are all known. Here they are assumed equal to the corresponding estimates V_j^2. The second set of assumptions, that all the σ_j^2 are equal, is not applied to this data. The fact that the 90% confidence interval for σ_1^2 does not intersect any of the confidence intervals for the remaining σ_j^2 leaves the validity of this assumption open to serious question in this case.

5. RELAXATION OF INDEPENDENCE CONDITIONS

In the results presented this far, independence has been a necessary condition. The principal result in Theorem 2.1, however, is a special case of a more general theorem where the independence assumption can be replaced with one of multivariate lognormality. For this, some additional notation is necessary.

Denote by $\mathbf{R}^{m\times n}$ the set of matrices with m rows and n columns, having real entries. Following Aitchison and Brown (see [5], p. 11) a random variable $\vec{X} \in \mathbf{R}^{n\times 1}$ is said to have a multivariate lognormal distribution with parameters $\vec{\mu} \in \mathbf{R}^{n\times 1}$ and $\boldsymbol{\Sigma} \in \mathbf{R}^{n\times n}$ if the variable $\vec{Y} = \ln(\vec{X}) = (\ln(X_1), \ldots, \ln(X_n))'$ has a multivariate normal distribution with mean vector $\vec{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. It is assumed that $\boldsymbol{\Sigma}$ is symmetric (i.e. $\sigma_{i,j} = \sigma_{j,i}$) and positive definite, thus assuring that its inverse, $\boldsymbol{\Sigma}^{-1}$, exists. If $\boldsymbol{A}$ is a matrix, denote its transpose by $\boldsymbol{A}'$. The following result then holds:

<u>THEOREM 5.1</u>

If the age-to-age development factors $\vec{D} = (D_1, \ldots, D_n)' \in \mathbf{R}^{n\times 1}$ have a multivariate lognormal distribution with parameters $\vec{\mu} = (\mu_1, \ldots, \mu_n)' \in \mathbf{R}^{n\times 1}$ and $\boldsymbol{\Sigma} = (\sigma_{i,j}) \in \mathbf{R}^{n\times n}$, symmetric and positive definite, then, each age-to-ultimate factor

$$D_j^* = \prod_{k=j}^{n} D_k, \quad j = 1, 2, \ldots, n \tag{5.1}$$

is lognormally distributed (with a single variate lognormal distribution) with parameters given by:

$$\sum_{k=j}^{n} \mu_k \text{ and } \sum_{i,k=j}^{n} \sigma_{i,k} \tag{5.2}$$

Proof: By definition $\vec{Y} = \ln(\vec{D})$ is normally distributed with parameters $\vec{\mu}$ and $\boldsymbol{\Sigma}$.

By a well known result in multivariate normal analysis (see [6], p. 383), the sum $Y_j + Y_{j+1} + \ldots + Y_n$ is normally distributed with mean and variance given respectively by the parameters in (5.2). Since $Y_j + Y_{j+1} + \ldots + Y_n = \ln(D_j) + \ln(D_{j+1}) + \ldots + \ln(D_n) = \ln(D_j \times D_{j+1} \times \ldots \times D_n) = \ln(D_j^*)$, it follows that $\ln(D_j^*)$ is normally distributed and thus D_j^* is lognormally distributed. The parameters for that distribution are then given by the sums in (5.2). This completes the proof.

From multivariate normal theory ([6], p. 382), each of the above Y_j is normally distributed with mean μ_j and variance $\sigma_{j,j}$. Thus, each D_j has a lognormal distribution with parameters μ_j and $\sigma_{j,j}$. Hence, once $\vec{\mu}$ and Σ are known, confidence intervals for the D_j can be determined as in the case when independence is assumed. Similarly, confidence intervals for the D_j^* can be determined.

Parameter estimation, however, is not as simply generalized. The author is unaware of any method to estimate the parameters $\vec{\mu}$ and Σ in the general case based on the usual triangular form of historical development factors arrays.

If $d_{i,j}$ denotes the historical age-to-age factor for accident year i from age j to age $j + 1$ and the collection of such factors is based on $n + 1$ years of annual experience, ending in the current year, then $d_{i,j}$ is not defined if $i + j$ exceeds $n + 1$. Thus, the usual estimators for $\vec{\mu}$ and Σ, which would require data for all allowable i,j values, cannot be applied. If it is assumed that the age-to-age factors are independent for $j \geq m$ for some m then the previously stated results apply to each column for which $j \geq m$.

If, now, $m \leq (n + 1)/2$ then the array of factors $d_{i,j}$, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$; $i + j \leq n + 1$ will have at least m observations in each column. In this case the results of Bhargava [7] are applicable. In that paper, Bhargava derives maximum likelihood estimators for $\vec{\mu}$ and Σ for normal distributions based on samples with data missing in particular patterns. The available data from the first m columns of a development triangle form such an array if $m < (n + 1)/2$.

Following Bhargava, set

$$\mu_k = \nu_k + \sum_{j=1}^{k-1} \beta_j^{(k)} \mu_j \tag{5.3}$$

$$\sigma_{i,k} = \sigma_{k,i} = \sum_{j=1}^{k-1} \beta_j^{(k)} \sigma_{i,j};\ i = 1, 2, \ldots, k-1 \tag{5.4}$$

$$\sigma_{k,k} = \sigma_{k(0)}^2 + \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} \beta_i^{(k)} \beta_j^{(k)} \sigma_{i,j} \tag{5.5}$$

Given the parameters $\vec{\mu}$ and $\mathbf{\Sigma}$, the equations in (5.4) form a set of $k - 1$ linear equations in $k - 1$ unknowns, $\beta_1^{(k)}$, $\beta_2^{(k)}$, ..., $\beta_{k-1}^{(k)}$. Once these values are determined, ν_k and $\sigma_{k(\theta)}^2$ can be found from (5.3) and (5.5), respectively. Conversely, given ν_k, $\vec{\beta}^{(k)}$ and $\sigma_{k(\theta)}^2$, these equations give the parameters $\vec{\mu}$ and $\mathbf{\Sigma}$. Bhargava then determined maximum likelihood estimates for the parameters ν_k, $\vec{\beta}^{(k)}$, and $\sigma_{k(\theta)}^2$, and, using (5.3), (5.4) and (5.5), derived maximum likelihood estimates for $\vec{\mu}$ and $\mathbf{\Sigma}$.

To state those results some further notation is necessary. Suppose that $\vec{y}_1$, $\vec{y}_2$, $\vec{y}_3$, ..., $\vec{y}_n$ are n independent observations from a population that has a multivariate normal distribution with m variates with $m \leq (n + 1)/2$. The sample will satisfy Bhargava's definition of a monotone sample if observations for the i^{th} variate are available in $\vec{y}_1$, $\vec{y}_2$, ..., $\vec{y}_{n+1-i}$. Since $m \leq (n + 1)/2$, there will be at least m complete vectors. Note, this merely formalizes the situation that exists in a development factor matrix showing annual development for $n + 1$ accident years if the vector y_j is thought of as the first m elements of the j^{th} row. Though independence of the various age-to-age factors is no longer assumed, independence of the rows (accident year observations) is.

Given this sample, define the matrix $\mathbf{y}_{(1,k-1)}$ as the matrix composed of a column of 1's, followed by the first $k - 1$ elements of the observations $\vec{y}_1$, $\vec{y}_2$, ..., $\vec{y}_{n+k-1}$. This is a matrix with a column of 1's followed by the first $k - 1$ columns of the largest matrix containing observations for all of the first k variates in the data triangle. Let $\vec{y}_{(k)}$ denote the column matrix composed of the $n + k - 1$ observations of the k^{th} variate.

With this notation, Bhargava presents the following result:

<u>THEOREM 5.2</u>

Assume that $\vec{\mu} \in \mathbf{R}^{m\times 1}$, $\mathbf{\Sigma} \in \mathbf{R}^{m\times m}$ is symmetric and positive definite, and that $\vec{y}_1$, $\vec{y}_2$, ..., $\vec{y}_n$ is an independent, monotone sample from the multivariate normal population with mean vector $\vec{\mu}$ and covariance matrix $\mathbf{\Sigma}$. If ν_k, $\vec{\beta}^{(k)}$ and $\sigma_{k(\theta)}^2$ are defined as in (5.3), (5.4) and (5.5) then the maximum likelihood estimators for ν_k, $\vec{\beta}^{(k)}$ and $\sigma_{k(\theta)}^2$ are given by:

$$(\hat{\nu}_k, \hat{\vec{\beta}}^{(k)})' = (\mathbf{y}'_{(1,k-1)}\mathbf{y}_{(1,k-1)})^{-1}\mathbf{y}'_{(1,k-1)}\vec{y}_{(k)} \tag{5.6}$$

$$(n + 1 - k)\hat{\sigma}_{k(\theta)}^2 = \sum_{j=1}^{n+1-k} (y_{j,k} - \hat{\nu}_k - \sum_{i=1}^{k-1} \hat{\beta}_i^{(k)} Y_{j,i})^2 \tag{5.7}$$

$$= \vec{y}'_{(k)}(\mathbf{I} - \vec{y}_{(k)}(\mathbf{y}'_{(1,k-1)}\mathbf{y}_{(1,k-1)})^{-1}\vec{y}_{(k)})\vec{y}_{(k)}$$

Though not immediately obvious, the value of $\hat{\nu}_k$ is the constant coefficient of the least squares multiple linear regression of $\vec{y}_{(k)}$ against the first $k - 1$ variates, based on the observations in the first $n + 1 - k$ rows of the matrix. Similarly $\hat{\vec{\beta}}^{(k)}$ are the coefficients of each of the first $k - 1$ variates. Finally, $\hat{\sigma}^2_{k(\theta)}$ is the conditional variance of the fit. It denotes the amount of variance which remains unexplained by the regression. Thus, estimation of the ν_k, $\vec{\beta}^{(k)}$ and $\sigma^2_{k(\theta)}$ can be accomplished using multiple regression for $k = 2, 3, \ldots, m$ while $\hat{\nu}_1$ and $\hat{\sigma}^2_{1(\theta)}$ are the sample mean and variance of the first column of the matrix.

If, now, it is assumed that the first m columns of the development factor matrix have a multivariate lognormal distribution with parameters $\vec{\mu} \in \mathbf{R}^{m \times 1}$ and $\mathbf{\Sigma} \in \mathbf{R}^{m \times m}$, symmetric and positive definite, then the above procedures, applied to $y_{i,j} = \ln(d_{i,j})$, will produce the maximum likelihood estimates for ν_k, $\vec{\beta}^{(k)}$ and $\sigma^2_{k(\theta)}$ and thus $\vec{\mu}$ and $\mathbf{\Sigma}$. This result follows since, under this assumption, the values of $\ln(d_{i,j})$ form a monotone sample from a multivariate normal distribution.

As an example of these methods, Exhibit 6 shows the estimators $\hat{\nu}_k$, $\hat{\vec{\beta}}^{(k)}$ and $\hat{\sigma}^2_{k(\theta)}$ along with estimators for $\vec{\mu}$ and $\mathbf{\Sigma}$ based on the hypothetical development factors shown in Exhibit 4. In this case, the matrix is 6×6 ($n = 6$). Here it is assumed that $m = 3$, that there is no development after the sixth year, that is, $D_6 = D_7 = D_8 = \ldots = 1$, and that D_4 through D_6 are all independent and independent of the first three factors. Finally, it is assumed that D_1 through D_3 have a multivariate lognormal distribution with parameters $\vec{\mu} \in \mathbf{R}^{3 \times 1}$ and $\mathbf{\Sigma} \in \mathbf{R}^{3 \times 3}$, symmetric and positive definite.

If it is assumed that the parameters of the distributions for D_1 through D_6 are equal to their maximum likelihood estimates then Exhibit 7 shows the resulting confidence intervals for the resulting age-to-age and age-to-ultimate factors. The intervals for D_1 through D_3 are based on the fact that the natural logarithm of each is normal with mean μ_k and variance $\sigma_{k,k}$. This exhibit also compares the intervals with those derived under the assumption of independence, assuming that the parameters equal the values of Y_i and V_i^2 in Exhibit 4.

Correlation among the age-to-age development factors will, of course, impact the marginal variance of any given factor and also the variance of the resulting age-to-ultimate factors. If the various age-to-age factors are positively correlated then the resulting age-to-ultimate factors will have wider variation (and hence wider confidence intervals) when derived using the multivariate estimation than those derived using the assumption of independence. Conversely, if there is negative correlation among the age-to-age factors then the resulting estimates of the age-to-ultimate factors derived using the multivariate techniques will probably have less variation than those derived assuming independence. This follows from the variance formula given in (5.2).

Parameter uncertainty is not as straightforward as in the completely independent case. Though the author does not know the distributions of the various estimates, Bhargava does provide likelihood ratio tests to test the hypothesis H_0: $\vec{\mu} = 0$ against H: $\vec{\mu} \in \mathbf{R}^{m \times 1}$. Those results are sufficiently complex, however, that they will not be presented here. One interesting result mentioned by Bhargava, however, is that the distribution of $(n + 1 - k)\ \hat{\sigma}^2_{k(\theta)}/\sigma^2_{k(\theta)}$, given the observations in the first $n + 1 - k$ rows and k columns, has a chi-squared distribution with $n + 1 - 2k$ degrees of freedom.

6. OBSERVATIONS

The usefulness of any theory lies in the nearness of the hypothesis of that theory to reality. In this regard, the first question that comes to mind is that of the lognormality of development factors in actual practice. The lognormal distribution has the benefit of being defined for only positive values of the random variable and does not impose an upper bound on those values. This corresponds to development factors which are generally positive and are unbounded. In practice, statistical tests such as the Kolmogorov-Smirnov Test as presented by Gary Patrik ([8], p. 65) may help in assessing the validity of the assumption of lognormality.

The independence of the various columns may also be able to be tested. Since $\ln(d_{i,j})$ are assumed to be normally distributed for $i = 1, 2, \ldots, n_j$ a test based on the sample correlation coefficient between the natural logarithms of two columns may give some insight as to the validity of this assumption. In addition, these results require the independence of the development factors of a given age from each other. Again, usual statistical tests, applied to the natural logarithms of the development factors, may be helpful in assessing the validity of this hypothesis. In any case, in actual applications, actuarial judgment is required to detect any patterns which may appear in the data (for example, correlation between columns, trend in age-to-age factors over time, etc.). Such patterns often add to the variation apparent in the factors. Actuarial judgment will thus decrease statistical variability.

In order to compare the results of various loss projection methods, the age-to-ultimate development factors must be multiplied by the appropriate loss amount to date. To draw statistical conclusions about the resulting loss projections, the age-to-ultimate factors must then be assumed to be independent from the amounts recorded to date.

The methods presented here can only provide estimates of *statistical* variability under very explicit assumptions. They should be looked on as providing a "range of reasonableness" of loss projections, based on such variability, rather than as a confidence interval about any specific ultimate loss estimate. In the latter case, the actuary's judgment is used to narrow a large range of possible choices (as presented by the historical development factors) in light of his or her knowledge of the underlying data.

7. CONCLUSIONS AND BEGINNINGS

This paper is presented more as an opening to further investigation than as a definitive solution to a problem. The model selected for study, that of development factor projection, is one of the simplest of the projection techniques in use by casualty actuaries and any actuary with experience in applying this technique knows its limitations and weaknesses. Hopefully the results presented here help in assessing the variability inherent in this method.

The larger challenge still facing casualty actuaries is to devise estimates of the amount of variation to be expected in the more complex projection methods used. However, a precise estimate of variability inherent in an actuary's "best estimate" probably is not possible. Actuarial judgment used to interpret diverse results of various methods, in light of the actuary's knowledge of events that may impact the patterns to be expected in the data, cannot be statistically quantified. This judgment is usually the most important aspect of the estimation of ultimate losses, but any further insight that can be gained from these techniques can be helpful in forming that judgment.

REFERENCES

[1] D. Skurnick, "A Survey of Loss Reserving Methods," *PCAS* LX, 1973, p. 16.

[2] D. R. Bickerstaff, "Automobile Collision Deductibles and Repair Cost Groups: The Lognormal Model," *PCAS* LIX, 1972, p. 68.

[3] C. Hewitt, "Credibility for Severity," *PCAS* LVII, 1970, p. 148.

[4] R. J. Finger, "Estimating Pure Premiums by Layer—An Approach," *PCAS* LXIII, 1976, p. 34.

[5] J. Aitchison and J. A. C. Brown, *The Lognormal Distribution*, Cambridge University Press, 1969.

[6] R. V. Hogg and A. T. Craig, *Introduction to Mathematical Statistics*, Third Edition, Macmillan Publishing Co., New York, 1970.

[7] R. Bhargava, "Multivariate Tests of Hypothesis with Incomplete Data," Technical Report No. 3, Applied Mathematics and Statistics Laboratories, Stanford University, California, August 1962.

[8] G. Patrik, "Estimating Casualty Insurance Loss Amount Distributions," *PCAS* LXVII, 1980, p. 57.

EXHIBIT 1

Some Choices as to Data Arrangement for Development Factor Techniques

Type of Data (L)	Aggregation Type
Paid Losses	Report Period
Incurred Losses	Accident Period
Paid (Closed) Claim Counts	Policy Period
Reported Claim Counts	

Exposure Period (i)	Valuation Period (j)
Year	Year
Half-year	Half-year
Quarter	Quarter

EXHIBIT 2

Simplified Example Development Factors

Parameters for the Age-to-Age Factors D_1, D_2, and D_3

j	μ_j	σ_j^2
1	0.175	0.075
2	0.045	0.005
3	0.005	0.001

Parameters for the Age-to-Ultimate Factors D_1^*, D_2^*, and D_3^*

j	μ_j^*	σ_j^{*2}
1	0.225	0.081
2	0.050	0.006
3	0.005	0.001

EXHIBIT 3

EXAMPLE PERCENTILES BASED ON SIMPLIFIED DEVELOPMENT FACTOR DATA

	Percentile				
Age	10% ($t = -1.282$)	25% ($t = -0.674$)	50% ($t = 0.000$)	75% ($t = 0.674$)	90% ($t = 1.282$)
	Percentiles for Age-to-Age Development Factors				
1	0.839	0.990	1.191	1.433	1.692
2	0.955	0.997	1.046	1.097	1.145
3	0.965	0.984	1.005	1.027	1.047
	Percentiles for Age-to-Ultimate Development Factors				
1	0.869	1.034	1.252	1.517	1.804
2	0.952	0.998	1.051	1.108	1.161
3	0.965	0.984	1.005	1.027	1.047

EXHIBIT 4

Hypothetical Development Factors

Accident Year	Stage of Development (years) 2/1	3/2	4/3	5/4	6/5	7/6
1	1.932	1.036	1.009	1.003	1.002	1.000
2	1.975	1.038	1.013	1.006	1.001	
3	1.809	1.041	1.011	1.005		
4	1.954	1.043	1.009			
5	1.997	1.035				
6	1.932					
Estimators:						
Y	6.59×10^{-1}	3.79×10^{-2}	1.04×10^{-2}	4.66×10^{-3}	1.50×10^{-3}	
V^2	1.21×10^{-3}	1.05×10^{-5}	3.59×10^{-6}	2.31×10^{-6}	4.99×10^{-7}	
90% Confidence Intervals for:						
μ	6.30×10^{-1} to 6.88×10^{-1}	3.48×10^{-2} to 4.10×10^{-2}	8.20×10^{-3} to 1.26×10^{-2}	2.10×10^{-3} to 7.22×10^{-3}	-1.65×10^{-3} to 4.65×10^{-3}	
σ^2	5.45×10^{-4} to 5.26×10^{-3}	4.41×10^{-6} to 5.89×10^{-5}	1.38×10^{-6} to 3.06×10^{-5}	7.71×10^{-7} to 4.49×10^{-5}	1.30×10^{-7} to 1.25×10^{-4}	

EXHIBIT 5

EXAMPLE CONFIDENCE INTERVALS FOR THE PARAMETERS OF THE AGE-TO-ULTIMATE DEVELOPMENT FACTORS

Assumptions:

$\sigma_1^2 = 1.21 \times 10^{-3}$ $\sigma_2^2 = 1.05 \times 10^{-5}$

$\sigma_3^2 = 3.59 \times 10^{-6}$ $\sigma_4^2 = 2.31 \times 10^{-6}$

$\sigma_5^2 = 4.99 \times 10^{-7}$

$$\sigma_j^2 = \mu_j = 0 \text{ for } j \geq 6$$

Age	Estimator for μ_j^*	90% Confidence Interval for for μ_j^*	
1	0.714	0.690	to 0.737
2	0.0554	0.0511	to 0.0578
3	0.0166	0.0143	to 0.0189
4	0.00616	0.00450	to 0.00782
5	0.00150	0.000678	to 0.00232

EXHIBIT 6

Example Parameter Estimates Using Multivariate Sample Estimation

Variable:		D_1	D_2	D_3
Estimators:				
	$\hat{\nu}_k$	6.59×10^{-1}	6.31×10^{-2}	1.53×10^{-2}
	$\hat{\boldsymbol{\beta}}^{(k)}$		-3.84×10^{-2}	-1.00×10^{-3} -1.09×10^{-1}
	$\hat{\sigma}^2_{k(\theta)}$	1.01×10^{-3}	6.60×10^{-6}	2.61×10^{-6}
	$\hat{\boldsymbol{\mu}}$	6.59×10^{-1}	3.79×10^{-2}	1.05×10^{-2}
	$\hat{\boldsymbol{\Sigma}}$	1.01×10^{-3} -3.86×10^{-5} 3.20×10^{-6}	-3.86×10^{-5} 8.08×10^{-6} -8.42×10^{-7}	3.20×10^{-6} -8.42×10^{-7} 2.70×10^{-6}

EXHIBIT 7

Example 90% Confidence Intervals Based on Multivariate Parameter Estimation

Intervals For Age-to-Age Factors:

D_1	D_2	D_3	D_4	D_5
Assuming Independence:				
1.825	1.033	1.007	1.002	1.000
to	to	to	to	to
2.046	1.044	1.014	1.007	1.003
Using Multivariate Estimators:				
1.834	1.034	1.008	1.002	1.000
to	to	to	to	to
2.036	1.044	1.013	1.007	1.003

Intervals for Age-to-Ultimate Factors:

D_1^*	D_2^*	D_3^*	D_4^*	D_5^*
Assuming Independence:				
1.926	1.049	1.013	1.003	1.000
to	to	to	to	to
2.161	1.063	1.021	1.009	1.003
Using Multivariate Estimators:				
1.904	1.050	1.013	1.003	1.000
to	to	to	to	to
2.147	1.062	1.021	1.009	1.003

A PRACTICAL GUIDE TO THE SINGLE PARAMETER PARETO DISTRIBUTION

STEPHEN W. PHILBRICK

Abstract

The actuarial literature has discussed several candidates for size-of-loss distributions—log normal, Weibull, multi-parameter Pareto, gamma, as well as others. However, despite the demonstrated success of these distributions, there is a dependence on techniques such as empirical data, judgment, or at times some unwieldy formulae. This suggests that there may be a need for a size-of-loss distribution that is relatively easy to apply in practice.

The one-parameter Pareto is an example of such a distribution. Its use may be restricted to the tail of a distribution, but it is easy to apply. The formulae for the mean, variance, and the variance of the aggregate loss distributions are simple in form and may be used as quick approximations in many cases.

I. INTRODUCTION

"The ultimate goal of model-building is either as a tool for communicating . . . or for predicting and making decisions . . ."

—William S. Jewell

Although model-building is common to many branches of science, there are important distinctions among the properties of various models. The laws of physics such as Newton's laws are attempts at mathematical models of reality. These efforts have been particularly successful because the major forces at work are few in number and often constant over time and position. Although technically there are many forces involved in, say, the movement of the planetary bodies, the dominant force of gravity dwarfs the other forces such as friction. Models can be developed based solely on the properties of the gravitational force which describe the motion of the planets to a very high degree of precision. In these situations, it is common to find mathematical models with few parameters that are highly accurate models of reality. It is appropriate, even if technically incorrect, to speak of the search for *the* correct mathematical model.

In the social sciences the situation is quite different. The forces involved in economics, for example, are numerous and usually not constant over time. Many forces exist that have the same order of magnitude; hence they cannot be ignored. Furthermore, in the social sciences it is often more difficult to do controlled experiments where one force is allowed to vary while all others are held constant. For these reasons, it is less appropriate to think of a search for *the* model in the social sciences than in the physical sciences. Although we might talk about such a concept theoretically, the practical reality is that any parameter-based model that completely describes an existing situation will require so many parameters as to make it unusable. In these situations, model-building requires a trade-off between accuracy and practicality.

Thus, the question "What is the appropriate loss distribution?" does not have a unique answer. It depends on the intended use of the distribution and the available data.

The question requires a cost-benefit analysis. Different models will have various costs related to:

- Mathematical complexity,
- Availability of computer/calculator software routines,
- Computer processing time requirements,
- Conceptual simplicity (ease of explanation to others), and
- Availability and accuracy of data.

Generally speaking, increasing sophistication of the model produces more accurate results. The selection of an appropriate model for a particular problem requires deciding whether the increased accuracy of the more complex model justifies the increased costs associated with it. Furthermore, in many situations the available data may be sparse or subject to inaccuracies. In these instances, a simple model may be preferred because the accuracy of results will not be materially improved by the use of a more complex model.

For example, suppose an actuary is trying to solve a typical risk management problem: the projection of losses for an individual risk. A common procedure in this analysis is to separate the projections of the large or excess losses from the projections of the more stable primary portions of the losses. Several characteristics of this situation make a simple model particularly appropriate.

- The projection of the limited losses may be accomplished without the need for a specific size-of-loss distribution. The moments of the data, using a frequency/severity or total loss approach, may be sufficient for a reasonable projection. It will then be necessary to fit a model only above a particular loss amount. Fitting a distribution to only a portion of the range will reduce the required complexity of the model.

- Inaccuracy of estimates of expected losses arises from a number of sources. Two major ones are:
  —Oversimplified models, and
  —Misestimated parameters.
  In a situation involving an individual risk, the number of large losses used to estimate the parameters will typically be less than the number involved in an insurance company or industry analysis. The errors arising from the sample size may dominate those arising from a less complex model. As a consequence, the simplicity of the less complex model may be preferred because the possible loss of accuracy is more than offset by the benefits of a simpler model.

- There may be a need to explain the loss projection process to people without extensive actuarial or statistical training. Although techniques should not, in general, be dictated by the sophistication of the audience, if competing models produce almost identical results, the ease of explanation of one may be an important consideration.

The remainder of this paper will be organized as follows:

- Section II—A discussion of the way distributions are depicted. An alternative to the "standard" representation will be presented.

- Section III—A discussion of the basic properties of the single parameter Pareto distribution.

- Section IV—Various methods of parameter estimation using empirical data.

- Section V—The results of trend on losses when a Pareto distribution is assumed.

- Section VI—A method to simulate Pareto losses.

- Section VII—Specific applications using a Pareto distribution.

The author would like to acknowledge the help of several people, including Jerry Miccolis and Claudia Leme, who reviewed early drafts of this paper; Kurt Reichle and John Yonkunas, who provided many helpful suggestions and encouragement; and Jerry Jurschak, who contributed some of the concepts embodied in this paper.

II. SIZE-OF-LOSS REPRESENTATIONS

Most texts on probability and statistics portray distributions (as well as density functions) using similar conventions. That is, the horizontal axis represents the value of the observations and the vertical axis represents the relative frequency (for density functions) or the cumulative frequency (for distributions). It is clear, from a mathematical point of view, that this choice is arbitrary. The axes could be switched without violating or changing any of the statistical concepts.

Despite the almost universal acceptance of this "standard" representation, the alternative representation turns out to be a clearer choice for size-of-loss distributions in some situations. The reason for this preference is that this representation can be developed in a "natural" way and will allow a number of concepts, such as loss limitation (truncation and censorship), to be applied in a more intuitive fashion. Appendix D contains a more detailed discussion of reasons for preferring this orientation.

In the following discussion, we will develop a size-of-loss representation where the y-axis is the horizontal axis and the x-axis is the vertical axis. We will refer to this representation as the "alternative" representation.

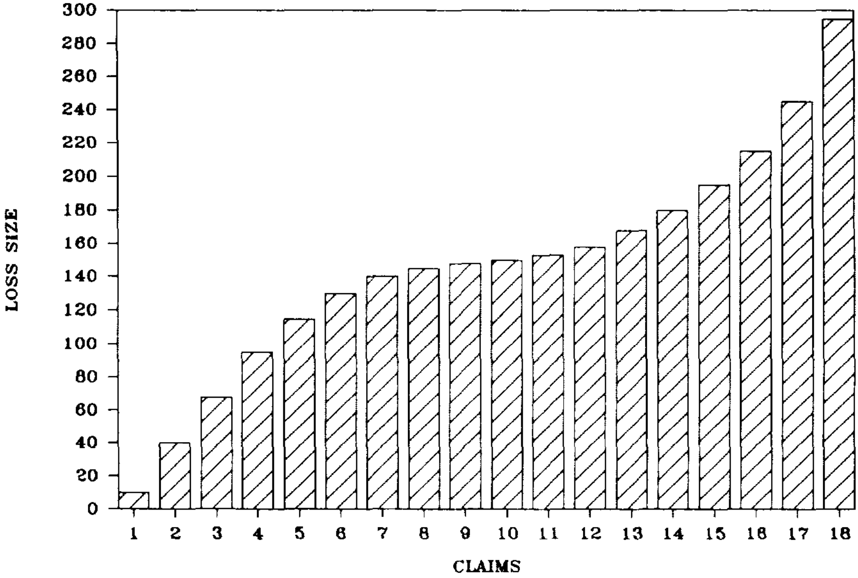
Because we have switched the axes rather than redefined them, the definitions of x and y will remain unchanged; that is, x refers to loss amounts and y refers to cumulative frequency.

Discrete Case

Consider a set of n losses from some arbitrary size-of-loss distribution where each loss has size S_i, $i = 1, 2, \ldots, n$. Represent each loss by a rectangle with

FIGURE 1

SIZE-OF-LOSS DISTRIBUTION



width one and height S_i. Arrange these losses from smallest to largest, each perpendicular to the y-axis. Figure 1 displays a typical example of such a procedure.

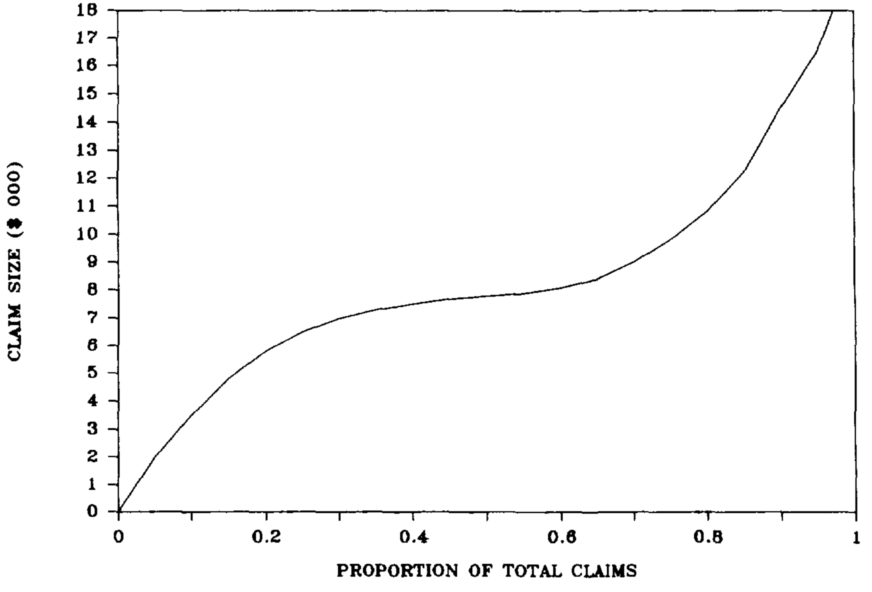
Define $G(y)$ to be the curve represented by the tops of each of the rectangles. Then, $G(y) = S_i$ for $i-1 < y \leq i$. Note that the interpretation of the random variable Y is the number of losses less than or equal to $G(y)$ (for integral values of y).

Continuous Case

When we consider the continuous case, the width of each loss is dy. The value of y ranges from 0 to 1 and represents the percentage of losses less than or equal to $G(y)$. A typical example would look like Figure 2.

FIGURE 2

CLAIM SIZE DISTRIBUTION



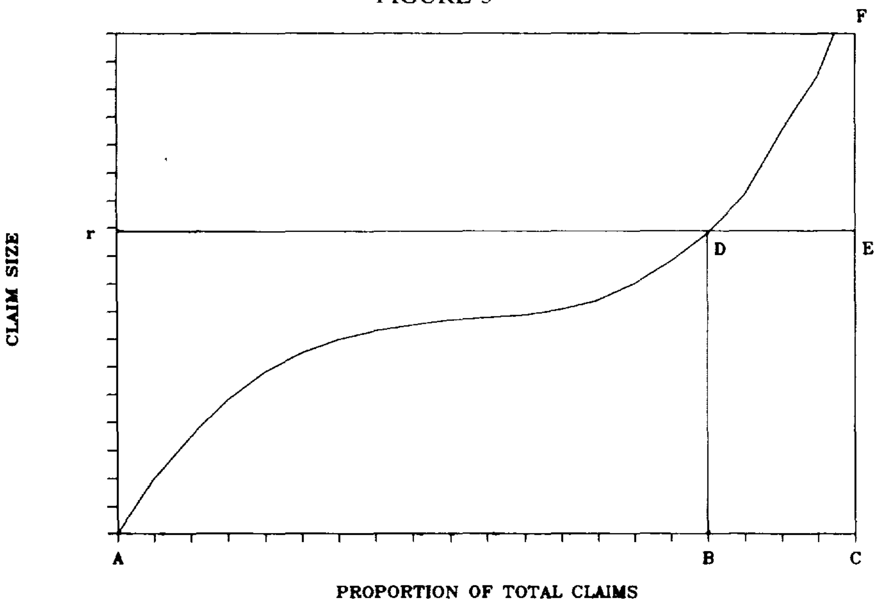
From this point on, the continuous version of the representation will be used. However, some of the concepts may be better understood if the original motivation of this representation is recalled, namely "stacking" individual losses along the y-axis.

When we work with a set of losses (whether actual or theoretical), we generally wish to partition these losses in some way. The most common partitions are "large" versus "small" and primary versus excess. These partitions can be graphically represented by defining areas under the curve $X = G(y)$.[1]

Generally, we will indicate the losses of interest by defining one or more straight lines on the graph (see Figure 3). When we define an area by a pair of lines parallel to the horizontal axis, we will refer to these losses as a "layer" of losses. Alternatively, if we use a pair of lines that are parallel to the vertical

[1] A third type of partition is described in Hewitt and Lefkowitz [H1]. That partition cannot be handled in this way.

FIGURE 3



axis, these will be referred to as "interval" losses. In this case, we are referring to those losses that correspond to an interval specified on the (horizontal) axis.

We could define the areas we are interested in by directly writing the integral over the appropriate limits. However, we can keep the notational complexity to a minimum if we adopt symbols for the areas that will be used most often.

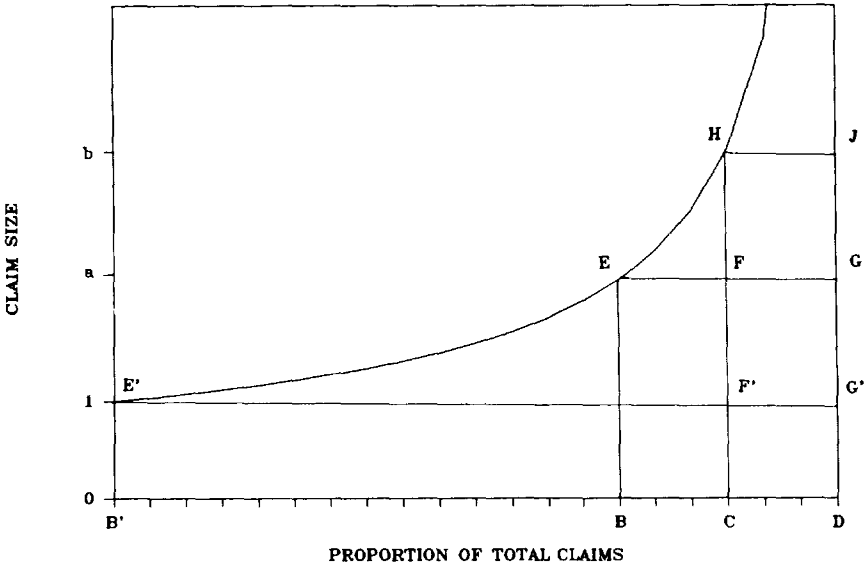
Given loss amount, r, we define[2]:

	Verbal	Mathematical
$T(r)$ —	The average claim size of all losses less than or equal to r; i.e., losses are *truncated* at amount r.	$\dfrac{\int_0^r x\, dF(x)}{\int_0^r dF(x)}$

[2] The reader may note that the notation used here is not entirely consistent with that developed in a discussion of LaRose [L1]. The notation developed by LaRose calculates claim amounts as percentages of the *average* claim. Unfortunately, the average claim size is not always well-defined, so a more general notation is required.

FIGURE 4

LARGE LOSS DISTRIBUTION



$C(r)$ — The average claim size of all claims where the amount of each claim is limited to size r; i.e., losses are *censored* at amount r.

$$\int_0^r x \, dF(x) + r[1 - F(r)]$$

For example, if r is \$100,000, then $T(\$100{,}000)$ represents the average of all losses less than or equal to \$100,000. In Figure 3, this average would be represented by the ratio of the area bounded by ABD divided by the number of claims in the interval. The quantity $C(\$100{,}000)$ is the average of all claims where amounts greater than \$100,000 are capped or limited to \$100,000.

The preceding discussion is applicable to any size-of-loss distribution. Figure 3 applies to any distribution that is used to model the entire range of losses. In the remainder of this paper, we will work with the tail of a loss distribution that is applicable to "large" losses. Consequently, we will truncate the loss distribution at some value and remove each loss less than that value. A typical distribution representing the remaining "large" losses is shown in Figure 4.

Figure 4 is derived from Figure 3 by truncating the loss distribution at loss amount r. Physically, we remove the portion of the graph to the left of the vertical line BD, then renormalize our axes so that the y-axis is the cumulative percentage of the "large" losses, that is, losses greater than or equal to r. It should be emphasized that Figure 3 is not drawn to scale for typical loss distributions. If we select a lower limit r such as \$25,000, the cumulative probability that a claim is less than \$25,000 (which is represented by point B) is typically in excess of 90%. We will work only with the large losses in the remainder of this paper, so Figure 4 is the important figure to keep in mind.

III. BASIC PROPERTIES OF THE SINGLE PARAMETER PARETO

The Pareto distribution as described in Johnson and Kotz [J1] has cumulative distribution function:

$$F(x) = 1 - \left(\frac{k}{x}\right)^a \qquad k > 0;\ a > 0;\ x \geqslant k$$

This is also known as the "Pareto distribution of the first kind." Strictly speaking, this distribution has two parameters, k and a. In general, both k and a may be estimated from the data. However, the verbal definition of k is the lower bound of the data in question. Although there may be situations where this value must be estimated, in virtually all insurance applications this value will be selected in advance. The typical insurance application will be to model losses whose value is in excess of some pre-selected size, such as \$25,000 or \$100,000.

Furthermore, if we "normalize" our losses, that is, divide each loss by the selected lower bound, then the normalized lower bound is 1, and the parameter does not need to be stated explicitly. Finally, we will use q as the parameter, rather than a, to be consistent with ISO usage (ISO [I1], p. 34). The distribution can then be written as:

$$F(x) = 1 - x^{-q} \tag{1}$$

and the density function is

$$f(x) = qx^{-(q+1)} \tag{2}$$

This is the distribution that will be discussed in the remainder of this paper.[3] Typical values for q can range from .7 to 2.0, although values outside this

[3] See Appendix C for a discussion of alternative forms of the Pareto.

range are possible. A typical value for q of property losses is 1.0, while a typical value for casualty losses is 1.5 (based upon empirical evidence). Note that a *low* value of q corresponds to a distribution with high severity. Fire may not be thought of as a line with high severity, but that is because there are so many very small claims. Considering only larger claims, e.g., claims greater than $25,000, fire claims have a fairly "thick" tail. The density function for a Pareto with parameter $q = 1.5$ is shown in Figure 5; the corresponding c.d.f. is shown in Figure 6.

If we "flip" the x- and y-axes of the cumulative distribution, we will produce Figure 4. Note that the curve intersects the x-axis at $x=1$, because we have normalized the losses. The curve is asymptotic to $y=1$. As mentioned earlier, we can visualize the area under the curve as being made up of thin vertical rectangles whose height corresponds to the size of loss. Thus the total area under the curve represents the total losses, and the losses associated with various retentions or policy limits can be described by different areas under the curve.

The distribution as shown in Figure 4 is based upon the assumption that the lower limit is 1 and the expected frequency of claims greater than or equal to this value is also 1. Formulae will be derived under these assumptions. The necessary conversions to real problems are simple and straightforward. Examples of conversions will be given in most cases. Although it may seem awkward at first to work with a normalized distribution, it will soon become very natural. The motivation for using the normalized distribution should become clear when we analyze the losses contained in different layers.

Unlimited Claims

The formula for the average claim size is as follows:

$$\text{Unlimited Mean Claim Size} = \frac{q}{q-1} \qquad q > 1 \tag{3}$$

Note that this formula also represents the expected total losses when the expected frequency is 1 (assuming independence of frequency and severity).

If the data being analyzed has a lower limit of $\$K$ per claim, then the mean size in "real" dollars is:

$$\text{Unlimited Mean Claim Size} = K\left(\frac{q}{q-1}\right) \tag{3a}$$

FIGURE 5

PARETO DENSITY

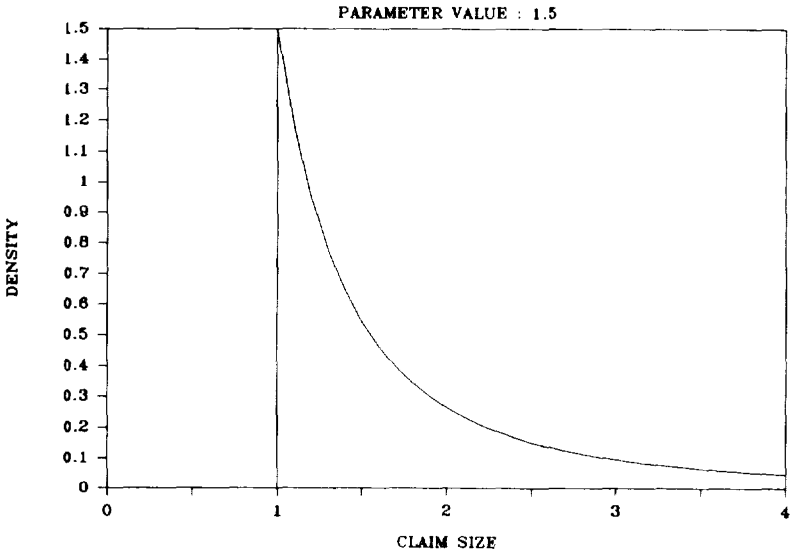
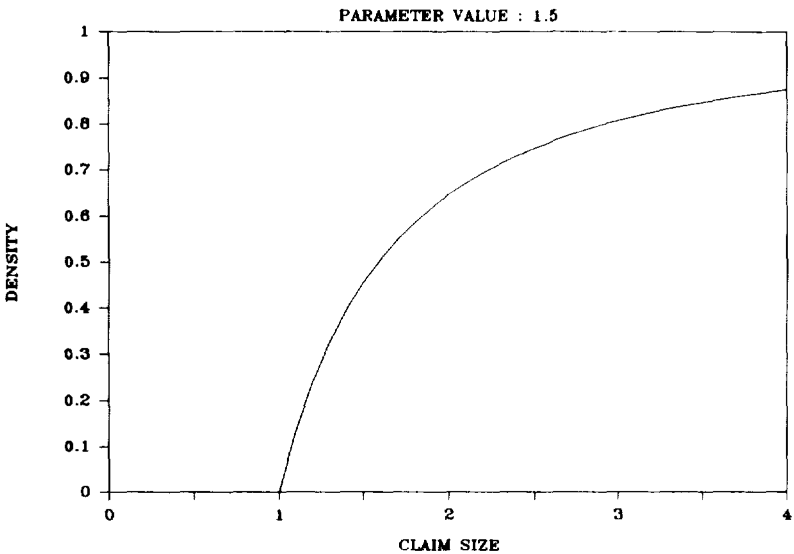


FIGURE 6

PARETO CUMULATIVE DISTRIBUTION



If we anticipate n claims greater than or equal to $\$K$ per claim, then the total expected losses are:

$$\text{Unlimited Expected Losses} = n\,K\left(\frac{q}{q-1}\right) \tag{3b}$$

For example, suppose we are analyzing claims where the lower limit is \$25,000. That is, all claims are greater than or equal to \$25,000. After normalizing our losses (dividing each by \$25,000) we conclude that a parameter value of $q = 1.5$ is appropriate. (A later section will discuss parameter estimation.) Then the normalized gross mean claim size is $1.5/(1.5 - 1) = 3$. In terms of "real" dollars, where $K = \$25{,}000$, the gross mean claim size is $3 \times \$25{,}000 = \$75{,}000$.

If we expect 7 claims to exceed \$25,000, then our gross expected losses are $7 \times \$75{,}000 = \$525{,}000$. (Again, it should be remembered that we are analyzing the large claims only. The expected losses arising from claims less than \$25,000 are assumed to be estimated separately.)

We may wish to calculate the net losses, for example, if we have a \$25,000 deductible. The formula for the net mean claim size is derived from the gross mean claim size simply by subtracting 1:

$$\text{Net Mean Claim Size}\left(\frac{q}{q-1}\right) - 1 = \frac{1}{q-1} \qquad q > 1 \tag{4}$$

The conversion to "real" dollars and total losses follows the same approach as above. For example, with $q = 1.5$, $K = \$25{,}000$ and a frequency of 7, the net expected losses above \$25,000 are $7 \times \$25{,}000 \times 1/(1.5 - 1) = \$350{,}000$.

Censored Claims

If we impose an upper limit (such as a policy limit) with value b, then the formula for the average loss limited to $\$b$ per claim is:

$$C(b) = \frac{q - b^{1-q}}{q-1} \qquad q \neq 1 \tag{5}$$

If $q = 1$ we can calculate the formula using L'Hôpital's Rule:

$$C(b) = 1 + \ln b \qquad q = 1 \tag{6}$$

In the case where we want net losses, we can simply subtract 1 from each formula. Note especially, when $q = 1$ the average loss with upper limit b is simply $\ln b$.

Continuing our previous example (with $q = 1.5$ and lower limit of \$25,000), if we impose an upper limit of \$500,000, then $b = 20 \times (500{,}000/25{,}000)$. The average claim whose value is greater than \$25,000 but limited to \$500,000 can be calculated using (5):

$$\frac{1.5 - 20^{1-1.5}}{1.5 - 1} = 2.553$$

In "real" dollars, the average claim is 2.553× \$25,000 = \$63,820 (calculations here and subsequently are performed without rounding at intermediate steps). If we are pricing reinsurance for the net layer (\$475,000 xs \$25,000), then we would subtract 1 first to calculate the net claim size: 1.553 × \$25,000 = \$38,820. Assuming we expect 7 claims over \$25,000, the expected losses in the layer are 7 × \$38,820 = \$271,738.

Truncated Claims

The situation described above (with an upper *censorship* limit) arises naturally in practice because of the existence of policy limits and the way companies commonly write excess of loss reinsurance. Another way to limit losses is to *truncate* the losses at some value. This means that losses in excess of the truncation point are "thrown away," rather than simply "capped" at the limit. Note that this is different from censorship in two ways:

1. More dollars are removed when losses are truncated at a value because the entire loss above the limit is removed.
2. Truncation affects the frequency. Censorship removes the excess portion of a claim, but does not affect the number of claims. Truncation removes the *entire* claim, so the formulae for average values must reflect the reduced claim count.

The concept of truncation arises rarely in property/casualty policy language (with the rare exception of franchise deductibles). However, the concept may arise in the analysis of experience. For example, it might be appropriate to separate losses into large versus small (rather than primary versus excess). In this case, the limit chosen to distinguish between large and small losses will be a *truncation* point, rather than a *censorship* point.

The formula for the average claim size with lower value 1 and truncation point b is:

$$T(b) = \frac{q(1 - b^{1-q})}{(q - 1)(1 - b^{-q})} \qquad q \neq 1 \qquad (7)$$

The comparable formula for the case $q = 1$ is:

$$T(b) = \frac{b \ln b}{b - 1} \qquad q = 1 \qquad (8)$$

Continuing our example, suppose we are interested in the losses larger than \$25,000 but ignoring all losses greater than \$500,000 (rather than including the first \$500,000 of those claims greater than \$500,000). With $q = 1.5$ and $b = 20$, the average claim size is calculated with (7):

$$\frac{1.5\,(1 - 20^{-.5})}{.5\,(1 - 20^{-1.5})} = 2.356$$

In "real" dollars, the average claim is 2.356 × \$25,000 = \$58,888. If we expect 7 claims over \$25,000 we can calculate the total dollars for the interval. Given 7 claims over \$25,000, we expect $7 \times F(20) = 7 \times .9888 = 6.922$ claims in the interval between \$25,000 and \$500,000. We have already calculated the average of those claims, so we multiply the frequency by the average claim size to yield the total dollars: 6.922 × \$58,888 = \$407,606.

The formula for the average truncated size follows directly from the definition of truncated claims given earlier. However, we can simplify the formula and reduce the amount of calculation by adopting a slightly non-standard convention. Note that, in our example, one of the terms in the denominator for the average truncated claim size is $1 - b^{-q}$. Note also that the number of claims in the *interval* is calculated by multiplying the expected frequency (above the lower limit) by $F(b)$ which is $1 - b^{-q}$. Obviously, these terms cancel out when the total dollars in the interval are calculated.

Define $T'(b)$ to be the average claim size, where the denominator is not just the claims in the interval, but the number of claims above the lower limit. In other words, use the same denominator as in the censored situation. The motivation is two-fold:

(1) The formulae will be simpler.
(2) It is more likely that we will have an estimate of the total number of claims above a limit than that we will have an estimate of the number of claims in an interval.

The formula for the revised "average" truncated claims size is:

$$T'(b) = \frac{q(1 - b^{1-q})}{q - 1} \qquad q \neq 1 \qquad (9)$$

When $q = 1$, the formula simplifies to:

$$T'(b) = \ln b \qquad q = 1 \tag{10}$$

Redoing the example above, the "average" claim size is calculated using:

$$\frac{1.5\,(1 - 20^{-.5})}{.5} = 2.329$$

In "real" dollars, the "average" claim size is $2.329 \times \$25{,}000 = \$58{,}229$. Multiplying this by the number of claims expected over \$25,000 yields $7 \times \$58{,}229 = \$407{,}606$.

In summary, if we are interested in the true average claim size, we use formula (7) or (8). However, if the calculation of the average claim size is simply an intermediate step in the calculation of the total dollars, we may prefer to use alternative formula (9) or (10).

Next we will look at the excess portion of the distribution. In this case, we are interested in the total losses or average claim size of claims greater than some limit b. In terms of Figure 4, the area of interest is bounded by HJK. Rather than directly calculate the total losses and average losses in this layer, we will exploit a powerful property of the Pareto distribution. If we renormalize the excess portion by dividing each loss by b and dividing the excess frequency by $1 - F(b)$, the resulting distribution will have a shape identical to that in Figure 4. (This renormalization is the result of a scale change to both axes. For more discussion of scale changes, see Venter [V1].) Thus, we may use the formulae already calculated, although keeping careful track of the appropriate factors to convert back to "real" dollars.

The average gross claim size is still $q/(q - 1)$ and the average net claim size is $1/(q - 1)$. In terms of our first renormalization, the average gross claim size is $b\,(q/(q - 1))$ and in "real" dollars, the average is $bK\,(q/(q - 1))$. The total dollars involved in claims greater than b can be calculated by multiplying by the frequency of claims greater than b which is $1 - F(b) = b^{-q}$.

In practice this works out easier than the formulae would indicate. Continuing our example ($q = 1.5$, $K = \$25{,}000$, frequency over $\$25{,}000 = 7$), suppose we are interested in the losses in excess of \$100,000 per claim. We don't actually perform the renormalization; we simply use the formula for net average claim size ($1/(q - 1)$) and substitute $q = 1.5$ into the formula yielding a net claim size of 2. Multiply by \$100,000 (it isn't necessary to multiply first by \$25,000, then by 4) to produce the average net claim size of \$200,000. To calculate the total dollars, recall that the ratio of claims exceeding \$100,000 is calculated by using the cumulative distribution $1 - F(b) = b^{-q} = 4^{-1.5} = .125$. Multiply this by the expected frequency over \$25,000 of 7 yielding .875 claims expected to exceed \$100,000. Thus, the expected excess losses are $.875 \times \$200{,}000 = \$175{,}000$.

This concept is important, as it allows us to quickly calculate the total losses and average claim sizes for arbitrary layers and intervals. As another example, suppose we continue our assumption that losses over \$25,000 have a Pareto distribution with $q = 1.5$ and the expected frequency of claims over \$25,000 is 7. Suppose we are asked to analyze the layer between \$75,000 and \$187,500 (i.e., \$112,500 xs \$75,000). The first step is to calculate the value of b, which is simply $187{,}500/75{,}000 = 2.5$. We can use (5) to calculate the gross average (censored) claim sizes:

$$\frac{1.5 - 2.5^{-.5}}{.5} = 1.735$$

The net average claim size is .735 or $.735 \times \$75{,}000 = \$55{,}132$ in "real" dollars. The frequency of claims is $7 \times (1 - F(75{,}000/25{,}000)) \times F(187{,}500/75{,}000) = 7 \times (3^{-1.5}) \times (1 - 2.5^{-1.5}) = 7 \times (.192) \times (.747) = 1.006$, so the expected losses in the layer are $1.006 \times \$55{,}132 = \$55{,}482$.

Next, we will calculate the variance of the individual claim amounts as well as the total loss variance. The formulae shown above for expected values are sufficient for pricing on an expected value basis or some function of the expected value. However, there are methods of pricing that include risk loading based upon variance, as well as other risk theoretic analyses that require the calculation of variances. (See Gerber [G1] for a discussion of various pricing approaches.)

Again, this is one of the motivations for the use of the Pareto. The calculation of total loss variance is a fundamental issue in risk theory, yet the procedures necessary to calculate the variance generally involve complex formulae or, more likely, computerized estimation techniques. The formulae associated with the single parameter Pareto are often easy to evaluate and may provide, at the very least, a reasonable first approximation.

Recall that the variance can be calculated as the second moment minus the square of the mean. The formula for the n^{th} moment of the Pareto distribution with no upper limit is

$$n^{\text{th}} \text{ moment} = \frac{q}{q + n} \tag{11}$$

Thus, the second moment is $q/(q + 2)$ and the formula for the variance of a single claim is:

$$\text{Variance} = \left(\frac{q}{q-2}\right) - \left(\frac{q}{q-1}\right)^2 \qquad q > 2 \qquad (12)$$

Again, we have the problem that the variance is undefined for typical values of q. But if we restrict ourselves to reasonable upper limits, the variance will always be finite. If we impose upper limit b, then the variance of losses within the layer is:

$$\text{Variance} = \frac{q - 2b^{2-q}}{q-2} - \left(\frac{q - b^{1-q}}{q-1}\right)^2 \qquad \begin{matrix} q \neq 1 \\ q \neq 2 \end{matrix} \qquad (12a)$$

The formula simplifies in the cases where $q = 1$ or 2 as follows:

$$\text{Variance} = 2b - 1 - (1 + \ln b)^2 \qquad q = 1 \qquad (12b)$$

$$\text{Variance} = 1 + 2 \ln b - ((2b - 1)/b)^2 \qquad q = 2 \qquad (12c)$$

These formulae apply in either the net or unlimited layer cases.

To convert these results to "real" dollars, multiply by K^2 where K is the lower bound of the losses. It is important to realize that these formulae reflect only the variance associated with the loss severity. The total loss variance also reflects the variability of frequency, which will be covered shortly.

We will continue the example where the lower limit is \$25,000 and $q = 1.5$. As we have shown earlier, the gross mean claim size is 3 and the net mean claim size is 2 when no upper limit is imposed. However, the variance is not defined in this case. With an upper limit of \$500,000, $b = 20$ and the variance of a single claim is calculated by substituting into (12a) with $q = 1.5$ and $b = 20$. The result is 8.372. In "real" dollars, the variance is $8.372 \times (\$25{,}000)^2 = 5.23 \times 10^9$. This means that the standard deviation is \$72,335.

The claim size variance is rarely useful by itself. The major motivation for calculating this formula is because it is needed in the formula for the total loss variance. This refers to the variability of total losses, arising either from frequency or severity. The variance we will calculate is also sometimes called "process variance," because it relates to the possible variations in results arising from the loss causing process. This is to be distinguished from "parameter variance," which relates to the variations arising from the possibility that the parameters used differ from the "true" parameters. Parameter variance is beyond the scope of this paper.

Calculation of the total loss variance is necessary if a risk loading will be used that is a function of either the total loss variance or standard deviation. In addition, the variance can be used to specify percentiles of the total loss distribution using the Cornish-Fisher expansion [M1] or other techniques [L3]. For example, we may wish to determine the probability that total losses will exceed $1,000,000 when the expected losses are $600,000.

The general formula for the total loss variance is given in various sources including Mayerson, Jones and Bowers [M3]:

$$\sigma^2 = M_f\,\sigma_s^2 + M_s^2\,\sigma_f^2 \tag{13}$$

where M_f, σ_f^2, M_s, and σ_s^2 represent the mean and variance of the frequency and severity distributions respectively.

If we make the reasonable assumption that the claim frequency follows a Poisson distribution, then $M_f = \sigma_f^2$ and we can simplify (13):

$$\sigma^2 = M_f(\sigma_s^2 + M_s^2) \tag{14}$$

Again, recalling that the variance can be expressed as the second moment less the square of the mean, we note that the expression in parentheses above simplifies to the second moment of the severity. Thus, the total loss variance can be simply calculated as the product of the expected claim frequency and the second moment (mean of the squares) of the loss severity.

We have seen the formula for the second moment of the severity in the case of no upper limit earlier (11). In this case, the total loss variance is:

$$\sigma^2 = M_f\,\frac{q}{q+2} \qquad q > 2 \tag{15}$$

where M_f is the expected claim frequency.

We have seen earlier that the severity variance is the same in the case of the unlimited and net layers. This is not the case for the total loss variance. If we have upper censorship point b, the total loss variance for the unlimited layer is:

$$\sigma^2 = M_f\,\frac{q - 2b^{2-q}}{q-2} \qquad q \neq 2 \tag{16}$$

In the case of the net layer, the total loss variance is:

$$\sigma^2 = M_f\left\{\frac{q - 2b^{2-q}}{q-2} - 2\left(\frac{q - b^{1-q}}{q-1}\right) + 1\right\} \qquad \begin{matrix} q \neq 2 \\ q \neq 1 \end{matrix} \tag{17}$$

The expression in the parentheses may be recognized more quickly if we recall that $E[(X - 1)^2] = E[X^2] - 2\,E[X] + 1$. Formula (16) for the case where $q = 2$ is shown in Appendix A. As before, to convert the results to "real" dollars, we multiply by K^2 where K is the lower limit of the losses used to normalize the values.

If we have the truncated case, with truncation point b, the total loss variance for the unlimited layer:

$$\sigma^2 = M_f \frac{q\,(1 - b^{2-q})}{(q - 2)(1 - b^{-q})} \qquad q \neq 2 \qquad (18)$$

Note carefully: the definition of M_f in this case is the expected number of claims greater than the lower limit, not simply the number between the lower limit and b. The situation with truncation point b and a *net* layer is almost never seen in practice, so it will not be discussed.

Continuing our example, suppose we are pricing the losses in excess of \$25,000 but censored at \$500,000. As we have seen earlier, the expected losses in this layer are \$271,738 (assuming the expected number of claims is 7). Suppose we wish to add a risk loading that is a function of the total loss variance. We can calculate the total loss variance using (17). Substituting the parameters into the formula yields a variance of 75.48. In "real" terms, this is $75.48 \times \$25{,}000^2$. The standard deviation of this value is \$217,199. We won't go into methods for calculating a factor to multiply by the variance to arrive at an appropriate risk load, but, even without such methods, the total loss variance can be used to compare the relative risk on different treaties.

Finally, we note that the formulae derived in this paper are only applicable to the portion of losses above the selected lower limit. In practical situations, it is necessary to combine the results of the analysis of the large losses with the results of the analysis of the small losses. Clearly, the expected losses of the two portions of the analysis can simply be added together. The overall average claim cost can be calculated as the weighted average of the means of each portion, where the weights are the expected number of claims. The variances of the severity cannot be combined so easily, although, if one recalls that the second moments can be weighted by claim counts, the formula for the combined severity variance follows easily. If we assume a Poisson distribution for the frequency of the small losses, then the total loss variance of the small losses will be of the same form as the large losses, specifically, the mean claim frequency multiplied by the second moment of the severity, so the total loss variance of the entire distribution is simply the sum of the total loss variance of each portion.

IV. PARAMETER ESTIMATION

Numerous articles in the actuarial and statistical literature (e.g., Patrik [P1], p. 62.) discuss the attractive properties of the maximum likelihood estimate (MLE). However, the MLE is often difficult to calculate in practice.

One of the attractive properties of the Pareto distribution is the ease of calculation of the maximum likelihood estimate of the parameter. Consider a set of n losses, each greater than or equal to some value K, which are normalized by dividing each loss by K. Denote this set by $(X_i), i = 1,2, \ldots, n$. The MLE of q is

$$q = \frac{n}{\Sigma \ln X_i} \tag{19}$$

Note that an alternative formula is

$$q = \frac{n}{\ln \Pi X_i} \tag{20}$$

These formulae are equivalent, but the second is easier to calculate. Note also that the MLE of q is such that $e^{1/q}$ is the geometric mean of the X_i. If we use the 25 losses contained in Appendix B, the estimated parameter is $q = 25/26.16 = .955$.

Although the MLE has attractive properties and is easy to calculate, we will examine the use of alternative methods. Probably the most common method is matching of moments. We have shown that the mean of the unlimited Pareto distribution is $q/(q - 1)$. If this is equated to the sample mean of the values in Appendix B, we have

$$\frac{q}{q-1} = 6.202$$

$$q = 1.192$$

This value is not particularly close to the true value. The discrepancy arises, not because of the relatively small sample, but from the method itself. If the formula for the mean is examined, it will be clear that a value of 1.0 could never result. If the true value of the parameter of the distribution is 1.0 or smaller, the method of moments will always produce too high a result. Because in many situations the value of the parameter may be close to or even less than 1.0, the method of moments may not be an appropriate method.

Another method of parameter estimation is based on quantiles.[4] Using the formula for the c.d.f.,

$$F(x) = 1 - x^{-q}, \tag{21}$$

we can equate the sample values of $F(x)$ to their theoretical values. Although this method of estimation is somewhat less efficient[5] than MLE, it is much faster and may be used as a quick method for approximating the parameter when only a rough estimate is needed. In our example the median, or 13th largest loss, is $55,843 or 2.234 when normalized. Solving $.5 = 1 - 2.234^{-q}$ for q is straightforward yielding $q = 0.826$. If we look at the other two quartiles, which are approximately the 6th and 19th largest losses, we solve the equations

$$.25 = 1 - (1.311)^{-q}$$

$$.75 = 1 - (3.955)^{-q}$$

which yield estimates

$$q = 1.062$$

$$q = 1.008$$

A more important use of this method is when the individual claim sizes are not available (or not easily available), and only grouped statistics are available. Suppose that the losses in Appendix B had been incurred, but the only information was as follows:

Interval (000)	Frequency
25–100	20
100–1,000	4
1,000–∞	1

[4] Quantiles is the general term which includes the median, quartiles, and percentiles as special cases.

[5] For a discussion of efficiency, see Hoel, Port and Stone [H5], *Introduction to Statistical Theory*, page 16.

Using the information that 80% of the losses are no greater than $100,000, we solve the following:

$$.8 = 1 - 4^{-q}$$

yielding

$$q = 1.161$$

This estimate is remarkably good when one considers the limited information available.

Alternative methods of parameter estimation are discussed in Quandt [Q1].

To this point, in this section we have assumed that there is no upper limitation on the loss data either by an upper censorship point created by policy limits or an upper truncation limit where certain values may be missing.

There are several reasons for suspecting that actual data has some type of limitation. In the case of insurance company data, the losses may be censored due to reinsurance agreements. In some cases, gross losses are available, but in others only net losses may be available in a usable form. Even if the losses are gross to reinsurance, there may be limitations due to policy limits.[6] Most casualty coverages have policy limits.[7]

One of the advantages of working directly on an individual risk is that these limitations can be overcome. Although the primary source for data is usually insurance company records, it is usually possible to make the appropriate adjustments whenever losses have been limited.

This does not totally remove the problems of limitations. In the case of property insurance, there is an upper bound to the amount of loss, namely the total value of the property. There seems to be no useful upper bound to liability situations, but most actual data suggests that the tail of the Pareto is still somewhat too "thick" at extremely high loss amounts. In other words, the theoretical density at high loss amounts is larger than empirical experience tends to indicate. Rather than discard the Pareto, it is easier to postulate that the distribution is censored or truncated at some high, but finite, value. As we have seen earlier, any upper limitation (either censorship point or truncation point)

---

[6] A discussion of the impact of policy limits can be found in Patrik [P1].

[7] Exceptions include workers' compensation coverage A and no-fault PIP in some states.

will produce formulae for the mean claim size that are finite for all possible values of q.

If we assume a censorship point c, then the density function is unchanged between 1 and c but will have a mass point at c and will be zero for all values greater than c. Let $f(x)$ be the unlimited Pareto density, that is

$$f(x) = q\,x^{-(q+1)}$$

Let $f_c(x)$ be the density function censored at c. Then,

$$f_c(x) = f(x) \qquad 1 \le x < c$$

$$f_c(x) = \int_c^{\infty} f(x)dx \qquad x = c$$

$$f_c(x) = 0 \qquad x > c$$

If we wish to consider the distribution truncated (above) at t, then the density function at all points less than or equal to t will have to be proportionately increased so that the total area under the curve still equals one, and the new density function is zero for all values greater than t. Let $f_t(x)$ be the distribution truncated above at t. Then,

$$f_t(x) = \frac{f(x)}{\int_0^t f(x)dx} \qquad 1 \le x \le t$$

$$f_t(x) = 0 \qquad x > t$$

Assume that we have n losses, of which m are less than the censorship limit c and $n - m$ are equal to c. The maximum likelihood estimate is

$$q = \frac{n - m}{\sum^{n-m} \ln X_i + (m)\ln c}$$

Suppose we have the loss data in Appendix B except that each loss is censored at \$100,000. Then,

$$q = \frac{20}{13.104 + 5(1.386)} = .998$$

Note that the MLE approach produces the parameter of the unlimited distribution; censorship is handled through definition of the density function.

V. EFFECT OF TREND

One of the practical problems with fitting size-of-loss distributions is the proper way to handle adjustments for trend and development. With most distributions, inflation of losses will change one or more of the parameters. In Hogg and Klugman, [H2] page 180, there is a table that shows the parameters of various distributions after the application of a trend factor. In each case (including the Pareto and generalized Pareto), the parameters are changed due to inflation.

However, the parameter of the Pareto distribution in this paper is unchanged due to trend. This result appears counterintuitive. After all, each of the formulae for mean claim size is a function of the parameter. If the parameter is unchanged, then the estimated average claim sizes must be unchanged. This appears unreasonable for several reasons.

First, it is obvious that, under influence of trend, the overall average claim size increases. This is true, but note that the distribution in question does not apply to the entire range of losses. It is not simply *better* suited for modeling excess losses, it does not fit small losses well at all. The typical size-of-loss distribution starts out with a small frequency of very small losses, growing to a larger frequency of intermediate losses, then a decreasing frequency of larger losses. The maximum density for the Pareto is always at the leftmost value, and the density is always decreasing as we move to larger claim sizes. Thus, the fact that the overall average claim increases with trend is simply evidence that the single parameter Pareto is not likely to fit the entire range.

Second, it may be recalled that trend is assumed to have a *leveraged* effect on excess losses, where the Pareto is supposed to fit so well. This is true (see Miccolis [M1]), but the leveraged effect is on the total excess *dollars,* not necessarily on the *average* excess claim size. It may seem ironic, but the major effect of trend is to increase the *frequency* of an excess claim, rather than its severity. This may be more obvious if we recall that a size-of-loss distribution is, by definition, the distribution of the relative *frequencies* of various sizes of claims.

Third, and most imporant, a review of empirical excess average claim sizes will show that they have been increasing over time for most coverages. This point is conceded and is inconsistent with an assumption that a Pareto fits the entire excess distribution to infinity. As has been noted earlier, the Pareto has "too thick" a tail, and, in most practical applications, an upper bound should be used. If one looks at the average excess claim *with* a reasonable upper limit, the average claim size will *not* be materially increasing over time.

Because this point is important, we will explore it in more detail. Consider the losses contained in Appendix B. These have been generated from a Pareto distribution with $q = 1.0$. The appendix contains both the normalized values and the raw dollars, which indicate that the raw losses represent losses greater than or equal to \$25,000. If we calculate the MLE of these losses (assuming we did not know how they were generated), we would estimate the parameter to be .955. As can be verified, this value will produce average claim sizes for various layers of intervals (with reasonable upper bounds) that are reasonably close to the theoretical values. Specifically, we can use this parameter to estimate the average claim for layers or intervals where the lower limit is \$50,000. Thus, this parameter can be thought of as the appropriate parameter for the size-of-loss distribution for claims greater than \$50,000.

Suppose these losses were from year zero, and we wished to project losses for year n. Suppose further that the annual trend, $1 + i$, is such that $(1 + i)^n = 2.00$. If we were to trend each of our losses in Appendix B by this trend factor and use these losses to calculate a parameter to fit losses in excess of \$50,000, it should be obvious that the estimated parameter would still be exactly .955.

What may be less than obvious is the fact that this parameter can be used for losses in excess of \$25,000 in year n. This means that the losses between \$25,000 and \$50,000 in year n, which correspond to losses less than \$25,000 in year 0, must be distributed in such a way that the Pareto distribution will still fit the distribution above \$25,000 (to the upper limit) in year n.

As may be guessed, the requirement is that the Pareto distribution must fit the losses in year 0 as low as \$12,500 (\$25,000/2.00). In general, if we are using losses greater than K from year 0 to estimate a parameter to use in year n, we must assume that the Pareto distribution (with the same parameter) provides a reasonably good fit to losses in year 0 which are as small as $K/(1 + i)^n$. Experience has shown that this is typically true for casualty losses as low as \$5,000 to \$10,000 (higher for medical malpractice), so values of lower limits in the oldest year of experience greater than \$25,000 will typically work. Of course, it is prudent to check the fit at the lower end of the range if possible.

We have gone over this point in some detail because it leads to an extraordinary result: to calculate the MLE of the Pareto parameter, given individual losses greater than a single fixed value K arising from several different years, it is not necessary to adjust the losses for trend.

For example, suppose the following data are available:

1978	100,000,	150,000,	225,000		
1979	109,000,	140,000,	180,000,	240,000	
1980	105,000,	115,000,	170,000,	290,000	
1981	104,000,	121,000,	160,000,	200,000,	300,000

Suppose we are interested in projecting losses for 1984 and the annual trend, $1+i = 1.1$. Under typical methods of analysis, we would trend each of the losses to a common date. The trend factor for 1978 would be $(1.1)^6 = 1.77$. But if we did not have any data on losses less than \$100,000 for older years, we would have to use a lower limit of \$177,000. Several of the losses in more recent years would then have to be thrown out, because their trended value is less than \$177,000.

With the Pareto distribution, we can use all of the data points, *if* we have reason to believe that the Pareto distribution fits losses as low as \$100,000/1.77 in 1978. But note that if we assume that the Pareto will fit above \$100,000 in 1984, this is equivalent to assuming that it fits equally well above \$100,000/1.77 in 1978 (assuming trend affects all claim sizes approximately the same).

Of course, this will allow us to estimate the parameter of the distribution, which will allow us to calculate the *average* severities for 1984. This is only half the problem, as we also need to estimate the frequency of claims to arrive at estimates of the total loss dollars. We cannot simply use the raw historical frequencies of claims greater than \$100,000 to estimate our future frequency. We can, however, calculate an adjustment factor that will allow us to put each of the historical frequencies on a comparable basis.

The calculation of this factor can be shown most easily with a concrete example. Suppose we expect 10 claims greater than \$25,000 in year n, where $q = 1.5$. Recall the formula for the distribution is $F(x) = 1 - x^{-q}$. How many claims in year n are expected to exceed $1.1 \times \$25{,}000 = \$27{,}500$? We calculate this using the distribution function, $F(1.1) = 1 - 1.1^{-1.5} = .133$. This means 13.3% of the claims will be less than \$27,500, or 86.7% will be greater than \$27,500. Thus, we expect $10 \times .867 = 8.67$ claims greater than \$27,500. This means that we expect $1/.867 = 1.153$ claims over \$25,000 for every claim greater than \$27,500. If we now examine year $n - 1$, the \$27,500 claim in year n would be \$25,000 in year $n - 1$, and the \$25,000 claim would be $\$25{,}000/1.1 = \$22{,}727$ in year $n - 1$. Clearly, for every claim greater than \$25,000 in year $n - 1$, we would expect 1.153 claims greater than \$22,727.

So if we multiply the number of claims greater than \$25,000 in year $n - 1$ by 1.153, we have the best estimate of the number of claims greater than \$22,727 in year $n - 1$, which corresponds to the number of claims that, if trended, would exceed \$25,000 in year n.

Typically, the frequency of claims in each year will be related to an exposure base such as number of beds (hospital malpractice), number of employees (workers' compensation), etc. When using the Pareto distribution, the first step is to multiply the raw frequency of claims greater than the underlying limit by the adjustment factor, then divide through by the exposure. The resulting values may be averaged, or perhaps a regression analysis will be performed. (Note that inflation sensitive exposure bases such as sales or payroll must also be put on a comparable basis.) The adjustment factor for n years of trend at annual inflation rate $1 + i$ with parameter q is simply $(1 + i)^{nq}$. The following table displays the factors for various combinations of i and q. Each value in the table is the one-year adjustment factor.

	q				
$1 + i$	1.00	1.20	1.50	1.80	2.00
1.05	1.050	1.060	1.076	1.092	1.103
1.08	1.080	1.097	1.122	1.149	1.166
1.10	1.100	1.121	1.154	1.187	1.210
1.12	1.120	1.146	1.185	1.226	1.254
1.15	1.150	1.183	1.233	1.286	1.323

VI. SIMULATION OF LOSSES

One type of analysis frequently performed by actuaries involves Monte Carlo simulation of results based upon a particular model of the loss process. One advantage of this Pareto distribution is the ease with which it can be simulated.

One method for simulating values for a function involves inverting the cumulative distribution. This is not always possible with some functions, but it is particularly easy with the Pareto. The cumulative distibution is

$$F(X) = 1 - X^{-q}$$

Thus

$$F^{-1}(Y) = (1 - Y)^{-1/q}$$

where Y has the uniform distribution.

A moment's reflection will reveal that $(1 - Y)$ is symmetric when Y has the uniform distribution, so we can replace $1 - Y$ with Y. Thus, if we can generate a uniform random variable Y, then $Y^{-1/q}$ will have a Pareto distribution with parameter q.

Consequently, we find that even hand-held calculators, such as the HP-15, can be used to simulate Pareto losses. For example, the following values in the first column were generated from a calculator with a random number generator. The second column contains the normalized loss when $q = 1.5$, and the third column contains the "real" claim amount if the lower limit is \$25,000.

(1) Random Value from Uniform Distribution	(2) Normalized Pareto Value	(3) "Real" Dollars
.19875	2.93630	73,407
.73616	1.22655	30,664
.52174	1.54298	38,575
.97358	1.01801	25,450
.26635	2.41562	60,390
.54727	1.49462	37,366
.85879	1.10682	27,670
.31708	2.15058	53,764
.38295	1.89630	47,407
.23006	2.66341	66,585

VII. APPLICATIONS

In this section, we will discuss several applications of the Pareto distribution. In some of the cases, we will use actual data from published sources for two reasons: first, to demonstrate that this distribution works well with "real" data, and second, so that this distribution can be compared to those used in the original source of the data.

Application 1

Consider the OL&T BI claims for policy year 1976 contained in Appendix F of Patrik [P1]. We will fit the Pareto to losses greater than \$25,000. There are 90 losses in this exhibit. Individual losses are not shown, but the ranges are quite narrow, as they are \$5,000 ranges up to \$100,000, and \$10,000 thereafter. We can use the average claim size in the range as a reasonable proxy for the individual claim amounts (with wider ranges, we might need to make adjustments). The sum of the normalized logs (dividing each claim by \$25,000) is 81.2; thus, our estimate of q is 1.108. Note that there are no claims greater than \$500,000. We would expect $90\ (1 - F(20)) = 90 \times (20^{-1.108}) = 90 \times .0362 = 3.26$ claims greater than \$500,000 if the Pareto fit all the way to infinity. This is evidence that the theoretical tail overstates the actual tail. We can calculate the expected average claim size with an upper limit of \$500,000, using (5) with $b = 20$. This yields an estimate of $\$25{,}000 \times 3.559 = \$88{,}975$. The actual average claim size is \$89,703.

Application 2

Consider the 40 wind-related catastrophes in 1977 listed in Hogg and Klugman [H2] page 64. Only claims of \$2,000,000 or more were included. These values, recorded in millions, are as follows:

2,	2,	2,	2,	2,	2,	2,	2,	2,	2,
2,	2,	3,	3,	3,	3,	4,	4,	4,	5,
5,	5,	5,	6,	6,	6,	6,	8,	8,	9,
15,	17,	22,	23,	24,	24,	25,	27,	32,	43

If we calculate the MLE of the parameter using (19) or (20), the result is $q = .976$. This tends to confirm the statement made earlier that a typical parameter value for property is 1.0. In the same reference, on page 68, are 31 wind catastrophes over \$1 million for 1971. The MLE for these losses is $q = .959$.

Application 3

Suppose we have the following hypothetical information for the professional liability experience of a hospital. Assume that the hospital has a \$25,000 retention and that information on claims less than the retention is either unavailable or unreliable.

Accident Year	1978	1979	1980	1981
# Occupied Beds	200	200	260	260
Individual	127,000	71,000	34,000	55,000
Claims	28,000	119,000	26,000	43,000
Greater	32,000	135,000	38,000	40,000
than	103,000	42,000	93,000	42,000
\$25,000		37,000	40,000	50,000
		55,000	34,000	31,000
			228,000	30,000
			57,000	29,000
			27,000	29,000
			36,000	137,000
				61,000

Suppose we are interested in projecting the experience for 1984 for the layer \$225,000 excess of \$25,000. Assume that external data leads us to believe that the severity trend has been 20% annually between 1978 and 1981, but is projected to be 15% annually between 1981 and 1984. We also estimate 240 occupied beds in 1984.

First, as noted earlier, we can use all 31 losses in the analysis. Each loss is normalized by dividing by \$25,000. The MLE of the parameter is calculated using (19) or (20). The sum of the logs is 22.024, so the estimate of the parameter is 31/22.024 = 1.408.

We can calculate the average claim size in the layer \$225,000 xs \$25,000 using formula (1)

$$XC(b) = \frac{q - b^{1-q}}{q - 1}$$

With $b = 250/25 = 10$ and $q = 1.408$, the result is 2.493, which corresponds to an average claim size of \$25,000 × 2.493 = \$62,326. Thus, we expect that the average claim, greater than \$25,000 but limited to \$250,000, will be \$62,326. The amount within the insured layer will be \$62,326 − \$25,000 = \$37,326 per claim.

To estimate the frequency of claim within the layer, we first calculate the frequencies in terms of claims per 100 beds. The resulting ratios are:

Year	# Claims/100 Beds
1978	2.00
1979	3.00
1980	3.85
1981	4.23

We now have to adjust the frequency for trend, so that each year will be on a comparable basis. We will convert each frequency to the frequency that would be expected in 1981, using the adjustment factor in the section on trend, $(1 + i)^{nq}$ where $1 + i = 1.20$, $q = 1.408$, and n is the number of years between each year and 1981. For example, the adjustment factor for 1978 is $(1.20)^{3 \times 1.408} = 2.16$. This means that for every claim that exceeded $25,000 in 1978, we would expect 2.16 claims over $25,000 in 1981. The adjustment factors and the adjusted frequencies are shown in the following:

Year	Raw Frequency	Adjustment Factor	Adjusted Frequency
1978	2.00	2.16	4.32
1979	3.00	1.67	5.01
1980	3.85	1.29	4.98
1981	4.23	1.00	4.23

We can calculate a simple average of the adjusted frequencies to arrive at an estimate of the frequency of claims greater than $25,000 for 1981. This value is 4.63. Alternative methods to calculate an overall frequency could be used. For example, it might be appropriate to use the number of occupied beds as weights. If the adjusted frequencies show a pronounced trend over time, then the frequencies are being affected by something other than changes in claim sizes and further analysis is indicated.

We now calculate the frequency appropriate for 1984. Based upon the assumption of a 15% annual trend, the adjustment factor is $(1.15)^{3 \times 1.408} = 1.805$. Thus, our estimated frequency for 1984 is $1.805 \times 4.63 = 8.36$ claims per 100 occupied beds. Using our assumption that there will be 240 occupied beds in 1984, we expect $8.36 \times 2.4 = 20.07$ claims greater than $25,000 in

1984. Thus, our expected losses in the layer \$225,000 xs \$25,000 are 20.07 × \$37,326 = \$749,301.

Application 4

Finally, we note that the fact that a typical value of q for property losses is 1.0 and the formula for the average loss when $q = 1.0$ is so simple, allows us to easily provide a rough estimate of the average claim size for various layers. Suppose we are asked to quote a reinsurance cover on a book of property business for the layer \$2,750,000 xs \$250,000. The ratio of the upper limit to the lower limit is 3,000/250 = 12, so an estimate of the gross mean claim size is $1 + \ln 12 = 2.485$ or \$621,000. The net mean claim size would be \$371,000. This could be used as a rough estimate for discussion purposes. More refined analysis can be performed if both parties to the intended transaction are still interested.

BIBLIOGRAPHY

[B1] Benktander, G., "On the Correlation in Results from Different Layers in Excess Reinsurance," *Astin Colloquium*, 1961, pp. 203–209.

[G1] Gerber, Hans U., *An Introduction to Mathematical Risk Theory*, Huebner Foundation Monograph Number 8, Richard D. Irwin, Inc., 1979.

[H1] Hewitt, Charles C., Jr. and Lefkowitz, Benjamin, "Methods for Fitting Distributions to Insurance Loss Data," *PCAS*, LXVI, 1979, pp. 139–160.

[H2] Hogg, Robert V. and Klugman, Stuart A., *Loss Distributions*, J. Wiley and Sons, Inc., 1984.

[H3] Haung, J. S., "A Note on Order Statistics from Pareto Distribution," *Scandinavian Actuarial Journal*, 1975, pp. 187–190.

[H4] Hastings, N. A. J. and Peacock, J. B., *Statistical Distributions*, J. Wiley and Sons, Inc., 1974.

[H5] Hoel, Paul G., Port, Sidney C. and Stone, Charles, J., *Introduction to Statistical Theory*, Houghton Mifflin Co., 1971.

[I1] Insurance Services Office, "Report of the Increased Limits Subcommittee: A Review of Increased Limits Ratemaking," 1980.

[J1] Johnson, Norman L. and Kotz, Samuel, *Continuous Univariate Distributions-1*, Houghton Mifflin Co., 1970.

[K1] Kendall, M.G. and Stuart, A., *The Advanced Theory of Statistics, Vol. 2*, (2nd ed.) Mafner Publishing Co., 1967.

[L1] LaRose, J. Gary, "A Note on Loss Distributions," *PCAS*, LXIX, 1982, pp. 15–29, and discussion by Stephen W. Philbrick, *PCAS*, LXXI, 1984, pp. 1–7.

[L2] Lwin, Thaung, "Empirical Bayes Approach to Statistical Estimation in the Paretian Law," *Scandinavian Actuarial Journal*, 1974, pp. 221–236.

[L3] Lau, Hon-Shiang, "An Effective Approach for Estimating the Aggregate Loss of an Insurance Portfolio," *Journal of Risk and Insurance*, LI, No. 1, 1984, pp. 20–30.

[M1] Miccolis, Robert S., "On the Theory of Increased Limits and Excess of Loss Pricing" and discussion by Sheldon Rosenberg, *PCAS*, LXIV, 1977, pp. 27–73.

[M2] Malik, H. J., "Estimation of the Parameters of Pareto Distribution," *Metrika*, Vol. 15, 1970, pp. 126–132.

[M3] Mayerson, A. L., Jones, D. A., and Bowers, N. L., Jr., "On the Credibility of the Pure Premium." *PCAS*, LV, 1968, pp. 175–185.

[P1] Patrik, Gary, "Estimating Casualty Insurance Loss Amount Distributions," *PCAS*, LXVII, 1980, pp. 57–109.

[Q1] Quandt, R. E., "Old and New Methods of Estimation and the Pareto Distribution," *Metrika*, Vol. 10, 1966, pp. 55–82.

[S1] Snader, R. H., "Fundamentals of Individual Risk Rating and Related Topics," Study Note for Parts 9b & 10, Casualty Actuarial Society.

[V1] Venter, Gary, "Scale Adjustments to Excess Expected Losses," *PCAS*, LXIX, 1982, pp. 1–14.

APPENDIX A

SUMMARY OF FORMULAE

This appendix contains a summary of the most commonly used formulae. It begins with the formulae used to calculate the maximum likelihood estimates of the parameter. Formulae are shown later for the mean and total loss variance (under the assumption of a Poisson frequency). The formula for the variance of severity alone is not given, because the primary use for this formula is to derive the formula for the total loss variance.

It should be noted that "K" is used to represent the lower bound of the distribution in nominal or "real" dollars. This is the value used to normalize the distribution. The letter "n" is used in the formula for the MLE to denote the actual number of losses used in the calculation. In the calculation of the expected losses, "n" is used to denote the expected number of claims in the period of interest. The letter "b" is used to denote an upper limit to losses, either a censorship or truncation point.

Density $\quad f(x) = qx^{-1-q}$

Distribution $\quad F(x) = 1 - x^{-q}$

Maximum Likelihood Estimates

Unlimited $\quad q = n / \ln \prod^{n} x_i$

or

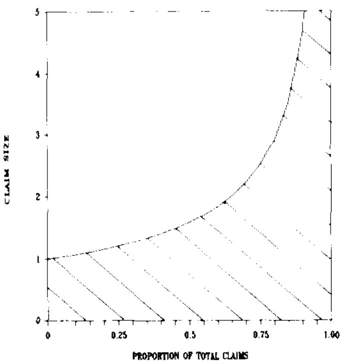
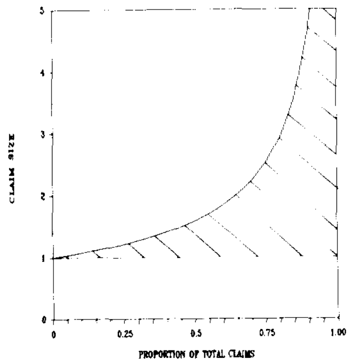
$$q = n / \sum^{n} \ln x_i$$

Censored at b $\quad q = (n-m) / \sum^{n-m} \ln x_i + m \ln b$

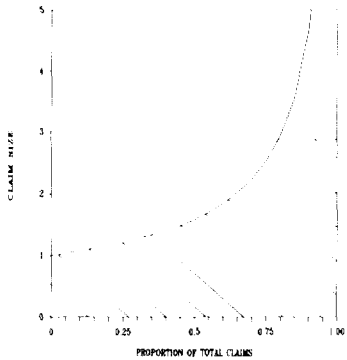
Truncated at b $\quad q = n / \sum^{n} \ln x_i - (n \ln b/(b^q - 1))$

Note that q is on both sides of the equation; thus, it must be solved using numerical methods.

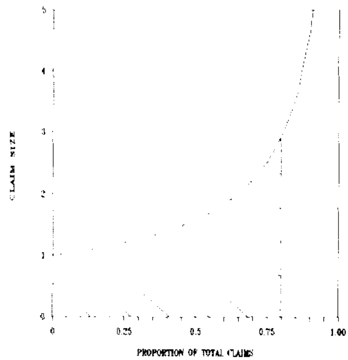
No Upper Limit

	Gross Layer	Net Layer
		
Mean Claim Size $q \neq 1$	$\dfrac{q}{q-1}$	$\dfrac{1}{q-1}$
"Real" Mean Claim Size	$K\dfrac{q}{q-1}$	$\dfrac{K}{q-1}$
"Real" Expected Losses	$nK\dfrac{q}{q-1}$	$\dfrac{nK}{q-1}$
Total Loss Variance	$\dfrac{q}{q-2}$	$\left(\dfrac{q}{q-2}\right)-\left(\dfrac{2q}{q-1}\right)+1$
Total Loss Variance in "Real" Dollars Where Expected Number of Claims is n	$nK^2\left(\dfrac{q}{q-2}\right)$	$nK^2\left(\dfrac{q}{q-2}\right)-\left(\dfrac{2q}{q-1}\right)+1$

Censored Distribution (Gross Layer)



Truncated Distribution (Gross Layer)



	Censored Distribution (Gross Layer)	Truncated Distribution (Gross Layer)
Mean Claim Size		
$q \neq 1$	$\dfrac{q - b^{1-q}}{q - 1}$	$\dfrac{q\,(1 - b^{1-q})}{(q - 1)\,(1 - b^{-q})}$
$q = 1$	$1 + \ln b$	$(\ln b) / (1 - b^{-1})$
"Real" Mean Claim Size		Multiply appropriate formula by K
Expected Losses		Multiply appropriate formula by nK
Total Loss Variance		
$q \neq 2$	$\dfrac{q - 2b^{2-q}}{q - 2}$	$\dfrac{q\,(1 - b^{2-q})}{(q - 2)\,(1 - b^{-q})}$
$q = 2$	$1 + 2 \ln b$	$(2 \ln b) / (1 - b^{-2})$
Total Loss Variance in Real Dollars		Multiply appropriate formula by nK^2

Net Layer—Mean formula can be calculated by observing that $E[X - 1] = E[X] - 1$. Variance formulae can be calculated by noting that the variance is equivalent to $E[X^2]$, and using the relationship $E[(X - 1)^2] = E[X^2] - 2E[X] + 1$.

APPENDIX B

SIMULATED PARETO LOSSES

25 pseudo-random losses from a Pareto distribution with $q = 1$ and a lower limit of \$25,000.

	Amount of Loss	Normalized Amount of Loss
1	69,976	2.799
2	62,913	2.517
3	25,766	1.031
4	39,800	1.592
5	97,739	3.910
6	36,356	1.454
7	139,665	5.587
8	34,749	1.390
9	45,716	1.829
10	96,353	3.854
11	1,847,213	73.889
12	25,231	1.009
13	48,057	1.922
14	31,744	1.270
15	98,882	3.955
16	209,031	8.361
17	214,700	8.588
18	396,323	15.853
19	32,772	1.311
20	45,190	1.808
21	32,044	1.282
22	55,843	2.234
23	99,601	3.984
24	29,900	1.196
25	60,463	2.419

APPENDIX C

VARIOUS FORMS OF THE PARETO

The Pareto distribution is mentioned in a large number of statistical texts and technical papers. Although many distributions (e.g., Poisson and normal) have a fairly standard notation, there is a wide variety of formulations of the Pareto distribution. This appendix will present a brief survey of some of the alternatives.

Johnson and Kotz [J1] contains one of the most thorough treatments of this distribution, as it devotes an entire chapter to the Pareto distribution. This reference includes a discussion of the history of the distribution, which can be traced to the Italian born, Swiss professor of economics, Vilfredo Pareto. Three main representations of the cumulative distribution are given:

Johnson and Kotz $F_x(x) = 1 - \left(\frac{K}{x}\right)^a \qquad K > 0;\ a > 0;\ x \geq K$

Johnson and Kotz $F_x(x) = 1 - \frac{K_1}{(x + c)^a}$

Johnson and Kotz $F_x(x) = 1 - \frac{K_2 e^{-bx}}{(x + c)^a}$

The first is referred to as the "Pareto distribution of the first kind," the second as the "Pareto distribution of the second kind," and the third as the "Pareto distribution of the third kind." Johnson and Kotz note that the first two formulations are Pearson Type VI distributions.

Patrik [P1] uses a form of the Pareto distribution of the second kind:

Patrik $F(x|\beta,\delta) = 1 - \left(\frac{\beta}{x + \beta}\right)^\delta$

Hogg and Klugman [H2] discuss two formulations. The first is referred to as the Pareto distribution and has the cumulative distribution:

Hogg and Klugman $F(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^\alpha \qquad \alpha > 0 \quad \lambda > 0$

The second is referred to as the generalized Pareto distribution and has a cumulative distribution as follows (where B $(\cdot)$ refers to the beta distribution):

Hogg and Klugman $F(x) = \mathrm{B}\left(K, \alpha: \dfrac{x}{\lambda + x}\right)$

The density function is as follows:

Hogg and Klugman $f(x) = \dfrac{\Gamma(\alpha + K)\, \lambda^{\alpha} \quad x^{K-1}}{\Gamma(\alpha)\, \Gamma(K)\, (\lambda + x)^{K+\alpha}}$

They note that the Pareto distribution is a special case of the generalized Pareto when $K = 1$.

Formulations by authors who work primarily with the cumulative distribution include:

Huang	$G(X; a, v) = 1 - a^{v} x^{-v}$	$x > a, a > 0, v > 0$
Benktander	$F(x) = 1 - x^{-\alpha}$	$x \geq 1$
Quandt	$F(x) = 1 - \left(\dfrac{K}{x}\right)^{a}$	$K > 0,\ a > 0,\ x \geq K$

Other authors present this distribution in terms of the density function:

Malik	$f(x) = v\, a^{v}\, x^{-v-1}$	$a > 0, v > 0, x \geq a$
Lwin	$f(x\|\lambda, a) = \lambda\, a^{\lambda}\, x^{-\lambda-1}$	$a > 0, \lambda > 0, x > a$
Kendall and Stuart	$dF = \dfrac{K}{x^{\alpha}}\, dx$	$0 < K \leq x \leq \infty, \quad \alpha > 1$
Hastings and Peacock	$f(x) = cx^{-c-1}$	$1 \leq x, \quad c > 0$

Finally, the ISO uses a Pareto distribution in the "Report of the Increased Limits Subcommittee: A Review of Increased Limits Ratemaking" [I1]. In that paper, they use "q" as a parameter. For that reason, "q" has been selected as the parameter in this paper.

APPENDIX D

REASONS FOR PREFERRING "ALTERNATIVE" REPRESENTATION

Although we typically portray density and distribution functions with the loss size along the horizontal axis and the density or cumulative probability along the vertical axis, there are a number of logical reasons for preferring the "alternative" representation, as portrayed in Figures 1, 2, 3, and 4.

1. In the standard representation, a loss limit is a vertical line and the excess losses lie to the right of the line. In my representation, a loss limit would be a horizontal line, and excess losses would lie *above* the line. It seems more intuitive to think of excess losses lying *above* a line.
2. In my representation, losses eliminated by a deductible would be *below* the line representing the deductible amount, rather than to the left of a line.
3. If we apply a trend factor to the cumulative distribution of losses, the new line is *below* the old line in the standard representation but *above* it in my representation. It makes more sense to think of inflation as producing a new curve *above* the old one.

Finally, I would note that this alternative representation is not new. It is essentially equivalent to that used in Snader [S1] to depict the insurance charge and savings.

DISCUSSION BY KURT A. REICHLE AND JOHN P. YONKUNAS

Once again, Steve Philbrick has taken a concept which makes many actuaries feel uncomfortable and, through lucid writing and clear examples, made it available to all who take the time to read him. Prior to Mr. Philbrick's paper, fitting size of loss distributions has been a tool primarily available only to the "pure actuary." This guide to the Pareto distribution provides all actuaries access to a powerful means of analysis. The strength of this tool is matched only by its simplicity as presented by Mr. Philbrick.

Using data prepared by the Actuarial Committee of the Insurance Services Office (ISO), this review will examine three facets of the single parameter Pareto distribution: the impact of development, the impact of trend, and evidence that the Pareto may overstate the tail of the distribution function. We also will suggest some practical guides for putting the Pareto distribution to use, including an analysis of the sensitivity of the parameter estimate to the number of claims available.

WHY THE PARETO?

Beginning in 1977, the Ad Hoc Increased Limits Subcommittee of ISO (subsequently the Increased Limits Committee) searched for the best-fitting continuous distribution for liability losses. Because no distribution seemed to fit both small and large claims well, the Subcommittee decided instead to look for the best-fitting curve for losses above a lower truncation point. After a good deal of research, the two parameter Pareto was selected as providing the best fit to liability losses. Although many enhancements have been made in the methodology used to derive increased limits factors since 1977, the Pareto curve remains ISO's favored distribution.

Implementation of the two parameter Pareto distribution does require complex formulas, including a set of Newton-Raphson equations used iteratively to solve for the Maximum Likelihood Estimates (MLE) of each parameter. These formulas are not solved easily without the use of a computer, and therefore require extensive programming and computing costs. While this complexity does not pose an insurmountable problem to the "pure actuary," it may hinder the efforts of the "lay actuary" to use models rather than empirical data directly.

The one parameter Pareto distribution is quite simple to use, as demonstrated in Mr. Philbrick's article. Estimates of various moments of the distribution are very simple to calculate, and the formulas are easily remembered. The same is true for many estimates of the parameter, including the MLE. However, to our knowledge, no extensive research has been published on how appropriate the one parameter Pareto is for loss distributions. The two parameter Pareto has a proven track record as an acceptable model for excess losses. A simple mathematical transformation will show that the parameter of the one parameter Pareto is equivalent to a parameter of the two parameter Pareto (see Appendix A). Hence, by using the one parameter version, we obtain much of the power of the two parameter Pareto without its accompanying complexity.

LOSS DEVELOPMENT

The change in the cumulative value of losses for a given accident period has been discussed extensively in the actuarial literature. But very little has been published on how the distribution of individual claims changes as losses mature, and in particular how the parameters underlying that distribution change. A full discussion of loss development and its effect on the Pareto is beyond the scope of this review. We will, however, cite some of our observations from examining data provided by ISO.

An analysis of losses usually begins by segregating the data into various time periods (report year, accident year, policy year, etc.). To put these periods on a comparable basis, two adjustments are commonly utilized: trend and loss development. Although the use of the one parameter Pareto implies that severity trend may be ignored (as discussed in a later section), loss development may not. An adjustment must be made for loss development prior to combining various periods for analysis.

In casualty lines of insurance, loss development generally has a positive impact on losses; i.e., average losses become more severe as the largest losses emerge most slowly. Remember that severity and the Pareto parameter are inversely related. Therefore, an *a priori* expectation is that the Pareto parameter should decrease as an accident period becomes more mature.

As will be seen, the value of the Pareto parameter varies substantially from one valuation to another. Property losses from several accident periods may be combined to derive the parameter with no recognition of the date of loss. Applying the same approach to casualty lines may severely overstate the parameter and understate the excess severity. An excellent example of this is inadvertently included in Mr. Philbrick's paper. In Application 3 in Section VII, Mr. Philbrick combines professional liability claims from four accident years with no adjustment for development, calculating a MLE of the Pareto parameter of 1.408. Deriving the maximum likelihood estimate of each year separately produces Pareto parameters of 1.176, 1.002, 1.570 and 1.746 for 1978 through 1981 respectively. The clear upward trend in these values is to be expected and most often results in an overstatement of the parameter if the claims are simply combined with no adjustment for development.

Additional evidence that the parameter is inversely related to maturity was found when we examined occurrence size distributions (OSD's) provided by ISO. A lower truncation point of $25,000 has been selected. The MLE of the Pareto parameter was calculated by policy year, by evaluation month. A table of parameters for Owners, Landlords, and Tenants (OL&T) Bodily Injury Liability follows.

Policy Year	Evaluation Month						
	27	39	51	63	75	87	99
1975	1.313	1.546	1.412	1.377	1.308	1.283	1.281
1976	1.547	1.467	1.407	1.309	1.258	1.225	
1977	1.539	1.578	1.482	1.389	1.347		
1978	1.644	1.578	1.460	1.364			
1979	1.688	1.518	1.443				
1980	1.700	1.590					
1981	1.717						

As expected, the parameter decreases as the policy year matures. Loss development must be accounted for prior to analysis. One could use a triangulation to adjust immature parameters to their ultimate values.

Note that with the exception of the 27 month evaluation, the parameter is relatively stable across policy years within a given evaluation. We found this to be true for other values of the truncation point and for Products Bodily Injury Liability data.

Why the parameters calculated at 27 months exhibit an upward trend is not clear. It may indicate that data as of 27 months is too immature for analyzing excess losses. It may also indicate a change in industry reserving practices. Such a change would affect the distribution most at the earliest evaluation and least at later maturities.

We suggest that more research be devoted to determining the impact of loss development on the Pareto parameter. We also recommend that the user of the one parameter Pareto not blindly combine data without adjusting for loss development.

TREND

Of all the implications of the Pareto distribution, the most vexing is that trend does not affect excess loss severity, only loss frequency. How can such a distribution be appropriate for casualty-property losses? It is "obvious" that trend changes severity values. The work of ISO in fitting Pareto distributions to excess liability losses provides us with much data to evaluate this property.

As shown in the section on loss development, the Pareto parameter has remained relatively stable across policy periods for a given evaluation, which provides solid evidence that the parameter may be unaffected by trend.

Another empirical test is to examine the value of the average excess claim size over time. We again turn to the OSD's for OL&T Bodily Injury as compiled by ISO. Note that this raw data has not been adjusted for trend or loss development.

It is readily apparent that the average claim sizes have remained stable over time: both across policy years and within policy years. Trend does not appear to affect the average size of loss within a specific excess layer.

A more direct approach is to examine the form of distribution after making a transformation for trend. Assuming uniform trend, the value of the parameter is preserved; that is, q remains unchanged. The mathematical details of this transformation can be found in Appendix B.

How does one explain that the average claim size within a given excess interval remains unaffected after trend (and development)? At first glance it is intuitively unappealing if not totally unacceptable. Is it possible that the Pareto simply is not a realistic model for size of loss distributions?

The explanation is that the forces of trend and development fall upon the frequency side of the equation. As Mr. Philbrick points out, trend and development merely act to shift claims from one layer to another without changing the average in the layer. Instead, the frequency by layer changes as losses develop and occur later in time. So we still are stuck with adjustments for trend and development when the objective is to forecast aggregate loss dollars.

AVERAGE CLAIM SIZE IN LAYER
$50,000 TO $100,000

Policy Year	Evaluation Month						
	27	39	51	63	75	87	99
1975	79,174	77,135	78,306	78,407	80,263	80,462	80,533
1976	77,039	75,303	76,920	78,484	79,264	79,864	
1977	77,742	76,496	76,373	77,540	78,278		
1978	75,247	76,994	78,765	79,026			
1979	73,067	76,232	77,827				
1980	73,789	75,733					
1981	75,011						

AVERAGE CLAIM SIZE IN LAYER
$100,000 TO $250,000

Policy Year	Evaluation Month						
	27	39	51	63	75	87	99
1975	172,059	165,249	163,967	166,215	167,825	171,715	173,931
1976	170,587	170,295	170,584	173,939	176,422	179,241	
1977	156,528	159,315	159,453	165,980	167,384		
1978	161,748	162,952	167,251	173,905			
1979	156,233	162,447	165,895				
1980	161,820	165,275					
1981	154,519						

In developing increased limits factors or excess loss premium factors, claim frequency drops out of the equation. All that remain are ratios of severities. Therefore, since the parameter is preserved after trend, an adjustment for trend may not be necessary. This could greatly simplify current procedures.

Data we have examined support the conclusion that trend does not affect excess severity. Hence, our preconceptions turned out to be significant stumbling blocks to accepting the Pareto. We hope that other readers will note the strength of the empirical evidence before accepting what appears to be "common sense."

GOODNESS OF FIT

In its initial consideration of the Pareto, the Increased Limits Subcommittee of ISO expressed concern that the Pareto may overstate the tail probabilities. Mr. Philbrick also refers to the fact that "most actual data suggests that the tail of the Pareto is still somewhat too 'thick' at extremely high loss amounts." Empirical evidence for casualty lines demonstrates the greater the truncation point, the larger the parameter estimate. That is, the indicated excess severity declines as one raises the truncation point when fitting the distribution. If excess claims were truly Pareto distributed, then one would obtain the same maximum likelihood estimate of the parameter independent of the truncation point chosen.

To demonstrate this overstatement, we look at Pareto parameters derived from ISO data for liability lines. These data are censored above at $500,000. The Workers' Compensation data are from a single insurer and are unlimited.

PARETO PARAMETERS

Line of Insurance	Lower Truncation Point (000)			
	25	50	100	250
OL&T Bodily Injury	1.281	1.330	1.447	1.508
Products Bodily Injury	.991	1.269	1.714	2.584
Workers' Compensation	1.454	1.715	2.316	2.086

It is clear from these data that, depending on the line of insurance, the Pareto parameter may be influenced greatly by the truncation point chosen. A significant implication of this upward trend is that parameters estimated with a low truncation point will generate conservative estimates of severities in the higher layers. For example, the estimate of the layer $1,000,000 excess of $1,000,000 may be greatly overstated if the truncation point for deriving the parameter is $25,000. The impact of this shortcoming is minimal if the excess layer estimated has a lower bound close to the truncation point. For instance, the estimate of the severity of any layer excess of $25,000 will be close to the actual severity in that layer if the truncation point for deriving the q parameter is close to $25,000. This is true even when the parameter increases rapidly with the truncation point.

ISO data provide evidence to support these conjectures. Displayed in the following table are comparisons of actual and fitted average severities for a selected group of gross layers.

OWNERS, LANDLORDS AND TENANTS BODILY INJURY
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1975 AS OF 99 MONTHS

Losses Limited to	Actual Severity	Difference Between Actual and Fitted Severities for Truncation Point of $25,000	Difference Between Actual and Fitted Severities for Truncation Point of $100,000
$125,000	$119,486	1.8%	1.5%
$150,000	$135,162	2.3%	1.4%
$175,000	$147,818	2.7%	1.1%
$200,000	$158,426	2.9%	0.7%
$250,000	$173,931	3.9%	0.7%
$300,000	$184,779	5.3%	1.1%
$350,000	$190,338	8.0%	2.9%
$400,000	$195,144	10.1%	4.2%
$450,000	$199,153	11.8%	5.2%
$500,000	$202,348	13.4%	6.1%

Two facts are readily apparent from this exhibit. First, the wider the layer being estimated, the greater the potential error. Second, the closer the truncation point is to the lower end of the layer, the smaller the error. For those interested, Appendix C contains similar data for other truncation points and evaluations.

In using the Pareto to derive increased limits factors, the magnitude of these errors is significantly reduced. Losses in excess of the lower truncation point generally represent a small percentage of the total claim count. Since increased limits factors incorporate claims of all sizes, the large percentage of losses below the truncation point reduces the impact of any error in the excess estimate and, therefore, any error in the increased limits factor.

PRACTICAL CONSIDERATIONS

The inability to correctly estimate the Pareto parameter will obviously affect the accuracy of the excess severity. As is commonly true when modelling, the error in the parameters is dependent upon the amount of data available. The Pareto is no exception.

A generally accepted way to express the potential errors in a parameter estimate is a classical credibility approach based on claim counts. Confidence intervals, although complex in their derivation, can be developed and used to indicate the number of claims required to achieve a given level of confidence for a given level of tolerance. For example, it can be shown that 310 claims are necessary to be 90% confident of being within 10% of the true value of the Pareto parameter. Confidence intervals in the following table were generated based on the MLE of the parameter. Formulas for the confidence intervals are developed in Appendix D.

Level of Tolerance	Level of Confidence				
	97.5%	95%	90%	85%	80%
± 5%	2160	1655	1165	890	710
±10%	580	445	310	240	190
±15%	275	210	150	115	90
±25%	115	85	60	45	40
±50%	40	30	20	15	10

This table can give the user an indication of the accuracy of the MLE. Clearly, a large number of excess claims is required for a high degree of accuracy. When sample data lack the credibility required, it is desirable to have available a source of parameters based on a larger volume of data. These parameters can then be used as the complement of credibility to the parameter derived from the data being analyzed. In Appendix E are Pareto parameters from ISO for various sublines of General Liability, Automobile Liability and Professional Liability. When either no data or limited volumes of data are available, these factors can provide reasonable estimates of excess severities.

An important question to answer before determining whether enough claims are available or whether to use ISO factors for credibility weighting is: "How sensitive is the estimate of an average net claim size to errors in the parameter estimate?" The following charts display the error in the estimate of the average net claim size for various layers of loss for a given error in the MLE.

ERROR IN AVERAGE CLAIM COST
PARETO PARAMETER = 1.00

	Error in MLE		
Net Layer	10%	25%	50%
\$400,000 excess of \$100,000	7.6%	17.7%	31.3%
\$900,000 excess of \$100,000	10.7%	24.0%	40.6%
\$1,900,000 excess of \$100,000	13.6%	29.6%	48.6%

ERROR IN AVERAGE CLAIM COST
PARETO PARAMETER = 1.50

	Error in MLE		
Net Layer	10%	25%	50%
\$400,000 excess of \$100,000	9.7%	21.9%	37.3%
\$900,000 excess of \$100,000	12.7%	27.6%	44.8%
\$1,900,000 excess of \$100,000	15.1%	31.8%	50.0%

Two generalizations can be drawn from this example. The percentage error varies with both the size of the parameter and the width of the layer being estimated. It is also interesting to note that the error in estimating an average net claim size for a specific layer can easily exceed the error in the MLE.

Because of the special properties of the Pareto, the errors for the layers shown above are dependent only on the relationship of the endpoints to the truncation point. Thus, the error in each of the two layers $400,000 xs $100,000 and $4,000,000 xs $1,000,000, with truncation points of $100,000 and $1,000,000, respectively, is the same, given an identical error in the underlying parameter.

Even though the percent error in a layer varies with the size of the parameter, the absolute dollar error decreases. This may be obvious since severity is inversely proportional to the Pareto parameter. Thus we might be more lenient with a lower degree of tolerance for a larger value of the parameter.

The following table displays absolute dollar errors in various net layers for a 10% error in the MLE.

DOLLAR ERROR IN NET LAYER

Net Layer	$q=1.00$	$q=1.50$
$400,000 excess $100,000	$12,284	$10,756
$900,000 excess $100,000	$24,587	$17,350
$1,900,000 excess $100,000	$40,708	$23,382

CONCLUSION

The empirical data we have examined indicate that the implications underlying the use of the one parameter Pareto are satisfied for casualty lines of insurance. This is not to say that limitations and restrictions on its use do not exist. It would be asking too much of any one parameter distribution to perfectly fit excess losses of all property and casualty lines. But the range of applications of the Pareto are substantial and, therefore, significant to anyone involved in excess pricing. This paper should provide encouragement to those who may have felt intimidated by the complexity of most modelling techniques available to actuaries. At the same time, it provides a powerful tool for those who regularly use more complex models but do not always need ten decimal point accuracy.

Most of the data referenced in this review is the product of the Increased Limits Committee and the staff of the Insurance Services Office. We wish to thank the ISO for allowing us the use of their data and analysis.

APPENDIX A

DERIVATION OF THE ONE PARAMETER PARETO FROM THE TWO PARAMETER PARETO

The one parameter Pareto is a special case of the two parameter Pareto. A common form of the two parameter Pareto and the one currently used by Insurance Services Office is:

$$f(x) = \frac{q \times (b)^q}{(x + b)^{(q+1)}} \qquad \text{for } 0 \leq x < \infty \tag{1}$$

In this formula, the value of x represents individual claim sizes. Generally, this form is fit to losses above some lower truncation point.

We wish to derive $g(y)$, where

$$y = (x + b)/b \qquad \text{for } 1 \leq y < \infty \tag{2}$$

and

$$dy = dx/b \tag{3}$$

Substituting (2) and (3) into (1) we have,

$$g(y) = f(x) \times (dx/dy)$$

$$g(y) = q \times y^{-(q+1)} \qquad \text{for } 1 \leq y < \infty$$

which is the general form of the one parameter Pareto.

APPENDIX B

PARETO AND TREND

This appendix will show that, assuming uniform trend (i.e., all claims sizes trend at the same rate), the value of the parameter q is preserved (remains unchanged). We start by restating the Pareto,

$$f(x) = q \times x^{(q+1)} \text{ for } 1 \leq x < \infty$$

Under uniform trend we have the following transformation,

$$y = a \times x \text{ for } a \leq y < \infty$$

and,

$$dy = a \times dx$$

Here the multiplicative factor a represents the impact of trend on individual claims.

Making this change of variable and solving for $g(y)$ we have,

$$g(y) = (1/a) \times q \times (y/a)^{(q+1)} \text{for } a \leq y < \infty$$

Renormalizing this density function by dividing all values of y by a, we have,

$$z = y/a;\ dz = dy/a \text{ for } 1 \leq z < \infty$$

The transformation then becomes,

$$h(z) = g(y) \times (dy/dz)$$

$$h(z) = q \times z^{-(q+1)}$$

The parameter q in all three density functions is the same and has not been affected by the transformation.

APPENDIX C

COMPARISON OF FITS

Presented in the following exhibits are comparisons of actual severities to data fitted by a one parameter Pareto. All data are from OL&T Bodily Injury as provided by ISO.

Exhibit C-1 displays fits of gross losses excess of $25,000 using a truncation point of $25,000. These fits produce an average absolute error of 1.3% and range from 0.0% to 3.8%.

Exhibit C-2 displays fits of gross losses excess of $100,000 using a truncation point of $100,000. The absolute errors in these fits average 2.0% and range from 0.0% to 6.1%.

Exhibit C-3 displays fits of gross losses excess of $100,000 using a truncation point of $25,000. The absolute errors are much greater in these fits. They average 5.6% and range from 1.9% to 13.4%.

In general, the wider the interval the greater the divergence. But these differences are relatively small when the lower bound of the layer is equal to the truncation point. Exhibit C-3 demonstrates that the error in predicting average severities can be quite large when the lower bound of a layer is much larger than the truncation point.

EXHIBIT C-1
SHEET 1

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $25,000
POLICY YEAR 1981 AS OF 27 MONTHS
MLE OF THE PARAMETER: 1.7172

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$50,000	$37,975	$38,655	1.8%
$75,000	$43,649	$44,005	0.8%
$100,000	$47,087	$46,960	−0.3%
$125,000	$48,962	$48,868	−0.2%
$150,000	$50,527	$50,215	−0.6%
$175,000	$51,605	$51,224	−0.7%
$200,000	$52,549	$52,013	−1.0%
$250,000	$53,946	$53,173	−1.4%
$300,000	$54,955	$53,992	−1.8%
$350,000	$55,548	$54,606	−1.7%
$400,000	$56,006	$55,086	−1.6%
$450,000	$56,370	$55,472	−1.6%
$500,000	$56,702	$55,791	−1.6%

EXHIBIT C-1
SHEET 2

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $25,000
POLICY YEAR 1980 AS OF 39 MONTHS
MLE OF THE PARAMETER: 1.5899

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$50,000	$38,695	$39,223	1.4%
$75,000	$44,758	$45,213	1.0%
$100,000	$48,499	$48,673	0.4%
$125,000	$50,850	$50,980	0.3%
$150,000	$52,765	$52,653	−0.2%
$175,000	$54,135	$53,933	−0.4%
$200,000	$55,335	$54,951	−0.7%
$250,000	$57,196	$56,484	−1.2%
$300,000	$58,550	$57,595	−1.6%
$350,000	$59,413	$58,446	−1.6%
$400,000	$60,143	$59,123	−1.7%
$450,000	$60,721	$59,677	−1.7%
$500,000	$61,207	$60,141	−1.7%

EXHIBIT C-1
SHEET 3

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $25,000
POLICY YEAR 1979 AS OF 51 MONTHS
MLE OF THE PARAMETER: 1.4427

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$50,000	$39,551	$39,922	0.9%
$75,000	$46,500	$46,749	0.5%
$100,000	$51,110	$50,901	−0.4%
$125,000	$53,954	$53,777	−0.3%
$150,000	$56,253	$55,924	−0.6%
$175,000	$58,053	$57,610	−0.8%
$200,000	$59,646	$58,979	−1.1%
$250,000	$61,993	$61,095	−1.4%
$300,000	$63,613	$62,675	−1.5%
$350,000	$64,614	$63,915	−1.1%
$400,000	$65,447	$64,923	−0.8%
$450,000	$66,088	$65,764	−0.5%
$500,000	$66,601	$66,479	−0.2%

EXHIBIT C-1
SHEET 4

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $25,000
POLICY YEAR 1978 AS OF 63 MONTHS
MLE OF THE PARAMETER: 1.3644

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$50,000	$39,696	$40,313	1.6%
$75,000	$47,088	$47,633	1.2%
$100,000	$52,090	$52,209	0.2%
$125,000	$55,529	$55,442	−0.2%
$150,000	$58,321	$57,895	−0.7%
$175,000	$60,545	$59,846	−1.2%
$200,000	$62,466	$61,449	−1.6%
$250,000	$65,349	$63,960	−2.1%
$300,000	$67,484	$65,866	−2.4%
$350,000	$68,749	$67,381	−2.0%
$400,000	$69,806	$68,627	−1.7%
$450,000	$70,706	$69,676	−1.5%
$500,000	$71,518	$70,577	−1.3%

EXHIBIT C-1
SHEET 5

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $25,000
POLICY YEAR 1977 AS OF 75 MONTHS
MLE OF THE PARAMETER: 1.3466

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$50,000	$39,895	$40,404	1.3%
$75,000	$47,363	$47,841	1.0%
$100,000	$52,203	$52,519	0.6%
$125,000	$55,322	$55,839	0.9%
$150,000	$57,747	$58,367	1.1%
$175,000	$59,602	$60,384	1.3%
$200,000	$61,216	$62,046	1.4%
$250,000	$63,890	$64,657	1.2%
$300,000	$66,044	$66,646	0.9%
$350,000	$67,240	$68,232	1.5%
$400,000	$68,175	$69,539	2.0%
$450,000	$68,979	$70,642	2.4%
$500,000	$69,716	$71,592	2.7%

EXHIBIT C-1
SHEET 6

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $25,000
POLICY YEAR 1976 AS OF 87 MONTHS
MLE OF THE PARAMETER: 1.2254

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$50,000	$40,414	$41,043	1.6%
$75,000	$48,555	$49,329	1.6%
$100,000	$53,923	$54,765	1.6%
$125,000	$57,602	$58,746	2.0%
$150,000	$60,653	$61,853	2.0%
$175,000	$63,165	$64,382	1.9%
$200,000	$65,374	$66,503	1.7%
$250,000	$68,936	$69,908	1.4%
$300,000	$71,613	$72,565	1.3%
$350,000	$73,314	$74,728	1.9%
$400,000	$74,685	$76,543	2.5%
$450,000	$75,807	$78,098	3.0%
$500,000	$76,844	$79,455	3.4%

EXHIBIT C-1
SHEET 7

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $25,000
POLICY YEAR 1975 AS OF 99 MONTHS
MLE OF THE PARAMETER: 1.2805

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$50,000	$40,137	$40,748	1.5%
$75,000	$48,060	$48,637	1.2%
$100,000	$53,525	$53,714	0.4%
$125,000	$57,217	$57,379	0.3%
$150,000	$60,188	$60,208	0.0%
$175,000	$62,586	$62,490	−0.2%
$200,000	$64,597	$64,388	−0.3%
$250,000	$67,535	$67,406	−0.2%
$300,000	$69,591	$69,735	0.2%
$350,000	$70,644	$71,614	1.4%
$400,000	$71,555	$73,177	2.3%
$450,000	$72,315	$74,508	3.0%
$500,000	$72,920	$75,661	3.8%

EXHIBIT C-2
SHEET 1

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1981 AS OF 27 MONTHS
MLE OF THE PARAMETER: 2.0623

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$114,905	$119,867	4.3%
$150,000	$127,343	$132,944	4.4%
$175,000	$135,917	$142,187	4.6%
$200,000	$143,421	$149,057	3.9%
$250,000	$154,519	$158,571	2.6%
$300,000	$162,540	$164,833	1.4%
$350,000	$167,257	$169,259	1.2%
$400,000	$170,897	$172,549	1.0%
$450,000	$173,788	$175,088	0.7%
$500,000	$176,430	$177,104	0.4%

EXHIBIT C-2

SHEET 2

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1980 AS OF 39 MONTHS
MLE OF THE PARAMETER: 1.6478

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$117,647	$120,777	2.7%
$150,000	$132,015	$135,659	2.8%
$175,000	$142,302	$146,940	3.3%
$200,000	$151,305	$155,842	3.0%
$250,000	$165,275	$169,103	2.3%
$300,000	$175,438	$178,602	1.8%
$350,000	$181,915	$185,802	2.1%
$400,000	$187,391	$191,484	2.2%
$450,000	$191,733	$196,104	2.3%
$500,000	$195,380	$199,948	2.3%

EXHIBIT C-2
SHEET 3

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1979 AS OF 51 MONTHS
MLE OF THE PARAMETER: 1.741

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$117,217	$120,567	2.9%
$150,000	$131,141	$135,022	3.0%
$175,000	$142,041	$145,809	2.7%
$200,000	$151,681	$154,207	1.7%
$250,000	$165,895	$166,513	0.4%
$300,000	$175,709	$175,162	−0.3%
$350,000	$181,765	$181,616	−0.1%
$400,000	$186,812	$186,641	−0.1%
$450,000	$190,693	$190,679	0.0%
$500,000	$193,797	$194,004	0.1%

EXHIBIT C-2
SHEET 4

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1978 AS OF 63 MONTHS
MLE OF THE PARAMETER: 1.5282

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$119,167	$121,049	1.6%
$150,000	$134,728	$136,499	1.3%
$175,000	$147,124	$148,449	0.9%
$200,000	$157,831	$158,043	0.1%
$250,000	$173,905	$172,639	−0.7%
$300,000	$182,802	$183,351	0.3%
$350,000	$192,851	$191,638	−0.6%
$400,000	$198,746	$198,290	−0.2%
$450,000	$203,759	$203,781	0.0%
$500,000	$208,289	$208,412	0.1%

EXHIBIT C-2
SHEET 5

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1977 AS OF 75 MONTHS
MLE OF THE PARAMETER: 1.5706

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$117,980	$120,952	2.5%
$150,000	$131,964	$136,198	3.2%
$175,000	$142,660	$147,907	3.7%
$200,000	$151,968	$157,249	3.5%
$250,000	$167,384	$171,357	2.4%
$300,000	$179,804	$181,622	1.0%
$350,000	$186,699	$189,506	1.5%
$400,000	$192,095	$195,797	1.9%
$450,000	$196,731	$200,962	2.2%
$500,000	$200,977	$205,296	2.1%

EXHIBIT C-2
SHEET 6

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1976 AS OF 87 MONTHS
MLE OF THE PARAMETER: 1.2489

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$119,416	$121,706	1.9%
$150,000	$135,523	$138,568	2.2%
$175,000	$148,781	$152,239	2.3%
$200,000	$160,440	$163,665	2.0%
$250,000	$179,341	$181,931	1.4%
$300,000	$193,369	$196,121	1.4%
$350,000	$202,346	$207,626	2.6%
$400,000	$209,584	$217,242	3.7%
$450,000	$215,508	$225,462	4.6%
$500,000	$220,981	$232,613	5.3%

EXHIBIT C-2
SHEET 7

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1975 AS OF 99 MONTHS
MLE OF THE PARAMETER: 1.4467

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$119,486	$121,238	1.5%
$150,000	$135,162	$137,087	1.4%
$175,000	$147,818	$149,515	1.1%
$200,000	$158,426	$159,611	0.7%
$250,000	$173,931	$175,194	0.7%
$300,000	$184,779	$186,822	1.1%
$350,000	$190,338	$195,941	2.9%
$400,000	$195,144	$203,348	4.2%
$450,000	$199,153	$209,525	5.2%
$500,000	$202,348	$214,782	6.1%

EXHIBIT C-3
SHEET 1

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1981 AS OF 27 MONTHS
MLE OF THE PARAMETER: 1.7172

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$114,905	$120,620	5.0%
$150,000	$127,343	$135,183	6.2%
$175,000	$135,917	$146,094	7.5%
$200,000	$143,421	$154,618	7.8%
$250,000	$154,519	$167,161	8.2%
$300,000	$162,540	$176,020	8.3%
$350,000	$167,257	$182,657	9.2%
$400,000	$170,897	$187,842	9.9%
$450,000	$173,788	$192,020	10.5%
$500,000	$176,430	$195,471	10.8%

EXHIBIT C-3
SHEET 2

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1980 AS OF 39 MONTHS
MLE OF THE PARAMETER: 1.5899

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$117,647	$120,908	2.8%
$150,000	$132,015	$136,062	3.1%
$175,000	$142,302	$147,662	3.8%
$200,000	$151,305	$156,893	3.7%
$250,000	$165,275	$170,784	3.3%
$300,000	$175,438	$180,852	3.1%
$350,000	$181,915	$188,559	3.7%
$400,000	$187,391	$194,692	3.9%
$450,000	$191,733	$199,714	4.2%
$500,000	$195,380	$203,921	4.4%

EXHIBIT C-3
SHEET 3

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1979 AS OF 51 MONTHS
MLE OF THE PARAMETER: 1.4427

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$117,217	$121,248	3.4%
$150,000	$131,141	$137,116	4.6%
$175,000	$142,041	$149,568	5.3%
$200,000	$151,681	$159,689	5.3%
$250,000	$165,895	$175,322	5.7%
$300,000	$175,709	$186,997	6.4%
$350,000	$181,765	$196,159	7.9%
$400,000	$186,812	$203,606	9.0%
$450,000	$190,693	$209,818	10.0%
$500,000	$193,797	$215,108	11.0%

EXHIBIT C-3
SHEET 4

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1978 AS OF 63 MONTHS
MLE OF THE PARAMETER: 1.3644

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$119,167	$121,431	1.9%
$150,000	$134,728	$137,693	2.2%
$175,000	$147,124	$150,624	2.4%
$200,000	$157,831	$161,254	2.2%
$250,000	$173,905	$177,901	2.3%
$300,000	$182,802	$190,533	4.2%
$350,000	$192,851	$200,578	4.0%
$400,000	$198,746	$208,835	5.1%
$450,000	$203,759	$215,792	5.9%
$500,000	$208,289	$221,767	6.5%

EXHIBIT C-3
SHEET 5

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1977 AS OF 75 MONTHS
MLE OF THE PARAMETER: 1.3466

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$117,980	$121,473	3.0%
$150,000	$131,964	$137,826	4.4%
$175,000	$142,660	$150,869	5.8%
$200,000	$151,968	$161,617	6.3%
$250,000	$167,384	$178,504	6.6%
$300,000	$179,804	$191,365	6.4%
$350,000	$186,699	$201,622	8.0%
$400,000	$192,095	$210,075	9.4%
$450,000	$196,731	$217,213	10.4%
$500,000	$200,977	$223,356	11.1%

EXHIBIT C-3

SHEET 6

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1976 AS OF 87 MONTHS
MLE OF THE PARAMETER: 1.2254

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$119,416	$121,762	2.0%
$150,000	$135,523	$138,749	2.4%
$175,000	$148,781	$152,576	2.6%
$200,000	$160,440	$164,171	2.3%
$250,000	$179,241	$182,786	2.0%
$300,000	$193,369	$197,315	2.0%
$350,000	$202,346	$209,143	3.4%
$400,000	$209,584	$219,061	4.5%
$450,000	$215,508	$227,565	5.6%
$500,000	$220,981	$234,983	6.3%

EXHIBIT C-3
SHEET 7

OWNERS, LANDLORDS, AND TENANTS
PARETO GOODNESS OF FIT
GROSS LOSSES IN EXCESS OF $100,000
POLICY YEAR 1975 AS OF 99 MONTHS
MLE OF THE PARAMETER: 1.2805

Losses Limited to	Actual Severity	Fitted Severity	Percent Difference
$125,000	$119,486	$121,630	1.8%
$150,000	$135,162	$138,326	2.3%
$175,000	$147,818	$151,790	2.7%
$200,000	$158,426	$162,993	2.9%
$250,000	$173,931	$180,801	3.9%
$300,000	$184,779	$194,546	5.3%
$350,000	$190,338	$205,632	8.0%
$400,000	$195,144	$214,855	10.1%
$450,000	$199,153	$222,708	11.8%
$500,000	$202,348	$229,517	13.4%

APPENDIX D

CONFIDENCE INTERVALS FOR THE PARETO PARAMETER

This appendix derives a formula that can be used to approximate the number of claims necessary to achieve a given level of confidence for a given level of tolerance in estimating the Pareto parameter. The results of this appendix are based upon the work of Jerry Jurschak in an unpublished paper entitled "The Pareto Distribution and Excess of Loss Reinsurance."

In Mr. Jurschak's paper he shows that the following formula represents a $100(1 - d)\%$ confidence interval for the Pareto parameter,

$$\left\{\frac{b \times q'}{n}, \frac{c \times q'}{n}\right\}$$

where

q' = MLE of the parameter

n = number of claims in the sample

$$b = \frac{1}{4} \times (Z_{d/2} + \sqrt{4n - 1})^2$$

$$c = \frac{1}{4} \times (Z_{(1-d/2)} + \sqrt{4n - 1})^2$$

Z = standard normal values.

Using a classical credibility approach, various values of n can be determined for a given level of confidence and a given level of tolerance (i.e. being within $\pm 10\%$ of the true value of q).

Assume that we wish to be within $100(k - 1)\%$ of the true value of the parameter $100(1 - d)\%$ of the time. The number of claims to comply with these constraints can be determined by solving the following confidence interval for n.

$$\left\{\frac{b \times q \times k}{n}, \frac{c \times q \times k}{n}\right\}$$

Substituting the above formulas into this confidence interval,

$$\frac{1}{4} \times (Z_{d/2} + \sqrt{4n - 1})^2 \times \frac{(q \times k)}{n} \geq q$$

$$q \geq \frac{1}{4} \times (Z_{(1-d/2)} + \sqrt{4n - 1})^2 \times \frac{(q \times k)}{n}$$

Since the absolute values of the standard normal numbers are equal, nothing is lost by dropping the right term. A few algebraic manipulations will produce

$$\frac{\sqrt{4n}}{\sqrt{k}} \geq (Z_{(1-d/2)} + \sqrt{4n-1})$$

and,

$$\sqrt{4n} - (\sqrt{4n-1} \times \sqrt{k}) \geq Z_{(1-d/2)} \times \sqrt{k}$$

For large n, we may question the necessity, bearing in mind the search for a simpler form, of subtracting the 1. In other words,

$$\sqrt{4n} \sim \sqrt{4n-1}$$

Using this simplifying assumption we have

$$\sqrt{4n} - (\sqrt{4n} \times \sqrt{k}) \geq Z_{(1-d/2)} \times \sqrt{k}$$

Solving this equation for n yields

$$n \geq \frac{Z^2_{(1-d/2)} \times k}{4 \times (1 - \sqrt{k})^2}$$

This formula is then used to generate the following table. Note that all figures have been rounded to the nearest multiple of five.

Level of Tolerance	Level of Confidence 97.5%	95%	90%	85%	80%
± 5%	2160	1655	1165	890	710
±10%	580	445	310	240	190
±15%	275	210	150	115	90
±25%	115	85	60	45	40
±50%	40	30	20	15	10

APPENDIX E

INDUSTRY VALUES OF THE PARETO PARAMETER q AS PRODUCED BY INSURANCE SERVICES OFFICE

Line of Insurance	Value of q	Truncation Point
GENERAL LIABILITY		
—Products		
—Bodily Injury		
—High Severity	0.938	$25,000
—Low Severity	0.848	$25,000
—Property Damage	1.144	$ 3,000
—Manufacturers and Contractors		
—Bodily Injury		
—All Classes	0.945	$40,000
—High Severity	0.825	$35,000
—Low Severity	1.031	$40,000
—Property Damage	0.987	$ 4,000
—Owners, Landlords, and Tenants		
—Bodily Injury		
—All Classes	1.245	$25,000
—High Severity	1.159	$30,000
—Low Severity	1.600	$30,000
PROFESSIONAL LIABILITY		
—Physicians	1.141	$22,000
—Surgeons	1.110	$22,000
—Hospitals	0.932	$ 1,000
—Dentists	1.527	$ 7,000
—Lawyers	2.098	$ 2,000
COMMERCIAL AUTOMOBILE LIABILITY		
—Zone Rated	0.882	$ 9,000
—Light/Medium Trucks	1.061	$ 9,000
—Heavy Trucks	0.941	$ 9,000
—Extra Heavy Trucks	0.949	$ 9,000
—Private Passenger, Publics, and Garages	1.080	$ 9,000

A SIMULATION TEST OF PREDICTION ERRORS OF LOSS RESERVE ESTIMATION TECHNIQUES

JAMES N. STANARD

Abstract

This paper uses a computer simulation model to measure the expected value and variance of prediction errors of four simple methods of estimating loss reserves. Two of these methods are new to the *Proceedings*. The simulated data triangles that are tested are meant to represent sample sizes typically found in individual risk rating situations.

The results indicate that the commonly used age-to-age factor approach gives biased estimates and is inferior to the three other methods tested. Theoretical arguments for the source of this bias and a comparison of two of the methods are presented in the Appendices.

I. INTRODUCTION

The purpose of this paper is to measure the expected value and variance of prediction errors of four simple methods of estimating loss reserves. This is done by using a computer simulation model to generate several thousand different sets of known loss data, applying each estimation method to predict ultimate losses, and then calculating the difference between the predicted and the actual (simulated) ultimate values.[1,2]

Various reserve estimation techniques based on accident year data triangles are described in [2], [5], [6], [7], [20], and [21]. [21] contains a very extensive bibliography. However, the only paper to test the efficiency of the technique it proposes is [6] (and a sample size of only 50 iterations was used).

---

[1] The expected value of the prediction error is referred to as the "bias"; the bias and variance of the prediction error are together referred to as the "efficiency" of the estimation technique.

[2] Results from a previous version of this simulation model were described in [19]. The new computer model is written in Forth and assembly language on an IBM-PC, and is over twenty times faster than the old version written in APL on an IBM 5110 (each iteration now takes 11 to 15 seconds). This allows many more iterations and, therefore, much higher precision in measurements of bias.

The simulated data triangles that are tested here are meant to represent an amount of data that is typically found in individual risk rating, either self insurance programs, or working excess reinsurance treaties (expected values of 40 claims per year and $10,400 per claim). For projecting loss reserves on much larger amounts of data, the statistical variations that are measured with this model will obviously be much less important.

II. AN OVERVIEW OF THE MODEL

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, $\underline{\theta}_a$. Let $\underline{\theta}$ represent a vector (of vectors) containing all the parameters from the first accident year of the experience period through the latest year under consideration (denoted y). So

$$\underline{\theta} = (\underline{\theta}_1,\ldots,\underline{\theta}_y).$$

Let $\underline{K}$ be a vector representing the insured's known loss experience during the experience period.[3] $\underline{K}$ is a random sample drawn from the distributions determined by $\underline{\theta}$.

Let the ultimate losses that a particular insured will have for accident year a be a random variable L_a. The loss reserving and rate making processes both seek to find the "best" estimate of $E(L_a)$.[4,5] $E(L_a)$ is some function of the $\underline{\theta}_a$, whereas the experience $\underline{K}$ was drawn from distributions determined by $\underline{\theta}_1,\ldots,\underline{\theta}_y$. In order for $\underline{K}$ to be useful in estimating $E(L_a)$, there must be some relationship between the $\underline{\theta}$'s for different accident years.

The simplest assumption would be that $\underline{\theta}_1 = \ldots = \underline{\theta}_y$, 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 components of the $\underline{\theta}$'s would be

---

[3] Later in the paper K will be used to denote the familiar loss development triangle matrix, which is a particular way of summarizing the information in $\underline{K}$. K_{aj} denotes the a,j element of K, where a is the accident year.

[4] This paper will only consider estimates of $E(L)$. One might also want to estimate other attributes of the distribution of L, such as $\mathrm{Var}(L)$ or 95% percentile of L.

[5] Actually the loss reserving process seeks to find $E(L_a - B_{a*}|K_{a*},B_{a*})$ where K_{a*} is the total known dollars of loss for accident year a (* denoting the latest known column) and B_{a*} is total paid dollars of loss. Footnote 7 shows that this distinction does not affect the methodology of this paper.

influenced by inflationary trends and by changes in a measurable exposure base, and that, after proper adjustments for these, the parameters would be stable over time. Examples of these type of adjustments are given in [2].

Any experience rating or reserving procedure is an estimator[6] of $E(\underline{L})$; it is some function R of the insured's past known loss and exposure information $\underline{K}$. A perfect reserve estimation procedure for accident year a would be a function R_a such that $R_a(\underline{K}) = E(L_a)$. However, $\underline{K}$ is also a random variable, so fulfilling this condition is not possible, except by chance. We can, however, hope that $R_a(\underline{K})$ is an unbiased estimator of $E(L_a)$, that is, that $E(R_a(\underline{K})) = E(L_a)$.

We would also like $R(\underline{K})$ to be close to $E(L)$, on the average. One common way of expressing this is to minimize $E((R(\underline{K}) - E(L))^2)$, the mean square error, which for an unbiased estimator is equivalent to minimizing $\mathrm{Var}(R(\underline{K}))$. For many simple statistical models, the form of estimator R that satisfies these criteria can be explicitly calculated. This is referred to as a Uniform Minimum Variance Unbiased (UMVU) estimator.[7]

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 L. 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

---

[6] 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. This paper will use the term prediction as a synonym for estimate. Also, note that $\underline{L}$ denotes the vector $(L_1, \ldots, L_N)$; $\underline{R}$ is defined similarly.

[7] In the computer model that follows, the quantities actually being measured are the expected value and variance of the prediction error $(R(\underline{K}) - L)$. Note that:

1. The error of any prediction $R_a(\underline{K})$ of ultimate losses L_a is identical to the error of using $R_a(\underline{K}) - B_a$ to predict necessary loss reserves $L_a - B_a$ so the expected values and variances measured in this paper apply equally well to loss reserves.
2. $E(R(\underline{K}) - L) = E(R(\underline{K})) - E(L) =$ Bias of $R(\underline{K})$
3. $\mathrm{Var}(R(\underline{K}) - L) = \mathrm{Var}(R(\underline{K})) + \mathrm{Var}(L) - 2\,\mathrm{Cov}(R(\underline{K}),L)$

   If L pertains to an accident year for which there is no known experience, then $\mathrm{Cov}\,(R(\underline{K}),L) = 0$ and we are measuring $\mathrm{Var}(R(\underline{K}))$ plus a constant that does not depend on R. If there is some known experience for the accident year—as is typical for loss reserving—then we are not actually measuring $\mathrm{Var}(R(\underline{K}))$; however the variance of the prediction error is actually what we are interested in.

Note that we have dropped the subscript a when not referring to a specific accident year.

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.

III. COMPUTER MODEL

The computer generates six accident years of known loss experience ($\underline{K}^{(i)}$ for the i^{th} iteration) from distributions with fixed parameters. It then applies four estimation techniques to this set of known losses, arriving at four different predictions of $\underline{L}^{(i)}$. The differences between each of the predictions and the actual ultimate losses are stored. This whole process (generating experience, then calculating predictions) is repeated several thousand times—using the same underlying distributions and parameters. It can then be determined how well the estimates $\underline{R}(\underline{K}^{(i)})$ fared as "guesses" of $\underline{L}^{(i)}$ and which estimator function R does the best.

Each iteration produced a set of loss experience for six accident years—($a = 0,\ldots,5$) where five years of development are known for accident year 0, four years of development for accident year 1, etc. 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.

For a single accident year a, 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 ($i=1,\ldots,N$):
M_i = Month of loss within accident year (uniform with minimum = 0, maximum = 11)

[8] The forms of the distributions chosen are somewhat arbitrary, but are consistent with actuarial literature. For negative binomial frequency see [1], [8] and [17]; for lognormal severity see [4], [10], [13], [14], [16] and [18]; for exponential report lags see [15] and [22]. However, it is important to note that, as demonstrated later in the paper, the conclusions are not particularly sensitive to the choice of the underlying loss generation model.

[9] The normal distribution was chosen as a good approximation for the negative binomial, which is more difficult to simulate. Also, N was restricted to be greater than zero.

Q_i = Report lag in months (waiting time between accident date and report date) (exponential with mean = 18 months)

All experience was viewed as being analyzed as of year-end, so a claim would first become known in $\lceil \frac{(M_i + Q_i - 1)}{12}$ years after the accident year.[10]

P_i = Payment lag in months (waiting time between report date and payment date) (exponential with mean = 12 months)[11]

Then the following dates are calculated:

m_i = accident month = $12a + M_i$
r_i = report month = $12a + M_i + Q_i$
p_i = payment month = $12a + M_i + Q_i + P_i$

Note that m_i, r_i, and p_i are fixed dates (where the first month of the first accident year is taken to be 0). M_i, Q_i, and P_i are lags relative to the accident year (a = 0,...,5) in which the simulated claim occurs, and relative to each other.

The random untrended payment amount, C_i, was drawn from a lognormal distribution with mean = \$10,400 and variance = $(\$34,800)^2$.

The final settlement value of the claim is calculated as $C_iT(m_i,p_i)$, where $T(m,p)$ is an inflation factor equal to $\left(\frac{I_{m+p}}{I_o}\right)^{\alpha}\left(\frac{I_m}{I_o}\right)^{1-\alpha}$ and I_k is an inflation index at month k. This inflation model was suggested by Robert Butsic in [9].

So far, the number of claims, and (for each of these claims) the report date, the payment date, and the 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 month for which the reserve was being set.

[10] The APL symbol $\lceil$, referred to as "ceiling," means "the smallest integer greater than or equal to." Note that if $M_i + Q_i < 12$ the claim is reported during the accident year, "zero" years after the accident year.

[11] These parameters for M, P and Q result in the following average age-to-ultimate factors:

	12-ult	24-ult	36-ult	48-ult	60-ult
Incurred	3.72	1.60	1.24	1.11	1.05
Paid	14.29	2.94	1.69	1.30	1.15

For each claim a random Reserve Error, V_i, was drawn from a lognormal distribution with mean = 1, and variance = 2. To calculate the reserve amount, this was multiplied by $C_iT(m_i,r_i)$ where r_i is the month that the claim was first reported (and therefore reserved). Two things should be noted about this model of case reserving: (1) the reserve error is only chosen once for each claim, regardless of how many years it remains open; and, (2) this system, on the average, leads to under-reserving—by (I_p/I_r), the amount of inflation between the report month and the payment month.[12]

The known loss amount at the end of year t on the i^{th} loss from accident year a is

$$k_i(a,t) = \begin{cases} 0 & \text{if } r_i > 12t + 11 \\ C_iV_iT(m_i,r_i) & \text{if } r_i \leq 12t + 11 < p_i \\ C_iT(m_i,p_i) & \text{if } p_i \leq 12t + 11 \end{cases}$$

So the actual ultimate losses are

$$L = \sum_{i=1}^{N} C_iT(m_i,p_i)$$

The full experience matrix known at the end of year four for an insured would be

$$\begin{pmatrix} \sum_{i=1}^{N_0} k_i(0,0) & \cdots & \sum_{i=1}^{N_0} k_i(0,4) \\ \vdots & & 0 \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{N_Y} k_i(4,4) & 0 \cdots & 0 \end{pmatrix}$$

This represents the familiar "loss development triangle." We will denote such an experience matrix by K (for known data).

---

[12] The author admits that this is a crude model of the case reserving process; however, it is unlikely that a more sophisticated model would significantly affect the results—unless it was one that allowed for changes in relative reserve adequacy along the diagonal. A method of setting reserves at V times the ultimate payment, which does not lead to under-reserving, was tested in [19], and it did not make a significant difference in the results. Also, see Section VI on sensitivity tests.

The matrix K is the statistic that we will use to estimate the vector of expected final loss amounts $E(\underline{L})$. Note that there are many other possible statistics we could have chosen (such as a triangle of claim counts, or a triangle of losses truncated at some "basic limits" point). Other such statistics would probably allow us to construct more efficient estimators—in fact, they definitely would unless K happened to be a "sufficient statistic" for $E(\underline{L})$, and there is no reason to believe that it is sufficient.

IV. RATING METHODS

Once the experience matrix K is calculated for one iteration, it is used as input for four different rating techniques (estimators of $E(L)$).

Let K_{aj} = Losses for accident year a known through period j (in other words, the aj element of matrix K)

K_{a*} = Latest known losses for accident year a

f_a = The age-to-ultimate factor for accident year a[13]

R_a = The estimate of expected ultimate losses, $E(L_a)$

1: Age-to-Age Factors

This is the very common procedure of projecting each accident year to its ultimate value by age-to-age factors (also known as the "chain ladder" method). So

$$R_a = K_{a*} f_a \qquad a = 0,\ldots,4$$

R_5 undefined (because $K_{5*} = 0$)

2: Modified Bornhuetter-Ferguson

This is a modified version of a commonly used method first presented in [5].[14]

$$R_a = K_{a*} + R_5\left(1 - \frac{1}{f_a}\right) \qquad a = 0,\ldots,4$$

[13] Age-to-age factors throughout this paper are calculated by summing corresponding elements in two adjacent columns of the triangle, then dividing these two sums. This is usually superior, as shown in [12], to taking a straight average of the individual age-to-age factors, which is likely to produce substantial additional bias.

[14] In [5] R_5 was obtained from external sources, rather than as shown here.

$$R_5 = (1/5) \sum_{h=0}^{4} K_{h*}f_h$$

3: Adjustment to Total Known Losses

This method (also referred to as the "Cape Cod method")is described in [7] and [19]. Appendix B presents a theoretical comparison of R_5 under this method with R_5 under method 2. It consists of averaging the known losses first, then applying an adjustment factor to the sum.

$$R_a = K_{a*} + R_5\left(1 - \frac{1}{f_a}\right) \qquad a = 0,\ldots,4$$

$$R_5 = \left(\sum_{a=0}^{4} K_{a*}\right) \div \left(\sum_{a=0}^{4} (1/f_a)\right)$$

4: Additive Model

Let K' denote the matrix of known loss experience where each cell is the losses incurred *during* a particular period (rather than cumulative losses *through* the period, as the matrix K denotes). The elements of K' are the differences of adjacent columns of K.

Project the unreported losses for an accident year as the sum of the expected unreported losses during each future period. Estimate the expected unreported losses by period as the average of the known losses by row. Specifically,

$$R_a = K_{a*} + \sum_{h=5-a}^{4} \frac{1}{a} \sum_{g=0}^{a-1} K'_{gh} \qquad a = 1,\ldots,5$$

$$R_0 = K_{0*}$$

This additive method is suggested by Hans Bühlmann [7]; he refers to it as the complementary loss ratio method.

V. RESULTS

Each of the four rating methods was tested under each of the following progressively more complex loss generation models. Exhibits I through V display the results for each model. These exhibits show the mean and standard deviation of the prediction error for each rating method for each accident year. The prediction error is $R_a - L_a$ (the estimated ultimate result minus the actual ultimate result). The "% of actual" is the prediction error divided by the true expected losses.

We would expect any rating technique based on known data to (on the average) under-predict by the expected amount of development between the most mature known data amount and ultimate $E(K_{a4} - L_a)$. Therefore, each of the expected prediction errors has been adjusted by this amount, so the exhibits actually show $E(R_a - L_a) - E(K_{a4} - L_a) = E(R_a - K_{a4})$. That is, we do not expect the estimation techniques to be able to predict beyond the triangle.[15]

EXHIBIT I—Claim Counts Only, No Inflation

In this version of the model, C_i was not randomly chosen, but was set at $1. The inflation index I_m was also held constant. The results show that simple age-to-age factors produced biased results and higher standard deviations. Methods 2 and 3 have very slight biases while method 4 is unbiased. Methods 3 and 4 have slightly smaller standard deviations than method 2.

What is interesting here is not the amount of the bias (which for practical purposes is negligible), but the fact that there *is* a bias. This fact was greeted with surprise and skepticism by many actuaries when it was first presented in [19]. Appendix A gives a technical argument to support this result.

EXHIBIT II—Random Claim Size, No Inflation

In this version, C_i is randomly chosen from a lognormal distribution with mean = $10,400 and variance = ($34,800)^2. The inflation index I_m was held constant. Here we see that method 1 is clearly inferior—it is significantly biased upward and has very high standard deviations in years 3 and 4. An interesting result from the older version of the model is that the median prediction error for method 1 was usually negative—that means that in over half of the cases method 1 under-predicted the actual (simulated) results, but a few cases of large over-predictions made the mean prediction error (the bias) positive. This is because the distribution of prediction errors for method 1 was very positively skewed. Method 3 has the lowest standard deviation. Methods 3 and 4 do not appear to have significant biases.

---

[15] A technique of estimating the parameters of the distribution of Q_i directly, such as described in [22], would allow prediction beyond the triangle.

EXHIBIT III—Constant 8% Inflation, $\alpha = 0.5$

In this version, an 8% per year inflation was assumed, with 50% applying to date of accident and 50% applying to date of settlement. Here we expect that methods 2, 3 and 4 will under-predict, because they all implicitly assume that expected losses by accident year are the same, which, with inflation, is not true. Method 1 does not rely on such an assumption.

The addition of inflation accentuates the bias in method 1, making its predictions 35% above the actual values (after the "tail adjustment"). Method 2 does very well on this example because the upward bias inherent in each age-to-ultimate prediction is balanced by the fact that the method assumes no inflation. Once again methods 3 and 4 do the best in terms of standard deviation, but, as expected, they are somewhat biased downward.

EXHIBIT IV—Constant 8% Inflation, $\alpha = 0.5$,
Adjust Rating Methods for 8% Inflation

This version was run with the same loss parameters and inflation assumptions as model III. However, each of the rating methods was modified as follows:

> Each element of each row, where an arbitrary row is row a, was divided by an assumed inflation index I'_a. The rating method was applied to the resulting triangle, then each projected ultimate result was multiplied by its respective I'_a. In this case I'_a was set as 1.08^a, $a = 0,\ldots,5$. This obviously represents perfect clairvoyance about the underlying past and future inflation rate.[16]

This slightly improves the standard deviation of method 1, but does not improve the bias, which is still quite high. However, this adjustment completely removes the bias on method 4, and leaves only a slight upward bias in method 3.

EXHIBIT V—10% Inflation Dropping to 6%, $\alpha = 0.5$,
Adjust Rating Methods for 10% Inflation

In this version, the actual inflation rate was 10% for 60 months (which covers the entire known claim period), then it drops to 6%. The index assumed by the rating methods is $(1.10)^a$.

[16] Note that a similar adjustment can be made when dealing with a triangle where the exposure varies by accident year, i.e., (1) divide each row by the corresponding exposure, (2) apply the rating method, then (3) multiply each estimate by its exposure. This could be further improved by using credibility weighted averages in the rating method, where a row's credibility was a function of its exposure; however, developing such a system is beyond the scope of this paper.

This results in only a slight bias in method 4, and a fairly small one in method 3.

VI. SENSITIVITY TESTING

As a test of the sensitivity of the results to the specific distributions used to generate loss experience, the following additional three scenarios were run. Note that these were all run with an assumption of no inflation, so they are meant to be compared with the results on Exhibit II (which will be referred to as the "standard model").

EXHIBIT VI—No Reserve Development

The standard model was used except that the reserve error, V, was always set equal to one.

EXHIBIT VII—Uniform Frequency and Severity

The standard model was used except that the frequency, N, was distributed discrete uniform [1,79] and severity, C, was distributed continuous uniform [0,20800]. This results in an ultimate aggregate loss distribution with about the same mean and variance, but much less skewness, than the standard model.

EXHIBIT VIII—Uniform Report and Payment Lags

The standard model was used except that the report lag, Q, was distributed discrete uniform [0,36] and payment lag, P, was distributed discrete uniform [0,24]. This results in the same average lags, but with a higher percentage of claims being reported and paid within the five columns of the experience triangle than the standard model.

Although the magnitudes of the biases and standard deviations differ in Exhibits VI through VIII from Exhibit II, conclusions about the existence of bias and about the relative efficiency of the four rating methods remain substantially unchanged.

VII. CONCLUSIONS

These results indicate that for data triangles of the size tested:

1. The common age-to-age factor approach (method 1) is clearly inferior to the other three methods.

2. The additive method 4 and the average-then-adjust method 3 have significantly lower variances than methods 1 and 2, and small biases (if adjusted for inflation). In fact, method 4 may be completely unbiased.

It is important to emphasize that the bias of the various methods is heavily influenced by a few large prediction errors. This means that in practical ratemaking situations it would usually be wrong to use method 1 and then do a judgment "bias adjustment"—doing so in most cases would result in under predicting. Instead, the practitioner simply should not put much credibility in predictions based on highly leveraged age-to-ultimate factors.

One may object that allowing accurate knowledge of the underlying inflation rate gives an unfair advantage to methods 2 through 4, because it allows all of the rows of the triangle to be used in estimating any particular row's ultimate value. However, one will normally have exogenous knowledge of past inflation rates and forecasts of future rates, and using this information should improve one's ability to predict. Also, in [19] it was shown that attempts to estimate the trend rate solely from data samples of this size by fitting lines to projected ultimate values produced terrible results—extreme bias, variance, and skewness.

The above major conclusions concern the relative ranking of techniques and the existence in some cases of bias. These conclusions were found to be robust to an extreme change in the form of the underlying distributions; this robustness was also found in [19]. Of course, the specific numerical results on Exhibits I through VIII should not be considered to be any more than examples—changing the parameters or the form of the loss generating model will change these in unpredictable ways.

One way in which numerical results from a model such as this would be of interest is if the parameters of the loss generation model were estimated from an actual data set which had been projected to ultimate by a specific loss reserving technique. Simulating the distribution of prediction error would give an estimate of the potential variability of the reserve estimate—which could be used to calculate confidence intervals (containing both "parameter" and "process" risk) for the loss reserve.

REFERENCES

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

[2] James R. Berquist and Richard E. Sherman, "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," *PCAS* LXIV, 1977, p. 123.

[3] Peter Bickel and Kjell Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics,* Holden-Day, 1977.

[4] David R. Bickerstaff, "Automobile Collision Deductibles and Repair Cost Groups: The Lognormal Model," *PCAS* LIX, 1972, p. 68.

[5] Ronald L. Bornhuetter and Ronald E. Ferguson, "The Actuary and IBNR," *PCAS* LIX, 1972, p. 181.

[6] Hans Bühlmann, Rene Schnieper, and Erwin Straub, "Claims Reserves in Casualty Insurance Based on a Probabilistic Model," *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker 80,* 1980.

[7] Hans Bühlmann, "Estimation of IBNR Reserves by the Methods Chain Ladder, Cape Cod and Complementary Loss Ratio," unpublished, 1983.

[8] Hans Bühlmann, *Mathematical Methods in Risk Theory,* Springer-Verlag, 1970.

[9] Robert Butsic, "The Effect of Inflation on Losses and Premiums for Property-Liability Insurers," *Inflation Implications for Property-Casualty Insurance,* 1981 Casualty Actuarial Society Discussion Paper Program, p. 58.

[10] Robert J. Finger, "Estimating Pure Premium by Layer—An Approach," *PCAS* LXIII, 1976, p. 34.

[11] George Fishman, *Concepts and Methods in Discrete Event Digital Simulation,* John Wiley & Sons, 1973.

[12] Charles A. Hachemeister and James N. Stanard, "IBNR Claims Count Estimation with Static Lag Functions," unpublished, 1975.

[13] Charles C. Hewitt, Jr. and Benjamin Lefkowitz, "Methods for Fitting Distributions to Insurance Loss Data," *PCAS* LXVI, 1979, p. 139.

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

[15] Charles L. McClenahan, "A Mathematical Model for Loss Reserve Analysis," *PCAS* LXII, 1975, p. 134.

[16] Robert S. Miccolis, "On the Theory of Increased Limits and Excess of Loss Pricing," *PCAS* LXIV, 1977, p. 27.

[17] Gary Patrik and Russell John, "Pricing Excess-of-Loss Casualty Working Cover Reinsurance Treaties," *Pricing Property and Casualty Insurance Products*, 1980 Casualty Actuarial Society Discussion Paper Program, p. 399.

[18] Sheldon Rosenberg and Aaron Halpert, "Adjusting Size of Loss Distributions for Trend," *Inflation Implications for Property-Casualty Insurance*, 1981 Casualty Actuarial Society Discussion Paper Program, p. 458.

[19] James N. Stanard, "Experience Rates as Estimators: A Simulation of their Bias and Variance," *Pricing Property and Casualty Insurance Products*, 1980 Casualty Actuarial Society Discussion Paper Program, and Review by John Robertson, p. 485.

[20] A. Tenenbein and D. Covall, "Determination of Development Factors for Automobile Insurance Rates," New York University Graduate Business School Working Paper #79-129, 1979.

[21] J. Van Eeghan and G. de Wit, *Loss Reserving Methods*, Surveys of Actuarial Studies No. 1, Nationale-Nederlanden N.V., 1981.

[22] Edward W. Weissner, "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood," *PCAS* LXV, 1978, p. 1.

EXHIBIT I

MODEL I—CLAIM COUNTS ONLY, NO INFLATION
5000 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
		Counts	% of Actual	Counts	% of Actual
1	0	0	0%	1.5	4%
	1	0.1	0	2.6	6
	2	0.2	1	3.7	9
	3	0.3	1	5.8	14
	4	0.9	2	11.6	29
	5	—	—	—	—
2	0	0	0%	1.5	4%
	1	0.1	0	2.5	6
	2	0.2	1	3.6	9
	3	0.2	1	5.1	13
	4	0.3	1	7.2	18
	5	0.4	1	8.8	22
3	0	0	0%	1.5	4%
	1	0.1	0	2.5	6
	2	0.2	0	3.5	9
	3	0.1	0	5.0	13
	4	0.1	0	7.1	18
	5	0.2	0	8.6	22
4	0	0	0%	1.5	4%
	1	0	0	2.5	6
	2	0.1	0	3.5	9
	3	0	0	5.0	13
	4	0	0	7.2	18
	5	0.1	0	8.6	22

EXHIBIT II

Model II—Random Claim Size, No Inflation
5000 Iterations

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$ — Mean: Dollars	Mean: % of Actual	Standard Deviation: Dollars	Standard Deviation: % of Actual
1	0	$ 0	0%	$ 88,600	21%
	1	9,892	2	182,206	44
	2	24,680	6	252,951	61
	3	49,766	12	392,435	94
	4	113,397	27	823,429	198
	5	—	—	—	
2	0	$ 0	0%	$ 88,600	21%
	1	9,354	2	177,605	43
	2	16,234	4	412,028	99
	3	29,183	7	303,322	73
	4	32,183	8	377,037	90
	5	36,314	9	372,499	89
3	0	$ 0	0%	$ 88,600	21%
	1	5,712	1	163,078	39
	2	13,138	3	212,171	51
	3	14,501	3	263,962	63
	4	4,662	1	320,142	77
	5	4,370	1	322,794	77
4	0	$ 0	0%	$ 88,600	21%
	1	−894	0	170,483	41
	2	−4,787	−1	438,705	105
	3	−3,986	−1	290,293	70
	4	−11,622	−3	338,545	81
	5	−7,490	−2	341,970	82

EXHIBIT III

MODEL III—8% INFLATION, $\alpha = 0.5$
15,000 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$ — Mean: Dollars	Mean: % of Actual	Standard Deviation: Dollars	Standard Deviation: % of Actual
1	0	$ 0	0%	$132,643	28%
	1	13,325	3	233,786	45
	2	40,012	7	528,989	95
	3	75,972	13	674,655	113
	4	225,406	35	1,636,846	254
	5	—	—	—	—
2	0	$ 0	0%	$132,643	28%
	1	18,162	4	281,171	54
	2	35,581	6	376,524	68
	3	37,095	6	498,673	83
	4	15,500	2	639,790	99
	5	−62,654	−9	609,556	87
3	0	$ 0	0%	$132,643	28%
	1	9,766	2	194,158	38
	2	15,783	3	280,995	51
	3	−607	0	385,999	65
	4	−49,904	−8	451,253	70
	5	−138,589	−20	450,961	64
4	0	$ 0	0%	$132,643	28%
	1	−2,462	−1	185,358	36
	2	−8,613	−2	273,372	49
	3	−32,982	−6	363,169	61
	4	−80,318	−13	423,457	66
	5	−158,472	−23	441,974	63

EXHIBIT IV

MODEL IV—8% INFLATION, $\alpha = 0.5$, 8% INDEX USED IN RATING
12,750 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
		Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$108,748	23%
	1	17,024	3	226,684	44
	2	41,313	7	368,332	66
	3	81,257	14	627,023	104
	4	214,678	33	1,545,088	240
	5	—	—	—	—
2	0	$ 0	0%	$108,748	23%
	1	17,021	3	242,467	47
	2	34,663	6	328,942	59
	3	56,512	9	486,315	81
	4	75,782	12	597,070	93
	5	83,162	12	640,675	92
3	0	$ 0	0%	$108,748	23%
	1	12,228	2	209,716	41
	2	22,240	4	284,919	51
	3	30,927	5	398,680	66
	4	27,951	4	451,586	70
	5	24,978	4	496,771	71
4	0	$ 0	0%	$108,748	23%
	1	1,546	0	228,070	44
	2	3,571	0	289,117	52
	3	6,014	1	405,582	68
	4	4,862	1	433,385	67
	5	6,569	1	492,804	71

EXHIBIT V

Model V—10% Inflation Dropping To 6%, $\alpha = 0.5$, 10% Index Used in Rating
8000 Iterations

		Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
Rating Method	Accident Year	Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$120,911	25%
	1	13,748	3	260,039	48
	2	34,538	6	397,028	69
	3	79,547	13	569,751	90
	4	227,292	33	1,331,666	193
	5	—	—	—	—
2	0	$ 0	0%	$120,911	25%
	1	12,627	2	243,327	45
	2	27,992	5	344,815	60
	3	54,273	9	446,620	70
	4	89,787	13	577,771	84
	5	108,456	15	617,252	83
3	0	$ 0	0%	$120,911	25%
	1	8,522	2	225,986	42
	2	17,345	3	310,728	54
	3	30,093	5	393,830	62
	4	42,842	6	481,436	70
	5	49,802	7	519,365	70
4	0	$ 0	0%	$120,911	25%
	1	−1,386	0	230,404	43
	2	−672	0	320,637	56
	3	6,152	1	393,860	62
	4	19,540	3	469,071	68
	5	31,185	4	516,946	69

EXHIBIT VI

SENSITIVITY TEST—PERFECT CASE RESERVING
6522 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$ Mean Dollars	Mean % of Actual	Standard Deviation Dollars	Standard Deviation % of Actual
1	0	$ 0	0%	$ 50,557	12%
	1	5,092	1	117,095	28
	2	12,428	3	147,644	35
	3	19,602	5	231,091	56
	4	67,650	16	504,934	122
	5	—	—	—	—
2	0	$ 0	0%	$ 50,557	12%
	1	4,726	1	109,725	26
	2	10,880	3	130,699	31
	3	14,206	3	186,497	45
	4	22,904	6	251,845	61
	5	17,081	4	305,529	73
3	0	$ 0	0%	$ 50,557	12%
	1	2,898	1	104,576	25
	2	6,574	2	120,845	29
	3	6,057	1	172,281	41
	4	8,546	2	229,368	56
	5	942	0	281,352	67
4	0	$ 0	0%	$ 50,557	12%
	1	−542	0	101,240	24
	2	899	0	116,160	28
	3	−658	0	168,233	41
	4	2,361	1	225,834	55
	5	−3,461	−1	278,649	65

EXHIBIT VII

SENSITIVITY TEST—UNIFORM FREQUENCY AND SEVERITY
3049 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
		Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$ 34,704	8%
	1	9,824	2	175,539	43
	2	8,593	2	155,670	37
	3	14,083	3	195,329	46
	4	44,232	11	331,371	79
	5	—	—	—	—
2	0	$ 0	0%	$ 34,704	8%
	1	7,410	2	128,228	31
	2	7,579	2	143,738	35
	3	8,791	2	168,244	40
	4	15,059	4	222,905	53
	5	20,915	5	287,287	70
3	0	$ 0	0%	$ 34,704	8%
	1	3,690	1	74,962	18
	2	2,208	1	97,500	23
	3	1,123	0	130,512	31
	4	3,353	1	204,091	49
	5	7,643	2	263,368	64
4	0	$ 0	0%	$ 34,704	8%
	1	342	0	65,336	16
	2	−2,937	−1	94,379	23
	3	−4,874	−1	133,405	32
	4	−2,227	−1	209,730	50
	5	3,629	1	264,368	65

EXHIBIT VIII

SENSITIVITY TEST—UNIFORM REPORT AND PAYMENT LAGS
6000 ITERATIONS

		Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
Rating Method	Accident Year	Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$ 33,898	0%
	1	8,958	2	154,927	38
	2	40,911	10	371,052	89
	3	94,371	23	667,695	162
	4	261,076	63	1,627,441	393
	5	—	—	—	—
2	0	$ 0	0%	$ 33,898	8%
	1	7,452	2	183,330	45
	2	39,655	10	348,887	84
	3	66,107	16	447,484	109
	4	79,065	19	499,807	121
	5	81,733	20	481,770	117
3	0	$ 0	0%	$ 33,898	8%
	1	5,097	1	153,015	37
	2	21,913	5	279,507	67
	3	23,616	6	321,873	78
	4	14,848	4	301,852	73
	5	13,765	3	297,400	72
4	0	$ 0	0%	$ 33,898	8%
	1	−3,670	−1	195,110	47
	2	4,794	1	276,092	66
	3	3,478	1	306,744	75
	4	−1,347	0	292,257	71
	5	1,321	0	293,734	71

APPENDIX A

AN ANALYTICAL ARGUMENT FOR BIAS OF AGE-TO-AGE FACTORS

Consider Model 1, (i.e., claim counts only and no inflation). Each row (accident year) of the data triangle K is independently and identically distributed with each other row.

This implies that $E[g(X_{ij+1},X_{ij})] = E[g(X_{kj+1},X_{kj})]$ $\forall i,k$ for any function g. However $E[g(X_{ij+1},X_{ij})] \neq g(E[X_{ij+1}],E[X_{ij}])$ unless g is linear.

Let $g(X_{ij+1},X_{ij}) = \dfrac{X_{ij+1}}{X_{ij}}$

Let f_{ij} be an age-to-age factor estimated from row i. Age to age factors attempt to estimate $E[X_{kj+1}|X_{kj}]$ with $X_{kj}f_{ij}$. If this estimate were unbiased it would mean that

$$E[E[X_{kj+1}|X_{kj}]] = E[X_{kj}f_{ij}]$$

But this becomes

$$E[X_{kj+1}] = E[X_{kj}]E[f_{ij}]$$

$$E[X_{kj+1}] = E[X_{kj}]E\left[\frac{X_{ij+1}}{X_{ij}}\right]$$

$$E[X_{kj+1}] = E[X_{kj}]E\left[\frac{X_{kj+1}}{X_{kj}}\right]$$

or

$$\frac{E[X_{kj+1}]}{E[X_{kj}]} = E\left[\frac{X_{kj+1}}{X_{kj}}\right]$$

which is not true in general.[17]

[17] A similar derivation was arrived at independently by John Robertson.

APPENDIX B

COMPARISON OF "ADJUSTING, THEN AVERAGING" VERSUS "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 $X_i = \dfrac{\mu}{a_i} + e_i \qquad i = 1, \ldots, n$

where μ = underlying expected losses
a_i = non-random adjustment factor
e_i = random error $E(e_i) = 0$, $\mathrm{Var}(e_i) = \sigma_i^2$

We wish to estimate μ.

Let $\hat{\mu}_1 = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i a_i$

This represents trending (and/or developing) known losses for each year and averaging the results.

Let $\hat{\mu}_2 = \left(\displaystyle\sum_{i=1}^{n} X_i\right) \div \left(\displaystyle\sum_{i=1}^{n} \frac{1}{a_i}\right)$.

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, i.e. $E(\hat{\mu}_1) = E(\hat{\mu}_2) = \mu$. (It is important to note that this only holds if a_i is non-random, which is not the case in real estimation problems.)

Calculate the Best Linear Unbiased Estimate (B.L.U.E.)[18] of μ. That is, find weights c_i, such that $\hat{\mu}\left(= \displaystyle\sum_{i=1}^{n} c_i X_i\right)$ is unbiased and has minimum variance. So, minimize $\mathrm{Var}\left(\displaystyle\sum_{i=1}^{n} c_i X_i\right)$ subject to $E\left[\displaystyle\sum_{i=1}^{n} c_i X_i\right] = \mu$

$$\mathrm{Var}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i^2 \sigma_i^2$$

[18] The approach of calculating the B.L.U.E. was suggested by Aaron Tenenbein.

$$E\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} \frac{c_i}{a_i}\mu = \mu \Rightarrow \sum_{i=1}^{n} \frac{c_i}{a_i} = 1$$

Let

$$L = \sum_{i=1}^{n} c_i^2 \sigma_i^2 + \lambda \left(1 - \sum_{i=1}^{n} \frac{c_i}{a_i}\right)$$

$$\frac{\partial L}{\partial c_i} = 2c_i\sigma_i^2 - \frac{\lambda}{a_i} = 0 \quad i = 1,\ldots,n$$

So $c_i = \dfrac{\lambda}{2a_i\sigma_i^2}$, $\quad \displaystyle\sum_{i=1}^{n} \frac{c_i}{a_i} = \sum_{i=1}^{n} \frac{\lambda}{2a_i^2\sigma_i^2}$

So $\lambda = \dfrac{2}{\displaystyle\sum_{i=1}^{n} \frac{1}{a_i^2\sigma_i^2}}$

So $c_i = \dfrac{1}{a_i\sigma_i^2} \dfrac{1}{\displaystyle\sum_{j=1}^{n} \frac{1}{a_j^2\sigma_j^2}}$

Now consider various possibilities for σ_i^2:

1. Let $X_i a_i = \mu + \epsilon_i$ where $\text{Var}(\epsilon_i) = \sigma^2 \; \forall_i$

   This means that $e_i = \dfrac{\epsilon_i}{a_i}$, so $\sigma_i^2 = \dfrac{1}{a_i^2}\sigma^2$, so $c_i = a_i/n$

   Therefore $\hat{\mu}_1$ is the BLUE.

2. Let $\dfrac{\text{Var}(X_i)}{E[X_i]} = k \quad \forall_i$

   So $\dfrac{\sigma_i^2}{\mu/a_i} = \dfrac{\sigma_i^2 a_i}{\mu} \Rightarrow \sigma_i^2 a_i = k\mu$

   This means that $c_i = 1 \Big/ \displaystyle\sum_{i=1}^{n} \frac{1}{a_i}$

   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.

DISCUSSION BY JOHN P. ROBERTSON

Mr. Stanard's paper offers the reader three things:
1) reserving techniques;
2) a methodology for assessing reserving techniques; and
3) conclusions about the reserving techniques.

Of these three, the methodology for assessing reserving techniques is the most significant. This methodology consists of developing a model of the loss emergence process and then simulating this process, applying the various reserving techniques, and keeping score of the results. This methodology is important because it is the most scientific system yet presented for assessing the validity and the accuracy of alternative reserving techniques. It is a general method, as readily applied to other models of the claim emergence process as to the model used in the paper.

The reserving methods Mr. Stanard presents are fundamental to casualty actuarial work. He is "filling out" familiar loss triangles and forecasting the next year's result. This is obviously the basis for most reserving methods and is also a key part of most ratemaking.

Previous literature on reserving techniques generally has concentrated on overcoming the effects of changes in the underlying mix of business, changes in the individual claim reserving and settling policies, and changes in claims reporting systems. Most of this prior literature assumes that once these changes are accounted for and the data has been restated so as to have relatively constant underlying conditions, then any number of loss development methods can be applied to obtain valid forecasts.

For instance, in Berquist and Sherman [1], examples are given of adjusting historical data to eliminate the effects of changes in the relative adequacy of case reserves and to eliminate the effects of changes in the rate of settlement of claims. Following these adjustments, standard loss development methods are applied with no question being raised as to the validity of these methods. Clearly, making adjustments for changes in the mix of business, etc., is an important part of reserve analysis; but the question of the validity of reserving methods, even in the face of completely uniform historical conditions, is also an important one.

Prior to this and Mr. Stanard's previous paper [2], there have only been a handful of attempts at evaluating reserving techniques. In one of these, Professor

Bühlmann, et al., sharply contrast the bases for development of reserving techniques between life and casualty actuaries [3]:

> "Since the early days of Life Insurance it has been understood that 'reserves for future payments of claims . . . had to be calculated from the probabilistic model describing the process of death within a specified population.'. . . . Strangely enough when actuaries were asked to put their skill to work in Non Life Insurance, they did not feel it necessary to have a probabilistic model for the setting of claims reserves. . . . The reason for the absence of probabilistic models leading to reserving techniques in Casualty Insurance may be explained (to some extent) by the common fashion in this field of assuming the individual claim amount to 'occur' suddenly even if in practice it is delayed portionwise over long periods of time. This paper takes exception to this fashion and models the individual claim amount as a random process over time."

Professor Bühlmann, et al., then proceed to develop a stochastic model of the claims process and to test several reserve estimation techniques against this model. They draw no conclusion about possible bias of the various methods, but do observe that the standard deviations of all the methods they consider seem quite high, and offer the opinion that the search for better methods should continue. They cite [4] and [5] as papers also exploring the validity of loss reserving methods based on stochastic models of the claims process.

It is easy to criticize Mr. Stanard's model of the loss development process as being too simple to be realistic. He only allows three sources of loss development: 1) late reporting of claims, 2) inflation from the the time a claim reserve is opened to the time the claim is settled, and 3) random variation between the estimated value of the claim and the final value of the claim. In particular, he does not allow for changes in the estimated value of a claim while the claim remains open, nor does he allow for any systematic development in the value of a claim, except for that due to inflation.

Does use of such a simple model invalidate Mr. Stanard's results? I think not. Any of the features which would make his model more realistic, i.e., more complicated, might just as well add to the biases and variances as they might subtract from them. If, for example, standard loss development methods really work so well, they should work in artificially simplified situations. The fact that Mr. Stanard has presented a situation where the standard loss development methods are biased may not quite prove that they fail in other more realistic situations, but it does show that they need to be tested and justified in relation to possible models of the claims development process they are used to forecast.

I continue to find the "Adjustment to Total Known Losses" or "Cape Cod" technique to be of interest. In addition to the possible advantages pointed out as a result of the simulations and in Appendix B, this technique complements the Bornhuetter-Ferguson technique in a way no other technique can, as discussed below.

Consider the case where there is no change in real exposure from year to year and there is no inflation (or past years' losses have been adjusted to eliminate these effects). Then an obvious estimator for R_5 is the average of R_0 to R_4, or $(1/5)(R_0 + \ldots + R_4)$. In Mr. Stanard's paper, both the "Modified Bornhuetter-Ferguson" method and the "Adjustment to Total Known Losses" method start by computing R_5. The former uses the formula:

$$R_5 = \frac{1}{5} \sum_{h=0}^{4} K_h \times f_h.$$

The latter computes R_5 by:

$$R_5 = \left(\sum_{a=0}^{4} K_{a*} \right) \div \left(\sum_{a=0}^{4} \frac{1}{f_a} \right) .$$

In each method, this value of R_5 is used to calculate R_0 through R_4. Once R_0 to R_4 are computed, their average can be compared to R_5. Under the "Adjustment to Total Known Losses" method, this average will always be exactly R_5. A proof of this is given in the Appendix to this discussion. Under the "Modified Bornhuetter-Ferguson" method, this average will not necessarily equal R_5. The consistency between the original estimate of R_5 and the average of R_0 to R_4 in the "Adjustment to Total Known Losses" method indicates, I believe, that this method makes the best use of loss information from all the years in order to project any given year. If the average of R_0 to R_4 is less than R_5 then one could argue that too high an R_5 had been selected, as reported development would appear to be occurring at a lower rate than predicted by R_5. The converse argument could be made if the average were higher than R_5. This inconsistency cannot happen under the "Adjustment to Total Known Losses" method.

It may be that there is reason to choose an R_5 from external sources or by some other method when the Bornhuetter-Ferguson method is being used. But in situations where one is estimating R_5 from the loss information, the consistency discussed above argues strongly for the use of the "Adjustment to Total Known Losses" method.

In conclusion, I believe Mr. Stanard's paper offers a valuable method for assessing whether common actuarial methods are accurate and reliable. As actuaries are called upon to look at smaller and smaller insurance, reinsurance, and self-insurance programs, and as determination of confidence levels for reserves becomes more important, then the usefulness of the methods in this paper should become more apparent. Additionally, the conclusions reached should spur development of improved models of the loss development process and improved reserving techniques.

REFERENCES

[1] James R. Berquist and Richard E. Sherman, "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," *PCAS* LXIV, 1977.

[2] James N. Stanard, "Experience Rates as Estimators: A Simulation of Their Bias Variance," *Pricing Property and Casualty Insurance Products*, Casualty Actuarial Society, 1980 Discussion Paper Program, p. 485.

[3] Hans Bühlmann, René Schnieper, and Erwin Straub, "Claims Reserves in Casualty Insurance Based on a Probabilistic Model," *Mitterlungen der Vereinigung Schweizerischer Verucherungsmathematiker* 80, 1980.

[4] F. DeVylder, "Estimation of IBNR Claims by Least Squares," *Bulletin of the Association of Swiss Actuaries*, Vol 78/2, 1978.

[5] Charles Hachemeister, "A Stochastic Model for Loss Reserving," *Transactions of the 21st International Congress of Actuaries*, Vol 1, 1980.

APPENDIX

Purpose

This table will show that $R_5 = \frac{1}{5}(R_0 + \ldots + R_4)$ for the "Adjustment to Total Known Losses" method, as claimed in the review.

Proof

Given:

$$(1)\ R_5 = \left(\sum_{a=0}^{4} K_{a*}\right) \Big/ \left(\sum_{a=0}^{4} \frac{1}{f_a}\right)$$

$$(2)\ R_a = K_{a*} + R_5\left(1 - \frac{1}{f_a}\right); \qquad a = 0 \text{ to } 4$$

Then:

$$\frac{1}{5}(R_0 + \ldots + R_4)$$

$$= \frac{1}{5}\left[K_{0*} + R_5\left(1 - \frac{1}{f_0}\right) + \ldots + K_{4*} + R_5\left(1 - \frac{1}{f_4}\right)\right] \qquad \text{(By (2))}$$

$$= \frac{1}{5}\left[K_{0*} + \ldots\ K_{4*} + 5R_5 - R_5\left(\frac{1}{f_0} + \ldots + \frac{1}{f_4}\right)\right] \qquad \text{(Rearranging)}$$

$$= R_5 + \frac{1}{5}\left(K_{0*} + \ldots + K_{4*} - R_5\left(\frac{1}{f_0} + \ldots + \frac{1}{f_4}\right)\right) \qquad \text{(Rearranging)}$$

$$= R_5 + \frac{1}{5}\left(K_{0*} + \ldots + K_{4*} - \frac{K_{0*} + \ldots\ + K_{4*}}{(1/f_0) + \ldots\ (1/f_4)} \times \left(\frac{1}{f_0} + \ldots + \frac{1}{f_4}\right)\right) \qquad \text{(By (1))}$$

$$= R_5 \qquad \text{(Cancelling)}$$

Q.E.D.

LOSS PORTFOLIOS: FINANCIAL REINSURANCE*

LEE R. STEENECK

Abstract

The property-casualty insurance operating environment has changed dramatically. Total return is more a function of investment results than ever before. Competition has pressured rate levels. And a greater proportion of total premiums is coming from "long tail" lines, making reserving more difficult.

Reinsurance is becoming somewhat more financially oriented. Loss portfolio transfer reinsurance is becoming popular for a variety of reasons, not the least of which involves poor operating results. This paper surveys loss portfolio transfer reinsurance from a benefit-cost standpoint and includes actuarial, tax, accounting and contractual aspects necessary to the evaluation process.

With the advent of high interest rates and cash flow underwriting, composite ratios have skyrocketed to unprecedented high levels. For a variety of reasons, insurance executives seeking to improve results are investigating loss portfolio transfer reinsurance.

In the simplest terms, this form of financial reinsurance involves the transfer of a portfolio of loss liabilities from a cedent to a reinsurer at a price. The cedent extinguishes his liability with a favorable cash (or equivalent) outlay. The consideration is generally based on a discounted cash flow analysis of loss reserves plus a reinsurer loading. The amount by which the extinguished liability exceeds the consideration becomes a financial benefit to the cedent. The loss liability may be for case reserves only, case reserves plus development, or case reserves plus development and IBNR losses. The transfer can include allocated and sometimes unallocated loss adjustment expenses. Transferred liabilities may belong to a single class of business, a territory, a policyholder, or an accident year. The transfer may apply to all net (of other valid reinsurance, collectible or not) losses, or depend on an aggregate attachment level (in dollars or days) or size per occurrence.

---

* Mr. Steeneck's paper was first submitted to the Casualty Actuarial Society in 1983. References in this paper to future events and to future time periods should be interpreted accordingly.

— *Editor*

Understanding financial reinsurance is becoming a top priority among insurance executives, regulators, stock security analysts, and others in the property-casualty insurance field.[1] Balance sheets and income statements can become less meaningful. And the volume of loss portfolio transfers is increasing. Industry observers and participants estimate that over $1 billion of such transactions occurred during 1983. The following example, albeit a dramatic one, shows the effect loss portfolio reinsurance can have on a company's accounting position.

A New York based reinsurance company recently sold a loss portfolio at about the same time as a regular triennial examination (12/31/80) found liabilities exceeded assets by $12,400,000. The result of the loss portfolio reinsurance transfer left the company with a healthy income statement and a statutory surplus of $10,800,000! Although the details are unknown, we can speculate that assets could have been $200 million and examined liabilities estimated at $212.4 million. Suppose $50 million of loss reserves were sold for $26.8 million. The net resulting liabilities would be $162.4 million with assets of $173.2 million. Statutory surplus would be the difference or $10,800,000. The $23.2 million gain could be reflected in the income statement.

The following two lists outline business purposes served by loss portfolio transfers and the costs the cedent must consider. The paper then treats the actuarial, tax, and accounting aspects. Then contractual and pricing considerations are mentioned. Finally, the uncertain regulatory environment is noted.

BUSINESS PURPOSES LOSS PORTFOLIO TRANSFERS SERVE

Depending on the financial position of the insurance company, several but not all of the following nine purposes may be attractive.

1. Improve underwriting results. By converting future investment income into current underwriting income, the composite ratio and income statement are improved. The case of the New York based reinsurance company loss portfolio transaction exemplifies these effects.
2. Increase GAAP earnings. In the case of the New York reinsurer, the loss portfolio transaction increased GAAP earnings directly by $23.2 million.
3. Improve GAAP deferred tax position. By raising GAAP taxable income the validity of the tax deduction for other underwriting losses is demonstrated.

[1] Mary Rowland, "Games insurers play with loss reserves," *Institutional Investor,* November 1983.

4. Increase surplus. The after-tax benefit goes directly into statutory surplus. Later in this paper, examples of the accounting treatment for loss portfolio transfers will illustrate the generation of underwriting income which flows into the surplus account.
5. Strengthen loss reserves. A cushion between carried loss reserves and possible adversely developing loss reserves will strengthen the cedent's balance sheet implicitly.
6. Improve NAIC IRIS Test results. This type of reinsurance is not penalized as are surplus relief treaties. Favorable Best's ratings may be retained.
7. Maintain premium volume. Ceded premium need not be affected but could be if controls on premium to surplus ratios are required.
8. Terminate a segment of business instantly. This was the original purpose of loss portfolio transfers.[2] Certain medical malpractice occurrence form writers may be considering a rapid exit from the business. Rather than running off the associated liabilities, they may sell their complete books.
9. Discount reserves. Without setting a precedent and changing accounting methods, the cedent effectively can discount reserves. Other industries have recently received SEC endorsement of accounting treatments termed "insubstance defeasance." The balance sheet is strengthened as a large amount of old debt is retired while paying for it with a smaller face amount of new debt at a higher rate. The ceding insurer's large debt (loss reserves) is replaced with a smaller debt (loss portfolio transfer payment) reflecting a higher interest rate.

COST CONSIDERATIONS TO LOSS PORTFOLIO TRANSFERS[3]

1. Decreases future GAAP earnings and surplus increases. The current year surge in GAAP earnings (see purpose #2) is at the cost of future investment income. Recall that assets are reduced by the transfer payment amount.
2. Adds reinsurance costs not budgeted. This includes reinsurer expenses, profit, and risk charge.

---

[2] For example, in 1595 "one Roemer Visscher of Amsterdam took over the insurance of certain marine risks, because the original insurer, Jacob Bruynsen Smallinck, had gone bankrupt." (Excerpted from a speech by Mr. Michael Felts, CAS Special Interest Seminar on Reinsurance, 1982.)

[3] Some of these business purposes and cost considerations come from a speech by Mr. John Murad at the American Academy of Actuaries Loss Reserve Seminar, 1983.

3. Reduces the liquidity of assets. The purchase may cause a cedent to keep taxable bonds (with higher coupons, effectively shielded from taxes if in a non-taxable position) and sell tax-exempts (generally considered to be more easily marketed at favorable prices). Other liquid assets may also be sold, leaving the less liquid ones.
4. May subject cedent to future taxes. If the cedent gets into a future taxable position and has retained less liquid taxable bonds, the after-tax investment returns may not be optimized.
5. Can create a capital loss by the sale of bonds to purchase the reinsurance.
6. Can lose tax deferred status. If cedent is in a taxable position, actual payment of taxes can occur.
7. Will likely distort schedules O and P. The abrupt decrease in loss and expense reserves and surge in payments can distort any loss ratio, loss development, or triangle projection analysis.
8. Creates dependence on reinsurer security. The possible non-collectibility of the reinsurance (by insolvency or dispute in coverage) has a cost which is difficult to quantify.
9. May create future costs.
   a. The transaction may prove unacceptable to regulators, tax authorities, and auditors from a risk transfer perspective. The consideration paid by the cedent may overfund the loss transfer especially if payment schedules are imposed on recoveries. Open ended retrospective adjustments of the consideration will also fail the risk transfer test.
   b. The company's accounting may have to be restructured as the accounting profession and regulators establish stricter guidelines. Beginning in 1984, the NAIC blank will require disclosure of loss reserves ceded, the consideration, the effect on underwriting results and statutory surplus, and limiting schedules of actual recovery. Insurance departments may require different accounting treatments than registered by the companies.
   c. There is a potential loss of company stature in the insurance community. On the other hand, the sale may be judicious.

The predominant statutory and tax accounting requirement of a loss portfolio transfer agreement is that it exhibit legitimate risk transfer. Without it, the transaction is voided and the accounting and financial effects must be unraveled.[4]

[4] See AICPA SBAS #5 paragraph 44 and #60 paragraph 40.

The cedent must use an authorized reinsurer to get credit for the reserves taken down. If it is using an unauthorized reinsurer, that reinsurer should post a letter of credit on the cedent's behalf or place assets equal to the transfer liability in escrow.

ACTUARIAL ASPECTS

In order to accelerate the greatest amount of investment income and place it in the underwriting account, "long tail" business (from a payment profile perspective) is required. If a company has little long tail business to cede, it cannot gain much financially from loss portfolio transfer reinsurance. Lines generally considered to give maximum effect are medical malpractice, workers' compensation, and products liability or other liability.

Basic actuarial loss data is required for a quantitative analysis leading to a responsible reinsurance offer. Payment and reported loss development triangles for the subject business are essential. In this author's experience, all too often data is not supplied with sufficient detail for scrutiny.

Large loss "outliers" and under represented losses must be normalized. Large losses may be over represented or under represented. These may also have loss development characteristics camouflaging the underlying loss process. Certain hazards may have produced too few losses to date. The actuary must use intuition and observed or postulated continuous size of loss distributions to adjust history so projections are accurate.

Allocated loss adjustment expense reserves are analyzed and included where necessary. Unallocated loss adjustment expense reserves may be analyzed and included as well (but in practice this is seldom done). Certain annual statement schedule P expense data may prove useful if both allocated and unallocated loss adjustment expense reserves are to be included in the portfolio transfer.

This author knows of no completely stochastic process in viewing potential outcomes regarding ultimate loss and payment profiles. Many reinsurance actuaries look at "best case-expected case-worst case" scenarios in determining outgoing cash flow. The various present values of those outgoing cash flows are calculated.

It is likely that in costing coverage reinsurers will attempt to match bond maturities with expected cash requirements. Unlike the single reinsured policy where loss payments are totally unpredictable, the loss portfolio has expected cash outflows. Coupons and maturing bonds can be matched to expected cash requirements. Bonds lock in specific returns (as opposed to many other investment vehicles). The reinsurer's management specifies the quality and type of securities acceptable in loss portfolio transfer reinsurance arrangements. Depending on secured rates of return and the reinsurer's tax position, a variety of corporate and government bonds with taxable and tax exempt status are available for a dedicated portfolio. Reinvestment risks on coupons can be of staggering importance.[5] Currently, there are a variety of "felines" on the market to eliminate this risk. For example, Merrill-Lynch has TIGR, or Treasury Investment Growth Receipts, which repackages T-Bonds to act like zero coupon bonds. Other felines include CATS and LIONS. Felines offer somewhat lower yields than non-stripped bonds as investment houses require a hedge on reinvestment.

TAX ASPECTS

The reinsurer's tax position is critical in the choice of taxable or tax exempt bond purchases and the resulting present value (market value) cost of the bond portfolio. Insurers are taxed like other corporations except as noted in Parts 2 and 3 subchapter L of the Internal Revenue Code. They use a modified accrual accounting system and have two classes of income: underwriting and investment. If the reinsurer or its consolidating parent has taxable income, the underwriting loss it will assume will effectively shield federal income taxes and a higher rate than tax exempt interest *may* be credited in the pricing. If the reinsurer has no taxable income and expects none in the foreseeable future, then marginal expected results suggest taxable bonds are most advantageous as an investment vehicle.

My understanding is that an over-structured transaction may be viewed by the IRS as, in essence, a single premium immediate annuity purchase. In that event, the ceding company would include in taxable income a portion of each payment recovered from the reinsurer. Over structuring may be hazardous.

The most competitive quotes combine high risk-adjusted yields with low reinsurer margin. Since bond yields vary day to day, today's quoted consideration must expire quickly and be subject to requotation. Changes in interest rates have a leveraged effect on cost.

The following exhibit demonstrates the effect tax position has on a reinsurer's net present value calculation. Suppose, for simplicity, that a portfolio consists of two $1,000 liabilities to be paid in 12 and 24 months. The reinsurer can purchase taxable bonds with 7% coupons semi-annually or tax exempt bonds with 5% coupons payable semi-annually. Tax exempts may prove preferable if the reinsurer is in a taxable position.

---

[5] Ronald Ferguson, "Duration," *PCAS* LXX, 1983, pp. 265–288.

ILLUSTRATION OF EFFECT OF REINSURER TAX POSITION

Loss portfolio transfer date 1/1/85
$1,000 loss payment expected 1/1/86
$1,000 loss payment epected 1/1/87

#1 Taxable Bonds—7% per coupon semi-annually (including reinvestment).

Present Value* = $1,000 ($a_{\overline{2}|.07}/s_{\overline{2}|.07}$) = $1,636.33

	1985		1986	
	Underwriting Income	Investment Income	Underwriting Income	Investment Income
Net Premium Earned	$1,636.33		$0.00	
Incurred Loss	2,000.00		0.00	
Result	$ (363.67)	$237.10	$0.00	$126.57
Carry forward to 1986	$ (126.57)		$126.57	

#2 Tax Exempt Bonds—5% per coupon semi-annually (including reinvestments).

Present Value* = $1,000 ($a_{\overline{2}|.05}/s_{\overline{2}|.05}$) = $1,729.73** (before tax effect)

	1985		1986	
	Underwriting Income	Investment Income	Underwriting Income	Investment Income
Net Premium Earned	$1,565.20		$0.00	
Incurred Loss	2,000.00		0.00	
Result	$ (434.80)	$160.43	$0.00	$74.37
Recouped Taxes (46%)	$ 200.00			
Net	$ (234.80)	$160.43	$74.37	

Depending on the reinsurer's tax position, the net present value is between $1,565.20 and $1,636.33. Reinsurer expenses and profit/risk charge have not been included in Net P.E. (premium earned).

*This standard actuarial notation represents the present value of 2 annual $1,000 payments when interest is credited on a semi-annual basis at 7% and 5%.

**Net Premium Earned + Recouped Taxes/1.05^2 = $1,729.73
Also, .46 (2,000 − Net Premium Earned) = Recouped Taxes.
Solving the equations: Net Premium Earned = $1,565.20 (after tax).

The challenge facing the pricing actuary is to meet the financial objectives of the cedent while at the same time offering a risk product which has the expectation of reinsurer profit. This is frequently difficult since potential payment profiles, possible runoff liabilities, and unanticipated "shock" disturbances play havoc. The following scenarios demonstrate the effect payment profile and quality of carried reserves have on the present value (at 10%) of potential outcomes.

Scenario (A) describes the present value of the complete 4 and 7 year possible payouts of possible runoff liabilities. Cessions under scenario (A) have large cash flow consequences. With slightly less benefit (reserve less present value), the portfolio ceded may be structured more effectively in scenarios (B) and (C).

Also notice that, if the cedent believes the likely outcome to be 2a and the reinsurer believes the likely outcome to be 3b, a deal may be struck. They may agree to cede/accept company paid losses after 24 months but with an overall limitation in recoveries of $3.5 million. This cedent releases $3.4 million of carried reserves. The reinsurer's net present value is $2.0 million/$1.10^3$ + $1.5 million/$1.10^4$ or $2,527,150. If the reinsurer prices this at $2.7 million the cedent will generate $700,000 of income.

ACCOUNTING ASPECTS

At this point, there is no standard accounting treatment for these transactions. The simplest accounting treatment, however, from the cedent's perspective is to note, following the last example, that $3.4 million of loss reserves is offset by a negative $3.4 million reinsurance recoverable. Further, $2.7 million of paid losses are registered and the gain flows through the balance sheet, income statement, and schedule O or P (as appropriate). This accounting treatment can be called the "loss method".

There is also a "premium method." The treatment calls for premium reduction of $2.7 million. Paid losses remain unchanged. Reserves are reduced by $3.4 million. Implied by this treatment, the cedent's loss ratio goes down and his expense ratio goes up. If the reinsurer offers a 10% commission ($300,000), ceded premium goes up by $300,000 and the net loss ratio increases. But the

EFFECT OF LOSS AND PAYMENT CHARACTERISTICS IN VARIOUS LOSS PORTFOLIO SCENARIOS

Loss Portfolios
(in 000's)

	Current O/S plus IBNR	Payment Profile (in months) + 12	+ 24	+ 36	+ 48	+ 60	+ 72	+ 84
1. Optimistic	$6,000							
a. Fast pay		2,000	1,500	1,500	1,000			
b. Slow pay		1,500	1,500	1,000	1,000	500	300	200
2. Expected	$8,000							
a. Fast pay		2,600	2,000	2,000	1,400			
b. Slow pay		2,000	2,000	1,300	1,300	700	400	300
3. Pessimistic	$12,000							
a. Fast pay		4,000	3,000	3,000	2,000			
b. Slow pay		3,000	3,000	2,000	2,000	1,000	600	400

Present or Current values at 10% interest
Payments in 12 month increments

	Scenario (A) Entire Portfolio	Scenario (B) Over 24 Months	Scenario (C) Over $5,000,000 Retained
1a.	$4,867,837	$1,809,986	$ 683,013
1b.	4,620,069	2,016,763	582,434
2a.	6,475,377	2,458,848	2,158,323
2b.	6,150,083	2,679,009	1,927,694
3a.	9,735,674	3,619,971	5,272,864
3b.	9,240,137	4,033,525	4,859,972

commission will offset insurer operating expenses and, therefore, the expense ratio will decline. To illustrate (statutory accounting):

Ceding Company Marginal Effects	Loss Method	Premium Method: No Commission	Premium Method: Commission
Earned Premium (+)	\$ 0	\$ 2,700,000	\$ −3,000,000
Operating Expenses (−)	\$ 0	\$ 0	\$ −300,000
Paid Losses (−)	\$ 2,700,000	\$ 0	\$ 0
Change in O/S Losses (−)	\$ −3,400,000	\$ −3,400,000	\$ −3,400,000
Underwriting Gain	\$ 700,000	\$ 700,000	\$ 700,000

The reinsurer could mirror these accounting entries by merely changing signs and penalizing surplus. As yet it is not necessary, but regulators, auditors, etc. and advise it.

As an aside, for GAAP accounting purposes, the reinsurer might book the present value (\$2.7 million) of expected payments to escape this "surplus hit". He could do this if he normally discounts for GAAP purposes, and he makes adequate disclosure. The same is not true for statutory accounting purposes. Full reserves must be established, otherwise general interrogatories 16 and 28 of the convention blank must be answered so as to invite criticism.

"Another (reinsurer) statutory alternative is to consider the transaction as other income as opposed to underwriting income. A rationale here is that the investment income to be earned to offset future loss payments does not flow through underwriting income either, and the effect of the transactions still impacts statutory results. This treatment has not received broad acceptance."[6]

There are other considerations to be made in the pricing of loss portfolios. Some are contractual. Others deal with reinsurer margin requirements.

CONTRACTUAL ASPECTS

Extra contractual obligations (ECO) can be defined as punitive and/or compensatory damages assessed against an insurer as a consequence of his tortious acts. ECO's do not fall under the auspices of the original subject insurance

[6] Taken from a speech by Mr. James Faber at the American Academy of Actuaries Loss Reserve Seminar, 1983.

policy. Historic data is not generally available nor projectable so cedent payments from this hazard should be excluded.

The reinsurer will also insist on some verbal, if not written, understanding on the use of structured settlements. The commutation of a claim by the purchase of a life annuity changes both the expected liability transfer amount and the payment timing. Special treatment is required such as substituting an actuarial equivalent payment stream for the commuted value.

The insurance industry now faces more "common cause" losses than ever before. These are the "asbestos type", unpredictable from one exposure over time. But when they occur, they create a flood of individual claimants demanding tremendous aggregate sums of money. For certain classes of insurance, the reinsurer will consider the likelihood of common cause events, charge for it, limit it in some way contractually, or both.

Claims handling is also important. Loss portfolio transfers are frequently sought by self-insureds wishing to extricate themselves from their developing insurance experience or are being acquired and, therefore, in need of a fully insured program. It is difficult to properly run off liabilities without continuity in claims handling.

A front company may be necessary to issue a primary insurance policy which is then reinsured.

EXPENSE/PROFIT ASPECTS

Having considered all of the above, the reinsurer now must decide how much to charge in excess of the bond portfolio cost. The reinsurer will have expenses, both current in the marketing and initial set up of administration, as well as on going administrative costs. In addition, it will desire a profit and risk charge dependent on the following:

1. Predictability of results—Investment risk and underwriting risk can be significant. Actual runoff could be heavier and/or faster than expected. To the extent a structure of reimbursements exists as to timing and amount, this lessens risk.
2. The surplus rent—The reinsurer's charge against surplus will restrict its writings for potentially 5–30 years. This requires substantial profit loading. It can be measured as a percentage of the first year charge against surplus, or the present value of the annual charges.

3. The contractual risk—Structured settlements, common cause losses, and claims handling features have some bearing on the value the reinsurer places on the proposed transaction. Other features such as a contingent commission or an experience rating scheme will cause the reinsurer to change profit and risk charge expectations.

FUTURE EXPECTATIONS

It should be evident loss portfolios involve more than "shelf technology." Only the educated professional can and will be successful. But what does the future hold?

The ultimate destiny of loss portfolio transfer reinsurance may be in the hands of the taxing authorities and accountants.

The AICPA is studying the issue of loss reserve discounting. It sees four types of claims:

1. Short term claims closing in one or at most two years. Discounting may not be economically justified here.
2. Long term uncertain claims like medical malpractice and auto bodily injury having reserves which earn investment income but are not subject to rigorous loss payment schedule. It is possibly impractical to discount here since conservative interest rates are indicated.
3. Long term reasonably certain claims like periodic medical payments for life under worker's compensation pension cases.
4. Long term claims with fixed payment like some workers' compensation fixed periodic indemnity for life claims. These annuities or near annuities are subject to accurate discounting procedures.

Discounting has some negative connotations including the publication of unstable and potentially unreadable insurer results. These could confuse regulators, analysts, and the public. Loss reserve evaluations and tests would prove difficult. Some actuaries observe that in recent years, reserve shortfalls are generally offset by investment earnings. As the crutch is removed, the lame patient must fall. The pressure on companies to set adequate reserves would intensify if the investment crutch were removed.

To more closely monitor the financial effects of loss portfolio transfers, many states are requiring disclosure. The NAIC is adopting a disclosure note to first appear on the 1984 Blank. The SEC is also concerned about the ability of investors to evaluate the financial condition and results of companies with P/C operations.[7]

Finally, the government has a large stake in the loss reserve discounting arena. There are current attempts to restructure life and health and property-casualty insurance company taxation regulations in order to generate substantially greater tax revenues. It is quite conceivable that the 1983 proposal to discount liabilities for schedule P lines at a 5% rate of interest could be eventually adopted. This would generate taxable income. Accountants would likely endorse this, I believe.

Some companies discount loss reserves on a GAAP basis already (but these are largely offshore companies). If the definition of taxable income changes to embrace discounted loss reserves, can a change in statutory accounting principles be far behind? The market for loss portfolio financial reinsurance would largely evaporate. There are some very unhealthy implications currently under investigation and discussion. Until the final outcome is known, loss portfolio transfer reinsurance will continue to be a valuable tool for insurance and self-insured company managements.

---

[7] See "SEC Seek Loss-Reserve Disclosure Rule To Assist Investors in Property Insurers," *The Wall Street Journal*, March 12, 1984, p. 10.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXX

REINSURING THE CAPTIVE/SPECIALTY COMPANY

LEE R. STEENECK
VOLUME LXX

DISCUSSION BY OAKLEY E. VAN SLYKE

Mr. Steeneck has presented the basic principles of applying utility theory in reinsurance pricing in an admirable fashion. His article is straightforward and comprehensive. The footnotes provide an excellent bibliography of the current literature on the subject.

The interested reader is particularly directed to the monograph by Leonard Freifelder (Freifelder (1976)). These two works complement each other well.

Utility theory has been useful to this reviewer as a means of achieving a fresh viewpoint on a problem, rather than as a simplistic solution to the problem of finding numerical results (e.g., rates) that adequately reflect one's risk aversion. If the user avoids the pursuit of simple answers through abstract formulas, he can find much of practical value in the methods discussed by Mr. Steeneck. This is especially true for exponential utility functions.

Several of Mr. Steeneck's points merit discussion. This review also provides an opportunity to show two minor results of the reviewer's investigations into utility theory.

PRACTICALITY

Utility theory is practical. We are all familiar with the inadequacy of the simple calculation of expected value. Using utility theory only requires that we shift our mental framework from calculating the $E(X)$ and $E(X - u)^2$ to include the calculation of $E(U(X))$.

This mental shift will be clear if a train of thought developed by Steeneck on page 257 is followed, together with a change in terminology. Let the utility function be

$RAC(X,c) = c(1 - \exp(x/c))$.

$RAC(X,c)$ is the "risk adjusted cost" of outflow X with utility scale c.

Then for an aggregate loss distribution

$F(X) = \Pr(x < X)$
$RAC_F(c) = c \ln E_F (\exp (x/c))$

In other words, the risk adjusted cost of an aggregate loss distribution is a scale adjustment in c of the calculation of the expected value of the aggregate loss distribution. The adjustment in c scales down each possible loss to its multiple of c, inflates it using the exponential function, takes an average, and then backs out the scale adjustments by taking the log and multiplying by c.

In the example given by Steeneck on page 259, c was \$4,000,000.

The use of c instead of $1/r$ makes sense. It puts the constant in real units, dollars, instead of imaginary ones, $(\text{dollars})^{-1}$.

$RAC(c)$ is a simple concept which includes a great deal of information about $F(X)$. (As we shall see, it includes all of the information about $F(X)$.) We expect to see the day when $RAC(c)$ will be computed as routinely as Var (X) is today, particularly when it can be expressed in closed form.

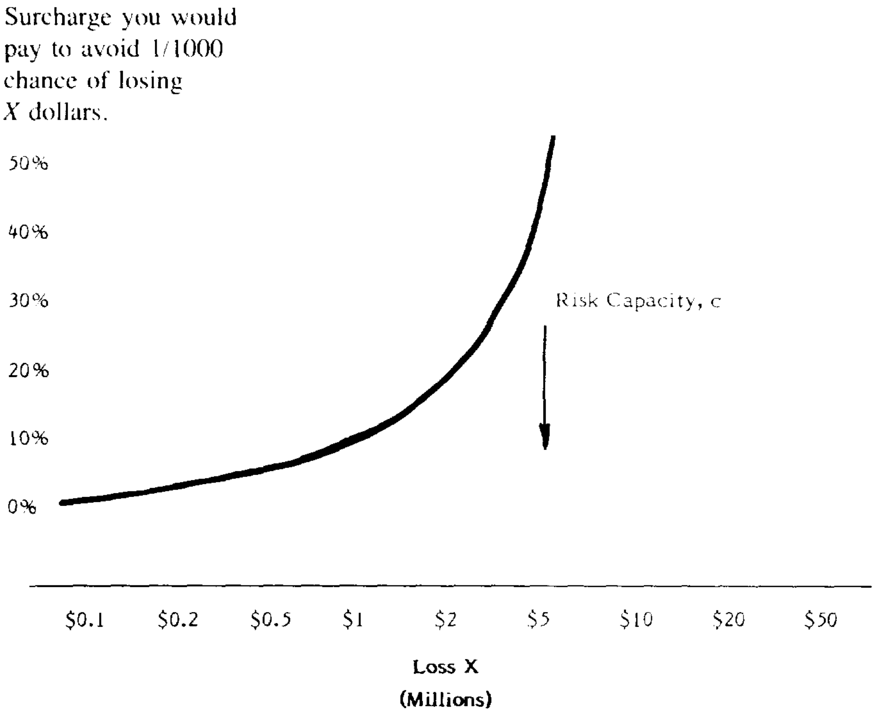
ESTIMATING RISK CAPACITY

It is easy to estimate the risk capacity, c, well enough for practical applications. Reinsurance exists because all insurance companies have a limited capacity to bear risk. In some cases, risk aversion is so high that the firm will do whatever possible to ensure that a catastrophic loss will not bankrupt the firm. In practical applications, however, the level of loss at which management begins to get really concerned is quite a bit less than the level of loss that would bankrupt the firm. In insurance jargon, we call this level of loss the firm's "risk capacity."

Exhibit 1 shows an example of the risk capacity of a particular firm. In this example, the height of the line shows the surcharge the reinsured would be willing to pay to avoid a 0.1% chance of losing a sum of money. The reinsured would be willing to pay only about 0.1% of the sum of money if this sum, X, were not very great. It would not pay a significant surcharge to avoid a 0.1% chance of losing \$10,000. If the amount were much greater than the reinsured's risk capacity, however, then the firm would be willing to pay much more than 0.1% of the possible loss.

EXHIBIT I

Surcharge for Risk Curve



Move the bottom scale left or right until it is in the right place for your decision. Your risk capacity, c, will be below the vertical arrow.

Because of the reinsured's limited capacity to bear risk, management is willing to pay a surcharge (risk charge) to avoid financial fluctuations. To avoid a 0.1% chance of paying out $1,000,000, for example, management is willing to pay something in excess of 0.1% of $1,000,000. The additional amount is called a "risk charge." The total amount management is willing to pay, perhaps $1,100 in this example, is called the "risk adjusted cost" (*RAC*) of the risk's probability distribution.

As Cozzolino (1979) has pointed out, the selection of risk capacity c is not even necessary to make a decision. All that it is generally necessary is that one's risk capacity is known to be in a certain range.

The technique suggested by Cozzolino is to show the risk adjusted cost for one's own aggregate loss distribution with and without the inclusion of the reinsurance contract being evaluated. Each net aggregate loss distribution leads to a unique risk adjusted cost profile. Exhibit 2 shows the profile for a reinsurance decision about a possible cession that involves a considerable amount of risk. In this example, if the reinsurer's risk capacity is less than about $2,000,000, he will not accept the retrocession.

The success of this technique hinges on the fact that more risky alternatives will always have curves that slope downward more steeply than less risky alternatives. As a result, different options will produce risk profile curves that intersect one another if there are significant differences in the uncertainty of results for the options. Obviously, if the risk profile curve for one option is lower than the risk profile curve for another option regardless of one's risk capacity, it is the more attractive alternative.

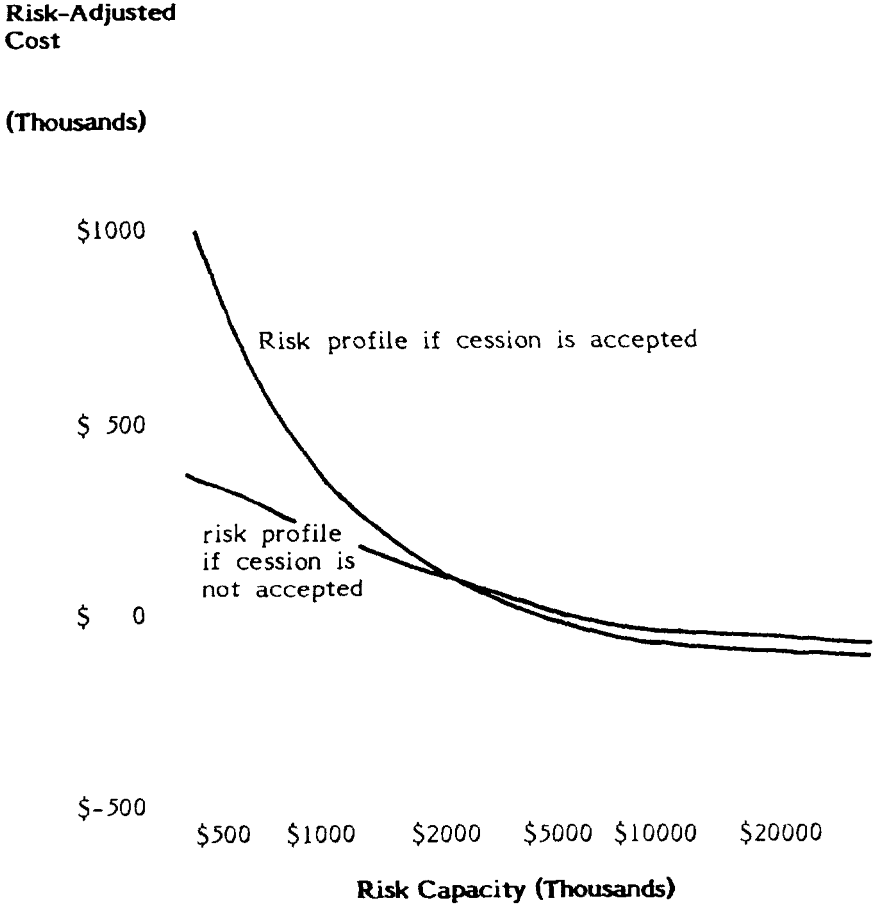
REINSURANCE NEGOTIATIONS

Reinsurance makes sense even when the reinsurer is more risk averse than the reinsured. Steeneck's statement, "If the reinsurer has the same utility function or is less risk averse, a deal can be struck" is unnecessarily restrictive. This is seen in practice as small reinsurers take small pieces of treaties reinsuring large primary companies.

The reason is simple: The reinsured losses are not correlated with the reinsurer's losses; they are correlated with the reinsured's losses.

EXHIBIT 2

RISK PROFILE CURVES FOR THE REINSURANCE DECISION



A risk profile is a display of $RAC(X,c)$ for a range of c. A risk profile is a unique mapping of an aggregate loss distribution $F(X)$.

To the excess writer of, say, \$1,000,000 xs \$1,000,000, the risk looks like

$$RAC_e = c_e \ln E\left[\exp\left(\frac{\text{Max}(0, X^{**} - \$1 \text{ million})}{c_e}\right)\right]$$

where X^{**} is limited to \$2,000,000, and the excess writer's risk capacity is c_e.

To the primary writer, the cession is worth

$$RAC_p = c_p \ln E\left[\exp\left(\frac{X^{**}}{c_p}\right)\right] - c_p \ln E\left[\exp\left(\frac{X^{*}}{c_p}\right)\right]$$

$$= c_p \ln \frac{E\left[\exp\left(\frac{X^{**}}{c_p}\right)\right]}{E\left[\exp\left(\frac{X^{*}}{c_p}\right)\right]}$$

where X^* is limited to \$1,000,000, and the primary writer's risk capacity is c_p.

Reinsurance makes sense when

$$RAC_p > RAC_e$$

This leads to two thoughts:

— One's own risk profile and estimates of the risk profiles of the potential players in a reinsurance deal can help one create a negotiating strategy. Changes in the terms of the reinsurance arrangement can be reflected in changes in the risk profiles. This will identify ways to change the deal to improve it for all parties.

— This analysis makes it clear why new entries always appear in the reinsurance market. Reinsurers have portfolios of losses that are correlated with the potential cession. In workers' compensation, for example, losses in various contracts may be correlated through inflation, benefit level changes, or loss of statutory immunities or defenses. The new entries have risk capacity arising from their own cash flow, but do not have existing portfolios of losses that are correlated with the new cession. (Of course, presumably, they do not have the underwriting expertise of the experienced writer, either.)

EXPONENTIAL UTILITY VS. THE VARIANCE PRINCIPLE

Two advantages of using exponential utility instead of the popular variance principle are:

— Exponential utility provides the correct asymptotic behavior as the loss being considered gets large and its probability gets small. This is illustrated in Exhibit 3.

In contrast, the variance principle leads to premiums greater than the loss itself.

— Exponential utility leads to a more distinct concept of risk capacity. Exhibit 4 shows that the disutility associated with a loss in excess of one's capacity (as defined above) reflects a marked aversion to losses greater than one's risk capacity. This agrees with our intuitive understanding of how we accept and cede risks. The variance principle, in contrast, does not show such a distinct "flinch point."

ESTIMATION

New methods of estimating aggregate loss distributions make practical application much easier. Monte Carlo simulation is readily available, although somewhat costly in terms of computer time. Monte Carlo simulation handles virtually all practical problems including multiline contracts. Monte Carlo simulation also gives the flexibility to break apart workers' compensation losses by type of injury, distinguish various sublines of liability coverage, and so on.

Aggregate distributions are receiving more attention recently. Heckman and Meyers (1983) describe a method of calculating aggregate loss distributions by a method of characteristic functions. Venter (1983a) shows an application of a method of numerical estimation developed by Panjer. Jewell (1983) extends Panjer's work to a dynamic risk portfolio. Each of these authors shows how to calculate the expected value of an excess premium as well as first dollar losses. We can expand Venter's conclusion (from page 69) to read:

> "By approximating the severity distribution with discrete probabilities, the aggregate distribution and excess premium functions and the risk adjusted cost can thus be estimated recursively."

Venter (1983c) has dicussed the advantages of modeling aggregate loss distributions with transformed Gamma distributions. Distributional models may lead directly to general formulas for the risk adjusted cost.

EXHIBIT 3

Asymptotic Behavior

Surcharge to avoid 1/1000 chance of losing x units of risk capacity, expressed as a multiple of X.

$$S = (\ln(.999 + .001 \exp(x/c)) \div .001\, x/c) - 1$$

(exponential utility principle)

or

$$S = .999X/c$$

(variance principle)

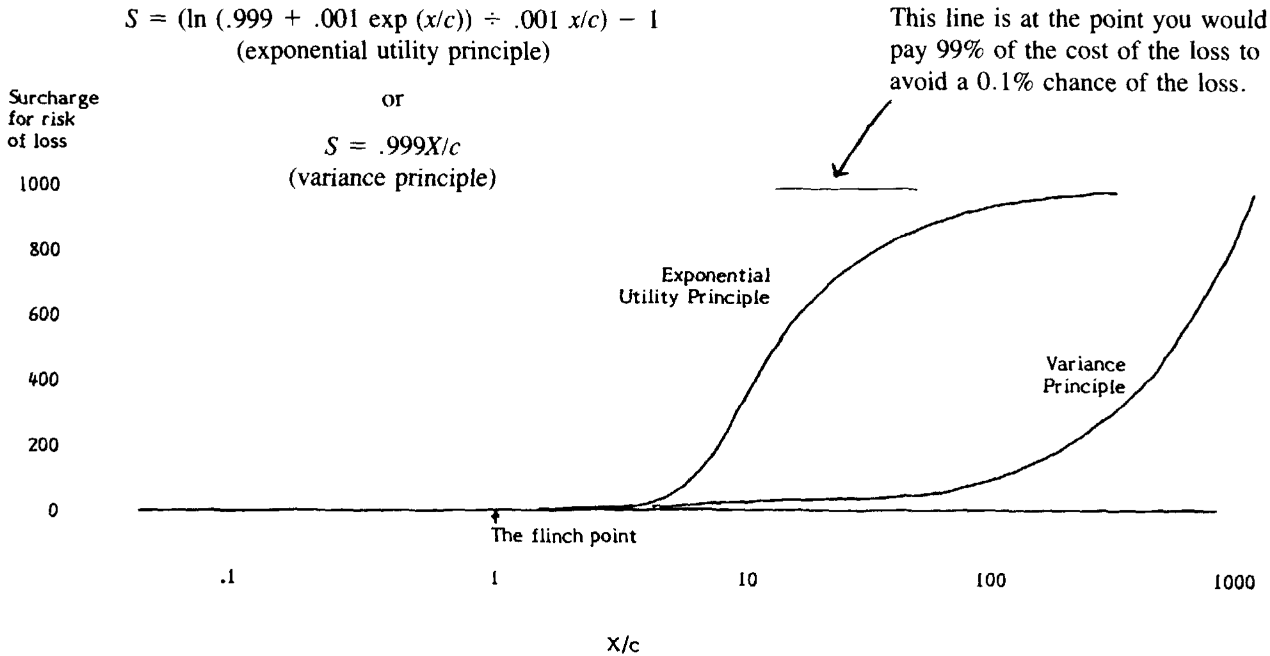
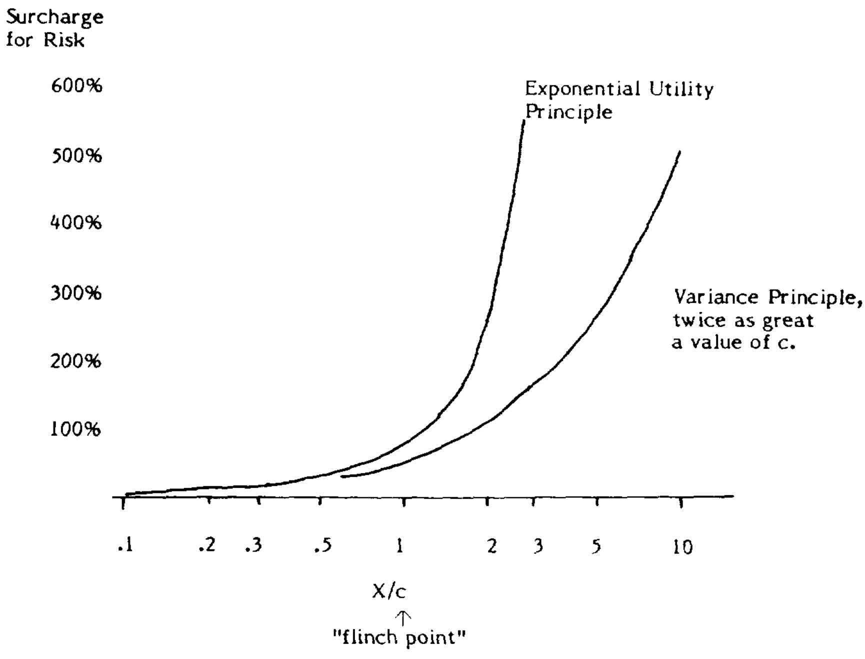


EXHIBIT 4

THE FLINCH POINT



DECREASING AVERSION TO RISK

Venter (1983b) has pointed out the theoretical advantages of

$$RAC\ (X,c,p) = c\left(1 - \exp\left(-\frac{X^p}{c}\right)\right)$$

or some other utility function with decreasing aversion to risk.

This may be a valuable point, but in practice a reinsurer is not likely to vary its risk capacity significantly in response to a loss under a single treaty. It is more realistic to expect a reinsurer to become more or less aggressive in response to a series of losses, a change in the competitive marketplace, or some other factor affecting many treaties. In short, the refinement will not matter in most practical applications.

Indeed, as we have seen, it is easy to explain the search for one's risk aversion if risk aversion is taken to be constant. It is difficult to develop such a procedure if one's risk aversion is supposed to be expressed as a function of the surplus left *after* the loss.

Most importantly, using an exponential utility function does not necessarily result in a misstatement of our utility function. We can be correct if we can correctly see the utility of $(a - X)$ from our vantage point at a. We can be as averse to $(a - X)$ as we wish.

DISTRIBUTIONAL STATISTICS

Characteristic functions and moment generating functions (m.g.f.'s) can be used in tandem to derive simple results for frequently used models. As Heckman and Meyers (1983) showed,

$$\phi_F\ (t) = E\ [\exp(itx)] = \int_0^\infty \exp\ (itx)\ dF(x).$$

where ϕ_F is the characteristic function of $F(x)$.

This leads directly to

$$RAC_F(c) = c \ln E_F\ [\exp\ (x/c)] = c \ln \int_0^\infty \exp\ (x/c)\ dF(x)$$

$$= c \ln \phi_F\left(\frac{1}{ic}\right)$$

where the subscript F refers to the aggregate loss distribution.

They also showed that

$$\phi_{F*G}(t) = \phi_F(t)\ \phi_G(t)$$

and that the characteristic function for an aggregate loss distribution F (with claim severity distribution S) is:

$$\phi_F(t) = \sum_{n=0}^{\infty} p(n)(\phi_S(t))^n.$$

This leads directly to

$$RAC_F(c) = c \ln \sum_{n=0}^{\infty} p(n) \exp\,[n \cdot (RAC_s(c)/c)]$$

where RAC_s is the risk adjusted cost of a single claim.

The $RAC_s\ (c)$ is also closely related to the moment generating function of the severity distribution

$$RAC_s(c) = c \ln M_s \left(\frac{1}{c}\right) .$$

We mentioned earlier that the $RAC_F(c)$ contained all the information in $F(X)$. This is now clear because the m.g.f. of a probability distribution is unique (Hogg and Klugman (1984), page 19). Hogg and Klugman have shown (page 50) that if the moment generating function of the severity distribution, $M_s(t)$, is known, and the claim frequency distribution is Poisson, the moment generating function of the aggregate loss distribution is

$$M_F(t) = \exp[\lambda(M_s(t) - 1)]$$

The risk adjusted cost is therefore

$$\begin{aligned} RAC_F(c) &= c \ln\,[M_F(1/c)] \\ &= c \ln\,[M_s(1/c) - 1]. \end{aligned}$$

For example, if the claim size distribution is exponential

$$p(x) = \frac{1}{\sigma} \exp\left(-\frac{x-\theta}{\sigma}\right)$$

then

$$M_s(t) = \frac{\exp\,(t\theta)}{1 - \sigma t}$$

$$M_s\left(\frac{1}{c}\right) = \frac{\exp(\theta/c)}{1 - (\sigma/c)}$$

The roles of σ, θ and c as scale adjustments are clear. This leads to the following risk adjusted costs:

$$RAC_F(c) = c\lambda\left[\frac{\exp(\theta/c)}{1 - \sigma/c} - 1\right]$$

$$= \frac{c\lambda}{1 - \sigma/c}\left[\frac{\sigma}{c} + \exp(\theta/c) - 1\right].$$

If $\theta = 0$,

$$RAC_F(c) = c\lambda\,\frac{\sigma}{c - \sigma}.$$

If θ/c is close to 0,

$$RAC_F(c) \doteqdot c\lambda\,\frac{\sigma + \theta}{c - \sigma}.$$

This development suggests several obvious extensions to be pursued:

—To determine the risk adjusted cost if the claim frequency distribution is negative binomial.

—To determine the risk adjusted costs for other severity distributions for which the m.g.f. is known in closed form.

—To determine the risk adjusted costs for truncated versions of distributions for which the m.g.f. is infinite.

—Numerical approximations based on m.g.f.'s, characteristic functions, or recursive methods.

PROBABILITY, UTILITY, AND PRESENT VALUE

The time value of money is important in many practical problems. In these problems a present value factor $v(i)$ can be associated with each event that produces a loss $X(i)$. The functions v and X may be continuous or discrete.

Interest should be handled in such a way that the distributive property applies to the function RAC. That is, the risk adjusted cost of a possible set of events should be independent of how fine a description one makes of the set of possible events.

The function

$$RAC(c) = c \ln \Sigma\, p(i)\, v(i) \exp (X_i/c)$$

meets this criterion. So does its continuous counterpart

$$RAC(c) = c \ln \int_0^{\infty} v(x) \exp (X/c)\, dF(X).$$

With this definition, the total risk adjusted cost *RAC* of a set of possible events with risk adjusted costs *RAC*(i) is the risk adjusted cost of all possible events, with each taken at its present value:

$$RAC_0 = c \ln \sum_i p(i)\, v(i) \exp (RAC(i)/c)$$

$$= c \ln \sum_i p(i)\, v(i) \left[\sum_{\substack{within\\ i}} p(x)\, v(x) \exp (x/c)\right]$$

$$= c \ln \sum_i \sum_{\substack{within\\ i}} (p(i)p(x))(v(i)v(x)) \exp (x/c).$$

In practice, then, probability and present value are almost interchangeable concepts. Present value and utility are not interchangeable concepts. This surprising result follows from the distributive property.[1]

CONCLUSION

The reader is encouraged to try the utility-user's viewpoint in practical problems. Starting perhaps with a discrete decision (such as whether or not to underwrite a particular risk or block of risks), decide on your risk capacity using Exhibit 1 or Exhibit 4. Sketch the risk profile curves for the decision by calculating a few points on each. Think about the interplay between your risk capacity and the decision you prefer (yes or no). Are you being consistent? Have you learned anything about the decision you didn't know before? With use, this additional viewpoint may begin to feel as natural as considering both probability and time in the decision.

[1] It would be reasonable to postulate that the multiplicative associativity of $p(i)$ with $p(x)$ and $v(i)$ with $v(x)$ follows directly from a distributive property on *RAC*. I have not been able to prove this, nor find an exception. A friend of mine says he proved it on a popcorn box at an Oilers game, but lost the proof. I would like a demonstration of whether or not "Oilers postulate" is true.

REFERENCES

John M. Cozzolino, (1979), "A New Method for Risk Analysis," *Sloan Management Review,* Massachusetts Institute of Technology, Spring, 1979.

Leonard R. Freifelder, (1976), *A Decision Theoretic Approach to Insurance Ratemaking,* Heubner Foundation Monograph No. 4, Richard D. Irwin, Inc.

Philip E. Heckman and Glenn C. Meyers, (1983), "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *PCAS* LXX, 1983.

William S. Jewell, (1983), "Approximating the Distribution of a Dynamic Risk Portfolio," ORC 83-15, Operations Research Center, University of California, Berkeley, California.

Robert V. Hogg and Stuart A. Klugman, (1984), *Loss Distributions,* John Wiley & Sons.

Gary G. Venter, (1983a), discussion of "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions" by Philip E. Heckman and Glenn C. Meyers, *PCAS* LXX, 1983.

Gary G. Venter, (1983b), "Utility with Decreasing Risk Aversion," *PCAS* LXX, 1983.

Gary G. Venter, (1983c), "Transformed Beta and Gamma Distributions and Aggregate Losses," *PCAS* LXX, 1983.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXI

EXTRAPOLATING, SMOOTHING, AND INTERPOLATING DEVELOPMENT FACTORS

RICHARD E. SHERMAN

VOLUME LXXI

DISCUSSION BY STEPHEN P. LOWE AND DAVID F. MOHRMAN

Mr. Sherman's paper presents a potpourri of practical applications involving the fitting of a parametric equation to loss development factor data. The particular equation utilized is called the inverse power curve, the form of which is

$$f(t) = 1 + \frac{a}{(t + c)^b} \tag{1}$$

where a, b and c are parameters to be estimated, and t represents time as it relates to the maturity of the body of claims.

It should be readily seen that the parameter c provides a linear transformation of the time variable t, and is therefore somewhat extraneous to the formulation. The definition of t is arbitrary; $f(t)$ can be the development factor from t to $t + 1$, or alternatively $f(t)$ can be the development factor from $t - 1$ to t. Similarly, the beginning of the accident year can be $t = 0$ or $t = 1$ (or even $t = -1$ or $t = 1.7275$).

The above comment is not intended to suggest that the selection of the time scale embodied in the variable t is trivial; a different result will be obtained for each scale chosen. However, to simplify discussion, we can express Mr. Sherman's equation as

$$f(t) = 1 + \frac{a}{t^b} \tag{2}$$

where we are searching for the best a, b, and scale t that fits the data.

Like the author, these reviewers have found it useful in many circumstances to fit parametric equations to incomplete, erratic or irregular loss development data. This review will expand slightly on Mr. Sherman's paper by offering some alternative equations, and discussing some desirable characteristics for loss development models of this kind. In addition, we will offer some specific comments and point out some pitfalls associated with Mr. Sherman's approach.

ALTERNATIVE MODELS

The parametric equation in (2) above is referred to by the author as the inverse power curve. We refer to this equation as the polynomial decay model. As the author points out in Section II, this equation has the property that the initial development, a, decays at a rate of

$$1 - \left(1 - \frac{1}{t}\right)^{b} \tag{3}$$

over the interval from $(t - 1)$ to t. For example, if $b = 1$, then the following decay rates would apply.

t	Development from $t - 1$ to t	Rate of Decay
1	$1 + a$	
2	$1 + a \times 1/2$	50%
3	$1 + a \times 1/2 \times 2/3$	33%
4	$1 + a \times 1/2 \times 2/3 \times 3/4$	25%
5	$1 + a \times 1/2 \times 2/3 \times 3/4 \times 4/5$	20%

An alternative model to the polynomial decay is one involving exponential decay:

$$f(t) = 1 + \frac{a}{b^t} \tag{4}$$

In this model the initial development, a, decays at a constant rate, $1 - b$, over each interval.

Viewing loss development as a decay process is intuitively appealing. It is certainly reasonable that, as an ever increasing proportion of losses are paid, their propensity to develop must decline.

Both the polynomial and the exponential decay models can be expanded by the addition of a third parameter involving a squared term.

$$f(t) = 1 + \frac{a}{t^b} + \frac{c}{t^{b^2}} \tag{5}$$

$$f(t) = 1 + \frac{a}{b^t} + \frac{c}{b^{t^2}} \tag{6}$$

There are also a variety of mixed models that might prove useful.

$$f(t) = 1 + \frac{at^c}{b^t} \tag{7}$$

$$f(t) = 1 + \frac{a + ct}{t^b} \tag{8}$$

All of these six models have been used by the reviewers to fit emergence data of one form or another.

Equations (5) and (6) are interesting as they can be conceptualized as modelling two different kinds of development taking place simultaneously, but decaying at different rates. For example, if the data were accident year reported losses, the a term might represent development caused by newly reported claims, while the c term might represent development on existing claims.

A specific instance where this approach is very useful is in the case of subrogation and salvage. The following table compares actual loss development factors for Auto Physical Damage to those obtained using the three parameter polynomial decay model, presented in equation (5).

Year of Development	Actual Development	Model Development
2:1	1.240	1.240
3:2	.993	.992
4:3	.996	.997
5:4	.998	.999
6:5	.999	.999
7:6	1.000	1.000

In this instance the parameters of the model are $a = -.07$, $b = 3$ and $c = .31$. In this instance the model has a nice intuitive appeal. The positive development of losses embodied by the c term decays very quickly, leaving the slower negative development of subrogation and salvage embodied by the a term.

CHOOSING A MODEL

As we have noted, all of the models described previously have proven useful in fitting various kinds of emergence data. We suspect that the reader could easily conjure up other models that would also prove useful.

Each of the models that we have described is "well-behaved", but only over a limited range of parameter values. It is worthwhile to consider what kinds of constraints on the parameters are necessary for a model to be reasonable.

In traditional applications, we want the development factors to be positive, decreasing, and approaching one. These can be expressed mathematically as

1. $f(t) \geq 1 \qquad \lim_{t \to \infty} f(t) = 1$

2. $f'(t) < 0 \qquad \lim_{t \to \infty} f'(t) = 0$

3. $f''(t) > 0 \qquad \lim_{t \to \infty} f''(t) = 0$

While the constraints on the limits are probably necessary in all situations, special circumstances may require the relaxing of one or more of the constraints on values of $f(t)$, $f'(t)$ or $f''(t)$. For example, to produce the auto physical damage factors cited earlier, it was necessary to violate the first constraint. Similarly, the third constraint restricts us to curves that are concave upward over the entire domain of t. In some instances a curve that starts out concave downward may be desired.

For Sherman's two parameter polynomial decay model

$$f'(t) = \frac{-ba}{t^{b+1}}$$

$$f''(t) = \frac{b(b + 1)a}{t^{b+2}}$$

We see that all conditions are satisfied when $a > 0$ and $b > 0$ (and $t > 0$).

For the two parameter exponential decay model

$$f'(t) = \frac{a}{b^t}\left(\ln \frac{1}{b}\right)$$

$$f''(t) = \frac{a}{b^t}\left(\ln \frac{1}{b}\right)^2$$

Here all conditions are satisfied when $a > 0$ and $b > 1$.

Similar calculations to these should be performed on any proposed model before its use, so that a clear understanding of the properties and limitations of the model is obtained.

A much more critical property of any model used to estimate report-to-report development factors is whether the product of the infinite series converges. While arbitrary truncation of the series at some point (such as 80 years) may be acceptable from a practical standpoint, it would be more desirable to restrict the model by requiring that it produces a less-than-infinite development factor to ultimate.

Unfortunately, testing for convergence of the product of an infinite series is often difficult, as it usually involves intractable series of logarithms.

Such is the case with Mr. Sherman's equation. Several quick attempts failed to produce an algebraic solution to the question of whether the product series converges for all values of a and b, or some limited set. The reviewers are, however, convinced that with further effort (perhaps by someone more adept at real variable analysis) a solution to this question is obtainable.

Our investigations did lead us to the following conclusion, however. Consider the following hypothetical loss development data.

Maturity(t)	Reported Losses	Report-to-Report Development Factor
1	$ 100	2.000
2	200	1.500
3	300	1.333
4	400	1.250
5	300	1.200
6	600	1.167
7	700	1.143
8	800	1.125
9	900	1.111
10	1,000	.
.	.	.
.	.	.

The reader should readily recognize that if the loss development continues at its present rate of $100 per interval, losses will be infinite. It follows that the loss development factor product series must not converge.

However, it is also true that the development factors above can be produced identically using Mr. Sherman's equation by setting both a and b equal to one. This strongly suggests that the parameter b should be restricted to values greater than one in order to guarantee convergence.

We were led to raise the question of convergence by the discussion in Section II of Mr. Sherman's paper. In that section he derives the rate of decay for his model and points out that the rate of decay (as we have defined it) declines towards zero as t increases. (His "decay ratio" approaches unity.) This is in strong contrast to the exponential decay model, under which the rate of decay is constant for all values of t.

Upon initial reading of this section of the paper we were concerned that a declining rate of decay implied non-convergence of the ultimate development factor. However, upon reflection this does not appear to be the case.

It seems reasonable that there should be a relationship between the rate of decay of the development and the convergence or non-convergence of the development factor to ultimate. Clearly this question should be resolved before any model gains widespread use.

FITTING THE FUNCTION TO ACTUAL DATA

In Section I of his paper, Mr. Sherman suggests a simple procedure for fitting his equation to loss development factor data. The technique uses only natural logarithms, exponentials and linear regression, and therefore has the distinct advantage of requiring only a (reasonably sophisticated) pocket calculator to perform the calculations.

While the technique is handy, any prospective user should be aware that it does suffer from several problems. First, under the proposed transformation, an actual loss development factor of 1.000 is inadmissable because the natural logarithm of zero is undefined. What does one do under such circumstances? One possibility is to substitute a factor "sufficiently" close to 1.000.

A similar problem exists with observed development factors less than 1.000. These must be ignored or somehow smoothed out of the data.

Another problem is that the fitting technique minimizes the errors of $\ln(f(t) - 1)$ and not the errors of $f(t)$. The result is that, in the fitting process, differences between actual and fitted values are more significant when the development factors are close to 1.000 than when the development factors are significantly greater than 1.000. This bias in the errors is not necessarily bad; it simply needs to be understood as a part of the fitting process.

A related problem is that, since the measured errors are of the logs of the factors rather than the factors themselves, the coefficient of determination that results directly from the computation is inaccurate, and usually overstates the goodness of the fit.

For example, the coefficient of determination of the fit presented in Exhibit I is described as .99887. This is the coefficient of determination of a straight line through column (4) and not the coefficient of determination of the inverse power curve through column (2). This latter coefficient of determination is .971, which is still good, but less favorable than the author suggests (especially considering that there are only three data points being fit).

Obviously, the proper measurement of errors, and the decision as to what errors to minimize is key to any curve fitting procedure.

A particular problem with fitting Mr. Sherman's inverse power curve (or any of the other alternative curves that we have proposed) to the report-to-report development factors is that the resulting fitted factors will be multiplied together, compounding the errors. This can be a particular problem when the errors are not random. In such cases a significant error in the development factors to ultimate can accumulate.

For example, in Section II of his paper, Mr. Sherman uses his model to extrapolate general liability report-to-report development factors, using only the first few development factors to obtain the equation's parameters. While expressing some caution about the reliability of the resulting factors, the author does suggest that the extrapolated report-to-report development factors compare relatively favorably when compared to the actual factors over each interval.

The comparison is considerably less favorable if one compares the compounded, rather than the report-to-report, development factors. The errors in the IBNR reserve that would result from using the extrapolated factors range from 16% (1.667 vs. 1.575), to 112% (1.495 versus 1.234).

Year of Development	Extrapolated Development Factors Based On			Actual Factors
	First 2 Factors	First 3 Factors	First 4 Factors	
3 to 15	1.667			1.575
4 to 15	1.455	1.670		1.329
5 to 15	1.331	1.495	1.322	1.234

An alternative fitting approach that avoids the compounding of errors would be to fit the curve that results from compounding the factors to the actual loss emergence data, measuring the errors between actual and fitted losses reported at each valuation point. In essence, this alternative approach "dollar weights" the fitted factors.

An outline of this approach can be stated as follows. Minimize

$$\sum_{\forall(p,t)} (\hat{L}_{(p,t)} - L_{(p,t)})^2$$

Where $L_{(p,t)}$ is a valid point in the loss triangle, with p representing the exposure period of the losses (accident year, for example) and t representing the valuation point; and

$$\hat{L}_{(p,t)} = L^*_{(p,t^*)} \frac{\prod_{v=1}^{t} f_v}{\prod_{v=1}^{t^*} f_v}$$

where $L^*_{(p,t^*)}$ is some base value for the accident year in question at some time t^* (e.g., the latest valuation point), and f is the chosen decay model.

The problem so stated can be solved using partial derivatives and non-linear programming techniques.

CONCLUSION

Mr. Sherman's paper provides an excellent introduction to a timely topic. The paper presents practical ideas and approaches for the solution of problems encountered with increasing regularity in reserve analysis: incomplete, immature or fluctuating loss development data. We wholeheartedly agree with the author that the fitting of loss development data to curves such as the inverse power function often provides a practical solution to these problems.

AUTHOR'S REPLY TO DISCUSSION

One of my hopes in writing a paper on development factor analysis was that it would help to stimulate others in their research in this area. The subject is so important that if Charles Darwin were alive today, his contribution to link ratio analysis might be the discovery of the long-awaited missing link. In the absence of a re-vitalized Darwin, we are fortunate to have the review of Stephen P. Lowe and David F. Mohrman and the ideas and models they present.

Why is this subject so timely? Let us consider a commonly encountered situation. As we head out toward the more mature parts of our development triangles, and our data transition from the credible to the less credible, we are presented with several alternatives (presented in ascending order of preference):

1. Satisfy the actuarial craving to deploy a complex model which fits the given data points perfectly and wildly diverges as we attempt to use it to extrapolate beyond the historical experience.
2. Close your eyes, swallow hard, make an undocumentable selection, smile like a Cheshire cat, turn to the world at large and exclaim, "Trust me."
3. Rely on the indications of only two or three factors, each of which is often heavily impacted by large claims. The dictum, "When in doubt, throw it out," is often invoked here.
4. Use models which closely fit related data to extrapolate factors for later development factors based on earlier factors from more credible data. Some of the Lowe-Mohrman models could be very useful here.

It would have been helpful if the reviewers had provided some comparative tests of how well their models fit actual data, such as was provided for equation (5) and the salvage and subrogation data. I suspect that equation (5) often would represent a better fitting model than the basic inverse power function because each term behaves in the same manner as the inverse power curve and an extra parameter should increase accuracy. However, equations (6) through (8) present models which add complexity and may or may not increase accuracy. It is often true that more complex models improve accuracy within the range of historical data points. But, it is also true that they may tend to diverge from expected patterns when used for the purpose of extrapolation. The advantage of a simpler model is that its behavior for extrapolation purposes tends to be more reliable. This suggests that an important criterion in assessing different models is their ability to predict known factors for later periods of development based solely on earlier factors.

The reviewers' presentation of the hypothetical results obtained when a and b are set equal to unity in the two-parameter inverse power curve is quite interesting. It clearly illustrates the necessity of an eventual decline in the incremental amounts of development if convergence is to occur. This is generally not a problem as long as the historical data include later periods when the incremental amounts of development decline. If the incremental amounts are constant (as in Lowe's example) or increasing, the product series will diverge.

Lowe and Mohrman observe that the fitting method in the paper minimizes the errors in $\ln(f(t) - 1)$ and not the errors of $f(t)$. They further observe that differences between actual and fitted values are more significant when the development factors are close to 1.000 than when the development factors are significantly greater than 1.000. Thus, the fitting method puts more emphasis on factors for the more mature periods than for the earlier periods. This is usually a desirable result, since the estimated factors of consequence are those of the later development periods. For applications where greater accuracy is required for the earlier periods, the errors of $f(t)$ should be minimized instead of $\ln(f(t) - 1)$.

With regard to estimating a tail factor by multiplying together the extrapolated factors, the reviewers correctly note that this procedure results in a compounding of the errors. It should be noted, however, that a compounding of errors as the extrapolating proceeds further into the future is probably unavoidable as it would appear to be inherent in the process of foreseeing the distant future. The degree of uncertainty in our estimates will, in all likelihood, increase progressively as we forecast events occurring further away from the immediate present. Even if we are using the best possible model, the extrapolation is based on data of limited credibility and the results are very sensitive to statistical fluctuations in the historical experience.

In closing, it may be noted that the inverse power curve can easily be used for estimating the number of IBNR claims as an alternate method to that presented by Edward Weissner in his paper, "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood" (*PCAS* LXV, 1978). The procedure is much easier to apply and chi-square tests for goodness of fit indicate that a closer fit is obtained using the inverse power curve rather than maximum likelihood. A comparison of actual and fitted development factors for cumulative reported claims is included here as Exhibit 1.

EXHIBIT 1

COMPARISON OF ACTUAL AND FITTED REPORTED COUNT DEVELOPMENT FACTORS USING AN INVERSE POWER FUNCTION

Year of Develop-ment	Medical Malpractice		Other Bodily Injury Liability		Auto Bodily Injury Liability	
	Actual	Fitted	Actual	Fitted	Actual	Fitted
2:1	2.094	2.162	1.274	1.295	1.160	1.163
3:2	1.179	1.199	1.062	1.060	1.013	1.014
4:3	1.099	1.071	1.027	1.024	1.004	1.003
5:4	1.032	1.034	1.014	1.012	1.001	1.001
6:5	1.021	1.019	1.006	1.007	1.000	1.001
7:6	1.010	1.012	1.005	1.005	1.000	1.000
8:7	1.008	1.008	1.003	1.004	1.000	1.000
9:8	1.007	1.006	1.003	1.003	1.000	1.000
10:9	1.004	1.004	1.001	1.002	1.000	1.000
R^2		.98759		.99038		.99483
a =		1.16241		0.29501		0.16288
b =		2.54727		2.29051		3.56889
c =		−1.00000		−1.00000		−1.00000

ADDRESS TO NEW MEMBERS—MAY 9, 1985

THOMAS E. MURRIN

Twenty years ago last November in New York City, at the Fiftieth Anniversary meeting of the Casualty Actuarial Society, one of my duties as President of the Society was to admit the new members. At that time there were six Fellows admitted and Stan was one of the ten new Associates. To put our Society's growth in perspective, the membership today is more than two-and-one-half times the number at the time of that meeting. Since then, society and business have changed greatly; in the future they will change even more. Our challenge is to be prepared.

The warm applause of welcome you just heard was well-deserved, genuine, and not perfunctory. It has been so for the 71 years of our Society's history because the audience members know and remember well the effort, the difficulty, and the obstacles encountered along the way, as well as the sense of achievement, pride, and satisfaction that one feels toward having completed one of the two milestones that we note here this morning. Equally deserving of the applause of recognition and welcome are the spouses, the families, and the friends who shared the sacrifice and encouraged your endeavors. For Associates, it is the first major step and for Fellows a second, not a final one. While Associateship does confer membership, I urge all Associates not to slacken their efforts but to achieve Fellowship by concentrating on the remaining exams in the next few years. Incidentally, the proportion of Fellows and Associates is about the same as it was for 1984. Successful completion of these exams will significantly enhance your actuarial knowledge and effectiveness, as well as broaden your horizons of your dynamic business and the role the actuary plays in it. Additional experience gained in your employment between the Associateship and the Fellowship designations will also help improve your understanding of the profession, the business of insurance, and the society it serves.

I would equally urge the new Fellows to consider continuing education as part of their career development. Increased knowledge can be gained in many ways—through formal educational programs; actuarial, insurance, and financial literature; broader and new experiences that impart additional knowledge; as well as private study. Increasingly, seminars sponsored by different organizations, including the Casualty Actuarial Society and the Academy, offer opportunities for continuing education. On a related subject, I would urge the Fellows to use some of the time previously allocated to exam preparation to serving on our Society's committees—or by writing papers for the Society's *Proceedings* on topics where your experience, knowledge, and/or research has provided you with insights that are worthy of sharing with the members. This is particularly true if the topical area is one where a previous paper has opened up a new area for exploration, or one wherein the literature is thin.

To sum up before turning to another subject: Associates and Fellows, take pride and satisfaction in the achievements gained thus far, and in your own way enhance your knowledge and value to the profession and to the Society. Each of you probably has thirty or more years of your career ahead of you. It will pass quickly, more quickly than you think, and the enjoyment and satisfaction you receive from it will be in proportion to your efforts to make it so.

Now to a topic which is part of every profession. In a word, I am talking about professionalism. It is a subject which must be uppermost in your mind throughout your career because, in my opinion, no professional person—or even one whose occupation is not so designated—can be internally satisfied, content, or proud of perceived accomplishments or success (no matter how great in monetary terms) if they have been gained by compromising personal integrity, ethical principles, or truthfulness.

The introductory sentences of the "Guides to Professional Conduct" are worth repeating:

> *"Professional conduct involves the actuary's own sense of integrity and his professional relationship with those to whom he renders service, with his employer, with other members of the profession, and with the world at large. In all these relationships every member of the profession is concerned with his own behavior and, as the good name of the profession is the concern of all its members, with the behavior of his colleagues."*

These guides have been developed over many years and revised from time to time, but with great care and deliberation always to avoid infringement upon the personal nature of the actuary's work and to keep in mind the overriding importance of his or her professional duty and relationship to a client and to employers, as well as to colleagues. The guides are worthy of careful reading from time to time—several times a year—to keep them fresh in mind, as are the more detailed supporting interpretive opinions.

In the introduction to "Interpretive Opinion 1," Francis Bacon is quoted as follows:

> *"I hold every man a debtor to his profession, from which as men of course do seek to receive countenance and profit, so ought they of duty to endeavor themselves by way of amends to be a help and ornament thereunto."*

Each of our members should be a help and ornament to our profession and the Society, and reciprocally so to each member colleague.

For the first time, in my memory, we are having a panel discussion on actuarial malpractice. I am confident that through familiarity, understanding, and application of the "Guides to Professional Conduct" in all our endeavors we can avoid or minimize any exposure to allegations of actuarial malpractice. I urge all of you to attend and participate in the discussion.

In conclusion, I want to wish each of the new Fellows and Associates long, healthy, successful, and happy careers in the true meaning of each word. I also thank President Stan Khury for affording me this unique privilege of welcoming you.

KEYNOTE ADDRESS—MAY 9, 1985
ADAPTING TO AN ERA OF CHANGE

DR. MICHAEL J. KAMI

Today I am going to discuss the future. I read the Casualty Actuarial Society booklet about all your accomplishments and all the exams you have to take; obviously you know all about the past. The future, as you know from looking at the record of insurance companies, is a little bit more confusing.

There is nothing permanent except change. That observation was made some 500 years before Christ. An early papyrus dated 4000 B.C., the first written piece of information, started: "Alas, things are not what they used to be."

We live today in a new era. For five thousand years we were an agricultural society: the major source of power was human strength and a few animals. Our industrial revolution started only 250 years ago. Since the invention of the steam and then, of course, the combustion engine and nuclear power, we multiplied our force by a hundred, a thousand, a million, a trillion times.

And then we left the industrial society and entered the era of service. The exact date was 1956. That was then the proportion of employees in the service industry in the United States passed the 50 percent mark. And today, in 1985, 76 percent of workers are in the service industry; only 24 percent are in industries such as construction or manufacturing. The service era also was an evolutionary era.

I would like to suggest that we entered, in 1980, only five years ago, a brand new era of our society: the era of knowledge. I believe that is a very important transition for the future, and that a major psychological adaptation is required to really understand that future.

You probably took a lot of courses in economics. Forget them all. All the Nobel prizes for analysis of the economics of modern society are based on a premise that the standard of living of a country depends on the proportion and growth of energy consumption in that country. For 250 years economies could be modelled by the approximation that the consumption of energy and the GNP moved in parallel: the more energy consumed, the higher the standard of living.

However, energy consumption per dollar of GNP since 1973, which was the time of the OPEC oil embargo, has declined gradually. The real GNP since then has increased by 43 percent, while the energy consumption in the United States has decreased by 23 percent. Something new has been added to the entire structure of our society.

Furthermore, we are facing a brand new economy because of microminiaturization. This is not just an empty word. For example, if ten million cars built in 1974 and ten million cars built in 1985 are compared, the average difference is two thousand pounds less weight per car. Multiply this weight difference by ten million cars and you have how much? Twenty billion pound less . . . of what? Of steel, of aluminum, of tires, of rubber, of everything. Yet, the product performs the same function.

Another example is even more dramatic. Fifteen years ago, a telephone conversation to London relied on a transatlantic cable. That transatlantic cable weighed 275,000 tons, and required ships made out of steel to lay the cable. Today, the same communication, or better, uses a one-quarter ton satellite. This represents a difference of many magnitudes in materials needed, with the same results produced.

This is a revolutionary new world, with changes in economics, changes in perception, and a new ingredient—knowledge—that differs from energy, which always ruled us as a consumable commodity. Knowledge is a self-regenerative cumulative commodity and we really have to understand the change that has occurred.

Psychological adaptation to this new society is critical. Yet it is not accomplished easily. I deal with presidents of large corporations and small corporations, with members of the power structure. The power structure, in my opinion is like an ostrich with its head in the sand. And this ostrich does not look up.

How does one adapt? How does one change? How does one plan, establish strategies, and think of everything that is uncertain about this new revolutionary future, but in practical terms. Why don't corporations adapt? Corporations fail and new ones replace them. For example, Penn Square, which didn't understand that the oil boom was not forever. AM International: the company that still used electromechanical technology when everybody else was using electronics. Braniff, whose megalomaniac president decided to paint planes multi-color, thinking that people would fly Braniff because the plane was pink and yellow. And Lionel: who plays with trains today? Your kids play with electronic games like Atari. Atari started with seven people and in eighteen months grew to a $1.3 billion company with thousands of employees. And today, guess what Atari is:

seven people in South Korea. Today we have a compression of time: things don't last as long.

Now, why do corporations and institutions go through a long, wonderful period of growth and then decline very, very fast? There is an unlimited potential for losses. Let me offer a few examples. Beatrice Foods was voted in 1977 by Dunn's Review as one of the best managed companies. Today they are struggling to create an identity with some questionable advertising programs. Caterpillar was voted the best managed company in 1978. Today, in 1985, half of Caterpillar's employment in Peoria is gone, and Komatzu is beating Caterpillar throughout the world. In 1978, the bank voted the best managed bank in the United States was Continental of Illinois. In 1984, it was voted by peer banks as the most hated, least admired in the Fortune 500. Pan Am was great in the 1960s and is almost bankrupt today.

You probably have read the book by Thomas Peters and Robert Waterman, *In Search of Excellence*. And then you probably read the *Business Week* article "Oops." What was "Oops"? The book was researched four years ago; a current review of the companies selected as the most excellent companies is quite enlightening. Here then, is "excellence" revisted.

Texas Instruments: major failures in marketing consumer electronics and quality failures. Caterpillar: losing its market position in the world.

Fleur: a good example of corporate forecasting. This was a construction company that spent $2.5 billion to buy St. Johns minerals to have a supply of copper; six months later, the price of copper declined to the lowest level since 1932. That's a decision.

Levi-Strauss: remember denims and jeans? Levi-Strauss became arrogant and started trying to dictate the market.

Kodak is a company married to an obsolete technology of silver nitrate. In a few years cameras won't have film, they will have little chips. After taking 300 pictures, remove the chip, and play it through a television screen. You can buy it, by the way, next year. This is an excellent example of change.

Some companies—by which we really mean the companys' management, people, and talent—lose, while others succeed. IBM, Merrill-Lynch, Lincoln Electric, Cannon, American Express—except for the casualty insurance operation, Federal Express, and Delta are consistent winners. We also have consistent losers. Why?

One answer lies in a simple but complex observation: the economic life of a decision, whether it's a decision to build a plant, to pay a compensation plan, to establish a risk factor, to promote someone or to retire someone, or to get a product on the market, has shortened. It has shortened from ten years to seven years to four years to two or three years.

Changes are occurring faster. Technological breakthroughs are faster. Computers are now changing at the rate of every four years. Combined with two-and-a-half year delivery time, this means that your machine already is obsolete by the time it comes into your office.

In response to the shortened economic life of decisions, we must make decisions faster. Doing so tends to increase risk. But management doesn't want to increase risk; therefore we need faster decisions with the same risk factor. There is the simple—but complex—formula for the strategy of planning.

Let us examine this a little bit. I want to make decisions today and in the future twice as quickly as ever before. Therefore, I need better data twice as fast. I want better information about the external environment and the internal operation of the company organization. Twice as fast; twice as good. I need better communications. It's no use just having the data. I want twice-as-fast communications between people, between customers, between entities. Therefore, I want a flat organization, and I want a fast feedback of communications from the grassroots.

But there is no use having fantastic communications and fantastic data if you have people incapable of making the decisions. I also want faster decision-making.

In short, what I am saying is that we have entered an era of unpredictability in an era of knowledge; this is a paradox in itself. And that unpredictability has created fluctuations. We have entered an era where the fluctuations are going to be twice as big and occur in one-half the time. This is my formula for fluctuations.

What are some examples of these fluctuations?

In the past five or six years interest rates have been 6 percent, 10 percent, 20 percent, 24 percent, 10 percent, 20 percent, and then 10 percent. Inflation has been 6 percent, 18 percent, 2 percent. How about the U.S. dollar versus other currencies? In 1980 and 1981, it was 65 percent of parity; at the beginning of this year, it was 165 percent of parity. These figures were never predicted, predictable, or used. If you knew how to do so, you wouldn't be here: you would be on your 100-foot yacht in the Mediterranean with a crew of nine.

Remember that in 1980 our President—at that time slightly younger—was running on the promise of a balanced budget. And today we have a $200 billion a year unbalanced budget. I don't blame him; I'm just saying that's life.

When you talk about the Iran-Iraq War, you talk about a 6-day war that changed to a 6-month war that changed to a 6-year war. Every expert in the world said that if we experienced a 6-year war between Iran and Iraq the supply of oil would be curtailed and there would be a shortage of oil in the world. Today we have the biggest glut of oil in history. Those are the predictions. This is the type of forecasting that is common.

Therefore, I say that the era of unpredictability, because of our known adaptation to knowledge, will continue. That's my assumption. Assuming you share this assumption, then what can we do? The first thing is psychological adaptation, learning to expect the unexpected. For instance, a very interesting unexpected that just happened is that there is not a single country in the world—whether it's an oil exporting country, industrial country or non-developing country—that doesn't have a federal deficit.

The world, and I wish I had coined the phrase, is not going to be good or bad, better or worse; it's going to be different. We have to understand this.

Let me describe one major difference that is going to affect our planning, our understanding. The population of the whole world is only growing now at 1.7 percent per year. The population of the key industrial countries is growing at 0.3 percent: 0.9 in the United States but 0.0 in West Germany, 0.1 in France, and 0.2 in England. When the world population stops growing, particularly in the industrial countries, there will be fewer consumers to buy things. Fewer consumers to buy things does not necessarily imply a worse economy, but it will be a different economy. We have entered a period of replacement and substitution instead of growth and addition. For everything that goes up something has to go down. That is substitution. For every market or telecommunication or computer or electronics segment that grows at a rate of forty percent, some other other sector of the economy is going to shrink at a rate of forty percent.

When this happens in a slow growth world economy, the companies must compete. For every company that grows at a rate of 15 or 20 percent, one must decline at 15 to 20 percent. Therefore, this situation creates a very competitive environment. This is apparent even today in the manufacturing sector, particularly with regard to the imports that are coming to United States. So one of the very important considerations for the future, in addition to the fast fluctuations I described earlier, will be an extremely competitive situation between companies.

There are two generic strategies that sellers must consider along with all the risks and all the various events in this world. A seller dealing with commodities that only depend on price must be one of two things. That seller must be a low-cost producer worldwide, because worldwide the price comparisons are being made, and it doesn't matter whether it's a service or a product. Or, if not the low-cost producer, the seller must have a deep pocket. A deep pocket is very important because a larger company that doesn't have the low-cost production can reduce prices for quite a while to put the small company out of the business, at least for a while. The winning combination for a commodity product with no differentiation is worldwide low-cost with a deep pocket.

Such a combination forces competitors to seek specialty products or try for uniqueness. Service becomes very important because here you can charge more in that service provides a dimension of uniqueness in addition to the price.

Another approach to consider is the niche strategy. This is sometimes illustrated by the difference between a very small company and a large company. The large companies fight for an additional one percent market share. That one percent for Coca-Cola is worth $250 million; that one percent for a cigarette manufacturer is worth $300 million. The little companies fight for the one percent that the large competitor doesn't want. A company must decide whether it wants 90 percent of the one percent specialty market or one percent of the overall market. Only the company can decide where it fits.

You may say that I am exaggerating; I am not. For years I made a comment that it is useless to fight against Pepsi-Cola and Coca-Cola: they fight against each other and the little competitors lose. So I suggested a possible specialty—a cola for dogs. After all, that's a small market; I am sure Coca-Cola will not want it. However, ladies and gentlemen, as of two months ago, Canine Cola has been introduced in Phoenix and its sells for $4.99 a six-pack. I tasted it; it isn't bad.

The next thing you have to consider in this changing environment is the present tremendous merger-mania. And merger-mania will continue because as big corporations get bigger they want to consolidate their hold on the market. This is not nefarious; it's good business, at least if you are smart enough to do it right. Every merger is intended to combine two plus two and yield five.

That's synergy. The reality is that in 80 percent of the cases two plus two yields three.

Let me suggest something else that is going to happen. Because of the compression of time, because of the desire to be in the market fast, I recommend joint ventures. Very difficult to accomplish, but they allow "instant vertical integration." And you see this approach with increasing frequency. Companies take someone who has the marketing ability, someone who has the production ability, someone who has the money, and someone who has manufacturing or distribution capability: instant joint venture, instant entry into the market, instant everything. It's necessary because the current environment allows only four months or six months, not two and a half years.

Procter and Gamble, which was the paragon of marketing, is changing its entire marketing and merchandising approach. They use market forecasters. Why? You can't go to Squeedunk and test-market a product for two and a half years and then go national because by the time you go national, everybody else and their brother is already there. So even Procter and Gamble is trying to learn how to introduce new products, new ideas, and new expansions of their existing products. This applies to the insurance business, too. Decide on extension of products, or new products, by sampling in three months time.

Coca-Cola, which in seventy years introduced three new products, has a new management that I really admire. Coca-Cola introduced more than a dozen new products in three years and lately has made a major risk decision by changing a formula that was okay for seventy years.

"Speed" is the message that I'm trying to give you. To get data fast we must understand information management. We really have to manage the management of information, not the information itself. We are up to our ears in information. How do we manage information better?

I do believe that in a few years companies that don't have the executive communication station are going to be passe. But what do I mean by that? I mean an executive, president of a company, chairman of the board, who sits at his terminal and really uses it; not one who gives it to a secretary for word processing. I mean a manager who actually gets fast information, fast data, from distributive shared data bases and is able to make intelligent decisions. Meetings are no longer occasions to present the data for hours; instead the data are available ahead of time and the meeting is to discuss strategy and make decisions. Virtually no presidents of corporations have reached this stage.

Some companies pioneer. Flextime to flexplace. I don't know why actuaries have to be in an office in Hartford. They just as well could live in Colorado and have a nice computer and communications facilities and never see anybody. I know of four insurance companies that have saved a great amount of money by not having people commute but having them work at home. But that's breaking a mold, breaking tradition.

Uniqueness and innovation are important characteristics for success. The key ingredient to profitability is for a company to have uniqueness. The future depends on the creativity in the present. The mind has tremendous potential; creativity is something that really must be cultivated.

Coleco, for instance, lost $500 million on a computer, but they made it up. They made it up on what? On this strange Cabbage Patch doll. $500 million. Why? Not because of the doll but because someone said that the kids would like to sign the adoption papers. What I admire most is not the Cabbage Patch doll, but is the fact that someone at Coleco saw the doll somewhere in Arkansas or Georgia and said, "Yes, that is going to sell," and bought it.

Now, think of yourself, your management, your operation. How many managements really are that innovative to try something new? That's what is needed. And that newness is important because of the fast obsolescence of products.

However, there is a clever way of being unique and a stupid way. For example, a principle that was used by airlines was to fly passengers through a hub city: Atlanta for Delta; Chicago for United; Dallas for American. But remember the gentleman who said five years ago, "I'm going to ship a package from Los Angeles to San Francisco by taking it from Los Angeles to Memphis and then from there to San Francisco." The world's greatest planner said; "That's a stupid idea!" As you know, Federal Express has become an extremely innovative, multi-billion dollar corporation.

But you see that innovation is only possible through people. No machine can innovate. And therefore I say, take a talent inventory in your company. This is how to do it. Think of four people, other than yourself, who are so indispensable to the company that it would be a major catastrophe if they left. Now write down their names. If you can't visualize only four peole who are so fantastic that it would make a difference, your company is in trouble. We need talented, unusual people and they don't have to be president of the company. Actually, it sometimes is a disadvantage for the president of the company to be innovative because then everything depends on him.

A good president of the company, a CEO, should have only one characteristic: the ability to find, motivate, nurture, and encourage talent. That's all there is to it. Management must try to be receptive because new ideas, new ways of approaching things, faster ways, define entrepreneurs. Don't expect anything original from an echo.

The best of all worlds is to find individuals who combine several unique characteristics. Some people are detectors of change, they can see things changing. Other people, having been told of changes become the architects of change: they plan and they calculate. Then the agents of change take that plan and run with it. Successful companies need people who are detectors of change, architects of change, and agents of change, all combined in one person. Those are the talented, unique, indispensible people that are needed in a company.

Another important consideration is internal organization. For faster and better communication, eliminate the middle management of a company, seek a horizontal organization. An organization must be fast, fluid, and flexible.

Look for a lean, highly-paid, intelligent staff. Strive for an organized chaos. An organized chaos has the ability to change fast. Seek a company that has a chairman of the board or chief executive officer and eighty division managers reporting to him. Does that sound ridiculous? The eighty managers must be very smart because if they are not the boss is going to spend all his life talking to them. But if they're smart, they can work independently. There really are companies with eighty people reporting to one president that are doing very, very well.

Think about why we need a foreman on the factory floor. The foreman serves only two purposes: first, to tell the worker what to do; and second, to maintain discipline. If the worker is trained to read a cathode ray tube or terminal and knows what to do, he does not need the foreman. And if the worker is motivated by incentive, he does not need the discipline. And companies are taking this approach to eliminate middle management. Decentralized authority is effective not because it's nice but because the speed of communication and decision-making requires it. Fight bureaucracy.

Another thing that is very important, and that is a part of the knowledge of the external environment, is the urgent need for customer orientation and for understanding the present and future behavior of people. What do I mean by that? We live today in an era where the homogeneous mass market no longer exists. There is no homogeneous market because we are in an era of pluralism.

When I came to this country 45 years ago, I wanted to be assimilated. I wanted to be part of the masses. Today, in 1985, I see a set of subcultures in which the young don't want to be like the old, and the old don't want to be like the young. Women don't want to be like men; men don't want to be like women. The blacks don't want to be like whites; and vice versa. These are all niches. Why do you think that big companies no longer have one advertising agency? They have an Hispanic advertising agency, a black advertising agency, a West advertising agency, and an East advertising agency, et cetera. They understand the pluralism.

Today's consumer, lovable but irrational, may be more affluent because of two family members working. The consumer may be less loyal to a brand or a company or an insurance product or anything else. Today's consumer is more mature, as you know. The consumer is older, more impatient, and less gullible. There is a new female role. All these factors must be understood by a company, and incorporated into a company's changes.

We have a trend that is very interesting: a trend toward extremes, which I call polarization. That polarization stems from the fact that we now have a bimodal society. The bimodal society has fewer of the blue-collar $30-an-hour workers; but more of the service workers at $9,000 to $10,000 per year, and more of the technical, managerial, talent workers at $30,000, with few workers in the middle of these extremes.

I see a trend to extremes: we now need either no service or excellent service; we need cheap products or expensive products. The strategies of companies are going to be aimed at the extremes. The companies hedging in the middle are not going to succeed.

The principles of mass merchandising, specialized boutiques, and bargain basements (for example, Sears, Gucchi, and K-Mart, respectively) apply to financial services, too. At one extreme, the customer is provided with insurance, mortgages, financial accounts, and stock brokerage services. For example, American Express and Merrill-Lynch take this approach. On a bargain basement level, Schwab or the Bank of America offer no service, but discount prices. Then there's the specialized boutique, a salesman in East Cleveland sitting with an old lady discussing five shares of stock for three hours. Where does your company fit? What is your marketing, your merchandising, your understanding?

From what I have said already you know that strategic planning is becoming tougher all the time. But I would like to eliminate the word "strategic planning" from your vocabulary. I don't want planning; I want management. Planning is filling books; planning is making learned studies; planning is preparing presentations; planning is not action. Management is action. Therefore, let's combine planning with action. And I want you to put into your planning more thinking; not just extrapolation, calculation, regression, and so forth. I want you to put your imagination to work, to imagine future new concepts, to understand unpredictability.

One simple planning principle is to identify an important change you can make tomorrow. I want an action proposal to be effected next Monday morning. If you present to me a plan for the future that says in 1989 you are going to change something, I will not be interested because you have four years to change your mind. But I will be interested if you say that, towards that change in 1989, you are going to spend money on energy and people, Monday morning, then you are doing long-range planning and implementation. Very few companies actually do it well.

Avoid paralysis through analysis. Because of the compression of time, I need smarter people, better tools, and better data. But it's not absolute. All I want is to be five percent better than the other guy. Because I'm not using absolute planning, I only want be a little bit better than the competition and I want to compare myself to the leading edge. Who does the best calculations? Who does the best planning? Who has the best personnel department? Who has the best data processing department in the world? How do I compare? And if I'm the leader; how fast are they catching up to me?

A key element of your planning should occur each morning: identify five issues that are absolutely urgent—in your opinion—for your company, your profession, or your operations. Five issues only, in order of priority. You may have a hundred issues; select only five. Then for each issue identify the action you will take today. Let me make it a little bit more difficult for you by insisting that the action be innovative and imaginative. Bernard Shaw said many years ago, "For every complex problem there is a simple solution." He was wrong. Ladies and gentlemen, there are no simplistic solutions; there are complex, imaginative, good solutions to very complex problems.

Don't believe in forecasting. Believe in assumptions about the future. When we talk of tomorrow, the gods laugh. And when you think that you really know how to forecast, and you believe in your own magic, be careful.

Let me tell you one little story. One of the greatest miscalculations in IBM history was in 1957, when I predicted for IBM that there would be only 52 large computers in the world. And now there are hundreds of thousands. How did I make that famous forecast? There were 52 experts in IBM, one for wholesaling, one for electrical, one for insurance, one for banking, and so on. I went to each one of them and showed them something new, a set of plans they didn't understand. Every one of the 52 was afraid to say, "I can't sell any," and didn't want to commit themselves to sell two, so they each said they could sell one. Those experts provided the basis for my famous forecast, but that is not an illustration of how forecasts should be made.

I really believe that you have in the current environment a tremendous opportunity to work with the two sides of your brain, left and right. You have tremendous mathematical, analytical, technological abilities. And you also have the ability, hopefully, to look at the world that is changing, to understand that change, and to introduce this compression of time, technology, psychology, action, and reaction into your mathematical formulas through intuition or guesstimate. You can provide an unforgettable and extremely valuable combination.

I would like to close with my favorite prayer, by Niebuhr: "God give me serenity to accept things I cannot change, courage to change things I can change; and the wisdom to know the difference between the two."

MINUTES OF THE 1985 SPRING MEETING

May 8–11, 1985

BOCA RATON HOTEL AND CLUB, BOCA RATON, FLORIDA

Wednesday, May 8, 1985

The Board of Directors held their regular quarterly meeting from 1:00 p.m. to 4:00 p.m.

Registration was held from 4:00 p.m. to 6:00 p.m.

The Officers held a reception for the new Fellows and their spouses from 5:30 p.m. to 6:30 p.m.

A general reception for all members and guests was held from 6:30 p.m. to 7:30 p.m.

Thursday, May 9, 1985

Registration continued from 7:15 a.m. to 8:00 a.m.

President C. K. Khury opening the meeting at 8:00 a.m. The first order of business was the admission of members.

Mr. Khury recognized the sixty-eight new Associates and presented diplomas to the nineteen new Fellows, who were introduced by Mr. Wayne Fisher, Vice President—Membership. The names of these individuals follow.

FELLOWS

Francois Bertrand	John R. Forney, Jr.	Allan R. Neis
Raja R. Bhagavatula	Loyd L. Fueston, Jr.	Donald W. Palmer
William P. Biegaj	Alan J. Hapke	Lois A. Ross
Terry J. Biscoglia	Heidi E. Hutter	James Surrago
Jeffrey R. Carlson	Michael J. McSally	Diane M. Symnoski
Stephan L. Christiansen	Robert E. Meyer	David L. White
Warren S. Ehrlich		

ASSOCIATES

Mark S. Allaben
Leonard A. Bellafiore
David M. Bellusci
Joseph A. Boor
Brian Y. Brown
George R. Busche
William M. Carpenter
Andrew R. Cartmell
Daniel B. Clark
Frederick F. Cripe
Kathleen F. Curran
Janice Z. Cutler
Todd H. Dashoff
Thomas J. DeFalco
Robert V. DeLiberato
Jacques Dufresne
Bruce G. Earwaker
Kenneth Easlon
Kirk G. Fleming
Robert W. Gardner
Daniel F. Gogol
Kevin M. Greaney
Christy H. Gunn
Gregory L. Hayward
Wayne D. Holdredge
Jeanne M. Hollister
Ruth A. Howald
Charles D. Kline, Jr.
Frederick L. Klinker
Robert H. Lee
Martin A. Lewis
Barry C. Lipton
Mark W. Littmann
Rebecca B. Lyons
Brian P. Maguire
Eugene McGovern
David L. Menning
William J. Miller
Warren D. Montgomery
Robert V. Mucci
Robert G. Muller
Thomas G. Myers
James W. Noyce
Arthur C. Placek
Jeffrey H. Post
Richard A. Quintano
Daniel A. Reppert
Richard D. Robinson
Jeffrey C. Salton
Joseph F. Sarosi
Jeffrey R. Scheuing
Timothy L. Schilling
David C. Scholl
Roger A. Schultz
Arlyn G. Shapiro
John Slusarski
Michael B. Smith
Edward C. Somers
Kathleen W. Terrill
Joseph P. Theisen
Nancy R. Treitel
Jean Vaillancourt
Gerald R. Visintine
Joseph L. Volponi
Stacy J. Weinman
Robert G. Whitlock, Jr.
Robert L. Willsey
Susan K. Woerner

Mr. Khury then introduced Mr. Thomas Murrin, a past President of the Society, who addressed the new members concerning their professional responsibilities.

Mr. Khury announced the first winner of the Harold W. Schloss Scholarship: Steven Book at the University of Iowa.

Mr. Charles A. Bryan described the recent activities of the Committee on Review of Papers, and summarized the five new *Proceedings* papers.

Oakley E. Van Slyke presented a discussion of Lee Steeneck's paper, "Reinsuring the Captive/Specialty Company." Richard E. Sherman, author of "Extrapolating, Smoothing, and Interpolating Development Factors," responded to a discussion of that paper by Stephen Lowe and David Mohrman.

Mr. Michael Walters summarized the activities of the Discussion Paper Program Committee, which had led to the current set of Discussion Papers related to "Analysis of Results, Forecasting and Corporate Planning," and introduced the Discussion Paper topic for 1986, "Reinsurance."

Mr. Khury concluded the business session at 9:00 a.m. and introduced Dr. Michael J. Kami, President, Corporate Planning, Inc. who delivered a very stimulating Keynote Address. Dr. Kami stressed that change is the single greatest force with which managers must deal, and highlighted that point with a discussion of the progress, in the most recent four years, of those corporations singled out for their excellence in 1981.

A panel presentation, "Managing the Insurance Industry into the 1990's," followed. The panel was moderated by Dr. Edwin S. Overman, President of the Insurance Institute of America. The panel members were Mr. John E. Fisher, Chairman, Nationwide Insurance Company; and Mr. Peter B. Walker, Director, McKinsey and Company, Inc.

A buffet luncheon followed from noon to 1:30 p.m.

The afternoon was devoted to presentations of the sixteen Discussion Papers, five new *Proceedings* papers, and a workshop presentation by the Membership Committees.

The Discussion Papers were presented in eight sessions, listed below.

Session 1. *Moderator:* James E. Biller
Chubb Group

"Corporate Planning: An Approach for an Emerging Company"
Authors: Irene K. Bass and Larry D. Carr
Crum & Forster Personal Insurance

"Budget Variances in Insurance Company Operations"
Author: George M. Levine
National Council on Compensation Insurance

Session 2. *Moderator:* Eric F. Gottheim
GEICO

"Branch Office Profit Measurement for Property-Liability Insurers"
Author: Robert P. Butsic
Fireman's Fund Insurance Company

"Measuring Division Operating Profit"
Author: David Skurnick
Argonaut Insurance Company

Session 3. *Moderator:* Stephen W. Philbrick
Tillinghast, Nelson & Warren, Inc.

"A Formal Approach to Catastrophe Risk Assessment and Management"
Author: Karen M. Clark
Commercial Union Insurance Companies

"An Econometric Model of Private Passenger Liability Underwriting Results"
Authors: Richard M. Jaeger and Christopher J. Wachter
Insurance Services Office

Session 4. *Moderator:* Bruce C. Anderson
General Reinsurance

"Measuring the Impact of Unreported Premiums on a Reinsurer's Financial Results"
Author: Douglas J. Collins
Tillinghast, Nelson & Warren, Inc.

"Projecting Calender Period IBNR and Known Loss Using Reserve Study Results"
Authors: Edward W. Weissner and Arthur Beaudoin
Prudential Reinsurance Company

Session 5. *Moderator:* Richard I. Fein
Insurance Technical & Actuarial Consulting Corp.

"Pricing, Planning and Monitoring of Results: an Integrated View"
Author: Stephen P. Lowe
Tillinghast, Nelson & Warren, Inc.

"Application of Principles, Philosophies and Procedures of Corporate Planning to Insurance Companies"
Author: Mary Lou O'Neil
Department of Insurance-New Jersey

Session 6. *Moderator:* Frank Harwayne
National Council on
Compensation Insurance

"The Cash Flow of a Retrospective Rating Plan"
Author: Glenn G. Meyers
The University of Iowa

"Bank Accounts as a Tool for Retrospective Analysis of Experience on Long-Tail Coverages"
Authors: Claudia S. Forde and W. James MacGinnitie, Jr.
Tillinghast, Nelson & Warren, Inc.

Session 7. *Moderator:* Sanford R. Squires
Commercial Union Insurance Companies

"Actuarial Aspects of Financial Reporting"
Author: Lee M. Smith
Ernst & Whinney

"Contingency Margins in Rate Calculations"
Author: Steven G. Lehmann
State Farm Mutual Automobile Insurance Company

Session 8. *Moderator:* Leroy A. Boison, Jr.
Insurance Services Office

"Interaction of Total Return Pricing and Asset Management in a Property/Casualty Company"
Author: Owen M. Gleeson
General Reinsurance Corporation

"Projections of Surplus for Underwriting Strategy"
Author: William R. Gillam
North American Reinsurance Corporation

The five new *Proceedings* Papers are listed below.

"An Estimate of Statistical Variation in Development Factors Methods"
Author: Roger M. Hayne
Milliman & Robertson, Inc.

"A Simulation of the Efficiency of Loss Reserve Estimation Techniques"
Author: James N. Stanard
F & G Reinsurance, Inc.

"On Stein Estimators: Inadmissibility of Admissibility as a Criterion for Selecting Estimators"
Author: James E. Buck, Jr.
Prudential Insurance Company

"Loss Portfolios: Financial Reinsurance"
Author: Lee R. Steeneck
General Reinsurance Corporation

"A Practical Guide to the Single Parameter Pareto Distribution"*
Author: Stephen W. Philbrick
Tillinghast, Nelson & Warren, Inc.

The CAS Membership Committees Workshop was presented in an effort to give Society members a better understanding of the structure and responsibilities of these committees. The Chair of each Committee described the committee's functions, reporting hierarchy and goals. The participants were:

Wayne H. Fisher	*Vice President—Membership*
Linda L. Bell	*Chairman—Education Policy Committee*
Allan Kaufman	*Chairman—Examination Committee*
David L. Miller	*Chairman—Syllabus Committee*

The President's Reception was held from 6:30 p.m. to 7:30 p.m.

Friday, May 10, 1985

Friday was devoted to a continuation of the Thursday afternoon sessions.

A reception was held from 6:30 p.m. to 7:30 p.m.

Saturday, May 11, 1985

A panel entitled "Actuarial Malpractice: How Can It Be Avoided?" was held from 8:30 a.m. to 10:15 a.m. The panelists were:

Moderator: Philip N. Ben Zvi
Senior Vice President
Continental Insurance Cos.

Panelists: William Hager
Principal
Hager & Associates

M. Stanley Hughey
Consulting Actuary
Tillinghast, Nelson & Warren, Inc.

A film from Peat, Marwick, Mitchell & Co. concerning accountants' negligence was shown, followed by a presentation by Mr. Hager on current case law, and a presentation by Mr. Hughey on professional standards.

At 10:45 a.m., Mr. Khury reconvened the business session. The Michelbacher Prize was jointly awarded to Robert Butsic and David Skurnick.

At 11:00 a.m., Mr. F. Lee Bailey addressed the membership concerning his career as a lawyer, with particular emphasis upon negligence suits in which he has been involved. This discussion included the Union Carbide plant in Bhopal, Johns-Manville asbestos litigation, and the DC-10 crash at O'Hare Airport.

Mr. Khury adjourned the meeting at 12:15 p.m.

In attendance by registration records were 261 Fellows; 146 Associates; and 32 guests and subscribers. The list follows.

FELLOWS

Addie, B. J.	Berens, R. M.	Brooks, D. L.
Adler, M.	Bertrand, F.	Brown, N. M., Jr.
Aldorisio, R. P.	Bhagavatula, R. R.	Brown, W. W., Jr.
Alfuth, T. J.	Biegaj, W. P.	Brubaker, R. E.
Arata, D. A.	Bill, R. A.	Bryan, C. A.
Asch, N. E.	Biller, J. E.	Buck, J. E., Jr.
Bartlett, W. N.	Biscoglia, T. J.	Burger, G.
Bashline, D. T.	Bland, W. H.	Byrne, H. T.
Bass, I. K.	Blivess, M. P.	Carlson, J. R.
Bassman, B. C.	Boccitto, B. L.	Chanzit, L. G.
Baum, E. J.	Boison, L. A., Jr.	Cheng, L. W.
Beer, A. J.	Bornhuetter, R. L.	Chernick, D. R.
Belden, S. A.	Bothwell, P. T.	Childs, D. M.
Bell, L. L.	Bouska, A. S.	Christiansen, S. L.
Ben-Zvi, P. N.	Bowen, D. S.	Clinton, R. K.

FELLOWS

Cohen, H. L.
Collins, D. J.
Conger, R. F.
Connell, E. C.
Conners, J. B.
Covney, M. D.
Crowe, P. J.
Cundy, R. M.
Curry, A. C.
Curry, H. E.
Daino, R. A.
Dawson, J.
Dean, C. G.
Demers, D.
Dempster, H. V.
Doepke, M. A.
Dolan, M. C.
Donaldson, J. P.
Drennan, J. P.
Easton, R. D.
Eddy, J. H.
Ehlert, D. W.
Ehrlich, W. S.
Engles, D.
Eyers, R. G.
Faber, J.A.
Fagan, J. L.
Fein, R. I.
Fiebrink, M. E.
Finger, R. J.
Fisher, R. S.
Fisher, W. H.
Flynn, D. P.
Foote, J. M.
Forney, J. R., Jr.
Foster, R. B.
Fresch, G. T.
Fueston, L. L., Jr.
Furst, P. A.
Fusco, M.
Gallagher, T. L.
Garand, C. P.
Ghezzi, T. L.
Gillespie, J. E.
Gleeson, O. M.
Gluck, S. M.
Goddard, D. C.
Goldberg, S. F.
Goldfarb, I. H.
Gottlieb, L. R.
Gorvett, K. P.
Gottheim, E. F.
Grady, D. J.
Graham, T. L.
Grippa, A. J.
Hachemeister, C. A.
Hafling, D. N.
Hall, J. A., III
Haner, W. J.
Hapke, A. J.
Hardy, H. R.
Hartman, D. G.
Harwayne, F.
Haseltine, D. S.
Hayne, R. M.
Henry, D. A.
Henzler, P. J.
Herder, J. M.
Hillhouse, J. A.
Hoffman, D. E.
Honebein, C. W.
Hoylman, D. J.
Hughey, M. S.
Hutter, H. E.
Ingco, A. M.
Jaeger, R. M.
Jameson, S.
John, R. T.
Johnson, L. D.
Johnson, M. A.
Johnston, T. S.
Jones, B. R.
Kaliski, A. E.
Kaufman, A. M.
Kelly, A. E.
Khury, C. K.
Kilbourne, F. W.
Kist, F. O.
Kooken, M. W.
Kozik, T. J.
Krause, G. A.
Kuehn, R. T.
Lamb, R. M.
LaRose, J. G.
Lehmann, S. G.
Levin, J. W.
Linden, O. M.
Lindquist, P. L.
Lombardo, J. S.
Lommele, J. A.
Lonergan, K. F.
Lotkowski, E. P.
Loucks, W. D., Jr.
Lowe, S. P.
MacGinnitie, W.J.
Mahler, H. C.
Makgill, S. S.
Marks, S. D.
Mathewson, S. B.
McAllister, K. C.
McCarter, M. G.
McClenahan, C. L.
McClure, R. D.
McConnell, C. W.

FELLOWS

McGuinness, J. S.
McManus, M. F.
McMurray, M. A.
McSally, M. J.
Meyer, R. E.
Meyers, G. G.
Miccolis, R. S.
Miller, D. L.
Miller, M. J.
Miller, P. D.
Miner, N. B.
Moody, R. A.
Moore, B. D.
Morell, R. K.
Munro, R. E.
Munt, D. S.
Murad, J. A.
Murdza, P. J., Jr.
Murrin, T. E.
Neale, C. L.
Neis, A. R.
Nichols, R. S.
Nickerson, G. V.
Nikstad, J. R.
Niswander, R. E., Jr.
Oakden, D. J.
Oien, R. G.
O'Neil, M. L.
Pagnozzi, R. D.
Palczynski, R. W.
Palmer, D. W.
Parker, C. M.
Pastor, G. H.
Pearl, M. B.
Perkins, W. J.
Petersen, B. A.
Petlick, S.
Philbrick, S. W.
Phillips, H. J.
Pierson, F. D.
Pinto, E.
Pratt, J. J.
Prevosto, V. R.
Purple, J. M.
Reichle, K. A.
Riddlesworth, W. A.
Robertson, J. P.
Rodermund, M.
Rogers, D. J.
Rosenberg, S.
Ross, L. A.
Rowland, W. J.
Ryan K. M.
Scheibl, J. A.
Schumi, J. R.
Schwartzman, J. A.
Sheppard, A. R.
Sherman, O. L., Jr.
Sherman, R. E.
Shoop, E. C.
Shrum, R. G.
Simon, L. J.
Skurnick, D.
Smith, L. M.
Snader, R. H.
Sobel, M. J.
Splitt, D. L.
Squires, S. R.
Standard, J. N.
Steeneck, L. R.
Steer, G. D.
Stewart, C. W.
Strug, E. J.
Surrago, J.
Symnoski, D. M.
Szkoda, S. T.
Taranto, J. V.
Tatge, R. L.
Thompson, K. B.
Tierney, J. P.
Tiller, M. W.
Toothman, M. L.
Van Slyke, O. E.
Venter, G. G.
Wacek, M. G.
Walker, R. D.
Walters, M. A.
Wasserman, D. L.
Weiland, W. T.
Weimer, W. F.
Weissner, E. W.
Weller, A. O.
Westerholm, D. C.
Whatley, P. L.
Whiting, D. R.
Whitman, M.
Wilson, J. C.
Winkleman, J. J.
Wiser, R. F.
Woll, R. G.
Yonkunas, J. P.
Zicarelli, J. D.

ASSOCIATES

Allaben, M. S.
Anderson, B. C.
Andler, J. A.
Bellafiore, L. A.
Bellusci, D. M.
Boor, J. A.
Brown, B. Y.
Bursley, K. H.
Campbell, K. A.
Carpenter, W. M.
Cartmell, A. R.
Chiang, J. D.
Chorpita, F. M.
Clark, D. B.
Connor, V. P.
Covitz, B.
Cripe, F. F.
Curran, K. F.
Currie, R. A.
Cutler, J. Z.
Dashoff, T. H.
Davis, R. C.
DeConti, M. A.
DeFalco, T. J.
Degarmo, L. W.
DeLiberato, R. V.
Deutsch, R. V.
Dufresne, J.
Duperreaut, B.
Dye, M. L.
Earwaker, B. G.
Easlon, K.
Edie, G. M.
Ellefson, T. J.
Elliott, P. L.
Eramo, R. P.
Esposito, D. L.
Flanagan, T. A.
Fleming, K. G.
Forde, C. S.
Friedman, H. H.
Gaillard, M. B.
Gardner, R. W.
Gillam, W. R.
Godbold, N. T.
Gogol, D. F.
Gould, D. E.
Grace, G. S.
Granoff, G.
Greaney, K. M.
Greco, R. E.
Gunn, C. H.
Hale, J. B.
Haskell, G. E.
Hayward, G. L.
Head, T. F.
Hofmann, R. A.
Hollister, J. M.
Howald, R. A.
Hurley, P. M.
Huyck, B. J.
Kline, C. D., Jr.
Koupf, G. I.
Krakowski, I.
LaFrenaye, C.
Lee, R. H.
Levine, G. M.
Lewis, M. A.
Licht, P. M.
Lipton, B. C.
Lis, R. S., Jr.
Littmann, M. W.
Loper, D. J.
Lyons, D. K.
Maguire, B. P.
McConnell, D. M.
McDonald, G. P.
McGovern, E.
Mendelssohn, G.
Menning, D. L.
Miller, W. J.
Montgomery, W. D.
Moody, A. W.
Morgan, W. S.
Morrow, J. B.
Mucci, R. V.
Mulder, E. T.
Muller, R. G.
Myers, T. G.
Nelson, J. K.
Neuhauser, F., Jr.
Nolan, J. D.
Ogden, D. F.
Pei, K.-J.
Penniman, K. T.
Petit, C. I.
Pilon, A.
Placek, A. C.
Port, R. D.
Post, J. H.
Potok, C. M.
Potter, J. A.
Pulis, R. S.
Quintano, R. A.
Ratnaswamy, R.
Reppert, D. A.
Rice, W. V.
Robbins, K. B.
Robinson, R. D.
Roth, R. J.
Ryan, J. F.
Salton, J. C.
Sandler, R. M.
Sansevero, M., Jr.

ASSOCIATES

Sarosi, J. F.
Scheuing, J. R.
Schilling, T. L.
Scholl, D. C.
Schultz, R. A.
Shapiro, A. G.
Silver, M. S.
Singer, P. E.
Skolnik, R. S.
Skrodenis, D. P.
Slusarski, J.
Smith, M. B.
Smith, R. A.
Somers, E. C.
Terrill, K. W.
Theisen, J. P.
Treitel, N. R.
Varca, J. J.
Visintine, G. R.
Visner, S. M.
Vogel, J. F.
Volponi, J. L.
Walker, L. M.
Webster, P. J.
Weinman, S. J.
Whatley, M. W.
Whitlock, R. G., Jr.
Willsey, R. L.
Wilson, W. F.
Winter, A. E.
Woerner, S. K.
Young, R. G.

GUESTS—STUDENTS—SUBSCRIBERS

Adams, M. H.
Altschuler, M.
Benson, D. W.
Brock, R.
Buhl, L.
Bühlmann, H.
Butsic, R.
Callahan, R. J.
Carr, L.
Chang, C. E.
Clark, K.
Clause, R.
Dibella, J.
Englander, J.
Eversmann, T.
Guarini, L.
Gutman, E.
Hager, G. A.
Jacobus, J.
Jensen, P.
Jordan, E.
Kedrow, W. M.
Kuizenga, D.
Leroux, D.
O'Farrell, R.
Spangler, J.
Stenmark, J.
Stephenson, M.
Sterling, J. A.
Thomas, A.
Wade, J. E.
Wootton, G.

Volume LXXII, Part 2 No. 138

PROCEEDINGS

November 10, 11, 12, 1985

PRESIDENTIAL ADDRESS—NOVEMBER 12, 1985
ANALYSIS AND SYNTHESIS

C. K. KHURY

THE ACTUARIAL METHOD

I would like to preface my comments on the *actuarial method* by quoting from the July-August, 1981 *Harvard Business Review*. In his "letter from the editor" column, Mr. Kenneth Andrews writes:

> "Many teachers and students find quantitative techniques and theoretical models easier to teach, intellectually fascinating, beguilingly self-contained, rigorous, and capable of being memorized and quickly applied, widely if not wisely. They forget to insist that for the most part only trivial management problems are neatly structured and quantifiable. All modeling and quantitative analysis directed at a decision are only preludes to subjective judgment. Vision then must transcend technique."

This quote neatly sums up the problem of those who look to canned methods for the solution of their business problems.

We're all familiar with the *scientific method*. One of the key aspects of the scientific method is that it can be applied with predictable outcomes. For example, two parts of hydrogen combined with one part oxygen, under certain conditions, will yield water plus some amount of heat. This process can be repeated at will, by anyone, with identical results. Today I'd like to explore the *actuarial method* in an attempt to obtain a perspective on the nature of actuarial practice: how much art and how much science.

Broadly speaking, given a specific actuarial question, the actuary goes through a number of steps in formulating an answer. For example, the actuary:

- determines the universe of available data that may be relevant,
- selects the types of data to be used,
- makes a number of assumptions,
- adjusts the data to recognize the special conditions associated with the specific problem (deseasonalizing historical data, projecting to recognize future cost changes, etc.),
- synthesizes the data with the aid of mathematical methods and judgment to produce a mathematical response to the question, and
- interprets the numerical result in the context of the original problem.

This sequence of activities, measured strictly against the criteria of the scientific method, would clearly render the actuarial method not a scientific method. If not a scientific method, then what is it?

A close examination of the steps listed above reveals two key points:

- Within the actuarial method, there are several applications of the scientific method. For example, all the mathematical computations, trend methods, deseasonalizations of data, etc., are 100% scientific exercises.
- The various applications of the scientific method are preceded by, connected together with, and followed by a host of applications of judgment.

At the risk of greatly oversimplifying, I could describe the actuarial method as *a process which consists of a number of scientific applications embedded in a collection of judgments*. In this sense, the actuarial method is neither pure art nor pure science; it is a synthesis. And different actuarial problems require different proportions of art and science. If the art-science mix is placed on a continuum where pure art is set at 0 and pure science is set at 100, then estimating next year's pure premium for the auto collision coverage would be closer to 100; but estimating next year's medical malpractice pure premium would be closer to 0.

Not infrequently two actuaries have produced vastly different answers to the same actuarial problem. For example, at a public hearing on medical malpractice rates, the rate level indications as calculated by two actuaries were more than 200 points apart (an *in*crease of 210 percent vs. a *de*crease of 5 percent). If one assumes that the scientific methods used were correctly applied, then the entire difference is attributable to the judgmental aspects of the actuaries' work.

A corollary question is, "To what should the buyer of actuarial services be entitled in terms of the standards that govern the selection and application of those judgments?" Not an easy question to answer; but answer it we must.

FULL DISCLOSURE

My view is that the buyer of actuarial services is entitled to *full disclosure* of the judgments made by the actuary in arriving at a solution to an actuarial problem. Although the precise meaning of full disclosure remains to be worked out, my idea of full disclosure consists of two aspects:

- disclosure of assumptions, judgments, interim conclusions, and whatever else influences the outcome by more than some preselected tolerance, and
- sensitivity analyses sufficient to illustrate the operation of these judgments.

In this manner, the buyer should be able to observe the pressure points governing the process and appreciate their relative impacts on the final outcome. If the buyer has a (rational or irrational) basis for differing with the actuary on any of the disclosed items, then the buyer would be free to make alterations to those judgments and accept the consequences.

One example might illustrate what I have in mind. Suppose an actuary, enroute to a conclusion, needed to select a trend line to be fitted to historical data. The disclosure and sensitivity analyses might include the following:

- Number of points actually used in deriving the line of best fit. Outline the rationale for this choice and demonstrate the effect of selecting fewer or more points on the final answer.
- Historical points omitted from the historical data. Outline the rationale for this choice and demonstrate the effect of restoring those points on the final answer.
- Seasonal adjustments. Outline the rationale for any seasonal adjustments and demonstrate the effect of "no seasonal adjustment" on the final answer.
- Tempering the projection of the line of best fit. Outline the rationale for tempering and demonstrate the effects of "no tempering," or other magnitudes of tempering, on the final answer.

And there are several others: type of line used (straight, exponential), length of period used (month, quarter, year), type of observations used (12 months moving averages, discrete time measurements), and so on.

One of the key ingredients of a profession is the existence of observed standards of practice. The actuarial profession needs rigorous standards of practice in order to accelerate the effort to obtain legal recognition. Also one can hear the footsteps of actuarial malpractice (to a few actuaries, it has already arrived). Adopting a universal standard of full disclosure accompanied by relevant sensitivity analyses can only strengthen the profession by separating fact from opinion in presenting the actuarial work product.

I should also note that all actuaries are included within the scope of my comments: actuaries who sell their skills to one client (employees) as well as actuaries who sell their skills to many clients (consultants).

Whenever I think of standards of practice, my mental reflexes tend to deal in terms of what is permitted and what is prohibited. And that reminds me of a wonderful quote from former FCC chairman, Newton Minow, on the results of his study of the legal systems of European countries:

> "In Germany, under the law everything is prohibited except that which is permitted. In France, under the law everything is permitted except that which is prohibited. In the Soviet Union, everything is prohibited, including that which is permitted. And in Italy, under the law everything is permitted, especially that which is prohibited."

It is important to be aware not only of the need for full disclosure, but of the implication of its absence.

The buyer of the actuarial product, a priori, does not know how to separate actuarial art from science, actuarial fact from opinion, and mathematical wizardry from pedantic applications of formulas. All of these ingredients may be mixed well, carefully packaged, and eloquently presented; without full disclosure, the buyer of the actuarial product is completely at the mercy of the actuary. This is an unnecessary jeopardy for both the actuary and the buyer of the work product. The actuary, if operating with professional integrity, has absolutely nothing to fear from exposing the assumptions and judgments that went into the final work product. The worst possible outcome is that the buyer can exercise his or her own judgment (instead of the actuary's) if he or she so desires. The point here is that failure to disclose the assumptions and judgments along with appropriate sensitivity analyses renders the actuarial work product incomplete.

ACTUARIAL CHALLENGES

Over the years a number of questions have lingered in my mind that, I believe, the actuary is particularly well suited to answer. Let me share some of these problems with you.

Functions of surplus.

What are the functions that surplus serves? Some functions are readily apparent: to provide a cushion for absorbing adverse investment fluctuations, to provide a cushion for absorbing the collective risk assumed by the insurer, to provide a cushion for absorbing adverse reserve fluctuations. And I can think of several others. The challenge is to define an exhaustive set of the functions of surplus.

How much surplus?

Having defined the functions of surplus, the next question is how much surplus is needed to support each function and how much surplus is needed to support all the functions combined? Interestingly enough, the only rule that has emerged over the years that is remotely related to this question has been the Kenney rule. Can we do better? I believe we can. The challenge is to devise a general model that uses a number of insurer measures as input and yields a range for required surplus as output.

Risk classification index.

We all know the extremes of risk classification for rating purposes. On one extreme we have the individual risk rate and on the other we have the average rate for the total subject risk population. Classification plans attempt to group risks somewhere in between. Is there an index that describes exactly where a risk classification plan falls between the two extremes? I believe there is. The challenge is to define such an index—to be used for management purposes as well as an aid to enlightened regulation of business.

Confidence intervals for loss reserves.

Loss reserves are stated as point estimates. This fact, we all recognize, is only part of the story. Every reserve estimate has a corresponding confidence interval, albeit one that is usually not known or, if known, not stated. The challenge is to define a general model for determining loss reserves confidence intervals so that a reserve is always stated as a point estimate together with a corresponding confidence interval.

Present value ratemaking.

With very few exceptions, ratemaking formulas generally use calendar/accident year loss ratio methods. The propriety and elegance of present value ratemaking can hardly be overstated. The challenge is to come up with a generalized model for making rates using present value methods that recognizes the peculiar characteristics of casualty insurance.

Sampling.

Much of our industry manages its affairs by using 100% samples. With few exceptions, sampling has not found its way onto center stage of actuarial science. The challenge here is to develop small sample models to derive answers to the two classic actuarial problems: ratemaking and reserving.

Solvency tests.

The current NAIC tests are, at best, pragmatic tests, lacking a sound theoretical basis. The challenge is to develop a set of actuarial tests of solvency that have a sound theoretical basis, not just empirical observation. There can be no overestimation of the value such a set of tests would have on a number of different fronts.

Inflation sensitive exposure bases.

With inflation already very much a part of the world economic fabric, the need for inflation sensitive exposure bases grows more acute on a daily basis. The problem has been substantially solved for property exposures. But for liability exposures the problem lingers. The result is a constantly recurring gap between the true rate and the rate actually in effect, causing insurers to engage in an endless game of catch-up. The challenge is to find a set of inflation sensitive exposure bases for liability exposures.

I invite each of you to reflect on these questions, to pick out a small part of anyone of them and adopt it. Let your mind engulf it, understand it, feel it, stalk it, own it, and then, ultimately, subdue it. And when you have thus solved it, write a paper and share it with us. Writing a paper is precisely the single most powerful method we have to expand the horizons of actuarial science. And I extend this invitation to each of you, whether an Associate, a new Fellow, or a thirty-year Fellow.

I am sure you have observed that all of the questions I have posed lie within a traditional insurance framework. The applicability of the actuarial method to areas outside the traditional insurance framework has been expanding. I will cite two examples. One is the rapidly growing involvement of actuaries in the field of risk management. Another is that one of our members has made it his life's work to extend actuarial applications to any question involving a contingency and a consequent transfer of money. My judgment is that these applications will continue to grow and will occupy a larger and larger proportion of our membership. My invitation to you in this regard is be alert to such opportunities to expand our horizons. Eventually we will likely have to amend and broaden the statement of objectives of the Casualty Actuarial Society.

PROFESSIONALISM

In *The Fountainhead*, Ayn Rand writes: "Throughout the centuries there were men who took first steps down new roads armed with nothing but their own vision."

Our founders had a vision in 1914. They saw a need. And they formed (what later became) the Casualty Actuarial Society for the purpose of meeting that need.

Over the years the CAS has grown in both numbers and stature. It continues to serve a useful purpose; and incidentally, that purpose has grown over time (as evidenced by the expansion in 1961 of the statement of objectives of the CAS). And, as I suggested earlier, our purpose will grow even more in the future.

The CAS today enjoys a very fine reputation. We are known as a learned association of professionals with rigorous entry requirements. The value of the education our members receive enroute to Fellowship is continually demonstrated by the incredible variety of functions actuaries are called upon to perform—as employees and as consultants.

I submit to you that the life force of our reputation derives from two sources, one from within and one from without:

- The contributions from within derive from those who write papers, develop the *Syllabus*, construct and grade examinations, put our programs together, serve on our program panels and workshops, publish our periodicals, question and challenge conventional wisdom, and otherwise conduct the business of the CAS.
- The contributions from without are accomplished by making sure that all the work we do is of uniformly high quality and by making sure that we conduct our business lives with impeccable professional integrity.

History gives us many accounts of great organizations that faltered because of complacency. They looked too much to past success and too little to current opportunities. The CAS, in order to continue to thrive and meet its objectives, requires each of us to contribute to each of the sources of its life force.

Every talent we have, every ability we possess, every skill we have acquired is a gift. It is a gift entrusted to us to put to good use. Also, it is well to remember, whether we are consciously aware of it or not, that in the course of achieving every success we have experienced, someone helped us.

Today you belong to a healthy, vibrant, and forward looking organization. In a very direct way, it has helped you. I'd like to ask you to ask yourselves the following question:

> "Is the Casualty Actuarial Society better and stronger for having me as a member?"

If we are the fulfillment of the vision of our founders and if we are to continue to keep the torch lit, your answer to this question must be a resounding *yes*. Your mission is to make sure the answer to this question will always be yes.

THE VALUATION OF AN INSURANCE COMPANY FOR AN ACQUISITION INVOLVING A SECTION 338 TAX ELECTION

JAMES A. HALL, III, ORIN M. LINDEN, STEPHEN GERARD, MICHAEL HEITZ

Abstract

One method of treating the acquisition of a stock company is a Section 338 election. This paper discusses such an election in the acquisition of a stock insurance company. The tax aspects are explored and the role of the casualty actuary in such an election is discussed.

INTRODUCTION

This paper discusses one possible tax treatment of an insurance company acquisition and the role of the casualty actuary in this process.

Traditionally, the casualty actuary has played a significant role in the accounting practices of the insurance industry through his work in analyzing loss reserves. However, actuarial input to other accounting areas has been more theoretical than applied. While casualty actuaries have written important papers and made valuable contributions to statutory accounting and annual statement accounting, most actuaries do not work closely with accountants.

The average company actuary is probably reasonably familiar with statutory accounting principles and has some idea of adjustments that must be made for purposes of generally accepted accounting principles (GAAP). Purchase accounting and tax accounting are probably much more foreign to most casualty actuaries working for insurance companies.

The purchase of an insurance company is a challenging opportunity for the casualty actuary, the insurance tax and valuation specialist, and the insurance accountant to cooperate in a multi-disciplinary team. The Section 338 election described herein is only one approach to the acquisition of an insurance company.

TAX ISSUES

Overview of Section 338

A significant provision of the Tax Equity and Fiscal Responsibility Act of 1982 ("TEFRA") concerns acquisitions of corporations and continues to afford an opportunity to partially "finance" such acquisitions through tax amortization of certain nonstatement assets. Under prior law, certain acquisitions of target corporations generally took the form of either an asset acquisition or a stock purchase. If the buyer acquired the assets from the target corporation, the buyer's basis in each purchased asset was that asset's share of the purchase price. The selling corporation would not recognize a gain on asset appreciation under Section 337 if it liquidated within a time frame provided by statute. The buyer who acquired stock instead of assets could allocate the purchase price to the underlying assets and liquidate the acquired corporation. If, instead, the buyer did not liquidate the acquired company, the target's tax attributes (e.g. net operating loss carryovers) continued and its assets retained their historical basis.

Congress, in what it perceived to be a correction of several areas of abuse as well as a simplification of existing law, repealed the stock purchase-liquidations of Section 334(b)(2) and added new Section 338. Among the abuses which Congress corrected was the buyer's ability to "pick and choose" in determining which assets received a favorable "step-up in basis" and which assets avoided a recapture tax. (A "step-up in basis" occurs when the buyer is allowed to increase the tax basis of the target company's assets (generally cost) to an amount equal to its cost (the current fair market value) of purchasing the target's stock). Further, under prior law a buyer was permitted to continue the target corporation's tax attributes for a period up to five years after the initial stock purchase while also treating the transaction as though assets had been purchased. This extended "survival" period led both to significant opportunities to combine the target corporation's tax attributes with those of the purchasing corporation as well as major complexities in determining the basis to be assigned to the target's assets on liquidation. Finally, if consolidated returns were filed by the acquiring corporation, the recapture tax liability could be deferred and in certain situations avoided.

In general, new Section 338 provides that with respect to certain stock acquisitions, the purchasing corporation may elect to treat the target corporation as having sold all of its assets on the stock purchase date and as having purchased those assets, acting as a new corporation, on the next day. This "sale" is generally considered tax free to the target corporation to the extent it would have been a sale under Section 337. Finally, the tax attributes of the target corporation, including net operating losses, are not carried over to the new successor corporation.

Under the new rules, it is no longer necessary to form a new company and liquidate the target corporation to get a stepped-up basis; the buyer merely needs to elect to have the stock purchase treated as a direct asset purchase. Thus, unlike certain instances under the prior law where a legal liquidation of the acquired corporation was required, outstanding contracts need not be amended, and permission from state insurance authorities to treat the stock purchase as an asset acquisition generally should not be needed. Under Section 338, the election applies only for tax purposes.

The election generally applies to "qualified stock purchases" (I.R.C. Section 338(d) (3)), of the target corporation's stock occurring after September 1, 1982. The term "qualified stock purchase" contains the same requirements previously provided in the prior law, i.e., a purchase within a 12-month period of 80% or more of the voting power and 80% or more of the nonvoting stock (except nonvoting, nonparticipating preferred stock) of the target corporation. For qualified stock purchases made after August 31, 1982, the purchasing corporation's affirmative election of Section 338 must be filed by the later of (1) the 15th day of the 9th month after the month in which the acquisition date occurs or (2) December 31, 1985. This means that decisions generally must be made more quickly now than in the past where, under a prior law, a buyer could wait up to two years before deciding whether a stepped-up basis was desirable. Once made, the election is irrevocable.

Tax Implications Resulting from the Election

A Section 338 election is particularly beneficial where the purchase price exceeds the book value of the target. Under this election, where the purchase price of the stock includes a "premium" over book value of the underlying assets, the basis in the acquired stock, including unsecured liabilities assumed, may be apportioned to all the acquired assets based on their relative net fair market values. As we discuss below, the actuary can be instrumental in helping to determine the fair market value of certain insurance-related intangible assets and their amortizable lives. In this process, appreciated property obtains a stepped-up basis, thereby providing an opportunity to obtain higher cost recovery and amortization deductions. The step-up also serves to reduce potential taxable gains on future dispositions of such assets. It should be noted that assets that have depreciated in value will be written down thereby producing the opposite results. Further, transactions between the purchasing corporation and the target or target affiliate for a period of one year both before and after the acquisition date must be treated as if included as part of the stock acquisition unless the sale by the target corporation is in the ordinary course of its trade or business, or one of several other limited exceptions are met.

The assets of the target corporation will be treated as sold (and purchased) for an amount equal to the grossed-up basis of the acquiring corporation in the stock of the target corporation on the acquisition date. "Grossed-up basis" is a tax concept and was devised for situations where less than 100 percent of the target's stock is purchased. If the purchasing corporation owns all of the target corporation's outstanding stock, the grossed-up basis of the target corporation's stock is its cost basis. If the purchasing corporation acquires less than 100 percent of the target corporation's stock, an adjustment must be made to the basis of assets acquired to reflect the continued interest of minority shareholders. The formula used to determine the grossed-up basis provides that this amount is to be adjusted under the regulations to be issued for liabilities of the target and other relevant items such as recapture tax liability.

Significant tax benefits are achieved where identifiable amortizable intangible assets are acquired. An intangible asset can be generally defined as a property or property right which does not have physical existence, but which can be expected to produce income in future years.

Intangible assets can be more narrowly categorized as identifiable and unidentifiable (e.g. see Revenue Ruling No. 74-456). To be considered an identifiable intangible asset for federal income tax purposes, the intangible asset should be "identifiable" with specific rights, properties, relationships, contracts, or other definable source of income potential. Common examples of identifiable intangible assets are patents, trademarks, franchises, equity in favorable contracts, and certain types of customer relationships.

Unidentifiable intangible assets, as the name implies, are valuable properties whose source of income potential cannot be pinpointed to a specific source. These assets are often referred to for tax purposes as goodwill (defined more narrowly as the propensity of satisfied customers to return to the old place of business resulting in an "excess earnings" potential) and going concern value (defined as a "turn key" premium for an established enterprise which can be expected to conduct a continuous and reasonably profitable business despite a change of ownership).

To justify that an intangible asset is amortizable, a taxpayer, moreover, must demonstrate that the asset has a life of limited duration and that this life can be estimated (see Rev. Rule 74-456). Examples of amortizable intangible assets may include customers and service lists, subscription lists, leaseholds, databases, future profits in existing company contracts, and the sales force.

With respect to property and liability insurance companies, amortizable intangible assets which the target corporation may possess include the following:

> *Value of future profits in the loss reserves.* This asset may exist because, for statutory purposes, companies are required to establish loss reserves at undiscounted values. It appears that the acquirer could take into consideration anticipated future investment income as an amortizable intangible asset.
>
> *Value of future profits in the unearned premium reserve.* This asset represents the potential future investment income and underwriting profits on the unearned portions of policies already written.
>
> *Future profits on renewal business.* Often referred to as either Book of Business, Expirations, or Dailies, this asset is the present value of the future profit stream associated with renewals of the current book. If the target company can accurately project its renewal business, underwriting profits on renewals, and future investment income associated with related reserves, then this may represent an amortizable intangible asset.

In the case of a life insurance company, identifiable amortizable intangible assets may include the following:

> *Future profits on business in force.* This represents the present value of future profits on current business. The determination of the value of this asset requires actuarial analysis of such key items as investment income, assumed rates of interest, lapse rate and mortality experience.
>
> *Policy loans.* Life insurance companies are required to make loans to policyholders at rates well below the current market rate. No rulings or decisions have dealt with the values to be assigned to this category of asset, but assigning a face value with an offset of an equivalent amount of reserves appears to be in accord with the statute rather than discounting the value of the loan with a possible increase in income when the loans are repaid.

In the case of both life and property and liability companies, the *agency force* may be an identifiable amortizable intangible asset. The value of the agency force is akin to the present value of future profits produced from new policyholder premiums. Profit margins from future sales may be based on actuarial assumptions similar to those made in valuing current business in force. However, an additional assumption must be made on the volume and product mix of future sales. Whether this asset is susceptible of being separately valued and amortized for tax purposes can be addressed only on a case-by-case basis.

Unresolved issues. In Rev. Proc. 83-57, the Service announced that it is extensively studying the consequences of an acquisition of a life insurer followed by a Section 338 election. Accordingly, the Service assumed (among other issues) a no ruling position regarding:

—whether life insurance reserves may be treated as unsecured liabilities for purposes of determining allocable basis, and
—whether a portion of the purchase price is properly allocable to insurance-in-force.

When the results of the Service study will be made known cannot be predicted at this time (November, 1985). However, suffice it to say that resolution of these issues may not occur in the foreseeable future.

Moreover, numerous issues as to the manner of making the election, allocation of the purchase price among others, remain unresolved. This is evidenced by the language of Section 338 that contains many references where Congress specifically authorizes the Treasury to promulgate regulations to amplify or implement this provision.

Evaluating the Trade-Offs. While an election pursuant to Section 338 may produce fairly significant tax advantages through the amortization of intangibles, the election is not without tax and economic costs. The extent to which net operating loss carryovers may be terminated should be included. Depreciation and investment tax credits claimed by a target corporation prior to the acquisition may be recaptured as of the date of the acquisition. Depreciation recapture is a limitation of the amount of long term capital gain arising on the sale of certain depreciable assets. Gain on the sale of such property is treated under recapture as ordinary income to the extent of depreciation taken as a deduction in prior years.

The assets with respect to which such recapture would arise would be valued at their fair market values as of the date of acquisition and such value would be used prospectively over several years in calculating depreciation and investment tax credit.

Recapture income and investment tax credit recapture from a Section 338 election cannot, except for limited exceptions, be included in a consolidated return of either the seller or the purchasing corporation. If the target corporation was not a member of an affiliated group, the recapture income is included in a short period return (i.e. a return for less than one year) which would also include the target corporation's income up through the date of acquisition. If the target corporation is a member of an affiliated group, a separate return, which reflects the recapture tax liabilities, is required.

Tax Planning. It is clear that opportunities exist to ascribe values to the amortizable intangibles not found on the statutory statement. Proper tax planning dictates that early consideration of these issues be incorporated into the negotiations. This planning must be done by a team of qualified tax professionals. As we shall see, there is ample opportunity for a casualty actuary to participate on this team.

THE ACTUARY'S ROLE

In order to comply with a Section 338 election one must value all assets as of the acquisition date. At first this does not seem to be an actuarial problem since most assets may be valued by auditors and appraisers. There are various methods based on cost, depreciation and market value that can be used. Actuaries, on the other hand, are typically concerned with future events. How many losses will occur next year? How will the reserves run off? But, as we'll see shortly, the same issues and techniques that an actuary deals with in resolving "standard actuarial problems" must be dealt with in valuing certain intangible assets not found on an insurance company's annual statement.

For purposes of this discussion we will consider two broad categories of assets. The first category consists of the assets usually found on the asset page of any company's annual statement. These include stocks, bonds, cash on hand, computers, accounts receivable, etc. They are reflected in policyholders' surplus. In addition, the annual statement discloses elsewhere certain non-admitted assets excluded from surplus. Considered together, we will refer to these assets as statement assets. Most of these assets the company or its auditor can value.

Since the annual statement is the basis for calculating taxable income, asset valuation for a Section 338 election logically should begin here. Of course, there are complications since the annual statement was not designed for this purpose. Statutory accounting requires certain types of valuations. In valuing the statement assets for tax purposes, adjustments must be made. Bonds at amortized values should reflect market values as of the acquisition date. (A word to the wise: The market value shown in Schedule D is often not a true market value.) Non-admitted assets also must be added back to the balance sheet at market values. Stocks and real estate must be set at market value as of the valuation date. All these adjustments are typically performed by auditors and appraisers.

The other category of assets is not displayed on the annual statement. These are the intangible assets where actuarial issues such as the run-off of past, present, or future business are a critical ingredient in the valuation process and have significant ramifications in regard to future tax treatment. Identifiable and unidentifiable intangible assets typically exist in any insurance company for three reasons:

—Companies sell insurance.
—Statutory accounting requires that insurance companies keep large sums of money, or liquid assets, available to pay claims.
—Funds held can earn money.

Consideration of intangible assets is critical for the buyer and seller in negotiating a purchase price.

For a Section 338 election, all assets whether tangible or intangible, amortizable or non-amortizable, should be valued. Furthermore, the IRS has acknowledged in numerous private rulings that certain intangible assets may be amortizable where the taxpayer meets his burden of proof of three critical requirements:

1. The asset must be severable from unidentifiable goodwill; that is, the specific source(s) of future income potential must be identifiable and capable of being separately valued;
2. The asset must be a "wasting asset"; that is, the economic viability of the identified asset must be of limited duration, such that its value declines over time; and
3. The remaining period of economic viability must be capable of being estimated within reasonable business accuracy.

It is here that an actuary can play a major role. Working closely with the tax specialist, actuarial expertise can be used in several ways. He can provide formulas to evaluate. He can review historical data to project runoff of current or future premiums and losses. He can analyze historical cash flows and project future contingencies.

EXAMPLES OF INTANGIBLE ASSETS OF A TYPICAL INSURANCE COMPANY

As mentioned, there are at least three potentially amortizable intangible assets common to every insurance company. The first of these deals with the loss reserves. We refer to it as "Future Profits in the Loss Reserves." Most loss reserves are carried at the full amount needed to settle all losses, reported or not, that have occurred to date. To the extent that the actuary can assist his client sustain the burden of proof that the reserves are identifiable with a specific group of insurance exposures; that the life of the reserves is of limited duration; that the runoff (or consumption) rate of the reserves can be estimated with reasonable accuracy and that appropriate projections of anticipated investment income can be allocated to the reserves in question, "Future Profits in the Loss Reserves" may be valued and amortized for tax purposes.

The second of these deals with the "Future Profits in the Unearned Premium Reserve." The unearned premium reserves will earn interest while they are being held by the insurance company. In addition, as the unearned premium reserve expires, the losses and expenses incurred may be less than the premium earned and produce an underwriting profit. The ability to earn investment income and underwriting profit form a valuable identifiable intangible asset. Again, to the extent that the actuary can assist with valuing this asset, and demonstrating that it is a wasting asset, where remaining life can be reasonably estimated, it may be amortized for tax purposes.

The third identifiable and potentially amortizable intangible asset is the "Future Profits on Renewals." A company can reasonably expect to renew a certain portion of its current book each year. These renewals will generate reserves and these reserves will possess the intangible assets described above. To see that this asset exists one need only note that often an insurance company will be bought solely to acquire its book of business.

The techniques applied in valuing these intangibles is beyond the scope of this paper. The literature abounds with such. Also, most companies need tailored methods. A clear understanding is necessary for evaluating the above assets as well as the unidentifiable intangible assets. Rather than try to give recipes we'll discuss concepts underlying these valuations.

The actuary is very familiar with the main tool needed to evaluate many intangible assets—projection of cash flow. Whether we are dealing with losses that have already occurred, or losses that will occur in the future, premium, or expenses, one must project not only the amount that a company will receive or pay but also the rate of payment. The next step is to determine an appropriate investment rate. This investment rate need not be based on a company's investment portfolio. Since stocks and bonds reflect market value for a Section 338 election, the historical interest rate is already removed. U.S. Government bonds give a reliable indicator of the available interest rate. A mix of U.S. Government securities, with appropriate durations, can be used for an average return. It is also possible to incorporate a mix of tax-free municipals. Other sources are available.

A different discount rate may be needed for projected profits. If projected investment income is to be used to make future payments, then the discount rate should equal the investment rate. However, profits available for stockholders should be discounted with an appropriate rate for the company.

Often the discount rate applied to future profits is greater than the projected investment yield, reflecting risk considerations. One choice of risk-related discount rate is given by the Capital Asset Pricing Model (CAPM). It's easy to use, it's objective, it can be tailored to a specific company, it's consistent with the government bond rate, and it is often used for tax valuation purposes. However, this choice is not without controversy. Other choices also exist.

SOURCES OF DATA

To project the above cash flows a large amount of data is needed. Federal income taxes are based on the annual statement, so this is the place to start. Schedules O and P offer information on the loss payment rates. The five-year history is a good source of calendar year premiums, earnings, and expenses. The other exhibits are also useful. The amount of additional detailed data necessary is an important consideration for the actuary and the tax specialist.

Knowledge of company operations is critical. As with any other actuarial area, changes in company operations may inhibit the usefulness of historical data. Here interviews are important. Senior company management will be helpful in pointing out problem areas. Annual reports, 10-K's, and even special data requests may also be helpful. The information developed may suggest adjustments be made to historical data to reflect new circumstances prior to making any projections. However, no such adjustment should be made unless evidence (e.g. new reinsurance contracts, etc.) is available to support the adjustment.

Valuing the future projects on renewals requires special care. The main problem here is to estimate the portion of business that will be renewed year to year. For this purpose, most companies rely on runs that start with a fixed block of business. The portions of this block renewed in succeeding years should be shown on a "dollar" and a "number of policies" basis. From these runs, future renewals of current business are predictable. There is, however, a pitfall. Due to intense competition in recent years, may companies have been canceling policies and rewriting them on different terms. Often a company will show this transaction as a cancellation, and a new writing. However, this is in effect a renewal and it is important to count it as such.

ADDITIONAL ANALYSIS

Before starting to evaluate intangibles some preliminary considerations should be addressed. It is important that current loss reserves be adequate and not redundant, so a reserve study may be necessary. The actuary must consider the effects of any restatements in choosing his parameters. For example, a severe reserve deficiency will not only require an increase in reserves. It will also be necessary to adjust historical loss ratios and earnings for this deficiency. If this isn't done projected loss ratios will probably be too low, and projected earnings will probably be too high. A redundancy has the opposite effect.

Once the valuation of the intangible asset is complete, an amortization schedule of the asset should be prepared. Usually amortization formulas, from compound interest theory, are sufficient. However, due to certain statutory accounting principles, straight amortization formulas may distort the depreciation schedule. For example, since statutory accounting requires an immediate writeoff of deferred expenses, a projected profit stream may start off negative in the early years and turn positive later on. In this case, use of a present value type of amortization would lead to results that aren't useful for balance sheet purposes. In these cases it might be better to choose a straight line depreciation schedule. Another alternative is to combine two or more profit streams from different intangible assets, so that the net profit stream is positive. Ongoing consultations between the actuary and the tax specialist may be necessary to select the most appropriate method.

Since projections of the future are used, and since insurance is a risky business, some sensitivity analysis may be required. Often an actuary estimates probabilities of different parameters in order to arrive at expected values. For documentation purposes it is probably better to work with an expected scenario and choose a discount rate that suitably adjusts for the risk involved.

OTHER ACTUARIAL CONSIDERATIONS

There are other factors to consider in a Section 338 election. Other intangible assets may exist. A company-owned agency force, relationships with independent agents, or the right to participate in a pool might generate future profits. In addition, some assets are hard to classify as either tangible or intangible but still should be valued. Computer software is a good example. One must also estimate the effect of income tax on projected future profits. Subsidiaries offer another complication. A similar analysis of each subsidiary may be needed.

FINAL ACTUARIAL REPORT

When all the analysis is completed a final report is a must. All computations should be carefully documented along with the selected methods and parameters. Documentation of data sources and parameter estimation must be included. If historical data has been adjusted this must also be cited. Schedules of amortization of intangible assets should also be included.

This documentation should be kept on file should the need for it arise. Again, consultation with the tax specialist is mandatory.

One cannot overemphasize the importance of this final actuarial report. Accountants and valuation specialists must have an explanation of all the factors contributing to the analysis in order to proceed with the Section 338 election. Equally important, in the event of an IRS audit, documentation of results is crucial to sustaining the taxpayer's burden of proof. Remember, what's obvious today will probably be incomprehensible three years from now. It is better to overdocument today than to not be able to reconstruct your thinking at an IRS audit.

CONCLUSION

In summary, the actuary can play a vital role in helping to quantify and support significant tax benefits in connection with the purchase of an insurance business. Working closely with tax professionals, he can use his traditional actuarial tools and professional expertise to help resolve complicated tax and valuation issues.

AN INTRODUCTION TO UNDERWRITING PROFIT MODELS

HOWARD C. MAHLER

Abstract

This paper will provide an introduction to the subject of underwriting profit models in order to provide actuaries with a basic framework for further study. This paper starts with the premise that the subject of underwriting profit provisions is an area in which actuaries can be of assistance in advancing knowledge and developing methods. While this paper will concentrate on the theoretical aspects, this subject has many potential practical applications.

The basic structure of the paper is to start off with an extremely simple model, and then add additional considerations. For clarity, this paper has focused on one basic method of calculating a provision for underwriting profits.

There are three basic ingredients used in these models. First, via a "cashflow" analysis, one estimates the length of time an insurer will have premium dollars on hand, prior to paying losses and expenses. Second, one estimates how much investment income an insurer will earn on this cashflow and the necessary equity backing up the policies. Finally, one sets the expected return on equity equal to a target return on equity. One can then solve this equation for the underwriting profit provision.

INTRODUCTION

The question of what provision for underwriting profits (or losses) to use has become a topic of increasing discussion over the last decade. Rather than use traditional numbers found in the actuarial literature, such as 5%, there have been attempts made to calculate profit provisions. These calculations have involved making certain assumptions and algebraic derivations. Thus they are commonly called underwriting profit "models." This paper will provide an introduction to the subject, in order to give actuaries a basic framework for further study.

In spite of the use in the title of the term underwriting profit, this paper concentrates on the total return on equity concept. In each particular case, one can calculate an underwriting margin (positive or negative) that can be expected to produce the desired or required total return on equity. From this point of view there is no fundamental difference between a positive and negative underwriting margin. Equivalently, there is no fundamental difference between a target combined ratio which is greater than 100% and one that is less than 100%. Rather, they are different points along the same continuum.

The basic structure of the paper will be to start off with an extremely simple model, and then add additional considerations. (Those readers already familiar with the subject may want to go directly to the third model or even the summary of that model.) Care has been taken to list all the assumptions made in each model. If, in a particular application, one or more of the assumptions are not reasonable, one can then make the appropriate change in the list of assumptions and derive modified equations. As with most actuarial calculations, the results produced by the models are dependent on the assumptions made and input values used. In actual applications, choosing the appropriate input values is usually a difficult task. (Examples of this are given in the numerical examples using model three and in Appendix II.)

For the reader's convenience, Appendix I contains the definitions of the various symbols used in this paper.

DEFINITION OF AN UNDERWRITING PROFIT PROVISION

Let P^* be the premiums loaded for profits. (The asterisk indicates that P has been loaded for profit. The author has found that this use of the asterisk to denote quantities that are loaded for profit avoids much confusion when working with underwriting profit models.)

In general, ignoring uncollected premium, the underwriting profit provision, u, is defined so that:

$$P^* = P^*u + \text{losses} + \text{expenses}.$$

This is the fundamental definition of an underwriting profit provision which will be used throughout this paper.

Let L be the losses paid by the insurer.

The expenses are made up of T^*, those expenses which are proportional to the premium, and E, the remaining expenses. We define $T^*/P^* = t$. (The use of the letter T comes from premium taxes, which vary with premiums.)

Thus:

$$P^* \, (1-u) = L + E + T^* = L + E + tP^*$$

$$u = 1 - (t + (L+E)/P^*)$$

In this paper we will usually solve for P^*, the premium loaded for profits, and then use the above equation to get the underwriting provision u.

THE FIRST MODEL

We make the following assumptions:

(0) An insurer writes a set of similar policies. Each policy is expected to be in effect for one year. (This assumption is labeled zero, since it is so basic that it is often left unstated.)

(1) The insurer receives premiums P^*. All premiums are received exactly at policy inception.

(2) The insurer pays losses L. All losses are paid exactly one year after policy inception.

(3) The insurer earns income on its investments at a rate r.

(4) The insurer wishes to break even. (We ignore any investment income the insurer may earn on its equity.)

For this very simple model, we have ignored expenses, equity, income taxes, and all the other complications that exist in the real world. Also, we have assumed that the insurer merely wishes to break even on average. (Under certain circumstances this might be true of a non-profit organization, such as Blue Cross or a Medical Malpractice Joint Underwriting Association.)

We will calculate the premium, P^*, the insurer should charge, so that it can be expected to break even. (Elsewhere in the paper, the insurer will desire a return on its equity.) Assumptions (1) and (2) imply that the insurer can invest a sum P^* for one year. During that time, rP^* investment income will be earned, due to assumption number (3). So the insurer will have $P^* + rP^*$ available at the end of the policy year. It will have to pay out L at that time, due to assumption (2). Assumption (4) is that the insurer wishes to break even. Therefore:

$0 = P^* + P^*r - L$

$P^* = L/(1+r)$

In our special case, $t = E = 0$. Thus:

$u = 1 - L/P^* = 1 - (1+r) = -r$

So even this very simple example demonstrates a basic feature. One can have a negative provision for underwriting "profit". This will occur when the target return is relatively small and/or when you can earn a large amount of investment income (either due to a high rate of return r, or due to a long period of time between when the premiums are received and losses are paid out.) In that case, you can achieve the desired total return, even though you have an underwriting loss. This basic feature has been noted by others, among them Ferrari [1].

THE SECOND MODEL

Until now, we have dealt with a very simple timing of transactions. The value of receiving one dollar depends on when one expects to receive it. See, for example, Kellison [2]. The further in the future one receives it, the less the dollar is worth to you now. In general, we wish to take the present value of the income received. (In taking present values in this paper, we will for convenience always discount to the end of the policy year. Why this is a convenient choice is explained below. In present value equations using a single discount rate, the choice of the point in time to which one discounts should not affect the answer, provided that all terms in the equation are discounted to the same point in time.)

If a dollar is to be received n years hence, and we discount to the end of the first year, using an interest rate i, then the present value is $(1+i)^{(1-n)}$.

We modify the assumptions of the first model, (1) and (2), in order to allow a general timing of the payment of premiums and losses.

(1′) The insurer receives premiums P^*. (The expected pattern of the timing of payments is known or can be estimated.)

(2′) The insurer pays losses L. (The expected pattern of the timing of payments is known or can be estimated.)

We modify or add the following assumptions:

(4)The insurer desires a target rate of return of R on the funds it supplies.

(5a) The insurer supplies funds, S, of its own. These funds exist throughout the entire policy year in a constant amount.

(5b) The required equity S is proportional to the premium P^*, with proportionality constant $s = P^*/S$.

We have assumed the insurer supplies S at the beginning of the policy year, and desires $(1 + R)S$ back at the end of the policy year.

For the purposes of this paper, these insurer-supplied funds are stockholder-supplied equity. However, the reader may find it helpful to think of them as surplus. These funds in some sense back up a group of policies so that even in the case of unexpected occurrences the insurer will be able to meet its obligation of paying claims.

We have yet to include expenses in the model. Examples of categories of expenses are loss adjustment expense, commissions, other acquisition expenses, general expenses, and premium taxes. (Investment expenses are presumably taken into account by subtracting them from the investment rate of return.) Generally, these expenses can be divided into three types: those that are fixed, those that vary with premium, and those that vary with losses. In this paper, the method by which the specific assignments were made will not be explored. One example of such an assignment is given in Snader [3].

In this paper we will make a slightly different division. First, we include in L those expenses that are assumed to have the same timing as the loss payments. (Alternatively they may have been included in the data from which we made our estimate of the timing of the loss payments.) Next, we separate out those expenses that vary with premiums, and call them T^*. (This almost always includes premium taxes, usually includes commissions, and sometimes includes all or part of other acquisition or general expenses.) Whatever expenses are left are called E. (See Appendix II for an example of such assignments. Although the calculations are not shown there, the expense-to-loss ratios depend on a determination of the relative amount of each type of expense. This depends in turn on a determination of which expenses are fixed, and which vary with losses.)

In order to include expenses, we make a minor revision to one of the prior five assumptions; otherwise, we retain them.

(2″) The insurer pays L, losses including those expenses whose timing is the same as the losses. (The expected pattern of the timing of such payments is known or can be estimated.)

We also add two additional assumptions.

(6) The insurer pays T^*, expenses that vary with premiums. (The expected pattern of the timing of such payments is known or can be estimated.)

(7) The insurer pays expenses E, other than those included in L and T^*. (The expected pattern of the timing of such payments is known or can be estimated.)

The insurer earns income from two sources. First, it earns rS investment income on the equity. Second, it earns income on the cashflow (premiums in, losses and expenses out.) The present value of the total income earned on the cashflow is $P^{*\prime} - (L'+E'+T^{*\prime})$. (The primes denote discounting by the rate of return on investments r.) This is a special case of a more general result. The present value of the total income on a cashflow is the present value of the inflows minus the present values of the outflows.

For this second model we have

present value of return on equity =
(present value of income earned on equity)
+ (present value of income earned on cashflow).

Setting the target return on equity equal to the present value of the return on equity, we get

$$RS = rS + P^{*\prime} - (L'+E'+T^{*\prime})$$

Note that RS and rS are assumed to be received at the end of the policy year, and thus they are already equal to their present values, since we are discounting to the end of the policy year. (If one had discounted to some other point in time, then this convenient relationship would no longer hold. This is why this point in time was chosen for use in these models.)

Let $g = P^{*\prime}/P^* = P'/P$

$h = T^{*\prime}/T^* = T'/T$

It should be noted that g and h depend only on the timing of the premium flow and the premium tax flow, and not on their overall magnitudes. Thus they can each be computed from quantities assuming a profit loading of zero for convenience. This is why we introduce g, h and other similar ratios into the model equations.

Then we have

$$RS = rS + P^{*\prime} - (L' + E' + T^{*\prime})$$

$$P^* = (L' + E')/(r/s + g - th - R/s)$$

We use the definition of the underwriting provision, u, when there are expenses

$$u = 1 - t - (E+L)/P^* = 1 - t - (r/s + g - th - R/s)/((L'+E')/(L+E))$$

Two points are worth making. First, we have not carefully distinguished here between surplus and equity. R is meant to represent the target return on equity, i.e., stockholder-supplied funds. Statutory surplus as defined in the Annual Statement is numerically different from the concept of equity used here, which can be thought of as net worth in accordance with Generally Accepted Accounting Principles (GAAP). It is important that the target return on equity R and the concept of equity S match each other. Adjustments might be appropriate for certain applications of the model. Among areas where adjustments might be appropriate are the treatment of prepaid expenses and the equity in the unearned premium reserves, estimated federal income taxes and policyholder dividends. The details would depend on the source of and the exact meaning attached to R and S. Unfortunately, the details are beyond the scope of this paper. See, for example, Section 1 of Appendix I to the Report of the Advisory Committee to the NAIC Task Force on Profitability and Investment Income [4], Measurement of Profitability and Treatment of Investment Income in Property Liability Insurance, pp. 783–799 [5], and Report of the NAIC Investment Income Task Force, p. 43 [6].

Also, the reader should notice that we have not distinguished between the rate of return on investments earned on equity and that assigned to the cashflows. Such a distinction may be appropriate in certain circumstances. For example, you may allocate different types of investments, different maturities of investments, etc., to the equity. Also some of the equity may be in fixed assets which can not be invested. These and other refinements could be reflected in the model.

This model reduces to the previous model if we take the special case where $E=T^*=0$, $L'=L$, $P'=(1+r)P$, and either $R=r$ or $S=0$.

FEDERAL INCOME TAXES

For the moment, let us assume that the insurer will pay income taxes at a rate *FIT*, one rate for all types of income. Also for the moment, let us ignore the question of the timing of the payment of these income taxes. Then simplistically our former equations would be changed, by multiplying all the income terms by $(1-FIT)$. The target rate of return, R, is the desired rate of return after the insurer pays Federal Income Taxes.

$$RS = (1-FIT)(rS + P^{*\prime} - L' - E' - T^{*\prime})$$

$$P^* = (L' + E')(1-FIT)/((r/s + g - th)(1-FIT) - R/s)$$

There is no inherent reason to divide the income into different types. However, different types of income are treated differently by the federal income tax system, as is explained in Beckman [7]. Income generally is divided into two types, underwriting income (or loss) and investment income. We will assume that the former is taxed at a rate *FITU*. In the case of an underwriting loss rather than a profit, *FITU* should be the rate at which the income that is offset by the underwriting loss would have been taxed. (One can usually assume, for modelling purposes, that the insurer will offset that income which is taxed at the maximum rate first, before using any remainder to offset income taxed at a lower rate.)

We will assume that the investment income is taxed at a rate *FITI*. (In many implementations, *FITI* will be some sort of weighted average of the tax rate on the different types of investment income. In a later section, an example of such a calculation is given.)

MODEL THREE

We make the following assumptions in addition to those in model two.

(8a) Underwriting income is taxed at a rate *FITU*. Underwriting income equals premiums minus losses and expenses $= P^* - L - E - T^*$.

(9a) Investment income is taxed at a rate *FITI*. Investment income is defined as the total income minus underwriting income.

The following assumptions concerning the timing of the payment of taxes have been found useful.

(8b) Federal income taxes on underwriting are paid at the end of the quarter in which the underwriting profit or loss is incurred.[1] (Ignoring any development of incurred losses, this leads to four equal payments at times $\frac{1}{4}$, $\frac{2}{4}$, $\frac{3}{4}$ and 1 year after policy inception.)

(9b) Federal taxes on investment income on the cashflow are paid at the time losses and expenses are paid.

It is common to make the following assumption, when one has made assumption (5) concerning the equity.

(9c) The Federal income taxes on the investment income earned on the equity are paid at the end of the policy year.

$$e = \frac{\text{present value of federal income taxes on underwriting}}{\text{federal income taxes on underwriting}}$$

The ratio e is dependent only on the timing of the payment of the federal income taxes on underwriting.

Assumption (8b) leads[2] to

$$e = ((1+r)^{3/4} + (1+r)^{2/4} + (1+r)^{1/4} + (1+r)^{0})/4$$

Thus, e is approximately $(1+r)^{3/8}$.

Define a ratio d, similar to the previously defined e.
Let d = (present value of *FITI* on cashflows)/(*FITI* on cashflows).

The present value of the total return after taxes can be broken up into five pieces. We have, where PV stands for present value,

PV(total return after taxes) =
PV(investment income on equity)
− PV(taxes on investment income on equity)
+ PV(total income on the cashflows)
− PV(taxes on underwriting income)
− PV(taxes on investment income on the cashflow)

---

[1] The timing of tax payments does not actually conform to this simplifying assumption. Expenses are deducted from income in accordance with Statutory Accounting Principles (SAP). This advances the recognition of expenses to an earlier time and makes the resulting tax credit more valuable than indicated by assumption (8b). On the other hand, incurred losses (including IBNR) generally develop upwards for long-tailed lines of insurance. This postpones the recognition of losses to a later time and generally makes the tax credit less valuable than indicated by assumption (8b).

[2] In a particular application, calculation of a more exact value of e may be appropriate.

We now need to write down expressions for each of these five pieces.

PV(investment income on equity) = rS

Due to assumptions (9a) and (9c),

PV(tax on investment income on equity) = $rSFITI$

Also, we have the general result that

PV(total income on the cashflows) = $P^{*\prime} - L' - E' - T^{*\prime}$

By definition,

tax on underwriting income = $(P^* - L - E - T^*)FITU$

By the definition of e we have

PV(tax on underwriting income) = $(P^* - L - E - T^*)\, FITUe$

We have four out of the five pieces we need. In order to get the fifth piece, first we will derive an expression for the investment income on the cashflow. From this will follow the taxes paid on this income and then the present value of these taxes. Unfortunately, this first step will be a little complicated. We know that

investment income on cashflow =
(total income on cashflow) − (underwriting income on cashflow).

We know that the present value of the total income on the cashflows is $P^{*\prime} - (L' + E' + T^{*\prime})$. In Appendix III it is demonstrated, given some not unreasonable assumptions, that we can remove the present value by dividing by a factor y, where $y = (L' + E' + T^{*\prime})/(L + E + T^*)$. (It is useful to think of this as follows. Multiplying by y would adjust the timing to when the losses and expenses are paid, which is the timing of the total income on the cashflow. Dividing by y backs that timing out.) Therefore

investment income on the cashflows =
$(P^{*\prime} - L' - E' - T^{*\prime})/y - (P^* - L - E - T^*)$

income tax on investment income on the cashflow =
$FITI((P^{*\prime} - L' - E' - T^{*\prime})/y - (P^* - L - E - T^*))$

Then we have from the definition of d

PV(income tax on investment income on the cashflow) =
$FITId((P^{*\prime} - L' - E' - T^{*\prime})/y - (P^* - L - E - T^*))$

Assumption (9b) leads to

$$d = (L' + E' + T^{*\prime})/(L + E + T^*) = y$$

Then we have

$$\begin{aligned}&\text{PV(income tax on investment income on cashflow)}\\&= FITIy((P^{*\prime} - L' - E' - T^{*\prime})/y + (L + E + T^* - P^*))\\&= FITI(P^{*\prime} - L' - E' - T^{*\prime} + L' + E' + T^{*\prime} - P^*y)\\&= FITI(P^{*\prime} - P^*y)\end{aligned}$$

This same result can be arrived at starting with a different approach. See Appendix VI for this approach, using a so-called investment balance for taxes.

Thus, the basic equation becomes

$$\begin{aligned}RS = rS - rSFITI + P^{*\prime} - L' - E' - T^{*\prime} - FITUe(P^* - L - E - T)\\ - FITI(P^{*\prime} - P^*y)\end{aligned}$$

We can solve for P^*

$$P^* = \frac{L' + E' - FITUe(L + E)}{(r/s + g)(1 - FITI) - th - R/s + FITIy - (1-t)FITUe}$$

However, y depends in turn on P^*, i.e., on the profit loading

$$y = (L' + E' + htP^*)/(L + E + tP^*)$$

Fortunately, y is usually relatively insensitive to the profit loading, since it is a weighted average of $(L'+E')/(L+E)$ and h, with weights $L+E$ and tP^*.

One can solve numerically via iteration on P^* and y. (For the usual range of input values, the iteration converges very quickly.)

As usual one now uses the defining equation to get the underwriting provision u

$$u = 1 - (t + (L+E)/P^*)$$

With this third model, we have reached a level of refinement which can be used for real world applications. We will later show how a few more refinements can be added, but of course at the cost of further complexity in the model. (As with any actuarial subject, the question of whether a particular technical refinement is worthwhile for a particular application is a matter of judgment. One has to compare the benefits of the extra precision with the extra complications introduced to the model and the cost of obtaining the additional data required.)

SUMMARY OF MODEL THREE

One can solve numerically via iteration on P^* and y

$$P^* = \frac{L' + E' - FITUe(L + E)}{(r/s + g)(1 - FITI) - th - R/s + FITIy - (1-t)FITUe}$$

$$y = (L' + E' + htP^*)/(L + E + tP^*)$$

Then the underwriting provision is given by

$$u = 1 - (t + (L+E)/P^*)$$

The following assumptions were used.

(0) An insurer writes a set of similar policies. Each policy is expected to be in effect for one year.

(1) The insurer receives premiums P^*. (The expected pattern of the timing of payments is known or can be estimated.)

(2) The insurer pays losses L, including those expenses whose timing is the same as the losses. (The expected pattern of the timing of such payments is known or can be estimated.)

(3) The insurer earns income on its investments at a rate r.

(4) The insurer desires a target rate of return on equity of R.

(5) The insurer supplies funds, S, of its own. This equity is around throughout the entire policy year in a constant amount. The required equity is proportional to the premium, with proportionality constant $s = P^*/S$.

(6) The insurer pays T^*, expenses that vary with premiums. (The expected pattern of the timing of such payments is known or can be estimated.)

(7) The insurer pays expenses E, other than those included in L and T^*. (The expected pattern of the timing of such payments is known or can be estimated.)

(8) Underwriting income is taxed at a rate $FITU$. Underwriting income equals premiums minus losses and expenses $= P^* - L - E - T^*$. Federal income taxes on underwriting are paid at the end of the quarter in which the underwriting profit or loss is incurred. (Ignoring any development of incurred losses, this leads to four equal payments at times ¼, ²⁄₄, ¾ and 1 year after policy inception.)

(9) Investment income is taxed at a rate *FITI*. Investment income is defined as the total income minus underwriting income. Federal income taxes on the investment income earned on the equity, are paid at the end of the policy year. Federal income taxes on investment income earned on the cashflow are paid at the time losses and expenses are paid.

NUMERICAL EXAMPLES USING THE THIRD MODEL

We will use all of the assumptions of this third model, including (8b), (9b), and (9c). We will choose input values which are not unreasonable for a real world insurer. However, these values are for illustrative purposes only. In any application it is very important to choose a consistent set of inputs. If the different input values are chosen independently of each other, one can get unusual results to say the least. Just as in ratemaking, the answer is only as valid as the assumptions of the method and the input values chosen.

For the target rate of return on equity after taxes, R, we will use 17%. This may have been given to the actuary by the president of the insurer, the commissioner of insurance, etc. It may have been estimated by looking at the rates of return earned by similar firms or industries. It may have been estimated by using an economic model such as the Capital Asset Pricing Model (CAPM), together with the observed "risk free" rate of return available on U.S. Treasury securities of the appropriate maturities. It may have been estimated by looking at the past results for that line of insurance in competitive markets. Of course, another method or combination of methods may have been used.

There are a number of questions of interest concerning the rate of return on equity. Should rates of return be measured with respect to book or market value of equity? Should the target rate of return differ by line of insurance? How does the target rate of return depend on the other inputs, among them the types of investments and the premium-to-equity ratio?

For the premium-to-equity ratio, s, we will use a value of 2. In the model, we are really interested in stockholder equity, rather than statutory surplus. Therefore, as stated previously, if one is trying to estimate s from data, various adjustments may have to be made to switch from the Annual Statement to Generally Accepted Accounting Principles (GAAP). As stated above, an important consideration is whether one should use the book or market value of equity. Another important consideration would be whether to adjust the equity for the effect of the discounting of loss reserves. Another question of interest is whether different lines of insurance have different acceptable or desirable premium-to-equity ratios.

For the rate of return on investments before taxes, r, we will use 10%. The rate of return, as well as the tax rate, depend on what types of investment the insurer will hold. Also, an insurer who takes more investment risk will generally expect a higher target rate of return, R. One of the questions of interest is whether to use the "imbedded" yields an insurer can be expected to earn on his current portfolio, or whether to use the current yields one could obtain by investing fresh funds. Generally, the rate of return on investments should be measured after taking into account necessary investment expense.

For the federal income tax rate on investment income, *FITI*, we will use 28%. As stated previously, the value of *FITI* would depend on the proportion of each type of asset held, and the rate of return expected on each type of asset.

For the federal income tax rate on underwriting income, *FITU*, we will use 46%. This is the current maximum corporate rate. In the event of an underwriting loss, the 46% tax rate would only be appropriate if there was sufficient income that would be taxed at the 46% rate, so as to be offset by the underwriting loss. As pointed out in Beckman [7], interest from tax-exempt bonds is not taxed and long term capital gains are taxed at less than the corporate rate. (While 85% of dividends on stocks can be deducted from net taxable income, the remaining 15% is taxable at the full corporate rate.)

We will use a ratio of variable expenses to premium, t, of 20%.

For simplicity, we will assume here that all the premium is collected at policy inception, and that all the variable expenses are paid out at policy inception. Also, we will assume that fixed expenses and losses are all paid precisely N years after policy inception. (These are unrealistic simplifications, but in Appendix II is a numerical example for private passenger automobile property damage liability, with more realistic timing assumptions.)

Assuming an arbitrary $800 for losses plus fixed expenses, we get

N(years)	P^*	y	Underwriting Profit Provisions
.5	$1,044	1.059	3.4%
1.0	980	1.020	−1.6%
1.5	916	.981	−7.3%
2.0	853	.943	−13.8%

Note: as we shall see later, the profit provision calculated above for $N=2$ assumes that there is income taxable at 46%, available from other than the investment income on this line of insurance and the equity backing it up, that can be offset by a portion of the projected underwriting loss.

We notice that all other things being equal, the larger N, i.e., the longer tailed the line of insurance, the more negative the profit provision. As has been mentioned above, there is no fundamental difference between positive and negative underwriting margins. We can see here that they are merely different points along the same continuum.

SENSITIVITY TO VARIOUS INPUTS

It is of interest to see how the underwriting profit provision changes as we vary one input at a time. Above we have already seen how the profit provision varies as the length of the cashflow changes. Let's now hold the length of the cashflow constant at $N=1$. For the set of inputs used above this gave a profit provision of -1.6%.

As expected, if you desire a higher target rate of return, you must have a more positive underwriting profit provision, all other things being equal. If you can earn a higher return on investments, you can afford a less positive underwriting profit provision. When you have a higher premium-to-equity ratio, you can afford a less positive profit provision. When you have a projected underwriting loss, the higher the federal income tax rate on underwriting, the more negative the profit provision, since the "tax shield" is worth more. The situation is reversed when you project an underwriting gain. The profit provision gets more sensitive to *FITU* as the profit provision gets further from zero. Finally, the higher the rate of federal income taxes on investments, the more positive the profit provision. The profit provision gets more sensitive to the value of *FITI* as the cashflow gets longer, and thus more investment income can be earned.

We have varied the different inputs one at a time. In actual practice, many of the inputs will depend on one another. Thus, one can not just vary them independently of each other. However, it is still enlightening to see how the profit provision varies, all other things being equal.

One could perform a similar analysis using differentiation. This is outlined in Appendix V.

SENSITIVITY TO VARIOUS INPUTS

Assumptions	Underwriting Profit Provision
target return on equity	
$R=16\%$	-2.6%
$R=17\%$	-1.6%
$R=18\%$	-0.7%
rate of return on investments	
$r=9\%$	0.1%
$r=10\%$	-1.6%
$r=11\%$	-3.4%
premium-to-equity ratio	
$s=1.5$	1.5%
$s=2.0$	-1.6%
$s=2.5$	-3.5%
federal income tax rate on underwriting	
$FITU=30\%$	-1.2%
$FITU=46\%$	-1.6%
federal income tax rate on investments	
$FITI=18\%$	-4.1%
$FITI=28\%$	-1.6%
$FITI=38\%$	0.8%

A COMPUTATION OF THE AVERAGE INCOME TAX RATE ON INVESTMENTS

In the previous numerical examples, a 10% rate of return and a 28% tax rate on investment income were used. Here is one possible source for these values.

Make the following specific assumptions as to the source of the projected invested income. Assume that the insurer will have his assets invested solely in bonds, one half taxable, and one half tax-exempt. Further, assume that the taxable bonds will return 12% before taxes, while the tax-exempts will earn 8%. Then the rate of return, r, and the federal income tax rate on investment income, *FITI*, can be computed as follows.

Type of Asset	Amount	Rate of Return	Income	Tax rate	Tax
Taxable Bond	.5	12%	.06	46%	.0276
Tax-Exempt Bond	.5	8%	.04	0%	0
	—		—		—
Combined	1.0		.10		.0276

Thus the combined rate of return is .10/1.0 = 10%. The combined tax rate is .0276/.10 = 27.6%, or 28% to the nearest percent. This matches the choices of $r = 10\%$ and $FITI = 28\%$, which were made for the numerical models above.

TAX SHIELD, UNDERWRITING LOSSES

We have seen that underwriting losses can be used to offset otherwise taxable income. As such they have a potential value, which can be only realized if there is taxable income available to be offset. In general, when one has a negative provision for underwriting profits, one should check whether income is available to be offset that would have been taxed at the value of *FITU* chosen.

Here we will check our numerical examples from above to see whether there is enough income taxed at 46%, so as to be offset by our projected underwriting loss. We will use the distribution of assets and rates of return on assets from the previous section.

How much taxable income is available to be offset by an underwriting loss? From the previous section, .06/.10 = 60% of the pre-tax investment income is taxable (at 46%).

We have for model three

investment income on the cashflows

$$= (P^{*\prime} - L' - E' - T^{*\prime})/y - (P^* - L - E - T^*)$$

$$= P^{*\prime}/y - P^* - (L' + E' + T^{*\prime})(L + E + T^*)/(L' + E' + T^{*\prime}) + (L + E + T^*)$$

$$= P^{*\prime}/y - P^* = P^*(g/y - 1)$$

However, we also have

investment income on the equity $= P^*r/s$

Therefore, adding the two sources of investment income gives

investment income $= P^*(r/s + g/y - 1)$

In this case, 60% of the investment income is taxable (at 46%).

taxable investment income $= .6P^*(r/s + g/y - 1)$

If our projected underwriting loss exceeded our projected income taxable at 46%, it might no longer be appropriate to take $FITU = 46\%$. It might still be appropriate if there is taxable income somewhere else which may be offset. For example, the use of Tax-Loss Carry-Overs allows interactions between separate calendar years, as explained in Beckman [7]. Also there may be taxable income generated elsewhere in the corporation. However, this gets into a complicated question of possible subsidies across lines of insurance or states, or even the question of the insurer being part of a larger corporate structure. While this subject is beyond the scope of this paper, the value of being able to use these tax credits available due to underwriting losses is far from merely theoretical. In part, it may explain some of the takeovers of property casualty insurers by firms outside the industry, as well as attempts at diversification by property casualty insurers.

Here we will assume that there is no taxable income available from other sources. Then the expected underwriting loss will exceed the income taxable at 46% if

$$L + E + T^* - P^* > .6\,P^*(r/s + g/y - 1)$$

$$(L + E + T^*)/P^* - 1 > .6(r/s + g/y - 1)$$

$$-u > .6(r/s + g/y - 1)$$

Note that, more generally, .6 would be that portion of investment income that is taxable at 46%.

In our numerical examples, the expected underwriting loss will exceed the income taxable at 46% if

$$-.6(.10/2 + 1.1/y - 1) > u$$

or, since y is approximately one for all our numerical examples,

$$-9\% > u$$

This is the case for our numerical example with $N = 2$. The calculated underwriting loss exceeds the taxable income available to offset it. (Here, for simplicity, we have assumed that we have income which is either taxed at 46% or is tax-exempt. In general there are other types of income. In certain cases the use of statutory tax rates may not be appropriate. See, for example, Report of the NAIC Investment Income Task Force, p. 23 [6].)

So unless one assumes that taxable income is available from somewhere else, the calculated underwriting provision for $N = 2$ is incorrect. In this case, a solution is to set $FITI = FITU = 0\%$. When we recalculate the profit provision it increases from -13.8% to -13.0%. This difference becomes more pronounced as N gets larger.

In general, a good check of any calculated profit provision is to rerun the calculation with $FITI = FITU = 0$. The profit provision in the former case should not be more negative than the latter case. However, even if this test is passed, you may still have a value for $FITU$ which is too large, if some of the income to be offset is taxed at a lower rate, e.g., long term capital gains.

NON-ITERATIVE APPROXIMATIONS TO MODEL THREE

Instead of the above iterative solution, one could solve for P^* in closed form, but the solution of the quadratic equation is less than illuminating. Except when dealing with long-tailed lines of insurance, (e.g. one in which loss payments take as long as for workers' compensation or longer), one can approximate the iterative solution fairly closely in either of two ways. One can either just set $y = 1$ in the above equation for P^*, or one can do so in the previous equation for the rate of return. In the latter case, we would get:

$$P^* = \frac{(L' + E')(1 - FITI) - (E + L)(FITUe - FITI)}{(r/s + g - th)(1 - FITI) - R/s - (1 - t)(FITUe - FITI)}$$

THE TIMING OF INVESTMENT TAXES ON THE CASHFLOWS

One can get slightly different equations from those in the third model depending on what timing assumptions you make concerning the timing of the federal income taxes on investment income earned on the cashflows. In Appendix IV a result is developed for a slightly different assumption than (9b).

When using these models for a specific case, it may be possible to more carefully determine when these taxes will be paid. Generally, interest income is taxed as accrued, but capital gains are only taxed as realized. While different assumptions about the timing of the payments of these taxes can have a large effect for a long tailed line of insurance, a further exploration of this subject is beyond the scope of this paper.

FOURTH MODEL
FINANCE CHARGE INCOME AND UNCOLLECTED PREMIUM

In calculating underwriting profit provisions two additional refinements have been found useful for certain applications. These will be presented as good examples of how additional refinements can be incorporated into the basic model. (In one actual application, finance charge income lowered the profit provision by about 1%, while earned but uncollected premium raised it by about 1/2%.)

Many insurers have finance plans under which the premium is paid in installments. The insured is often charged for this privilege. It seems appropriate to include separate consideration of this finance charge income, if it has not somehow already been included elsewhere, when such financing is responsible for a significant delay in the premium inflow, and the expenses relating to financing are included in the expenses used elsewhere in the ratemaking process.

Insurers usually do not collect all the premium that is "earned." Therefore, it seems appropriate to make the manual rate larger than otherwise determined, in order to end up collecting the desired premium. (This effect of the earned but uncollected premium can be incorporated somewhere else in the ratemaking process instead. However, it can be conveniently incorporated here.)

For this fourth model, we add the following two assumptions.

(10) The insurer receives finance charge income F^*. (The expected pattern of the timing of such payments is known or can be estimated.) Define $v = F/P = F^*/P^*$, the ratio of finance charge income to premium. Let $f = F'/F = F^{*\prime}/F^*$.

(11) The insurer will collect only a portion of the premiums which are earned. Define c = ratio of earned but uncollected premium to earned premium. (In the case of a cancelled policy, one should distinguish between any uncollected portion of the original written premium that was never earned and the uncollected portion of the earned premium.)

To include finance charge income in the equations from model three, one merely includes it as another inflow, similar to premium. The basic equation becomes

$$RS = rS - rSFITI + (P^{*\prime} + F^{*\prime} - L' - E' - T^{*\prime}) \\ - FITUe(P^* + F^* - L - E - T^*) - \mathrm{FITI}(P^{*\prime} + F^{*\prime} - P^*y - F^*y)$$

When we divide by P^* and solve for P^* we get

$$P^* = \frac{L' + E' - FITUe(L + E)}{\begin{bmatrix}(r/s + g + vf)(1 - FITI) - th - R/s \\ + FITIy(1 + v) - (1 + v - t)FITUe\end{bmatrix}}$$

where as before this can be solved by iteration on P^* and y, where y is

$$y = (L' + E' + T^{*\prime})/(L + E + T^*) = (L' + E' + htP^*)/(L + E + tP^*)$$

Now we wish to calculate the underwriting profit provision, taking into account earned but uncollected premium. The usual manner in which u would be used to construct manual rates is

(earned manual premium)$(1 - u)$ = losses + expenses.

The proper collected premium is by definition P^*. By the definition of c, $P^*/(1 - c)$ is the proper earned manual premium, since c = (earned manual premium $- P^*$)/earned manual premium. The variable expenses are assumed to be t times the collected premium P^*, rather than the earned manual premium. (This is true for the premium taxes, and is not an unreasonable asumption for other expense items which might be treated as variable, such as commissions.)

Then we would have

$$P^*(1 - u)/(1 - c) = L + E + tP^*$$

$$u = 1 - (1 - c)(t + (L + E)/P^*)$$

This differs from the equation in model three, by the addition of a factor of $1 - c$. It reduces to the prior case when $c = 0$.

FIFTH MODEL
EQUITY AS A FLOW

Starting with the second model, we have assumed in assumption (5) that the equity exists throughout the policy year in a constant amount. This simple assumption can be generalized, by thinking of equity as a flow.

(5′) The insurer supplies funds of its own, which we will call equity. The amount of equity backing up the policy varies over time. (It is zero in the distant past as well as in the far future.) Let W be the equity inflow and outflow, by quarter. Then the cumulative sum by quarter of W is the desired equity flow by quarter.

When treating equity as a flow, it has been found useful to introduce two new terms, the "cumulative premium-to-equity ratio" and the "initial premium-to-equity ratio." Depending on the equity flow chosen, one or the other concept is usually more readily applicable.

The cumulative premium-to-equity ratio is the usual concept of premium to equity as used elsewhere in insurance. Conceptually, it is the ratio one would observe if one looked at the insurer, or perhaps more abstractly, looked at just that portion of the insurer writing this line of business. Given a particular equity flow, the cumulative premium-to-equity ratio observed would usually depend on what growth rate one assumed for premium. Sometimes it is calculated using a zero growth rate, the so-called steady state case.

The initial premium-to-equity ratio is the ratio of premiums to equity at the inception of the policy.

We now assume

(5b′) Let S be either the cumulative equity or the initial equity, whichever concept is applicable. Then the required equity S is proportional to the premium P^*, with proportionality constant $s = P^*/S$.

For the basic equation we need the present value of the total return we wish to earn after taxes. As before, when dealing with the cashflows, this is just the present value of the inflows of equity minus the present value of the outflows of equity, at the target rate of return R.[3] The present value of the investment

[3] This assumes that the target return on equity is received at the time(s) the surplus flows out. While other assumptions could be made as to when the return on equity is received and/or paid out, a further discussion of this subject is beyond the scope of this paper.

income on the equity is similar, but instead uses the rate of return r. (There is nothing analogous to underwriting income as on the cashflows, since the sum of W is zero.)

Let's define $W' = W$ discounted by r
$W'' = W$ discounted by R.

Then, the desired present value of the total return is W''. PV(investment income on equity) $= W'$.

If we assume

(9b′) The federal income taxes on the investment income earned on the equity are paid as the equity flows out.

Then we have an analogy to the cashflow case

PV(tax on investment income on equity) $= FITI\ W'$

Thus the basic equation from model four becomes

$$W'' = W' - W'FITI + (P^* + F^{*\prime} - L' - E' - T^{*\prime}) - FITUe(P^* + F^* - L - E - T^*) - FITI(P^{*\prime} + F^{*\prime} - P^*y - F^*y)$$

Let $w = W/S$, then when we divide by P^* and solve for P^* we get

$$P^* = \frac{L' + E' - FITUe(L + E)}{\left[\begin{array}{l}(w'/s + g + vf)(1 - FITI) - th - w''/s \\ + FITIy(1 + v) - (1 + v - t)FITUe\end{array}\right]}$$

As before this can be solved by iteration on P^* and y, where y is

$$y = (L' + E' + T^{*\prime})/(L + E + T^*) = (L' + E' + htP^*)/(L + E + tP^*)$$

Our previous models are just special cases of this one. There we had the equity flow in at policy inception, and flow out one year later. This is sometimes referred to as the "block equity" assumption. In this case, w is a vector with value 1 at time = 0 and value −1 at time = 1 year. Thus

$$w' = (1 + r) - 1 = r.$$

Similarly

$$w'' = (1 + R) - 1 = R.$$

If one makes those substitutions in the equations here, and one uses the cumulative equity concept, the equations reduce to those in the fourth model.

For illustrative purposes, here is an example of an equity flow that varies over time. Set the equity backing up the policy at an initial value, and then have a decreasing balance as losses and expenses are paid. When the last payment is made, there is no longer any equity backing up this policy or group of similar policies. Let

$$w = 1,0,0,0,\ldots - (L + E + T^*)/\text{sum}(L + E + T^*)$$

where 1,0,0,0... is a vector by quarters, and represents an inflow of 1 at time equals 0. Here $L + E + T^*$ is also a vector of payments by quarter. (In the rest of the paper, this expression has represented their sum, which is a scalar rather than a vector quantity.) As we perform an iterative solution, T^* will vary with each iteration, and thus so will w. The sum of $w = 1 - 1 = 0$ as expected, since equity that flows in eventually flows out.

In this case, s represents the initial premium-to-equity ratio. If one used the same s, this flow would assign more cumulative equity to longer tailed lines than shorter tailed lines.

As an alternative, one could construct a flow based on when losses are incurred. One could of course come up with other timings of equity. One could have the desired amount of equity be determined in some manner other than as a proportion to premium.

In any case, it is important to remember that an insurer's entire equity is in theory available to back up each policy. So while the assignment of equity to a particular line or state may be a necessary assumption for the running of these profit models, one should not take it too literally. One must remember that an insurer who writes more than one line of insurance, in more than one state, would generally need less equity per dollar of premium, than one which wrote only a single line in a single state. When assigning equity for the purposes of these models, one should not ignore the spreading of risk available in multi-state and multi-line operations, since this goes to the very heart of the insurance process.

MISCELLANEOUS

Unless one thinks about it carefully, it is easy to misinterpret a negative underwriting profit provision, particularly a very negative one such as −50%. Since $P^* = (L + E + T^*)/(1 - u)$, if $u = -50\%$, the premium is two-thirds of the losses and expenses. Presumably, in this extreme case, one can earn enough investment income during the long time prior to paying the losses so that one will have the money available to pay the losses, as well as enough left over to earn the target return. For example, this might be the case for lifetime escalating benefits to widows under workers' compensation. An underwriting profit provision of -100% would mean that the premium was one half of the losses and expenses, something far from unheard of for annuities.

In this paper, the concept of return on equity has been used. This concept may not be appropriate for a mutual rather than stock insurer. One can adapt the methods presented here to deal with some other concept more appropriate for a mutual insurer. One example might be to substitute a target return on policyholder's surplus. This would relate to a desired growth rate in surplus. Another example might be to substitute a desired return on premiums, so as to cover "contingencies."

Dividends to policyholders have not been dealt with in this paper. However, anticipated or desired dividends could be incorporated into the models, as another outflow. If used in a ratemaking context, one must take care to be consistent with whatever ratemaking methodology has been used, i.e., one must not double count anticipated dividends.

Throughout this paper, we have assumed that one knows, or can make an unbiased estimate of, the input values to be used. Specifically, it is assumed, when using these methods in a ratemaking context, that some ratemaking method has been used in order to make an unbiased estimate of the expected value of losses and expenses. (If the estimation method is biased, the method should be changed so as to remove the bias. Methods of estimating future losses and expenses are dealt with extensively in the actuarial literature, and specifically on the Casualty Actuarial Society syllabus of the examination on the principles of ratemaking.) The fact that actual losses will vary around the prediction is an inherent feature of the insurance business. Such uncertainty should be taken into account either explicitly or implicitly when choosing a target rate of return for an insurer.

CONCLUDING REMARKS

There are a number of methods of reflecting the total return needs of an insurer. There is no single best procedure or method. However, for the sake of clarity, this paper has focused on one basic method of calculating a provision for underwriting profits. As with most actuarial questions, the choice of what method to use will depend on the peculiarities of the situation and the purpose for which it is to be used. One should carefully examine the assumptions underlying any model, as well as the choice of inputs, in order to see whether they are reasonable for the given situation.

In this paper, the author has been very careful to state all the assumptions used. The author feels that such a careful axiomatic approach is necessary, since it is very easy to get absurd results by mixing inconsistent assumptions or using input values which do not match the assumptions. Also, this approach allows one to examine the underlying assumptions and change those which may not hold for a particular application. For a particular application, it may be useful to modify a particular assumption in order to test the sensitivity of the result to this assumption.

This paper is not meant to address such controversial issues as whether investment income should be explicitly reflected when rate filings are submitted to state insurance departments. Rather this paper starts with the premise that the subject of underwriting profit provisions is an area in which actuaries can be of assistance in advancing knowledge and developing methods.

While this paper has concentrated on the theoretical aspects, this subject has many practical applications. A company actuary might use it to help price a product, or to estimate what rate of return on equity has been earned on a certain book of business. In regulated lines of insurance, these methods could be used by an actuary in regulation either to set rates or to examine the reasonableness of filed rates.

If one wants to employ these methods for some practical application, one runs into the usual problem with most actuarial methods: one must choose or determine the input values to use. In most cases the input values chosen will have an extremely large effect on the resulting answer. It is important to choose a consistent set of input values.

The input values should reflect the economic climate one expects during the relevant period of time. For example, as we have seen, the underwriting profit provision depends on the rate of return available from investments. A model allows one to adjust the profit provision for changing economic conditions. What may have been a proper claim cost trend in the 1950's, may no longer be appropriate for the 1980's. Similarly, a proper underwriting provision then may no longer be appropriate now.

ACKNOWLEDGEMENTS

Some of the ideas in this paper originated in work done jointly with Michael Kooken and Stefan Peters on the SRB Profit Model, which is listed below under unpublished material. The author wishes to thank Stefan Peters and Richard Derrig for reading an earlier version of this paper and providing helpful comments.

REFERENCES

[1] J. Robert Ferrari, "The Relationship of Underwriting, Investments, Leverage, and Exposure to Total Return On Owners Equity," *PCAS* LV, 1968.

[2] S. G. Kellison, *The Theory Of Interest*, Richard D. Irwin, Inc., pp. 8–10, pp. 36–38.

[3] R. H. Snader, "Fundamentals of Individual Risk Rating and Related Topics," A Study Note published by the Casualty Actuarial Society, pp. 61–62.

[4] Report of the Advisory Committee to the NAIC Task Force on Profitability and Investment Income. January, 1983.

[5] National Association of Insurance Commissioners. "Measurement of Profitability and Treatment of Investment Income in Property and Liability Insurance," June, 1970. 1970 *Proceedings of the NAIC*, Vol. IIa.

[6] National Association of Insurance Commissioners. Report of the Investment Income Task Force. June, 1984.

[7] R. W. Beckman, "Federal Income Taxes," *PCAS* LVIII, 1971.

SOME FURTHER READING

Arthur D. Little, Inc., "Prices and Profits in the Property and Liability Insurance Industry," Report to the American Insurance Association, November 1967.

Robert A. Bailey, "Underwriting Profit from Investments," *PCAS* LIV, 1967.

Casualty Actuarial Society, "Total Return Due a Property-Casualty Insurance Company," 1979 Call Paper Program, May 1979.

Robert W. Cooper, *Investment Return and Property—Liability Insurance Ratemaking*, Richard D. Irwin, Inc., 1974.

J. D. Cummins and S. E. Harrington, eds., *Fair Rate of Return on Property-Liability Insurance*, Hingham, Massachusetts, Kluwer-Nijhoff Publishing Co., 1986. (This source contains, among other materials, papers on the Fairley model, the Myers-Cohn model, the Hill-Modigliani model, and a history of the use of profit models in Massachusetts.)

William B. Fairley, "Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results," *Bell Journal* 10, No. 1 (Spring 1979).

J. D. Hammond, Arnold F. Shapiro, and N. Shilling, "The Regulation of Insurer Solidity Through Capital and Surplus Requirements," The Pennsylvania State University, University Park, Pennsylvania, 1978.

Paul Joskow, "Cartels, Competition and Regulation in the Property-Liability Insurance Industry," *Bell Journal* 4, No. 2 (Autumn 1973).

C. Arthur Williams, Jr., "Regulating Property and Liability Insurance Rates Through Excess Profits Statutes," *Journal of Risk and Insurance* 50, No. 3 (September 1983).

UNPUBLISHED MATERIAL OF INTEREST

While the reader may have difficulty obtaining copies of this material, those who can obtain copies should find them of interest.

Massachusetts Automobile Rating and Accident Prevention Bureau, Boston. Filing for 1984 Private Passenger Automobile Rates, Section 100H. Fall 1983.

Massachusetts Automobile Rating and Accident Prevention Bureau, Boston. Richard Derrig and Acheson Callaghan. "Position Paper on the Risk and Reward of Underwriting" and "Position Paper on Surplus." June 1982.

Massachusetts Automobile Rating and Accident Prevention Bureau, Boston. Richard Derrig, "The Effect of Federal Income Taxes on Investment Income in Property-Liability Ratemaking." February 1982.

State Rating Bureau, Massachusetts Division of Insurance, Boston. Filing for 1984 Private Passenger Automobile Rates, Section 100H. Fall 1983.

State Rating Bureau, Massachusetts Division of Insurance, Boston. Michael Kooken, Howard Mahler, and Stefan Peters. "A Profit Model Based on a Company's Actual Portfolio" and "A Profit Model Using a Statutory/Regulatory Company." Spring 1981.

James M. Stone, Decision on 1976 Automobile Rates. (Actually two separate decisions, one for BI and one for PD.) Fall 1975. Massachusetts Division of Insurance, Boston.

APPENDIX I

NOTATION AND VARIABLE NAMES

All discounting is to the end of the policy year.

The present value using the rate of return r is denoted by a single prime.

The present value using the rate of return R is denoted by a double prime.

An asterisk indicates that a quantity is loaded for profits.

u = provision for underwritng profit
r = rate of return on insurers investments (before taxes)
R = target rate of return on equity (after taxes)
$FITI$ = federal income tax rate on investments
$FITU$ = federal income tax rate on underwriting
S = stockholders' equity or insurer's net worth (although it is useful to think of this as surplus, the two concepts are not numerically equivalent)
P = premiums (based on 0% profit loading)
P^* = premiums loaded for profit.
$s = P^*/S$
L = losses, including those expenses whose timing is the same as losses
E = expenses which are not included in either L or T
T = expenses which are proportional to premium (based on premium with 0% profit loading)
F = finance charge income (based on premium with 0% profit loading)
$t = T/P = T^*/P^*$
$v = F/P = F^*/P^*$
$g = P'/P = P^{*\prime}/P^*$
$h = T'/T = T^{*\prime}/T^*$
$f = F'/F = F^{*\prime}/F^*$
e = (present value of federal income taxes on underwriting)/(federal income taxes on underwriting)
d = (present value of federal income taxes on investment)/(federal income taxes on investment)
$y = (L' + E' + T^{*\prime})/(L + E + T^*)$ = (present value of the outflows)/(outflows)
c = (earned but uncollected premium)/(earned premium)
W = surplus inflow and outflow
$w = W/S$

APPENDIX II

PRIVATE PASSENGER AUTO PDL

Here we will present, for illustrative purposes only, a numerical example using the third model. The timing of the cashflows presented here is similar to that which one might find for a real insurer. Many of the inputs are the same as the previous numerical examples we have presented for model three.

This example is for property damage liability coverage for private passenger automobile insurance. In practice it is not unreasonable to calculate separate profit provisions for different sublines of automobile insurance. One reasonable division is into bodily injury coverages, property damage liability, and physical damage coverages. The profit provision and the length of the loss flow for property damage liability is generally between the other two. Bodily injury coverages generally have the longest loss flow, and thus the smallest (least positive or most negative) profit provision of the three.

The timing of premium and loss payments used here is based on the timing of payments observed in one state in the recent past.

It is also necessary to estimate the magnitude and timing of the different expense payments. In the numerical example given here, certain assumptions have been made concerning expenses. (These particular assumptions are of little importance in and of themselves. However, they do serve to illustrate one method of estimating the timing of expense payments for modelling purposes.)

The allocated loss adjustment expense and one half of the unallocated loss adjustment expense have been assumed to be expended with the losses and are included in the loss flow. The remaining half of the unallocated loss adjustment expense is assumed to be expended evenly throughout the policy year. Other acquisition expense is assumed to be expended evenly over the five month period beginning with the first month prior to the policy effective date. General expenses are used here to mean expenses other than loss adjustment expense, commissions, other acquisition, and premium taxes. General expenses are assumed to be expended 30% in the three months prior to the policy effective date, while 70% is expended evenly during the policy year. We assume that general expenses and unallocated loss adjustment expense are equal in size, and other acquisition expense is half of these. (This assumption is a fair approximation for a typical agency company writing private passenger automobile insurance.) Also, let company expense be defined as general expense, plus other acquisition expense, plus one half of unallocated claims expense. Then our assumptions lead to a payment pattern for company expense of 20%, 30%, 20%, 15%, 15%, starting in the 0th quarter. (The policy effective date is the end of the 0th quarter and the beginning of the 1st quarter.) Commission expense is assumed to be paid as premiums are received. Premium taxes are assumed to be paid quarterly.

In general, the assignment of expenses to either the fixed category or the group that varies with premium should match the assumptions used elsewhere in the ratemaking methodology. This assignment has an important numerical impact on the calculated profit provision when the profit provision is far from zero, e.g., −10% or less. In the numerical example given here, only premium taxes are assumed to vary with premiums. This is why the ratio of variable expenses to premium, t, is only 2.3%.

Answers and Inputs

u = provision for underwriting profits = 3.7%
$P^* = 1039.7$
$P = 1000.000$
$T = 23.000$
$t = T/P = .023$
$g = P'/P = 1.0668$
$h = T'/T = 1.0492$
$E = 367.594$
$L = 609.406$
$E' = 392.373$
$L' = 610.700$
$e = 1.0368$
$R = 17\%$
$r = 10\%$
$FITU = 46\%$
$FITI = 28\%$
$s = 2$
$y = (L' + E' + htP^*)/(L + E + tP^*) = 1.0272$

Private Passenger Auto PDL Cashflows

Based on a company expense-to-loss ratio of .2545. Based on a commission expense-to-loss ratio of .3487. (Assume 0% profit loading for determining the weights of the various cashflows.)

Quarter	Premium	Premium Tax	Company Expense	Commission Expense	Loss
0	31.100		31.019	6.609	
1	334.000	5.750	46.528	70.975	29.046
2	467.900	5.750	31.019	99.429	92.869
3	143.200	5.750	23.264	30.430	110.577
4	23.800	5.750	23.264	5.057	125.722
5					111.860
6					56.510
7					27.011
8					16.311
9					9.562
10					7.676
11					4.630
12					3.447
13					3.928
14					3.235
15					1.928
16					1.692
17					1.455
18					0.851
19					0.669
20					0.427
Sum	1000.000	23.000	155.094	212.500	609.406

Note: one half of the unallocated loss adjustment expense is contained in "company expenses" and "losses." All the allocated loss adjustment expense is contained in the "losses."

All discounting of cashflows is to the end of the policy year. Cashflows are assumed to occur in the middle of the relevant quarter. For example

$$\begin{aligned} P' &= (31.1)(1.1)^{9/8} + (334)(1.1)^{7/8} + (467.9)(1.1)^{5/8} + \\ &\quad (143.2)(1.1)^{3/8} + (23.8)(1.1)^{1/8} \\ &= 1066.8 \end{aligned}$$

APPENDIX III

PRESENT VALUE OF INCOME ON THE CASHFLOW

In developing model three, we used a relationship between the total income on the cashflow and its present value. In this appendix we will show that, given certain assumptions, the present value of the income on the cashflow divided by the income on the cashflow is given by y, a similar quantity for the outflows.

Assume we have an outflow O, divided into payments $O\langle k\rangle$ by quarter. Assume we have an inflow I, divided into payments by quarter $I\langle j\rangle$. We wish to find out how much total income is earned on the cashflows. This depends on how long the inflow is invested.

This requires some assumptions. A not unreasonable assumption is to assume that the inflow is invested until the time of the outflow. With inflows and outflows occurring at various times, it is necessary to make a more precise assumption.

We assume that the inflow is divided up in proportion to the present values of the outflows. (This is neither a first-in first-out assumption, nor a last-in first-out assumption.) In other words, we assume I, or more precisely each $I\langle j\rangle$ is divided up into pieces using weights $O\langle k\rangle'/O'$. (We divide by O' so that the weights add up to one.) The kth piece of $I\langle j\rangle$ is invested until $O\langle k\rangle$ is paid. During the time it is invested each piece of I grows by a factor $(O\langle k\rangle/O\langle k\rangle')/(I\langle j\rangle/I\langle j\rangle')$. That this is the increase becomes clearer if one just puts each of the two ratios in terms of powers of $1 + r$. If $I\langle j\rangle$ occurs at time a and $O\langle k\rangle$ occurs at time b, then the ratio is just $(1 + r)^{(1-b)}$ divided by $(1 + r)^{(1-a)}$, or $(1 + r)^{(a-b)}$.

Thus we have that after growth, the kth piece of $I\langle j\rangle$, which was $I\langle j\rangle O\langle k\rangle'/O'$ has grown to:

$$(I\langle j\rangle O\langle k\rangle'/O')(O\langle k\rangle/O\langle k\rangle')/(I\langle j\rangle/I\langle j\rangle') = O\langle k\rangle\, I\langle j\rangle' \,/\, O'$$

Then the total income is

$$\begin{aligned}&\Sigma_j\,(\Sigma_k(O\langle k\rangle I\langle j\rangle'/O')) - O\\ &= I'\,O/\,O' - O\\ &= (I' - O')(O/O')\end{aligned}$$

Now $I' - O'$ is the present value of the total income on the cashflow. So we have that

income on cashflow = (present value of income on cashflow) / y

where y = (present value of outflows)/ outflows = O'/O

APPENDIX IV

ALTERNATIVE TIMING OF *FITI* ON CASHFLOWS

In this appendix we will develop further the work done in the previous appendix. We will use the same notation. We will see how an alternate assumption concerning the timing of federal income taxes on the investment income on the cashflows yields a different result than in model three.

We saw how the kth piece of $I\langle j\rangle$, which was $I\langle j\rangle\ O\langle k\rangle'/O'$, grew to $O\langle k\rangle I\langle j\rangle'/O'$. Thus the investment income is their difference

$$(O\langle k\rangle I\langle j\rangle' - I\langle j\rangle O\langle k\rangle')/O'$$

Let's assume that the income taxes on this investment income are paid at time k. Then one gets this piece of the federal income taxes on investment by multiplying by *FITI*. Since the tax payment has been assumed to be made at time k, we get the present value by multiplying by a factor $O\langle k\rangle'/O\langle k\rangle$. Thus the present value of the federal income taxes on this piece of the investment income is

$$\begin{aligned}&FITI(O\langle k\rangle'/O\langle k\rangle)(O\langle k\rangle I\langle j\rangle' - I\langle j\rangle O\langle k\rangle')/O'\\&= FITI(O\langle k\rangle' I\langle j\rangle'/O' - I\langle j\rangle O\langle k\rangle' O\langle k\rangle'/O\langle k\rangle O')\end{aligned}$$

When we sum over all i and j we get

$$\begin{aligned}&\text{PV}(FITI \text{ on cashflows})/FITI =\\&I' - (I/O')(\Sigma_k\ O\langle k\rangle' O\langle k\rangle'/O\langle k\rangle)\end{aligned}$$

This differs from model three where we had

$$\text{PV}(FITI \text{ on cashflows})/FITI = y(I'y - I) = I' - yI$$

If we define

$$z = (\Sigma_k\ O\langle k\rangle' O\langle k\rangle'/O\langle k\rangle)/O'$$

Then the result here can be rewritten as

$$\text{PV}(FITI \text{ on cashflows})/FITI = I' - zI$$

This is of the exact same form as the result used in model three, except we have z in place of y. Thus the equation for P^* would be the same, except we would replace y by z.

The resulting profit provisions are similar. For example, below are the results for the same numerical examples we calculated for model three, using the same inputs.

	As Per This Appendix			As Per Model Three		
N(years)	P^*	z	Profit Prov.	P^*	y	Profit Prov.
.5	$1,044	1.060	3.4%	$1,044	1.059	3.4%
1.0	979	1.021	−1.7%	980	1.020	−1.6%
1.5	914	.984	−7.5%	916	.981	−7.3%
2.0	850	.948	−14.2%	853	.943	−13.8%

APPENDIX V

DIFFERENTIATION OF THE FORMULA FOR PROFIT PROVISION

In this appendix will be shown the manner in which the underwriting profit provision varies with various important inputs. This will be done by differentiating the formula for the underwriting profit provision. (In the main text were shown some actual numerical results of varying inputs.) We will use the third model.

$$u = 1 - t - (E + L)/P^* = 1 - t - N/D$$

where $N = (r/s + g)(1 - FITI) - th - R/s + FITIy - (1 - t)FITUe$
and $D = (L' + E')/(L + E) - FITUe$
(N and D have been used for numerator and denominator, only in this appendix.)

Then we have

$$du/dR = 1/sD$$

Thus, du/dR is greater than zero, and is approximately 1 for the values used here.

$$du/ds = (r(1 - FITI) - R)/Ds^2$$

Thus, du/ds is less than zero and is approximately $-.05$ for the values used here.

$$du/dFITU = (1 - t)e/D - eN/D^2$$
$$= e(1 - t)/D - e(1 - u - t)/D$$
$$= eu/D$$

Thus, $du/dFITU$ has the same sign as u, and is approximately equal to $2u$.

$$du/dFITI = (r/s + g - y)/D$$

Thus, $du/dFITI$ is generally greater than zero. It is significantly larger the longer the cashflows.

The reason we don't give an algebraic result for du/dr is that the result of differentiating u by r would be quite a complex expression. Remember that variables which involve present values, such as g, h, y, E', and L', involve r, in a rather complicated manner.

APPENDIX VI

INVESTMENT BALANCE FOR TAXES

In this appendix we will explore an alternate way to get the expression for the present value of the taxes on investment income on the cashflow, which was used in the third model. We will set up something called the investment balance for taxes, *IBT* for short.

Assume there is an inflow $I\langle j\rangle$ and outflow $O\langle j\rangle$, each by quarter. Call the sums I and O. Deal with $I\langle j\rangle$ and $O\langle j\rangle I/O$, so that both vectors sum to the same value I. (Think of I as the premiums loaded for profits. Intuitively this manner of doing things prevents counting the underwriting profit or loss twice, since it is dealt with separately elsewhere in the model.)

Let $N\langle j\rangle = I\langle j\rangle - O\langle j\rangle I/O$. N is the net cashflow by quarter, but adjusted so that the outflows are loaded for profit.

Then the *IBT* is set up as follows. Take the cumulative sum by quarter of N. (Since we have set them up so that both vectors have the same sum, for large enough values of time *IBT* is 0.) $IBT\langle j\rangle = \Sigma_{k=1 \text{ to } j} N\langle k\rangle$. Then in each quarter this amount is available to earn investment income. So we multiply it by $q = (1 + r)^{.25} - 1$, the quarterly rate of investment return. Assume the income taxes are paid on this investment income the following quarter. Assume for convenience that the first element of the vectors has a discount factor of 1. (One can discount to any point in time. If another point in time is taken, an additional discount factor will appear, but make no difference in the result.)

$$\begin{aligned}&\mathrm{PV}(qIBT)FITI(1 + r)^{-.25}\\ &= FITIq\,(\Sigma_j\,\Sigma_{k=1 \text{ to } j}\,(N\langle k\rangle\,(1 + r)^{-j/4}))\end{aligned}$$

Now collect all the terms involving a given $N\langle j\rangle$. Each $N\langle j\rangle$ appears starting with a term in which it is multiplied by a factor of $(1 + r)^{-j/2}$. Then it appears in all the subsequent terms, except in the next term it is multiplied by $(1 + r)^{-(j+1)/4}$, in the one after that by $(1 + r)^{-(j+2)/4}$, etc. Thus,

$$= FITIq\,(\Sigma_j\,(N\langle j\rangle\,\Sigma_{i=j \text{ to } \infty}\,(1 + r)^{-i/4}))$$

Now take the sum of the infinite geometric series.

$$\begin{aligned}&= FITIq\,\Sigma_j\,N\langle j\rangle(1 + r)^{-j/4}/(1 - (1 + r)^{-.25})\\ &= FITI\,\Sigma_j\,N\langle j\rangle q(1 + r)^{-j/4}/q(1 + r)^{-.25}\\ &= FITI\,\Sigma_j\,N\langle j\rangle(1 + r)^{-(j-1)/4}\end{aligned}$$

But we have assumed for convenience that the $N\langle 1\rangle$ term has its present value given by a discount factor of 1. Each subsequent term of $N\langle j\rangle$ has an additional factor of $(1 + r)^{1/4}$ in order to get its present value, since it is one quarter later. Thus,

$$= FITI\ \Sigma_j\ \mathrm{PV}(N\langle j\rangle) = FITI\ (I' - O'I/O)$$

What we use in model three is

$$FITI\ (P^{*\prime} - L' - E' - T^{*\prime} - y(P^{*} - L' - E' - T^{*\prime}))$$
$$= FITI(I' - O' - y(I - O))$$
$$= FITI(I' - O' - (O'/O)(I - O)) = FITI(I' - O'I/O)$$

This is the same result as we got using the *IBT* method here.

AN ANALYSIS OF EXPERIENCE RATING

GLENN G. MEYERS

Abstract

Experience rating formulas that are currently in use have features that have no counterpart in the literature on Bayesian credibility. These features include the limiting of individual losses that go into the experience rating, separate treatment of primary and excess losses, and the gradual transition to self-rating. This paper analyzes the effect of these features using the collective risk model.

Most developments in Bayesian credibility assume that the variance of an individual insured's experience is inversely proportional to the size of the insured. This will not be the case if the parameters of the insured's loss distribution are changing over time. This paper analyzes the effect of this parameter uncertainty on the Bayesian credibility formulas.

Finally, Paul Dorweiler's method of testing experience rating formulas is updated using modern statistical methodology. The result is a very general method of evaluating the parameters of an experience rating formula.

1. INTRODUCTION

The passage of open competition laws for Workers' Compensation has indeed sparked a high degree of competition. Much of the competition is taking place on the individual insured level in the form of schedule and experience rating. In this new competitive environment the performance of these rating plans becomes crucial. The purpose of this paper is to examine the performance of some experience rating plans that are currently being used.

The predominant experience rating plan for Workers' Compensation is promulgated by the National Council On Compensation Insurance (NCCI). This plan is widely adhered to. In addition, the National Council performs the service of maintaining the experience and calculating the experience modification for each insured. These services relieve the insurance companies of considerable administrative expense.

For lines other than Workers' Compensation, an experience rating plan is promulgated by the Insurance Services Office (ISO). Variations from this plan by individual insurance companies are common. Also, ISO does not maintain experience for individual insureds. Getting reliable experience for new insureds is a real problem.

When designing experience rating plans, there are some administrative considerations that cannot be overlooked. The first is that experience ratings are done frequently and so simplicity is of paramount importance

A second consideration is that experience rating, as opposed to class rating, is very visible to the individual insured. A consequence of this is that the experience rating plans must give due consideration to what the insured perceives to be fair. Historically, see Snader [1], these considerations have included the following:

1. A single claim should change the experience modification by no more than a predetermined amount. This predetermined amount is known as the swing of the experience rating plan.
2. All insureds above a certain predetermined size are self-rated, that is they are rated entirely on the basis of their own experience.

In addition to the administrative considerations mentioned above, there are some mathematical considerations that should be made. The mathematical foundations of experience rating come from Bayesian estimation and credibility theory. As is the case with many other mathematical theories, a simplified mathematical model is proposed, and the optimal method of rating the insured is derived. The success of Bayesian estimation and credibility theory depend upon how closely the model represents reality.

The experience rating formulas derived from administrative considerations, hereafter referred to as "practical" formulas, may be different from those derived from the mathematical considerations, hereafter referred to as "theoretical" formulas. This paper investigates the compatibility of these two kinds of rating formulas. We would judge the formulas to be compatible if the accuracy of the "practical" formula is near that of the "theoretical" formula on the simplified models. While it is by no means certain that accuracy on simplified models implies accuracy in real life situations, inaccuracy on a simplified model should imply that something is wrong with the formula being tested.

Our first goal is to find "practical" formulas that perform well on simplified models. These formulas will depend upon unknown parameters which must be estimated from data. Our second goal is to show how these unknown parameters can be estimated. An example will be provided.

2. CURRENT EXPERIENCE RATING FORMULAS

We begin by briefly describing two experience rating plans that are currently in use. We will concentrate on the structure of the plans. The methods currently being used to derive the parameters of the plan are not really an issue at this time. In what follows, an experience modification will refer to the ratio of the premium after experience rating to the premium before experience rating.

2.1 The Workers' Compensation Experience Rating Plan

The Workers' Compensation Experience Rating Plan [2] has a long and rich history. Its development is described in detail by Perryman [3], Uhthoff [4] and Snader [1]. It is very much a "practical" experience rating plan and it has a strong appeal to common sense.

A feature of this plan is the partitioning of the actual losses into primary losses, denoted by A_p, and excess losses, denoted by A_e. In most states, the primary part, X_p, of a claim of amount X is given by the following formula:

$$X_p = X \qquad \text{if } X \leq 2000$$

$$X_p = \frac{10000 \times X}{X + 8000} \qquad \text{if } X > 2000.$$

The excess part of a claim, X_e, is equal to $X - X_p$. A_p is the total of the primary parts of all claims, and A_e is the total of the excess parts.

Let: E_p = expected primary loss;
E_e = expected excess loss; and
$E = E_p + E_e$.

Then the experience modification, *Mod*, is given by the following formula:

$$Mod = \frac{A_p + W \times A_e + (1 - W) \times E_e + (1 - W) \times K}{E + (1 - W) \times K}.$$

W is equal to zero for E less than some number Q, typically 25,000, and increases linearly to one as E increases to the self rating point S, which is usually around 500,000. K is generally set equal to 20,000.

E_p and E_e are products of expected loss rates and the amount of exposure for the insured. These expected loss rates are in the Workers' Compensation rating manual and are updated whenever there is a rate change.

This formula has some very appealing properties:

1. If $E \leq Q$, the formula simplifies to the following:

$$Mod = \frac{A_p + E_e + K}{E + K} . \qquad \text{(Equation 2.1)}$$

   This simplifies experience rating for small insureds.

2. Since X_p is always less than 10,000, the impact of a single large claim on the modification is limited.

3. The insured is self-rated for $E \geq S$. Also, the transition to self-rating is gradual.

4. It is generally believed that claim frequency rather than claim severity differentiates the good insured from the poor insured. The relatively greater impact of small claims is consistent with this belief.

2.2 The General Liability Experience Rating Plan

The General Liability Experience Rating Plan [5], like the Workers' Compensation plan, is very much a "practical" experience rating plan.

Let: ALR = adjusted actual loss ratio;
$AELR$ = adjusted expected loss ratio; and
Z = credibility factor.

Then the experience modification, *Mod*, is given by the following formula:

$$Mod = 1 + \frac{ALR - AELR}{AELR} \times Z.$$

The term "adjusted" refers to the fact that individual claim amounts are limited before entering the experience rating calculation. This limit increases with premium size. It is chosen so that a single large claim can change the experience modification by no more than .3.

Let: P = premium associated with the loss period; and
K = credibility constant (currently 100,000).

Then the credibility is given by the following formula:

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

If this formula were to apply for all values of P, no insured would ever be self-rated. Since self-rating is desired for very large insureds, the credibility formula changes to a linear function between a selected point, Q, and a selected self-rating point, S.

For $E > Q$:

$$Z = \frac{Q^2 + K \times E}{(Q + K)^2}\,. \qquad \text{(Equation 2.2)}$$

In the current General Liability experience rating plan, $K = 100{,}000$, $Q = 483{,}333$ and $S = 1{,}049{,}654$.

Currently, the premium used is collected basic limits premium. However, this is slated to be revised in 1985. The premium used in the adjusted expected loss ratio will be based on estimated prospective premium and adjusted for inflation and average exposure growth. Ideally, the premium should be based on the actual exposures of the experience period, but administrative considerations led to using estimated prospective premium. It should be noted that the plan contains optional provisions to use actual exposures if they are available.

When comparing the two experience rating plans, it should be noted that the Workers' Compensation plan is mandatory in most states. This includes many open competition states! The National Council can enforce the standards of their plan on all companies. They do this, of course, with the consent of the member companies.

3. MATHEMATICAL MODELS FOR EXPERIENCE RATING

Let X be a random variable which represents the total loss incurred by the insured. Let E be a measure of the size of the insured. E could be either the expected loss for the average insured or the premium of the insured which has been determined by a rating manual. Let $R = X/E$ and $\mu = \mathrm{E}[R]$, where E[ ] denotes expected value. R and μ are called the loss ratio and the expected loss ratio.

Experience rating is based on the premise that the expected loss ratio, μ, is different for each insured in a given classification. To model this, we assume that an insured has a loss ratio distribution, d, which is selected at random from a class of distributions, D. Each distribution d has its own mean, μ, and variance, V^2. Let $M = \mathrm{E}[\mu]$, $\tau^2 = \mathrm{Var}[\mu]$, and $\sigma^2 = \mathrm{E}[\nu^2]$, where these statistics are calculated over all distributions d in D.

This process is described by the following algorithm:

Algorithm 3.1

1. Select the distribution, d, along with μ and ν^2, at random from the class of distributions D.
2. Select the loss ratio, R, at random from the distribution d.

The goal of experience rating is to estimate the expected loss ratio, μ, given the loss ratio, R.

Two solutions to this problem are described by Bühlmann [6]. The first solution is the Bayesian solution:

$$B(R) = \mathrm{E}[\mu|R]$$

Bühlmann shows that the solution is optimal in the sense that

$$\mathrm{E}[(B(R) - \mu)^2]$$

is minimized.

A drawback to the Bayesian solution is that it requires knowledge of all the distributions d in D. The second solution, called the credibility solution, only requires knowledge of the quantities M, τ^2 and σ^2. It can be written in the form:

$$C(R) = Z \times R + (1 - Z) \times M.$$

Z is called the credibility factor. We want to choose Z so that

$$\mathrm{E}[(C(X) - \mu)^2]$$

is minimized. The solution, given by Bühlmann, is

$$Z = \frac{\tau^2}{\tau^2 + \sigma^2} \qquad \text{(Equation 3.1)}$$

Bühlmann goes on to show that the same choice of Z minimizes

$$\mathrm{E}[(C(R) - B(R))^2].$$

Thus the credibility solution can be characterized as the best linear approximation to the Bayesian solution. As Hewitt [7] and Mayerson [8] demonstrate, the Bayesian solution can be linear, and thus the credibility solution is identical to the Bayesian solution. However, Hewitt also gives an example where the Bayesian solution is different from the credibility solution. As we shall see below, the distinction can be important.

We shall use the collective risk model to describe the distribution of the losses. This model describes the total losses of an insured in terms of its claim count and claim severity distributions. This model has been described extensively by Meyers and Schenker [9], Heckman and Meyers [10], and Panjer [11].

Let N and S be random variables denoting the claim count and the claim severity for an insured, respectively. In its simplest form, the collective risk model can be described by the following algorithm:

Algorithm 3.2

1. Select the claim count, N, at random from a Poisson distribution.
2. Do the following N times:
   2.1 Select the claim severity, S, at random.
3. Set the total loss, X, equal to the sum of the claim amounts, S, selected in step 2.1.

Since credibility formulas are applied over a wide range of premium sizes, we need to be concerned with how the quantity σ^2 varies with premium. The usual assumption made is to let σ^2 vary inversely with premium. This is done mathematically by setting $\sigma^2 = \Sigma^2/E$, where Σ^2 is the constant of proportionality.

This assumption agrees with the intuition of many actuaries. One would certainly expect the variance of the loss ratio to decrease as E increases. This assumption can also be justified using collective risk theory. If we assume that the claim count distribution is Poisson for each insured and that the claim severity distribution is the same for all insureds, then it is demonstrated in Appendix A that σ^2 is inversely proportional to E.

Substituting Σ^2/E for σ^2 in Equation 3.1 yields the following expression for the credibility:

$$Z = \frac{E}{E + K} \qquad \text{(Equation 3.2)}$$

where $K = \Sigma^2/\tau^2$.

This formula for credibility is almost universally used in the actuarial literature on Bayesian credibility. An exception to this is in a paper by Robert A. Bailey and LeRoy J. Simon [12]. This exception is important and their demonstration is worth discussing in detail.

Using experience from the Canadian Merit Rating Plan, they were able to calculate empirical credibilities for the experience of a single private passenger car for one, two and three years of experience. These credibilities are given in the following table.

TABLE 3.1

EMPIRICAL CREDIBILITIES

Class	1 Year	2 Years	3 Years
1	.046	.068	.080
2	.045	.060	.068
3	.051	.068	.080
4	.071	.085	.099
5	.038	.050	.059

Let E denote the number of years in the merit rating period. Using the credibilities based on one year, the constant K in the credibility formula $Z = E/(E + K)$ is calculated. The credibilities for two and three years are then calculated using this value of K. The results are in the following table.

TABLE 3.2

DERIVED CREDIBILITIES $Z = E/(E + K)$

Class	K	2 Years	3 Years
1	20.7	.088	.126
2	21.2	.086	.124
3	18.6	.097	.139
4	13.1	.133	.187
5	25.3	.073	.106

We see, as Bailey and Simon observed, that the usual assumptions suggest that credibility should increase roughly in proportion to the number of years in the experience rating period. When comparing Tables 3.1 and 3.2 we see that the empirical credibilities are significantly less than what the usual assumptions would suggest!

Bailey and Simon attribute the failure of the usual assumptions to match the empirical credibilities, in part, to an "individual insured's chance for an accident changes from time to time within a year and from one year to the next." This phenomenon is very similar to that of parameter uncertainty, which is described by Meyers and Schenker [9]. In Appendix A it is demonstrated that the collective risk model with parameter uncertainty implies that σ^2 is of the form $\Sigma^2/E + \beta$, where $\beta > 0$. Substituting $\Sigma^2/E + \beta$ for σ^2 in Equation 3.1 yields the following expression for the credibility:

$$Z = \frac{E}{E \times J + K} \qquad \text{(Equation 3.3)}$$

where $J = 1 + \beta/\tau^2$ and $K = \Sigma^2/\tau^2$.

Using Equation 3.3, it is possible to take the credibilities for one and two years and solve for J and K. One can then calculate the credibility implied for three years. The results of these calculations are in the following table:

TABLE 3.3

DERIVED CREDIBILITIES $Z = E/(E \times J + K)$

Class	J	K	3 Years
1	7.7	14.1	.081
2	11.1	11.1	.068
3	9.8	9.8	.077
4	9.4	4.6	.091
5	13.7	12.6	.056

By comparing the above tables we see that the credibilities derived using Equation 3.3 come much closer to the empirical credibilities than those derived using Equation 3.2.

It should be noted that the maximum credibility obtainable in Equation 3.3 is $1/J$. Recall $J \geq 1$. Low maximum credibilities could be interpreted by saying that the insured is changing over time and that change is of a significant size when compared to differences between insureds.

Besides parameter uncertainty, there are other reasons why the usual assumptions may not be appropriate. The variable loss limit that is in the ISO experience rating plans is one such case. In Appendix A, it is demonstrated that the constant of proportionality, Σ^2, depends upon the second moment of the claim severity distribution. Since the effect of changing the loss limit is to change the claim severity distribution, one should not expect Σ^2 to be the same for all loss limits.

Since the loss limit increases with premium size, we would expect σ^2 to decrease slower than $1/E$ (See Appendix A.) Thus an attempt to impose a credibility formula of the form $Z = E/(E + K)$ will result in credibilities which are too small for the small insureds, and too large for the large insureds.

The formula $\sigma^2 = \Sigma^2/E + \beta$ also has the property that σ^2 decreases slower than $1/E$. Thus Equation 3.3 should provide a better estimate of the credibility. But the derivation of Equation 3.3 did not anticipate an increasing loss limit, and so one should not expect the estimated credibility to be perfect.

4. THE EFFICIENCY OF AN EXPERIENCE RATING PLAN

In the previous section we discussed optimal (for specific assumptions) experience rating plans. There are a number of reasons why an optimal plan might not be used. As discussed above, there may be several practical reasons for using some alternative plan. Another reason is that one must estimate the parameters M, τ^2 and σ^2. Estimation error will occur. The purpose of this section is to present a yardstick for comparing the performances of alternative experience rating plans.

The purpose of experience rating is to estimate the expected loss ratio, μ. If experience rating were not used, our estimate of μ would be M, which would be subject to error. A good measure, with historical precedent, would be to calculate the amount the expected error is reduced by a given experience rating formula.

Let F be an estimator of μ which results from an experience rating formula. F can be a function of any kind of loss experience of the insured such as total losses, claim count or limited losses. We then define the efficiency of F by the expression:

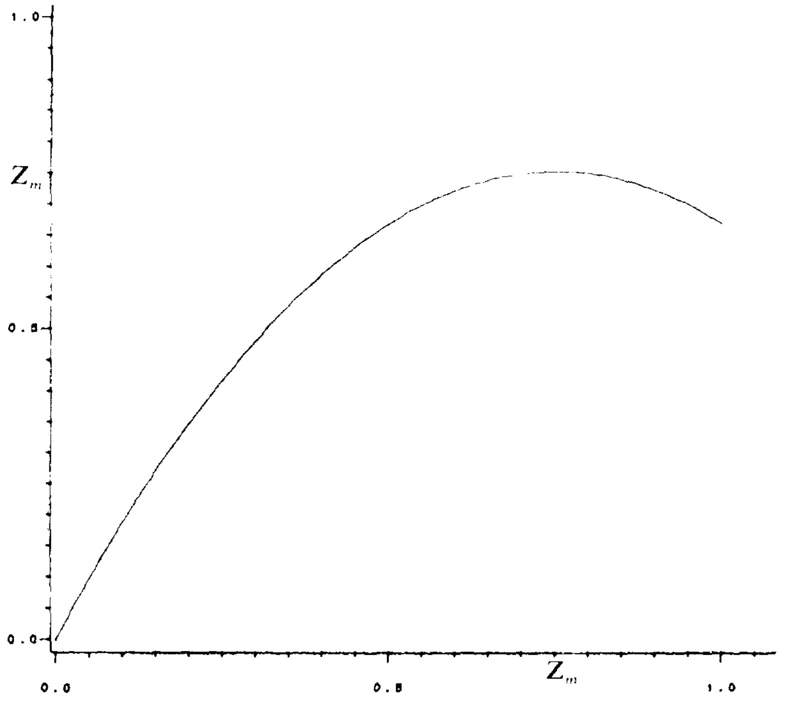
$$\frac{E[(\mu - M)^2] - E[(\mu - F)^2]}{E[(\mu - M)^2]}$$

If F is a perfect estimator for μ, its efficiency will be equal to 1. If $F = M$, its efficiency will be equal to 0. It is possible, as we shall soon see, for the efficiency to be negative for a poorly chosen F.

One should note the similarity of this measure of efficiency with the statistic R^2 that is used in regression analysis. It is different from R^2 in that it does not automatically assume that F was chosen in some optimal manner.

If F is a credibility estimator of the form $Z \times R + (1 - Z) \times M$, it is shown in Appendix B that the efficiency of F is given by the expression $2 \times Z - Z^2/Z_m$, where Z_m is the optimal credibility given by Equation 3.1. A graph of the efficiency as a function of S is shown in Figure 1. This expression has the following properties:

FIGURE 1



1. The efficiency is maximized when $Z = Z_m$. This is Bühlmann's [6] result.

2. As a function of Z, the efficiency starts at 0 when $Z = 0$, raises to a maximum of Z_m when $Z = Z_m$ and falls to 0 when $Z = 2 \times Z_m$. The efficiency is negative for $Z > 2 \times Z_m$.

It is not difficult to see why credibility , even the non-scientific version, has been so popular. If $Z < 2 \times Z_m$ then a credibility estimate using Z will be more accurate than no experience rating. If $Z_m > 0.5$ then any choice of $Z \leq 1$ will guarantee an improvement in accuracy.

It should be noted that Z_m is not the maximum efficiency obtainable by any experience rating formula. As noted above, a Bayesian formula could be more accurate. As we shall soon see, it is also possible that an experience rating formula that uses detailed information such as claim count and claim severity can be even more accurate than the Bayesian formula.

5. THE GENERAL LIABILITY EXPERIENCE RATING PLAN

We now use the concepts developed above to analyze the General Liability experience rating plan. In particular, we will discuss the effect of self-rating and loss limits. Also, credibility and Bayesian estimation will be compared.

Let us suppose, for the sake of discussion, that the credibility formula $Z_m = E/(E + K)$, with $K = 100{,}000$ is the "correct" formula. Now suppose that instead of using Z_m for credibility we use $Z = (Q^2 + K \times E)/(Q + K)^2$, where $Q = 483{,}333$. Then the following table shows the efficiency of the formula for Z.

TABLE 5.1

E	Z	Efficiency of Z	Z_m
500,000	.8335	.8333	.8333
600,000	.8629	.8571	.8571
700,000	.8922	.8748	.8750
800,000	.9216	.8877	.8889
900,000	.9510	.8971	.9000
1,000,000	.9804	.9035	.9091

Examination of this table shows that there is minimal loss of efficiency when using Z instead of Z_m. If one accepts the credibility formula $Z = E/(E + K)$, the gradual shift to self-rating should also be acceptable.

We now turn to loss limits. The collective risk model will be used to describe the loss distributions. The mathematics will be less cumbersome if there is a finite number of loss amounts. For this reason, the claim count distribution will be binomial with N trials and the probability of a claim equal to p. The claim severity distribution will be a discrete version of the shifted Pareto, which is used to describe claim severity in many lines of casualty insurance. The probability, $F(x)$, that a claim will be less than or equal to x is given by:

$$F(x) = 1 - (b/(x + b))^q \quad x = 1, 2, \ldots, 49.$$

The remaining probability will be at the basic limit, 50.

The parameter q will be set equal to 1.25 for all prior distributions. The parameter b, in the claim severity distribution and p, in the claim count distribution may be different for each prior. N will reflect the size of the insured.

For a selected loss limit, L, the total losses can vary anywhere from 0 to $L \times N$. Using Panjer's algorithm [11] one can calculate the probability of each total loss for each prior distribution. One then calculates credibility and Bayes estimates of the experience modification for basic limits losses, as well as the efficiency of each estimate. Detailed calculations are given for one case in Exhibit 5.1. Efficiencies for several cases are given in Tables 5.2–5.4.

We first consider the case where only the claim count distributions vary. The efficiency is at a maximum for both the credibility estimator and the Bayes estimator when the loss limit is equal to 1, and decreases as the loss limit increases. This should come as no surprise. When the loss limit is 1, there is no random element due to claim severity. As we increase the loss limit, we increase the randomness in our measurements.

As expected, the Bayes estimator is more accurate than the credibility estimator. It is worth noting that the Bayes estimator is less affected by the increasing loss limit. The accuracy of the credibility approximation to the Bayes estimator gets worse as the loss limit increases.

We now turn to the case where only the claim severity distributions are varying. In this case, information about the distribution, as well as the random element, increases as the loss limit increases. The efficiency of both the credibility and the Bayes estimators is near maximum at a loss limit of 8. After that point the increase in efficiency is, at best, marginal. In fact it can decrease.

When both the claim severity and claim count distributions vary, the efficiency first increases and then decreases as the loss limit increases. The best loss limit is 4 for this example.

Attempting to draw conclusions about real life experience rating plans from models can be a risky undertaking. But accuracy is important, and not attempting to draw conclusions can also be risky. With this in mind, we proceed.

The first conclusion is that limiting the loss for an individual claim is a good idea. A well chosen loss limit will be large enough to capture differences in claim severity distributions. If the loss limit is too large, increased randomness will wipe out any extra information gained by the higher loss limit. This has been the traditional argument in favor of loss limits. It is gratifying to see it verified on a mathematical model.

While the Bayes estimator is more accurate, in practice we do not have enough information to use it. An alternative is to create conditions where the credibility estimator is a good approximation to the Bayes estimator. A loss limit serves this purpose.

The negative effects of high loss limits appear to be less pronounced for larger insureds. Perhaps this could be taken as justification for varying the loss limit. However one should not raise the loss limit indefinitely. Once the loss limit reaches a sufficient level to capture enough information on the claim severity, it should go no higher.

EXHIBIT 5.1

Credibility and Bayes Estimates

$N = 4$ Loss limit = 4

Prior#	Weight	p	b	q	Limited Sev. Mean	Limited Std. Dev.	Basic Limits Sev. Mean
1	0.25	0.20	0.25	1.25	1.24	0.69	1.48
2	0.25	0.30	0.50	1.25	1.47	0.93	2.03
3	0.25	0.40	0.75	1.25	1.68	1.07	2.57
4	0.25	0.50	1.00	1.25	1.85	1.16	3.09

Aggregate Probabilities

X	Prior#1	Prior#2	Prior#3	Prior#4
0	0.40960000	0.24010000	0.12960000	0.06250000
1	0.35481700	0.30735000	0.22576000	0.14488800
2	0.14376800	0.19673800	0.19920200	0.16774800
3	0.04484430	0.09761280	0.13228600	0.14045300
4	0.02853980	0.07571090	0.11665700	0.13708900
5	0.01326000	0.04745030	0.09041290	0.12598400
6	0.00364740	0.02021520	0.05102120	0.08841880
7	0.00090941	0.00771968	0.02505340	0.05248160
8	0.00044414	0.00442423	0.01612450	0.03707380
9	0.00013752	0.00186283	0.00844632	0.02315500
10	0.00002474	0.00053879	0.00323913	0.01092810
11	0.00000535	0.00016668	0.00122928	0.00481558
12	0.00000245	0.00008447	0.00068151	0.00287956
13	0.00000042	0.00002078	0.00021562	0.00111993
14	0.00000004	0.00000341	0.00004769	0.00029946
15	0.00000001	0.00000101	0.00001552	0.00010565
16	0.00000000	0.00000048	0.00000819	0.00006104

Limited Grand Mean = 2.28643 Credibility = .216312

EXHIBIT 5.1 (continued)

X	Prob(X)	Credibility *Mod*	Bayes *Mod*	Difference
0	0.21045000	0.7837	0.7325	0.0512
1	0.25820400	0.8783	0.8629	0.0154
2	0.17686400	0.9729	1.0177	−0.0448
3	0.10379900	1.0675	1.1504	−0.0829
4	0.08949920	1.1621	1.2006	−0.0385
5	0.06927680	1.2567	1.2722	−0.0155
6	0.04082570	1.3513	1.3487	0.0026
7	0.02154100	1.4459	1.4012	0.0447
8	0.01451670	1.5405	1.4218	0.1188
9	0.00840042	1.6352	1.4552	0.1799
10	0.00368269	1.7298	1.4880	0.2417
11	0.00155422	1.8244	1.5071	0.3173
12	0.00091200	1.9190	1.5152	0.4037
13	0.00033919	2.0136	1.5347	0.4789
14	0.00008765	2.1082	1.5498	0.5584
15	0.00003055	2.2028	1.5551	0.6477
16	0.00001743	2.2974	1.5605	0.7369

Expected Error: Bayes = .226192 Credibility = .230147

Efficiency: Bayes = .222546 Credibility = .208954

TABLE 5.2

COUNT DISTRIBUTIONS VARY

Prior	p	b	q
#1	0.2	0.75	1.25
#2	0.3	0.75	1.25
#3	0.4	0.75	1.25
#4	0.5	0.75	1.25

Loss Limit	Credibility			Bayes		
	$N = 4$	$N = 8$	$N = 16$	$N = 4$	$N = 8$	$N = 16$
1	.189	.317	.482	.190	.323	.496
4	.123	.218	.359	.152	.237	.377
8	.087	.160	.275	.147	.214	.309
12	.069	.130	.230	.146	.210	.288
16	.059	.112	.201	.145	.209	.281

TABLE 5.3

Severity Distributions Vary

Prior	p	b	q
#1	0.4	0.25	1.25
#2	0.4	0.50	1.25
#3	0.4	0.75	1.25
#4	0.4	1.00	1.25

Loss Limit	Credibility			Bayes		
	$N = 4$	$N = 8$	$N = 16$	$N = 4$	$N = 8$	$N = 16$
4	.026	.051	.096	.038	.059	.101
8	.035	.068	.127	.046	.077	.134
12	.035	.068	.127	.048	.079	.137
16	.034	.065	.122	.048	.080	.137

TABLE 5.4

COUNT AND SEVERITY DISTRIBUTIONS VARY

Prior	p	b	q
#1	0.2	0.25	1.25
#2	0.3	0.50	1.25
#3	0.4	0.75	1.25
#4	0.5	1.00	1.25

Loss Limit	Credibility			Bayes		
	$N = 4$	$N = 8$	$N = 16$	$N = 4$	$N = 8$	$N = 16$
1	.189	.317	.482	.190	.323	.496
4	.209	.344	.507	.223	.365	.536
8	.178	.301	.461	.223	.354	.510
12	.154	.267	.421	.223	.351	.495
16	.138	.242	.389	.223	.350	.488

6. THE WORKERS' COMPENSATION EXPERIENCE RATING PLAN

It was demonstrated in the last section that a loss limit can increase the accuracy of an experience rating plan. However, the Workers' Compensation Experience Rating Plan gradually introduces excess losses as the size of the insured increases. We now analyze this treatment of excess losses using the-collective risk model.

We shall use the Weibull distribution to model claim severity. The probability, $F(x)$, that a claim will be less than or equal to x is given by:

$$F(x) = 1 - e^{-(x/b)^c}.$$

The Poisson distribution will be used to model claim count. The probability of n claims, $P(n)$, is given by:

$$P(n) = e^{-\lambda} \times \lambda^n/n!$$

The parameter c will be set equal to .25 for all prior distributions. The parameter b for the claim severity distribution and the parameter λ for the claim count distribution will be independently chosen at random from the following table. Each parameter value is equally likely to be chosen.

TABLE 6.1

b	λ
30	40
40	70
50	100
60	130
70	160

It was necessary to resort to Monte Carlo methods in order to properly treat primary and excess losses. The following algorithm was repeated 10,000 times.

Algorithm 6.1

1. Select the Poisson parameter, λ, at random from Table 6.1.
2. Select the number of claims, n, at random from a Poisson distribution with parameter λ.

3. Select the Weibull parameter, b, at random from Table 6.1.
4. Do the following n times.
   4.1 Select a claim value, x, at random from a Weibull distribution with parameter b and $c(=.25)$.
   4.2 From x, calculate the primary loss, x_p, and the excess loss, x_e.
5. A_p is the sum of all the x_p's and A_e is the sum of all the x_e's.

It can be demonstrated by numerical integration of the severity distribution that $E_p = 48{,}000$ and $E_e = 72{,}000$.

In addition to the standard Workers' Compensation experience modification formula, we want to consider a modification formula in which the excess losses are ignored. This formula will take the following form:

$$M = \frac{A_p + K}{E_p + K} \qquad \text{(Equation 6.1)}$$

One should note the difference between this formula and formula 2.1. Using Hewitt's formulas [7], it can be demonstrated that the optimal value for K in this formula is 22,900.

For each trial in the simulation it is possible to calculate the modification for various formulas involving primary and excess losses. By comparing the calculated modification with the "true" modification one can estimate the efficiency of each formula. The results are in the following table.

TABLE 6.2

EFFICIENCY

	W	$K = 18{,}000$	$K = 23{,}000$	$K = 28{,}000$
Formula 6.1		0.68	0.68	0.66
Standard Formula	0.0	0.48	0.46	0.45
" "	0.1	0.51	0.50	0.49
" "	0.2	0.50	0.50	0.49
" "	0.3	0.44	0.44	0.44
" "	0.4	0.32	0.33	0.34
" "	0.5	0.13	0.15	0.17
" "	0.6	−0.12	−0.09	−0.07

Formula 6.1 is a clear winner in this case. There are two possible reasons for this. First, as demonstrated in the previous section, the primary losses seem to capture most of the information about the severity distribution. Second, the structure of the experience rating formula may very well be wrong! The Bayesian and credibility formulas described above are optimal under certain specified conditions. This author does not know of any conditions where the standard formula is optimal. At the very least, a proposal to retain the present formula should include a plausible model in which the present formula outperforms the competing formulas.

7. CHOOSING AN EXPERIENCE RATING FORMULA

So far, we have seen how modeling can give some good hints for the right form of an experience rating formula. Since we rarely, if ever, have the distributional information to do a pure Bayesian analysis, it appears that a good choice of an experience rating formula would be a credibility formula. The credibility could be given by either Equation 3.2 or Equation 3.3. A loss limit of some kind should definitely be used.

As of this writing, there is no nice clean way to pick an optimal loss limit. This author has had good luck with the Weibull distribution for severity in Workers' Compensation and the shifted Pareto distribution, see Patrik [13], for the severity in other lines of insurance. By trial and error on various models, as was done in the previous sections, one might come up with a reasonable loss limit, or loss limit formula. There is room for improvement here.

Once a loss limit has been selected one then gathers the limited losses and the expected limited losses for individual insureds over a period of years. This information is absolutely essential. Experience rating depends upon how well the experience of one year predicts that of another. With data such as this one can use the empirical Bayesian credibility procedure as described originally by Bühlmann and Straub [14], and later by the ISO Credibility Subcommittee [15] and Meyers [16].

A problem with these procedures is that they all assume that the σ^2 is inversely proportional to the expected losses, which results in using Equation 3.2 for the credibility. While these procedures might well be modified to handle more general assumptions about σ^2, the author would like to propose a different approach. This approach has the advantage that: (1) it is easy to modify the parameter estimation to accommodate alternative assumptions about σ^2; and (2) one can test the assumption made about σ^2. This approach has its origins in a study done by Paul Dorweiler [17].

In what follows we shall take the term "loss ratio" to mean current losses divided by the modified premium, where the experience modification is calculated from prior years' loss experience. We assume that the expected losses used in the experience rating formula are correct. If the loss ratio is positively correlated with the experience modification, then the credibility factors used are too low. Conversely, if the loss ratio is negatively correlated with the experience modification, then the credibility factors used are too high.

This can be justified by the following. Suppose an insured had a low experience modification and tends to have a lower than average loss ratio. Then to raise his loss ratio, one can give the insured a lower experience modification by giving more credibility to the experience. A similar argument applies when the insured tends to have a higher than average loss ratio.

Dorweiler tested the performance of an experience rating plan by partitioning insureds by manual premium size and modification size. For each premium size group he calculated the trend in loss ratio as the modification increased. The idea was to compare the number of times a positive trend occurred with the number of times a negative trend occurred. This method of testing credibility formulas is very general. No assumptions about the nature of the experience rating formula are required.

During the past fifty years, our understanding of statistics has vastly improved. Our computing capability today was unthinkable in Dorweiler's time. Today, Dorweiler's method might well be similar to the following.

Assume we have the correct form of the experience rating formula and we want to know if we have selected the right parameters. That is, we want to test the hypothesis

H_0: The parameters of the experience rating formula are correct

against the alternative hypothesis

H_1: At least one of the parameters of the experience rating formula is incorrect.

To test this hypothesis, we proceed as follows.

1. Partition the insureds into groups with similar modified premium size. Modified premium is used rather than manual premium because we want all insureds in the group to have the same loss ratio distribution. It is felt that expected losses rather than exposure is a better indicator of the loss ratio distribution.

2. Calculate the correlation coefficient between the loss ratio and the experience modification for each group.

   Kendall's τ, see Conover [18], is the preferred measure in this case. This correlation coefficient compares the number of pairwise increases with the number of pairwise decreases. Let τ_i denote Kendall's correlation coefficient for group i and let n_i be the number of insureds in group i. Under the null hypothesis, $\tau_i = 0$ for each group, the distribution of τ_i is approximately normal with mean 0 and variance $n_i(n_i - 1)(2n_i + 5)/18$. This is a nonparametric result.

3. Calculate the normalized correlation coefficient, for each group, and a combined normalized correlation coefficient. These terms are defined as follows.

   For each group i, set T_i equal to τ_i divided by its standard deviation. Under the null hypothesis T_i is approximately normal with mean 0 and variance 1. We call T_i the normalized correlation coefficient for group i. Let m be the number of groups. Set T equal to the sum of all the T_i's divided by the square root of m. T also has mean 0 and variance 1 under the null hypothesis. We will call T the combined normalized correlation coefficient.

4. Reject H_0 at significance level α if the percentile of T is outside the interval $(\alpha/2, 1 - \alpha/2)$. The percentile of T can be determined from the standard normal distribution.

   By noting that the confidence region of the parameters is the set of all parameters for which one fails to reject H_0, one can find a confidence region of the parameters by testing several sets of parameters. Acceptable parameters are those for which the percentile of T falls within the interval $(\alpha/2, 1 - \alpha/2)$. A best estimate of the parameters is one for which the percentile of T is equal to .5.

Let's see how this test works on live data. During the late seventies, the Individual Risk Rating Plan Committee at ISO issued a special call for individual insured data from actual experience ratings. ISO supplied the author with the following data elements from this call. For each of three years there was given the basic limits premium and the basic limits losses (adjusted for the loss limit). In addition, the adjusted expected loss ratio (*AELR*) was given. Ideally, one would like to have the losses that resulted from the policy that was actually experience rated, but we did the next best thing. The first two years of data were used to predict the third year.

Before doing the analysis, two adjustments to the ISO data were made. First, all insureds which did not have a full three years of experience were deleted. Second, the AELR was adjusted so that the total expected losses were equal to the total actual losses for the first two years. In all, there were 1,980 insureds which form 33 groups of 60 insureds.

Let us first assume that the credibility formula given by $Z = P/(P + K)$ is correct. Hypothesis tests were performed for a set of K values, with the following results.

TABLE 7.1

K	T	Percentile
16,000	−2.9442	.0016
18,000	−2.1870	.0144
20,000	−1.1476	.1256
22,000	−0.2060	.4184
24,000	0.2449	.5967
26,000	0.5558	.7108
28,000	1.0665	.8569
30,000	1.6194	.9473
32,000	2.1078	.9825
34,000	2.5697	.9949
36,000	2.9850	.9986

The best estimate for K will be between 22,000 and 24,000. The 95 percent confidence interval for K will range from slightly over 18,000 to slightly less than 32,000. Table 7.2 shows the T_i's for each group when $K = 22{,}000$.

Close examination of Table 7.2 reveals that the correlations are predominantly positive for the smaller insureds and very definitely negative for the larger insureds. This indicates that the credibility is too low for the smaller insureds and too high for the larger insureds. Thus the formula $Z = P/(P + K)$ is not the correct form of the credibility formula. This can be explained in terms of the changing loss limit and parameter uncertainty as described in Section 3 above. If we have the correct form of the credibility formula, the hypothesis test described above should apply equally well for any subset of groups.

Let us now examine the credibility formula $Z = P/(P \times J + K)$. In addition to calculating the combined normalized correlation coefficient for all insureds, we calculate the combined normalized correlation coefficient for the five different subsets of groups. The rationale for selecting the subsets will be discussed below.

Before discussing the above tables one should note that there are some small reversals in what might seem to be a clear pattern. These are random fluctuations caused by the insureds shifting groups with each set of parameters. Recall that the groups were based on modified premium.

Let us first examine the subsets consisting of Groups 1 to 5, Groups 6 to 19 and Groups 20 to 33. It can be observed that when $J = 1.0$, no value of K is in the 95 percent confidence region for each subset. The following pairs (J,K) are in the 95 percent confidence region for each subset.

		(4.0, 1000)
		(4.0, 2000)
		(4.0, 3000)
		(4.0, 4000)
	(3.0, 5000)	(4.0, 5000)
	(3.0, 6000)	(4.0, 6000)
	(3.0, 7000)	(4.0, 7000)
	(3.0, 8000)	(4.0, 8000)
	(3.0, 9000)	(4.0, 9000)
	(3.0, 10000)	
	(3.0, 12000)	
(2.0, 14000)	(3.0, 14000)	
(2.0, 16000)		

The details of the calculations for $J = 4.0$ and K = 2,000 are given in Table 7.3. As mentioned above, the derivation of the credibility formula $Z = P/(P \times J + K)$ does not anticipate a loss limit which increases as the size of the insured increases. Thus we should not expect this credibility formula to be exactly right

TABLE 7.2

Headings

MINMBLP —Minimum modified basic limits premium
MAXMBLP —Maximum modified basic limits premium
N —Number of insureds
TAU —Kendall's tau correlation coefficient between the loss ratio and the experience modification
MODPCT10 —10th percentile of experience modifications
MODPCT50 —50th percentile of experience modifications
MODPCT90 —90th percentile of experience modifications
T —Normalized correlation coefficient

EXPERIENCE RATING ANALYSIS—GENERAL LIABILITY: $K = 22{,}000$ $J = 1.00$

OBS	MINMBLP	MAXMBLP	N	TAU	MODPCT10	MODPCT50	MODPCT90	T
1	10.0	250	60	0.10031	0.962409	0.98257	0.99638	1.1324
2	251.5	418	60	−0.01243	0.962809	0.97386	0.98161	−0.1403
3	422.4	604	60	0.10056	0.929858	0.95987	1.01518	1.1353
4	606.5	759	60	0.02147	0.908815	0.94638	0.97613	0.2424
5	762.8	884	60	0.01469	0.892809	0.93845	0.97723	0.1658
6	887.1	1032	60	−0.05537	0.879143	0.93260	1.01479	−0.6250
7	1036.8	1151	60	−0.04181	0.864093	0.92675	1.01257	−0.4720
8	1151.7	1279	60	0.10345	0.862320	0.92053	1.18874	1.1678
9	1286.5	1447	60	0.02318	0.824655	0.91998	1.10541	0.2616
10	1452.0	1603	60	−0.02373	0.839008	0.90577	1.06497	−0.2679
11	1609.1	1747	60	−0.05537	0.845583	0.90363	1.15102	−0.6250

TABLE 7.2 (continued)

OBS	MINMBLP	MAXMBLP	N	TAU	MODPCT10	MODPCT50	MODPCT90	T
12	1747.7	1910	60	0.20339	0.818599	0.89558	1.00920	2.2961
13	1913.8	2020	60	0.02938	0.803793	0.89252	1.17428	0.3317
14	2024.6	2169	60	0.14237	0.798667	0.90054	1.20914	1.6072
15	2169.8	2316	60	−0.02712	0.785614	0.86615	1.01141	−0.3061
16	2318.4	2495	60	0.14463	0.786776	0.86784	1.16782	1.6327
17	2498.1	2680	60	0.08475	0.743469	0.86263	1.23809	0.9567
18	2681.3	2885	60	0.08927	0.766698	0.85704	1.06101	1.0077
19	2885.0	3064	60	−0.00452	0.743238	0.84386	1.18246	−0.0510
20	3067.4	3346	60	0.07006	0.723105	0.84160	1.20551	0.7909
21	3352.6	3629	60	0.25085	0.717498	0.85007	1.24072	2.8318
22	3632.1	3880	60	0.02147	0.667270	0.85448	1.33079	0.2424
23	3883.3	4209	60	0.10734	0.713127	0.81707	1.26421	1.2118
24	4215.7	4580	60	0.01243	0.675236	0.80978	1.91797	0.1403
25	4581.6	5023	60	−0.04859	0.684537	0.82988	1.90822	−0.5485
26	5040.9	5529	60	0.05311	0.634504	0.77695	1.28748	0.5995
27	5533.5	6302	60	−0.08701	0.590834	0.83498	1.91602	−0.9822
28	6316.8	7390	60	−0.08023	0.641002	0.86242	1.95935	−0.9057
29	7405.5	8645	60	−0.35593	0.522315	0.80239	1.89673	−4.0181
30	8702.0	10808	60	−0.19955	0.468991	0.78968	1.97431	−2.2527
31	10847.9	15885	60	−0.20000	0.447977	1.06263	2.48548	−2.2578
32	16077.5	26102	60	−0.21695	0.622667	1.36827	2.34570	−2.4491
33	26118.2	297046	60	−0.26893	0.652677	1.52027	2.96620	−3.0359

TABLE 7.3

<u>Headings</u>

MINMBLP —Minimum modified basic limits premium
MAXMBLP —Maximum modified basic limits premium
N —Number of insureds
TAU —Kendall's tau correlation coefficient between the loss ratio and the experience modification
MODPCT10 —10^{th} percentile of experience modifications
MODPCT50 —50^{th} percentile of experience modifications
MODPCT90 —90^{th} percentile of experience modifications
T —Normalized correlation coefficient

EXPERIENCE RATING ANALYSIS—GENERAL LIABILITY: K = 2000 J = 4.00

OBS	MINMBLP	MAXMBLP	N	TAU	MODPCT10	MODPCT50	MODPCT90	T
1	9.9	230	60	0.07764	0.870207	0.90389	0.96442	0.8765
2	232.9	382	60	0.07910	0.862833	0.88592	0.90232	0.8929
3	389.4	572	60	0.01582	0.834214	0.86588	0.91637	0.1786
4	572.3	705	60	−0.07910	0.822935	0.85410	0.96478	−0.8929
5	709.2	821	60	−0.11073	0.812158	0.84499	0.89705	−1.2501
6	823.7	961	60	−0.26328	0.808248	0.83594	0.88607	−2.9721
7	971.3	1085	60	−0.02712	0.810787	0.83640	0.98953	−0.3061
8	1086.9	1232	60	−0.05483	0.810880	0.84006	1.03652	−0.6190
9	1243.2	1390	60	−0.11815	0.794511	0.83726	1.19594	−1.3337
10	1394.3	1553	60	−0.07910	0.799560	0.83418	1.06445	−0.8929
11	1555.2	1701	60	0.03277	0.803051	0.83198	1.09454	0.3699

TABLE 7.3 (continued)

OBS	MINMBLP	MAXMBLP	N	TAU	MODPCT10	MODPCT50	MODPCT90	T
12	1701.3	1830	60	−0.00339	0.799704	0.83866	1.35855	−0.0383
13	1833.6	1996	60	0.06893	0.795396	0.83288	1.33904	0.7781
14	1996.4	2142	60	0.11186	0.794848	0.82466	1.02667	1.2628
15	2146.2	2326	60	0.07797	0.788719	0.82349	1.01559	0.8802
16	2340.1	2518	60	0.10508	0.796236	0.82021	1.20622	1.1863
17	2518.6	2717	60	0.23616	0.792849	0.82674	1.27064	2.6660
18	2719.5	2924	60	−0.02486	0.787838	0.80931	1.04337	−0.2806
19	2927.5	3174	60	0.07232	0.785318	0.81290	1.21608	0.8164
20	3174.8	3446	60	0.13672	0.789134	0.82195	1.80638	1.5435
21	3450.3	3767	60	−0.06441	0.798985	0.85378	1.35756	−0.7271
22	3782.2	4143	60	0.15254	0.789817	0.90983	1.81165	1.7220
23	4147.5	4535	60	0.19096	0.782837	0.82072	1.32675	2.1557
24	4541.2	5042	60	−0.20791	0.790398	0.89006	2.10844	−2.3471
25	5043.4	5426	60	0.00452	0.783358	0.82616	1.84692	0.0510
26	5439.8	6073	60	0.05198	0.791312	0.84412	1.74888	0.5868
27	6084.6	6997	60	0.00339	0.788521	0.87782	2.34622	0.0383
28	7011.6	8064	60	−0.00339	0.785978	0.88043	2.35557	−0.0383
29	8115.5	9887	60	−0.11751	0.791988	0.98504	2.35900	−1.3266
30	9899.0	12103	60	−0.17467	0.783483	0.90444	2.02998	−1.9719
31	12344.9	17413	60	−0.04746	0.779974	0.90431	2.50209	−0.5357
32	17495.2	26177	60	−0.03164	0.810414	1.12274	1.73464	−0.3572
33	26221.3	206941	60	−0.00565	0.808315	1.05283	2.17655	−0.0638

TABLE 7.4

PERCENTILES OF T'S

$J = 1.0$

K	Groups 1:33	Groups 6:33	Groups 1:5	Groups 6:19	Groups 20:33
18000	.0144	.0051	.6795	.8646	.0000
20000	.1256	.0661	.7306	.9220	.0002
22000	.4184	.2411	.8716	.9677	.0022
24000	.5967	.3959	.8950	.9864	.0049
26000	.7108	.5197	.9050	.9875	.0149
28000	.8569	.7050	.9285	.9938	.0412
30000	.9473	.8665	.9374	.9976	.1063
32000	.9824	.9498	.9367	.9991	.2159
34000	.9949	.9795	.9609	.9994	.3594
36000	.9986	.9933	.9655	.9997	.5345

$J = 2.0$

K	Groups 1:33	Groups 6:33	Groups 1:5	Groups 6:19	Groups 20:33
8000	.0118	.0132	.2875	.6563	.0002
10000	.0843	.0863	.3785	.8020	.0027
12000	.2388	.2336	.4586	.8977	.0108
14000	.5767	.5186	.6505	.9662	.0391
16000	.7900	.7016	.7940	.9604	.1568
18000	.8964	.8358	.8235	.9809	.2449
20000	.9727	.9410	.8919	.9912	.4356
22000	.9893	.9754	.8951	.9954	.5696
24000	.9979	.9944	.9161	.9980	.7624

TABLE 7.4 (continued)

PERCENTILES OF T'S

$J = 3.0$

K	Groups 1:33	Groups 6:33	Groups 1:5	Groups 6:19	Groups 20:33
3000	.0136	.0152	.2913	.2652	.0075
4000	.0535	.0587	.3316	.4202	.0220
5000	.1626	.1677	.4024	.5796	.0590
6000	.3080	.2940	.4972	.7576	.0715
7000	.4104	.3837	.5471	.7854	.1134
8000	.5645	.5234	.6096	.9115	.1026
9000	.7601	.7388	.6183	.9435	.2481
10000	.8223	.8088	.6205	.9469	.3519
12000	.9430	.9229	.7546	.9597	.6054
14000	.9827	.9715	.8235	.9750	.7677
16000	.9963	.9935	.8421	.9931	.8535

$J = 4.0$

K	Groups 1:33	Groups 6:33	Groups 1:5	Groups 6:19	Groups 20:33
1000	.5953	.4443	.8293	.6717	.2602
2000	.5035	.5185	.4653	.6574	.3672
3000	.6111	.6415	.4471	.7501	.4356
4000	.7106	.7164	.5289	.8901	.3380
5000	.8044	.8108	.5472	.8704	.5467
6000	.8676	.8405	.6938	.7904	.7262
7000	.9215	.9317	.5450	.8689	.8373
8000	.9596	.9592	.6421	.9087	.8707
9000	.9878	.9886	.6526	.9732	.9014
10000	.9941	.9929	.7510	.9771	.9288

over the entire range of premium sizes. Examination of Table 7.3 reveals this to be the case. However, the results are superior to any that could be obtained with the credibility formula $Z = P/(P + K)$.

If loss limit increases with the size of the insured the credibility will increase more slowly than the formula $Z = P/(P \times J + K)$ would suggest. This is verified in Table 7.4 where the formula tends to assign too low a credibility to the medium size insureds in Groups 6 to 19 and too high a credibility to the large sized insureds in Groups 20 to 33. However, the formula tends to assign too high a credibility to the smaller sized insureds in Groups 1 to 5, and too low a credibility to the medium sized insureds in Groups 6 to 19. This is the opposite of what is expected.

We attempt to explain this reversal. We first note that there is a minimum premium size that qualifies an insured for experience rating. It is possible for an insured to have a sizeable decrease in exposure which will result in premiums which are below the minimum in the year being rated. But this happens rather infrequently. A far more common cause of insureds having a smaller size is for the insured to have low modification. This can be verified in Table 7.3 where over ninety percent of the insureds in Groups 1 to 5 have experience modifications which are less than 1.00.

Two possible explanations for the reversal can be given. First, since most insureds are those with good experience, the groups are more homogeneous (i.e. τ^2 is lower) and lower credibility is called for. Second, since the loss limit is assigned according to unmodified premium, σ^2 is not necessarily smaller for the smaller insureds. This would also have a tendency to lower credibility for the smaller insureds. If these explanations are correct, one should separate the very smallest insureds from the main part of the analysis. This is why the subsets were grouped in the above manner.

In summary, a very general way of analyzing data for experience rating has been proposed. Not only can it be used to determine parameters of an experience rating formula, but one can also test to see if the assumptions made in deriving the form of the credibility formula are valid.

8. SUMMARY AND CONCLUSIONS

This paper has attempted to use collective risk theory to analyze experience rating. Particular attention was paid to experience rating formulas used by the National Council on Compensation Insurance and Insurance Services Office. The first goal of this paper was to find an experience rating formula that worked well on mathematical models and would be easy to administer. An examination of the performance of experience rating plans on mathematical models led to the following conclusions.

1. A loss limit can be an effective tool for increasing the accuracy of an experience rating formula. Loss limits are particularly helpful when there are differences in claim frequency. Even if the only differences among the insureds are in claim severity, little accuracy will be lost with a loss limit.

2. The current formula in the Workers' Compensation Experience Rating Plan, which has a separate treatment of primary and excess losses, is less accurate than a formula which uses only primary losses.

3. There are some very plausible situations when the standard credibility formula $Z = E/(E + K)$ is not appropriate. These include parameter uncertainty over time and a loss limit which increases with the size of the insured. Failure to recognize this will result in overstating credibilities for larger insureds.

The author would recommend an experience rating formula based on the credibility formula $Z = E/(E \times J + K)$. A loss limit that does not vary by size of insured should be a part of the plan. Excess losses should not be a part of the plan. This formula is less complicated than current formulas and should be easier to administer.

It should be noted that the service performed by the NCCI in calculating the experience modification is probably more important than the choice of experience rating formulas. ISO would do well to perform a similar service, or at the very minimum, provide experience in a standard format so that individual insureds could calculate their own experience modifications.

A second goal was to show how the parameters of an experience rating plan could be estimated from data. This paper demonstrates, with live data, a very general procedure for testing the parameters of a proposed credibility formula. Systematic testing of various alternative parameters should enable one to derive a reasonably accurate formula. This method requires data with which one can compare actual losses with losses predicted by the proposed formula. The author considers this kind of data absolutely essential for accurate experience rating.

Experience rating has always been a combination of scientific and intuitive reasoning. While the intent of this paper is to put experience rating on a more scientific basis, it is hoped that the reader now has a better intuitive understanding of this very important subject.

9. ACKNOWLEDGMENTS

Research for this paper was supported by a grant from the Actuarial Education and Research Fund.

REFERENCES

[1] R. H. Snader, "Fundamentals of Individual Risk Rating and Related Topics," Part I of Study Note, Casualty Actuarial Society.

[2] National Council On Compensation Insurance: *Experience Rating Plan*

[3] F. S. Perryman, "Experience Rating Plan Credibilities," *PCAS* XXIV, 1937, p. 60.

[4] D. R. Uhthoff, "The Compensation Experience Rating Plan—A Current Review," *PCAS* XLVI, 1959, p. 285.

[5] Insurance Services Office: *General Liability Experience and Schedule Rating Plans.*

[6] H. Bühlmann, "Experience Rating and Credibility," *ASTIN Bulletin* IV, Part III, p. 199.

[7] C. C. Hewitt, "Credibility for Severity," *PCAS* LVII, 1970, p. 148.

[8] A. L. Mayerson, "A Bayesian View of Credibility," *PCAS* LI, 1964, p. 85.

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

[10] P. Heckman and G. Meyers, "The Calculation of Aggregate Loss Distributions from Claim Severity Distribution and Claim Count Distributions," *PCAS* LXX, 1983, p. 22.

[11] H. Panjer, "Recursive Evaluation of a Family of Compound Distributions," *ASTIN Bulletin* XII, 1981, p. 22.

[12] R. A. Bailey and L. J. Simon, "An Actuarial Note on the Credibility of a Single Private Passenger Car," *PCAS* LVII, 1959, p. 159.

[13] G. Patrik, "Estimating Casualty Insurance Loss Amount Distributions," *PCAS* LXVII, 1980, p. 57.

[14] H. Bühlmann and E. Straub, "Credibility for Loss Ratios," (Translated by C. E. Brooks), *ARCH*, 1972.

[15] Insurance Services Office, *Report of the Credibility Subcommittee: Development and Testing of Empirical Bayes Credibility Procedures for Classification Ratemaking*, September 1980.

[16] G. Meyers, "Empirical Bayesian Credibility for Workers' Compensation Classification Ratemaking," *PCAS* LXXI, 1984, p. 96.

[17] P. Dorweiler, "A Survey of Risk Credibility in Experience Rating," *PCAS* XXI, 1934, p. 1.

[18] W. J. Conover, *Practical Nonparametric Statistics*, 2nd Edition, John Wiley and Sons, 1980, p. 256.

APPENDIX A
WITHIN VARIANCE AND THE SIZE OF THE INSURED

In this appendix we discuss how the expected within variance, σ^2, depends upon the size of the insured.

Let: N be a random variable denoting the claim count;
λ be the expected number of claims;
χ be a random variable with $E[\chi] = 1$ and $Var[\chi] = c$;
S be a random variable denoting the claim severity; and
β be a random variable with $E[1/\beta] = 1$ and $Var[1/\beta] = b$.

The collective risk model with parameter uncertainty can be described by the following algorithm.

Algorithm A.1

1. Select χ at random.
2. Select the number of claims, N, at random from a Poisson distribution with parameter $\chi \times \lambda$.
3. Select β at random.
4. Do the following N times.
   4.1 Select the claim severity, S, at random.
5. Set the total loss, X, equal to the sum of the claim amounts, S, divided by β.

b and c measure uncertainty in the scale of the claim severity distribution and the mean of the claim count distribution, respectively.

Let $R_1 = X/E[X]$. Meyers and Schenker [9] show that

$$Var[R_1] = (1 + b) \times E[S^2]/(\lambda \times E^2[S]) + b + c + b \times c.$$

The size of the risk E is proportional to $E[X]$ and can be written as $C \times E[X]$. Thus we can then write:

$$Var[R] = \Sigma_1^2/E + \alpha_1 \qquad \text{(Equation A.1)}$$

where $\Sigma_1^2 = (1 + b) \times C \times E[S^2]/E[S]$ and
$\alpha_1 = b + c + b \times c$.

In keeping with the notation of Section 3, let d denote a distribution generated by the process described above. The linear relationship of Equation A.1

is preserved when taking expected values over all distributions, d. Thus we have

$$\sigma^2 = \Sigma^2/E + \alpha \qquad \text{(Equation A.2)}$$

where $\Sigma^2 = E[\Sigma_1^2]$ and

$$\alpha = E[\alpha_1].$$

Equation A.2 is used to derive the credibility formula 3.3. If b and c are equal to zero for each distribution d, then $\sigma^2 = \Sigma^2/E$. In this case the credibility formula 3.2 applies.

We see from Equation A.1 that Σ_1^2, and thus Σ^2, depend upon the severity distribution. An increase in the loss limit will increase Σ^2.

APPENDIX B
A FORMULA FOR THE EFFICIENCY

We prove that the efficiency is equal to $2 \times Z - Z^2/Z_m$. The proof is simply a rearrangement of concepts originated by Bühlmann [6] and discussed by ISO [15].

Lemma 1: $\mathrm{Cov}[X,\mu] = \tau^2$

Proof:
$$\begin{aligned}\mathrm{Cov}[X,\mu] &= \mathrm{E}[(X - M) \times (\mu - M)] \\ &= \mathrm{E}_\mu[\mathrm{E}[(X - M) \times (\mu - M)|\mu]] \\ &= \mathrm{E}_\mu[(\mu - M)^2] \\ &= \tau^2\end{aligned}$$

Lemma 2: $\mathrm{Var}[X] = \sigma^2 + \tau^2$

Proof:
$$\begin{aligned}\mathrm{Var}[X] &= \mathrm{E}_\mu[\mathrm{Var}[X|\mu]] + \mathrm{Var}_\mu[\mathrm{E}[X|\mu]] \\ &= \sigma^2 + \tau^2\end{aligned}$$

Theorem: Efficiency $= 2 \times Z - Z^2/Z_m$

Proof: Let F be an estimator for μ. By the definition of efficiency given in Section 4 we have:

Efficiency $= 1 - \mathrm{E}[(F - \mu)^2]/\tau^2$

In our case: $F = Z \times X + (1 - Z) \times M$. Thus we have:

Efficiency $= 1 - \mathrm{E}[(Z \times (X - M) - (\mu - M))^2]/\tau^2$

$= 1 - (Z^2 \times \mathrm{Var}(X) + \tau^2 - 2 \times Z \times \mathrm{Cov}(X,\mu))/\tau^2$

$= 1 - (Z^2 \times (\sigma^2 + \tau^2) + \tau^2 - 2 \times Z \times \tau^2)/\tau^2$

$= 2 \times Z - Z^2/Z_m$

Corollary: The efficiency is maximized when $Z = Z_m$.

Proof: d(Efficiency)/d$Z = 0$ when $Z = Z_m$.

This corollary is simply a restatement of Bühlmann's result.

ADDRESS TO NEW MEMBERS—NOVEMBER 11, 1985

GEORGE D. MORISON

First I would like to extend to all members of the Class of November 1985 congratulations on achieving the milestone which has just been acknowledged by this assemblage. While those of us who have experienced the thrill of such public recognition in the past can share with you the joy of this occasion, you alone know the particulars of the effort and sacrifice required to reach this stage of your professional careers. For that dedication to the accomplishment of the objective you have earned the respect of all of us.

Those largely unheralded supporters, such as spouses, relatives, and friends who have helped in their inimitable ways should not be overlooked in this celebration. The long hours of drudgery would no doubt have been even more intolerable without such loyal, interested backers.

In the belief that any comments in the nature of advice proffered today ought to be limited to matters that are substantially achievable, I would begin by suggesting that those who have just been admitted into the CAS as Associates unstintingly pursue their Fellowship designations. Among the less diabolical objectives of the ten-year old restructuring of the syllabus of examinations was that of enhancing new Associates' motivation to go the extra mile to achieve Fellowship. It was thought that a maximum of three more examinations to pass would be viewed as a rather modest hurdle. I therefore urge, in particular, the new Associates admitted today to bend every effort to clear that last hurdle that stands in the way of total involvement in the Casualty Actuarial Society. In today's argot, you ought to "go for it!" Experience has shown that sustained effort is more successful than is succumbing to the allure of even a brief hiatus which, all too often, becomes permanent.

The new Fellows, on the other hand, might devote some of their newly acquired free time to Society activities. One of the most rewarding experiences, I believe, comes from helping with vital education and examination work which has earned for us that valued designation as a learned society. The new Fellow's recent experience with the examining process can be used to keep the system responsive to the needs of the students. Here is offered that long-sought opportunity to introduce into the system those improvements which are best identified by recent exam-takers. Certainly the present size of the Education and Examination Committee, together with the ever-growing demands on its members, leaves room for all new Fellows willing to serve

And it is precisely such a "willingness to serve" that is the key ingredient in the voluntarism which has enabled the Casualty Actuarial Society to continue to attract bright, dedicated members into its ranks. In addition to the service rendered the organization and the camaraderie that attends joint efforts, the personal sense of fulfillment that results from voluntarily helping to achieve the objectives of a professional society is seldom found elsewhere. This reward is described as "inwardly satisfying" in an article by John Tierney in the May 1985 *Actuarial Review*—an article which I commend to your attention for some elaboration on the joys and benefits of voluntarism.

As new Fellows are volunteering their assistance with CAS activities and new Associates are pursuing their Fellowship status, like all members of the Society, today's graduates should also keep abreast of developments in the profession. One such development that is advancing toward fruition is the preparation of formal Standards of Actuarial Practice, not to be confused with the more general Guides to Professional Conduct. Those standards designed to govern the practice of casualty actuaries should be reviewed critically when exposed for comments and then observed with care when promulgated in final form. Guides to Professional Conduct and related Opinions should be reviewed frequently, not only for their substance, but also because of the tone they set for our dealings with others as well as their importance in setting us off as professionals. The remarkable level of acceptance we enjoy as members of the CAS depends on a continuing commitment to standards of professionalism.

On an occasion such as this it is not inappropriate to ponder also some longer range possibilities. If we actuaries weren't such a practical-minded lot, we might even describe this exercise as dreaming a bit about the future—a future which suddenly offers enhanced opportunities thanks to the professional designations conferred here this morning.

Whether the objectives of the individual members of the Class of November 1985 be to render service to the public at large, to their professional clients, employers or confreres is largely a personal decision. But success in achieving the selected goal will require creativity, flexibility and continuing analysis of the best available information. Only then can the actuary's unique contribution be brought to bear on those problems which cry out for solution in the years to come.

May you all find as much joy in the challenging interface of dreams and reality as has your grateful and honored speaker.

KEYNOTE ADDRESS—NOVEMBER 11, 1985
THE SEARCH FOR COMPETITIVE ADVANTAGE

WILLIAM A. SHERDEN

It is a pleasure to be here today to talk about competitive advantage, a subject of widespread interest throughout all industries and one that dominates current business literature. Why the sudden burst of interest in competitive advantage? History tells us much. After World War II, the U.S. economy was the only game around. We had an established industrial base that was undamaged by the war and fueled by the consumer demand pent up since the Great Depression. On top of that, Europe was rebuilding its economy and unable to fulfill its domestic market's demand for goods and services. This was a time when the countries of Europe, and perhaps most of the rest of the world, looked up to American management know-how and wondered whether they would ever catch up.

During the post-war period, it was difficult for American management *not* to excel. But now all of this is history. Stimulated by the entry of aggressive foreign firms, the deregulation of domestic industries, and the maturation of the U.S. market, competition has intensified to unprecedented levels. With so many U.S. firms struggling in this environment, it is not surprising that seekers of management technology look more to Japan than the United States.

To see how the competitive environment holds nothing sacrosanct, one need only look at the membership in the *Fortune 500* from 1955 to 1985. This list includes the largest and presumably strongest U.S. competitors. Yet of the 500 leading firms in 1955, only 219 survived to be included in the list in 1985. Of the 281 firms that dropped out of the 500, some went bankrupt, some dwindled in size, while others were subsumed into more successful corporations through acquisition. Indeed, of the 43 firms classified as "excellent" in the 1981 book *In Search of Excellence*, 13 were experiencing competitive difficulties in 1985.

With one-half of the top 100 property-casualty firms experiencing a negative cash flow in 1984, I need not acquaint you with the impact of competitive forces. Yet much of your industry's adverse performance has been shrouded by the underwriting cycle and other trends such as mounting litigation. Once the dust settles, however, you will find new competitive threats to deal with, especially from banks, retailers, and other parties that seek to enter niches of your business. As we see it, the strongest among you will struggle to thrive, while the weakest struggle to survive. There are even scenarios in which whole sectors of the insurance business diminish or change into unrecognizable forms. This is not to say that all of you will be working for Citicorp or Sears; but I bet some of you will.

It comes as no surprise that in this environment the literature and advice of consultants and academics are full of prescriptions for gaining a competitive advantage. Unfortunately, there has been too much advice, too much confusion, and too little demonstrated success. As a consultant working with financial service firms to develop competitive strategies, I want to put this advice in perspective, pointing out useful concepts and demonstrating how to apply them to your business.

All the current ideas about competitive strategy can be mapped on the grid in Figure 1, with the classifications of economic versus qualitative, and concept versus process. The economic ideas derive from the "rational" studies of industrial economics, whereas the qualitative, or "irrational," ideas spring more from the science of marketing. "Concept" refers to prescriptive ideas suggesting specific formulas for success, while "process" refers to frameworks for thinking about your business to gain a competitive advantage.

Let us first consider the economic concepts, since they have played such a major role in competitive strategy over the past 15 years that they are now almost synonymous with strategic planning. Economic concepts focus predominantly on costs and employ such phrases as "cost curves," "industry structure," "barriers," and "mobility." Many of these concepts were intellectually appealing, if not addictive, and came to be used throughout U.S. industry. Although there are too many concepts to cover here, I do want to give you a sense of their development and use today.

One of the earliest concepts was the product life cycle, which dictates that all products and businesses inexorably follow the path from development to decline shown in Figure 2. The concept identifies not only the nature of the competitive environment but also the specific market strategies a firm should pursue. For example, a firm should secure broad distribution capability in growth markets while emphasizing cost reduction and product differentiation in mature markets.

The concept of product life cycles, when properly applied, illustrates the falsity of the common assumption that financial services is a growth business.

Life cycles apply to insurance only if you use real growth measures, such as policies, instead of premiums, which tend to escalate with inflation and many other factors, obscuring the true maturity of the market. By definition, mature markets grow at a slower rate than real GNP. When real growth measures are used, as in Figure 3, it becomes clear that most major insurance lines are mature and exhibit many of the competitive characteristics predicted by the product life cycle.

Another early economic concept was the experience curve, which is shown in Figure 4. The concept dictates that a firm's production costs decline in a predictable pattern as production volume accumulates. Specifically, a percentage increase in accumulated production volume (or "experience") gives rise to a proportional percentage decrease in unit production cost. This concept is very appealing, since it provides a definite course of action to gain competitive advantage: increase volume faster than competitors in order to lower costs.

The experience, or learning, curve was little used in competitive strategy until it was incorporated into the growth-share matrix (Figure 5). The introduction of this new concept raised prescriptive formulas for competitive success to new heights of popularity. Very simply, the growth-share matrix uses the product life cycle as its vertical axis. Its horizontal axis is market share, a surrogate for the experience curve (the greater the market share, the greater the production experience and the lower the costs). The growth-share matrix prescribed a different course of action for the various business lines fitting within each of the four cells. For example, cash cows (low growth, high share lines) should be "milked" for the cash needed to convert question marks into stars. The stars, in turn, would eventually become new cash cows as their market matures. Above all, the growth-share concept prescribes cost reduction and market share expansion.

The growth-share matrix was the most elegant and popular competitive concept of its time. It has been estimated that by 1981 nearly one-half of the *Fortune 500* firms used the matrix in their planning processes. Yet it soon became apparent that the concept failed to explain many anomalies, such as why small businesses in mature markets have been able to displace larger, well-established firms. Further, it has been recognized that not all products sell on a price-commodity basis and that cost structures vary widely across industries and often do not conform to the experience curve. Recently, the growth-share matrix has been blamed for bringing to near ruin a number of leading firms that widely adopted it.

One does not have to look hard at the breakdown of insurance costs (Figure 6) to realize that the growth-share matrix does not apply to your business. Of all the cost elements, only the 10 percent administrative cost might be affected by the learning curve, and even that is driven much more by organizational structure, processing methods, and automation. If you really wanted to lower insurance costs, you would obviously focus on the 70 percent claims cost, which is affected by actuarial, underwriting, and claims methods. This may sound like a back to basics message. You have long competed by focusing on underwriting because that is where the largest costs are. In the future, much of the cost advantage for carriers, especially in personal lines, will come from reducing sales costs, which involves a consideration of distribution methods, not experience curves.

The heavy criticism of the growth-share matrix spawned many elaborate variations that attempted to overcome its weaknesses. The matrices grew larger, with more cells, vaguer definitions, and less straightforward prescriptions for gaining a competitive advantage. It is interesting to follow Michael Porter's exposition of these concepts in his first book. Ultimately, he proposes a new matrix called "strategy space." This is actually just an empty matrix, which you must label by determining the "key mobility barriers" in your business. Reading between the lines, I take this to mean that you have to determine the key success factors and you have to do so by yourself.

If you're getting the sense that these economic concepts do not offer an easy route to a competitive advantage, you're getting the right message. In fact, these prescriptive concepts recently have become so discredited as to cast a heavy pall over strategic planning in general. Many firms have disbanded corporate-level planning departments and pushed competitive strategy back down to the business units where managers know best how their markets function.

The major lesson from this review of prescriptive economic concepts is that there is no simple method for gaining a competitive advantage. Industries are too different and businesses too complex. Prescriptive economic concepts are reminiscent of stock market forecasts and models of the U.S. economy, both of which have been discredited for trying to explain the infinitely complex with simple rules. The only way to attain a competitive advantage is to take a hard, objective look at your own firm and the environment in which it exists.

Before leaving the subject of economic concepts, I should note Michael Porter's new book, which introduces the value chain (Figure 7). The value chain suggests that a firm is a collection of numerous activities and that to gain a competitive advantage one must isolate each activity, as well as the activities of upstream suppliers and downstream distributors, to determine how to reduce costs or add value to the end customer. The value chain is not in fact a prescriptive concept, but rather a framework for analyzing a firm which provides a lengthy checklist of ideas to consider. Again, the message is clear: there are no pat answers.

The insurance business holds many good examples of innovation to enhance value-added. Direct billing has eliminated the inefficient agency-billing method and reduced premium float. Agency automation and interface have reduced manual and redundant tasks. Expert systems have given rise to a new value-added approach to risk management. Computer access to claims data has enhanced the value of management reports. And finally, some carriers have replaced insured objects such as cars rather than reimburse customers for losses.

New competitors can also use the value chain framework to gain an advantage in your business. Banks, for example, believe that they will have a significant distribution advantage over agents, enabling them to heighten convenience and lower costs. In the group health business, hospital chains are the ultimate suppliers of health benefits and now believe they can increase consumer value by selling directly to corporations, bypassing the traditional intermediaries.

But the value chain is still an economic framework, not a qualitative concept. Qualitative concepts have existed for many years, upheld by adherents such as Theodore Levitt, and have recently been popularized by the two "Excellence" books. Rejecting the notion that all products are commodities sold on price alone, qualitative concepts are based on the belief that customers and employees are real human beings and must be properly cared for if a firm is to succeed.

The latest "Excellence" book suggests that a firm gains a competitive advantage by doing hundreds, if not thousands, of things well. The book's recipe for doing things well includes four ingredients: being close to customers, innovating constantly, having "turned-on" people, and having inspirational leadership. Though it may be hard to believe that such qualitative aspects lead to competitive advantage, we have seen it proved a number of times in the insurance industry.

In this regard, I want to share with you an interesting case study involving Third Party Administrators (TPAs), which are smallish firms that are making inroads into the self-insured niches of the industry. In particular, they have rapidly established themselves in the heavily self-insured group health business. In this segment alone, there are an estimated 3,000 TPAs, which have increased their share of the market from 5 percent in 1977 to 25 percent in 1985.

The success of these small firms at the expense of established carriers illustrates the value of the qualitative aspects of competition. At Temple, Barker & Sloane, we have undertaken extensive research on TPAs. We have found that consumers rate TPAs higher than carriers on systems and other capabilities, a startling fact since we know firsthand that carriers are much more sophisticated than TPAs in service offerings. There was only one explanation for this paradox: qualitative factors drive consumers' perception of quality. That is, the attitude and responsiveness of employees, as well as other service factors, cloud a customer's perception of fundamental product features.

We saw the same thing when we analyzed how commercial brokerage firms choose among carriers in competitive bids. We were surprised to find that brokers' decisions were often based on qualitative factors, such as personal relations, responsiveness, flexibility, competence, and authoritativeness, rather than strictly on price. As illustrated in the grid (Figure 8), brokers clearly judged carriers on these subjective criteria. Carriers A and B had highly skilled field people who were granted considerable authority and were flexible and responsive to brokers' requests. Carrier C ranked relatively low in this regard and won considerably fewer competitive bids than Carriers A and B.

Our experience indicates that qualitative factors are integral to achieving a competitive advantage. They are also more subtle than tangible economic advantages such as lower prices and superior claims systems. For example, the Japanese determined that car consumers place a higher priority on the quality of car interiors than exteriors. In response, the Japanese manufacturers focused on interiors, while U.S. manufacturers continued to focus on bodies. Perhaps we need a concept called "quality space," an empty matrix that requires one to label on the basis of how customers perceive quality. To determine how consumers define quality, you must ask them. It is particularly useful to inquire about why you won or lost a particular competitive bid or lost an established account.

Before ending my discussion of qualitative advantages, I want to mention market positioning, a concept that has its origins in advertising. According to the concept, illustrated in Figure 9, a customer can identify only one or, at most, a few suppliers that occupy certain niches, or positions, in the market; all other suppliers operate in a sea of anonymity. These positions are not necessarily traditional product or segment niches but rather perceptual niches that the customer can readily identify. For example, IBM occupies the positions of "industry standard" and "safety" in the computer market, while Intel occupies the position of "innovation" in the integrated circuit market. The competitive advantage of dominating a market position is clear: a dominant firm is more frequently included in competitive bids and wins them on bases other than price.

The concept of market positioning readily applies to the vast property-casualty industry, where hundreds of firms compete. Though some firms dominate various product or market niches, only a very few have recognizable positions in the major markets such as personal lines. Most carriers are afloat in the sea of anonymity.

As we have seen, the road to a competitive advantage is a twisting one; there are no simple rules, no foolproof prescriptions. Some general concepts are useful, but each industry practitioner must objectively analyze his or her own business to find the winning formulas. And, as we have also seen, a competitive advantage encompasses more than the rational economic model of industry structure; it also includes many of the qualitative factors presented here.

Let me now suggest a straightforward planning framework that we at TBS use to help clients apply these concepts and develop a competitive advantage. As shown in Figure 10, the framework involves four levels of analysis in the development of a strategic plan. The first step is to analyze three broad areas: (1) the environment, (2) the competition, and (3) your firm itself. The environmental analysis should include first and foremost the market. Where appropriate, the analysis should also include other factors such as legal and regulatory issues, which are particularly key to the property-casualty environment. In analyzing competitors, you must focus not just on traditional rivals but also on emerging firms with increasing market share, particularly non-traditional competitors. In analyzing your internal capabilities, you might use the value chain or other concepts, but it is most important that you apply them with objectivity.

By comparing developments in the environment with your analysis of the competition, you should be able to identify the threats and opportunities your firm faces. Similarly, by comparing your firm with its competitors, you should be able to define your competitive strengths and weaknesses. This analysis should produce a clear understanding of your overall competitive position in terms of the opportunities and challenges ahead and how well your firm is suited to deal with them. Once your competitive position is well defined, you will have a solid basis for identifying and evaluating alternative courses of action in the development of a strategic plan.

We would note, however, that even when armed with a brilliant strategy most firms have organizational and cultural barriers that will block their attempt to gain a competitive advantage. I recall a recent quote in *Business Week* that recommended spending 20 percent of your time on developing a competitive advantage and 80 percent on developing the culture to make it work. Judging from our experience, this might be understated: strategy, organization, and culture are inseparable aspects of gaining a competitive advantage.

The barriers to effective implementation are intrinsic to the typical organization with its many managerial layers and functional divisions. As Figure 11 shows, these organizations keep key decision makers too distant from the market. Market information filters out as it passes up through the myriad levels of management. Too often firms respond to this problem with a "bootstrap" marketing approach, attempting to compensate for the lack of direct market contact by feeding senior management quantitative market research. Market research is important, but it is a poor substitute for actual market exposure.

Organizational complexity often produces an ivory tower effect where senior managers have little or no contact with the customers who are the reason for their existence. We have, however, seen a number of cases of a firm breaking this barrier, gaining a significant competitive advantage by placing senior management in contact with customers. In one instance a major brokerage firm, which was consolidating its carriers, chose the winning firms largely because their CEOs were involved in the sale. The CEO is the most effective sales resource you have to gain an advantage, and one that is too often overlooked.

Remoteness from the market also breeds imitation, a common occurrence in the insurance industry. Firms often closely copy the innovations and procedures of their counterparts. Competitive advantage will again go to those carriers that buck this trend and start chasing the market rather than each other.

A second barrier to strategic implementation is tunnel vision, which is caused by the rigid functional structure that plagues many industries, especially insurance. A functional orientation may be necessary in a technical industry such as insurance, but when taken too far it breeds insularity. Employees know only their particular sphere of activity and are ignorant about the market's broader context. They are distant not only from the market but from each other. This structure results in functional specialists becoming CEOs or division heads, precluding any generalist perspective. It is illuminating to note that the development of Universal Life, the major product innovation in individual insurance, resulted when a securities executive was put in charge of an acquired life company.

The ponderous organization is the third barrier to strategic implementation. Even when a typical firm finds a competitive advantage, it often loses it by taking too long to exploit it in the market. The symptoms of the ponderous organization are an overreliance on committees, protracted decision making, slow response to the environment, a lack of anyone in charge, overlooked responsibilities, and conflict resolution residing in the CEO's office.

The fourth barrier to implementation is the risk-averse culture of most carriers. Carriers are in the risk management business by definition, but too often this cultural heritage spills over into many aspects of business. With regard to attitude toward customers, for example, there are often more people in a carrier that can say no to a customer than yes. Risk avoidance too often influences a carrier's attitude toward business risks, resulting in a paralyzing conservatism. Sacred cows block new innovations, and old, time-tested (and often outdated) ways are the only accepted ways. In this environment, innovation becomes frozen and competitive advantage becomes an illusion.

The risk-adverse culture also leads carriers to have an overly conservative human resources policy, built on the belief that ten lower-paid, mediocre employees are better than a few higher paid, highly motivated, skilled individuals. I know from experience this is false. Another manifestation of risk-aversity is a paternalistic attitude toward employees, which management takes in the hope that it will lead to loyalty when ultimately it leads only to a polarization of senior management and workers.

The development of insurance generalists is impeded by the typical organizational culture. It's been said that it is hard to find an industry that does so little job rotation as insurance. Judging from our observations of the talent at many leading-edge financial service firms, if you cannot attract high quality generalists or develop them internally, you are going to face stiff competition in the future.

The clear winners in tomorrow's financial services market will be those firms that make a successful transition from the traditional environment to the current competitive environment, as outlined in Figure 12. The winners will overcome the organizational and cultural barriers and will be able to identify and exploit competitive advantages in the market. Once again, strategy, organization, and culture are inseparable considerations in pursuing a competitive advantage.

I would like to conclude by discussing how my comments relate to your profession and the career paths you might take. The options before you are to pursue a technical career either within a firm or in consulting or to progress in the organization to the level of senior actuary or general manager. If you choose the latter path, you will be faced with the challenge of seeking a competitive advantage for your firm, and in this you must consider your own career development. Though actuaries are in some ways the intelligentsia of the firm, their long years of schooling and professional practice can lead to isolation. If you truly seek a general management career, you might consider advanced management programs, job rotation, or other vehicles to gain broad exposure. Although it is hard to imagine a flood of actuaries making sales calls, perhaps something along these lines would not be so bad.

Our experience has shown that directly exposing bright, intuitive executives to the market builds superior market intelligence and yields superior strategies. The development of a broadly skilled senior actuary who is market- and employee-oriented might be a large step a carrier can take toward attaining a competitive advantage.

Figure 1

HOW TO GAIN COMPETITIVE ADVANTAGE

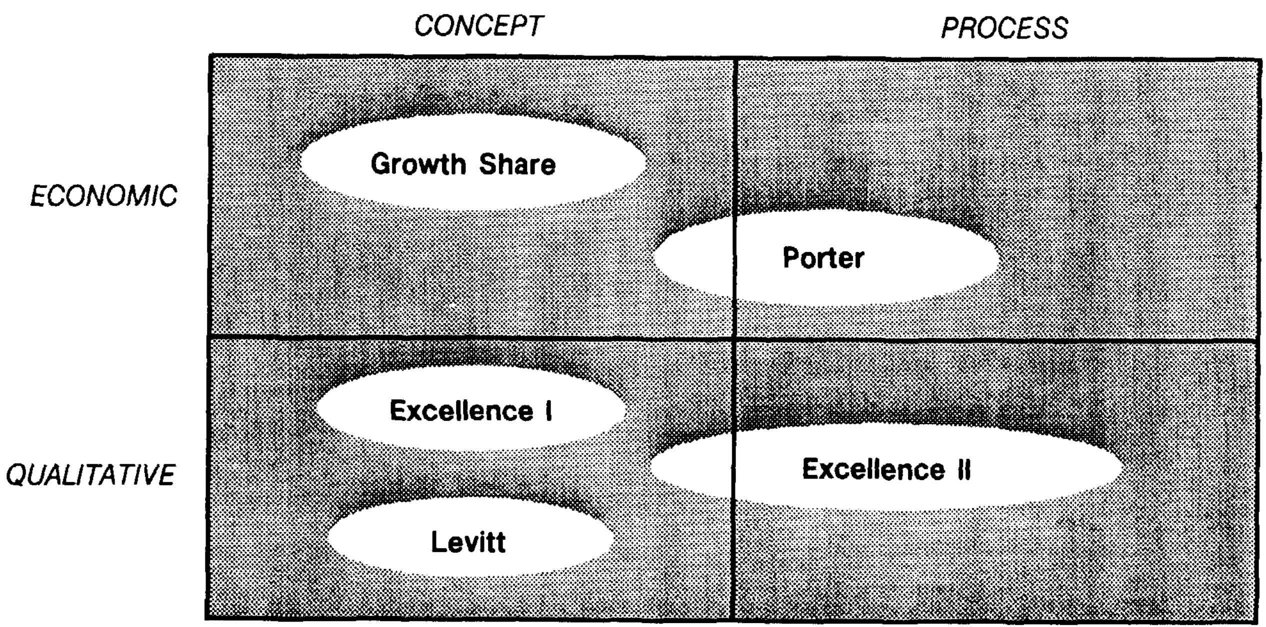


Figure 2

PRODUCT LIFE CYCLE CONCEPT FOR MARKET ASSESSMENT

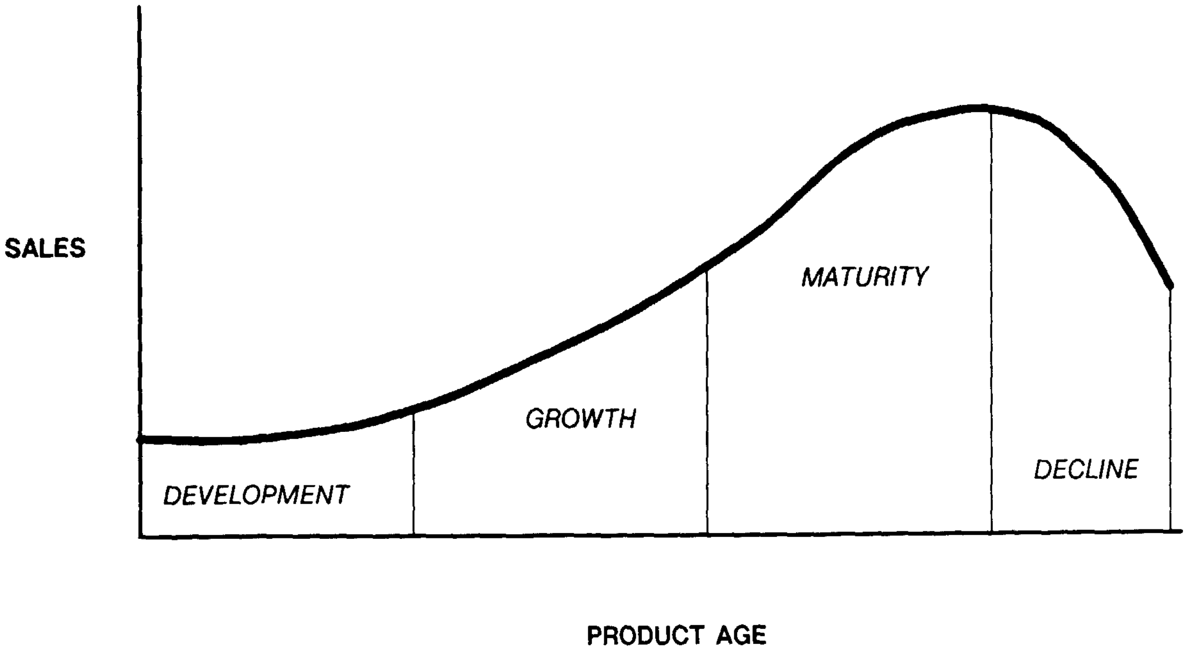


Figure 3

POLICY GROWTH IN MAJOR LINES

Compound Annual Growth

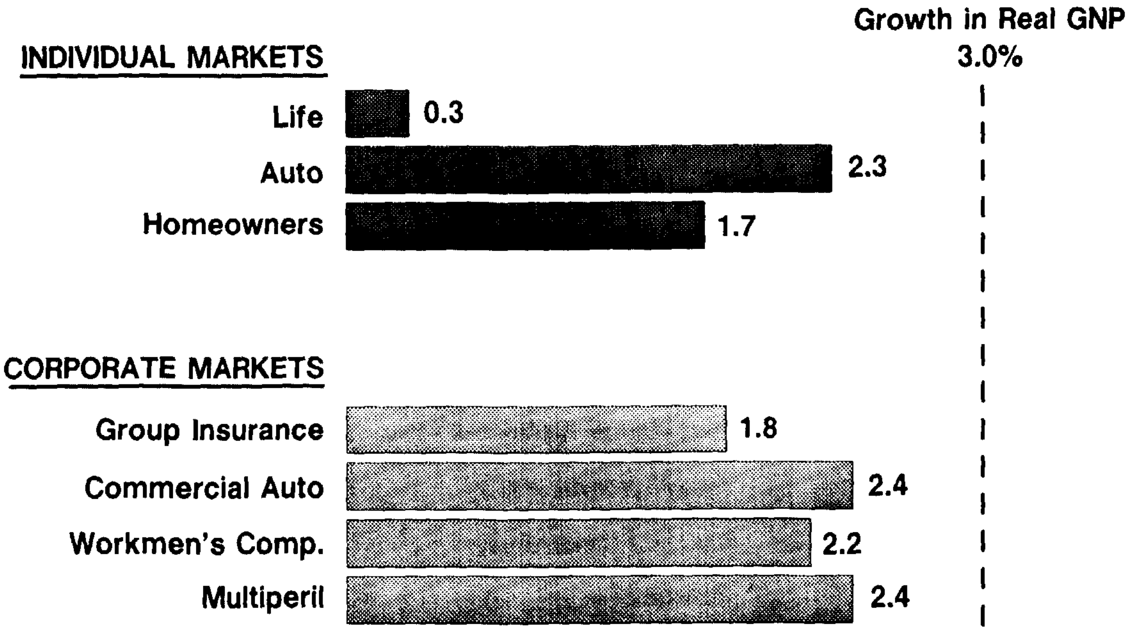


Figure 4

THE EXPERIENCE CURVE

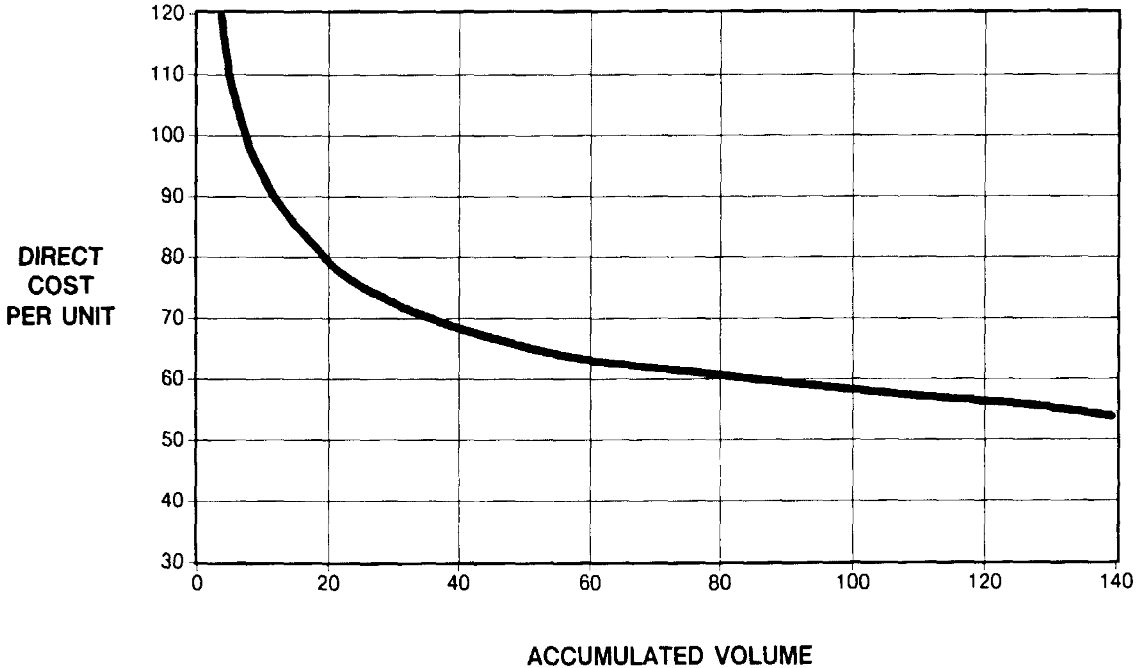


Figure 5

GROWTH-SHARE MATRIX

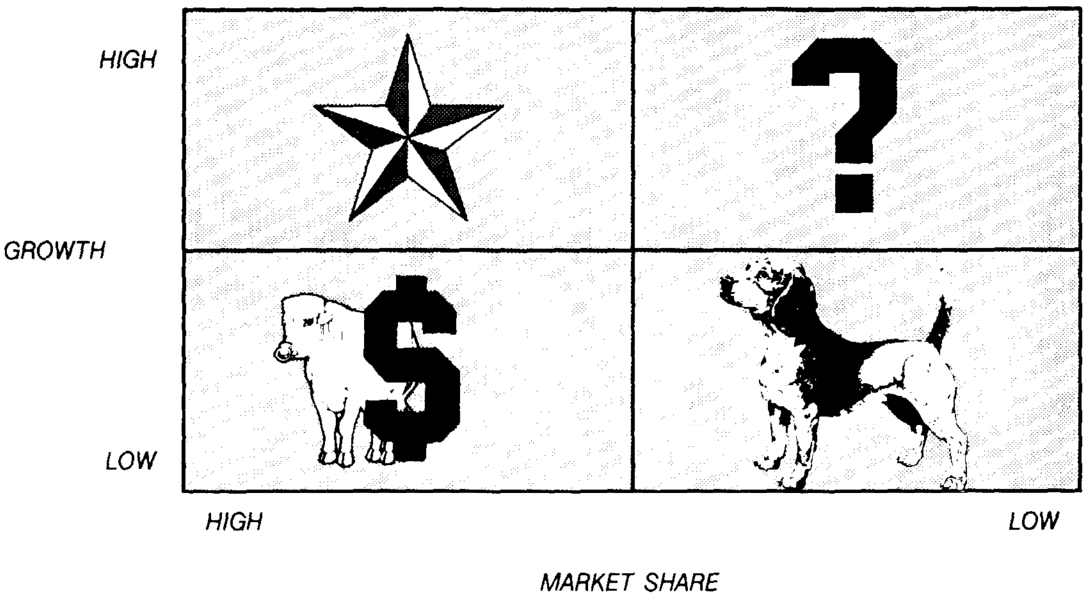


Figure 6

DETERMINANTS OF INSURANCE COSTS

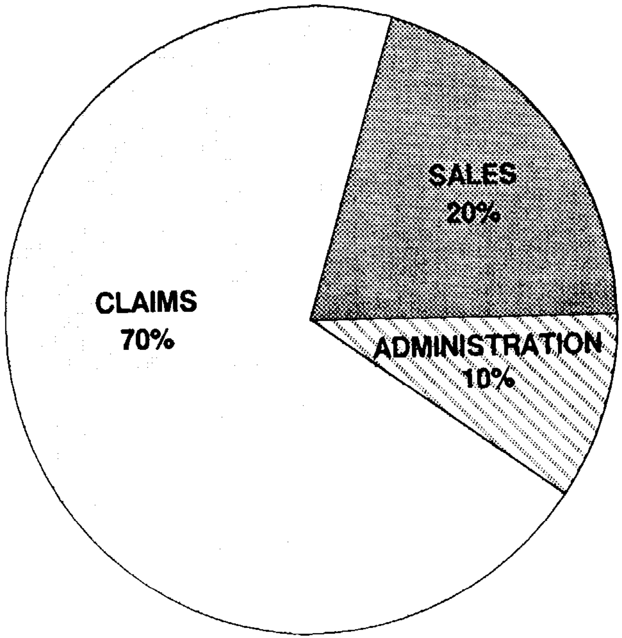


Figure 7

THE VALUE CHAIN

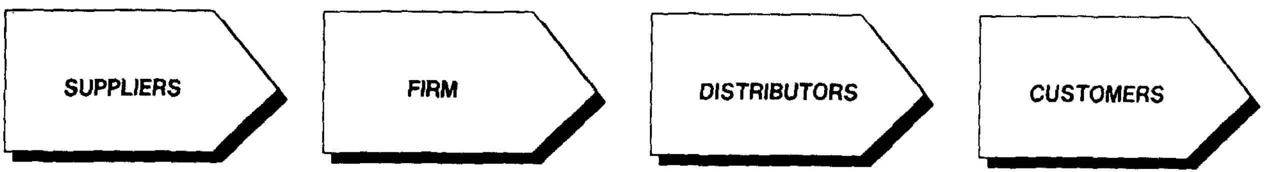


Figure 8

EXAMPLES OF QUALITATIVE FACTORS

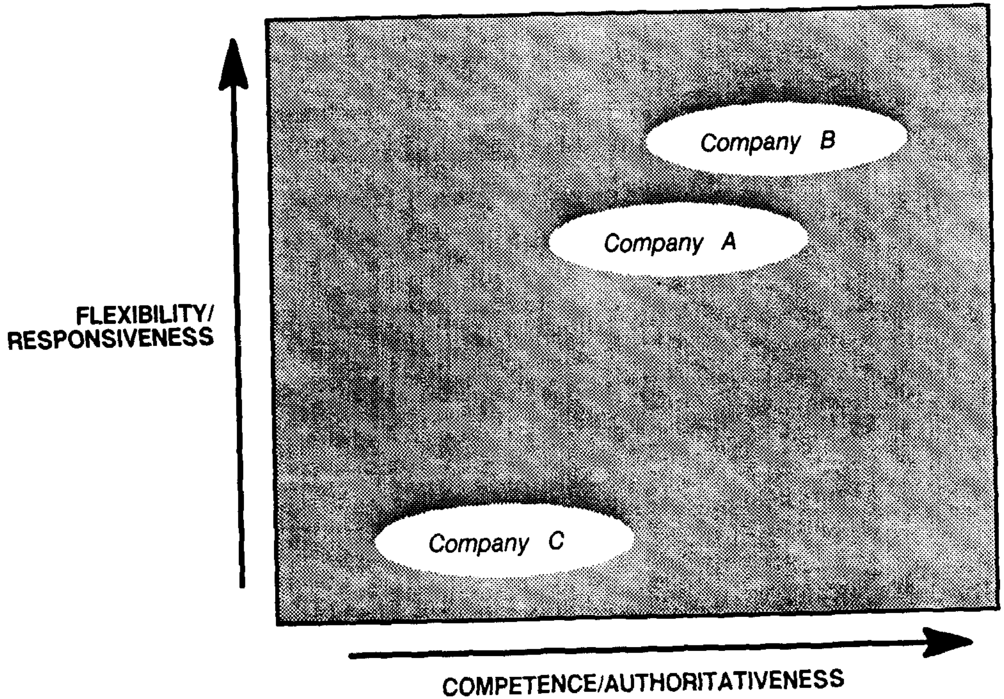


Figure 9

MARKET POSITIONING

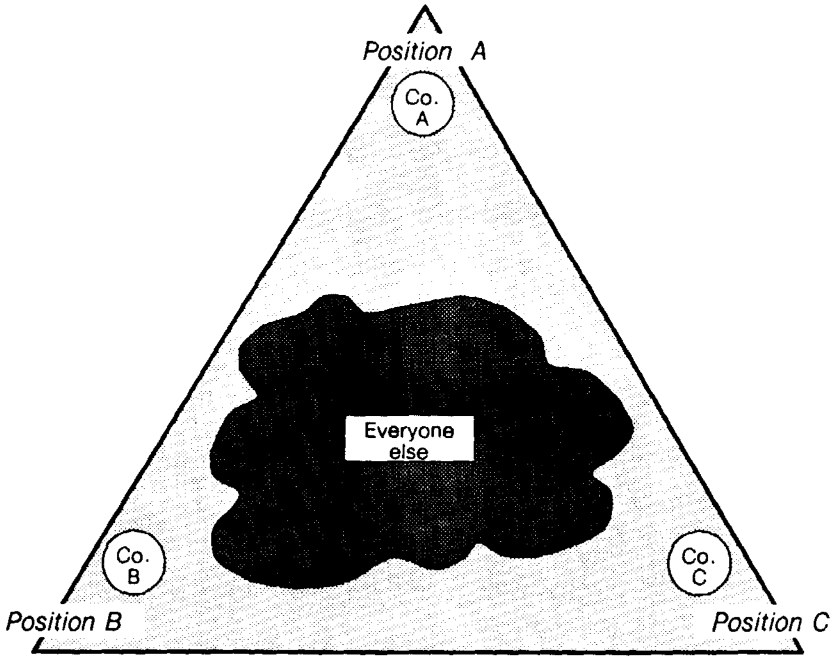


Figure 10

ANALYTICAL APPROACH

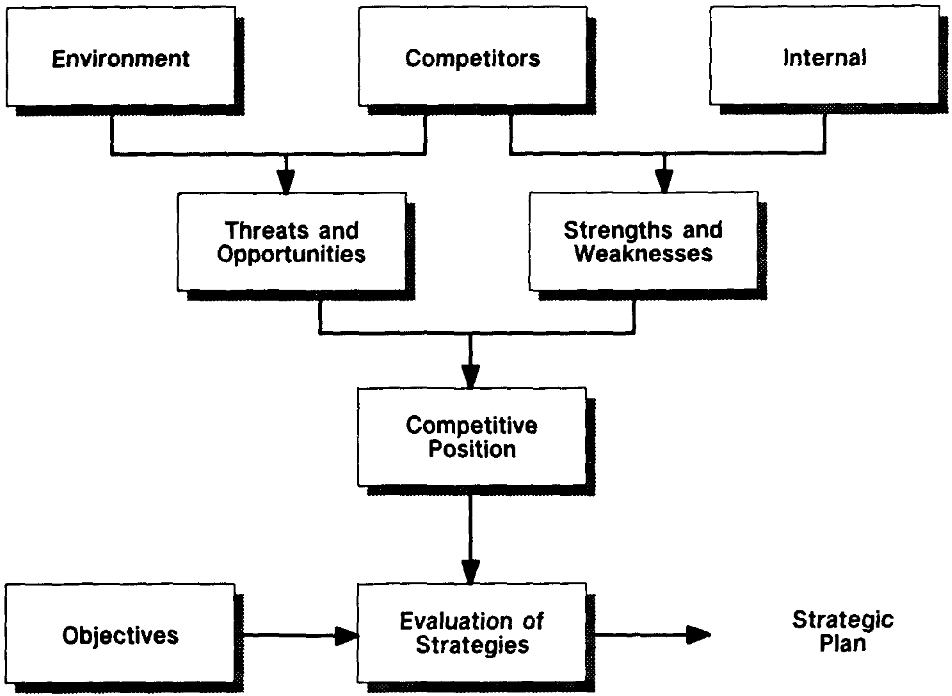


Figure 11

BARRIERS TO IMPLEMENTATION

Typical Organization

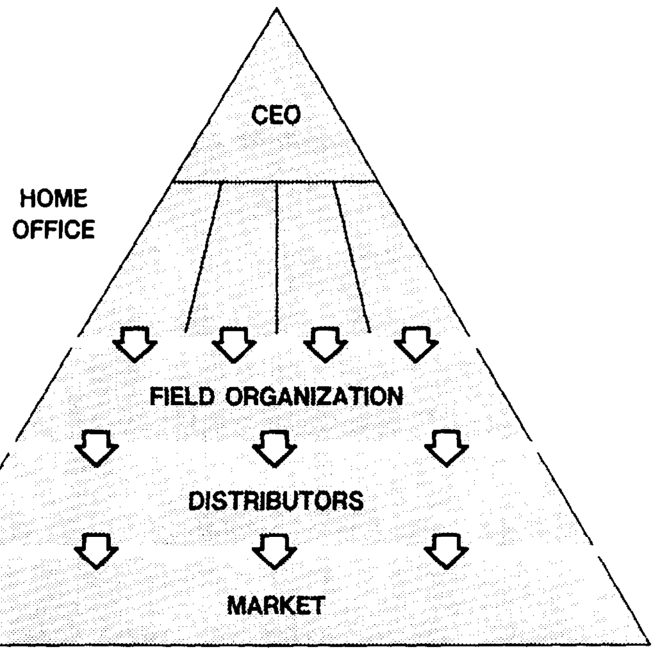


Figure 12

CHANGES IN CULTURE

TRADITIONAL	TODAY
STABILITY	MARKET-DRIVEN
MANAGEMENT HIERARCHY	ENTREPRENEURIAL
RISK-AVERSE	QUICK-RESPONSE
REGULATORY FOCUS	INNOVATIVE
BUREAUCRACY	P&L MANAGEMENT
TOP-DOWN DECISIONS	PERFORMANCE-ORIENTED
INWARD PERSPECTIVE	

MINUTES OF THE 1985 ANNUAL MEETING

November 10–12, 1985

WESTIN CROWN CENTER, KANSAS CITY, MISSOURI

Sunday, November 10, 1985

The Board of Directors held their regular quarterly meeting from 1:00 p.m. to 4:00 p.m.

Registration was held from 4:00 p.m. to 6:30 p.m.

The Officers hosted a reception for new Fellows and their spouses from 5:30 p.m. to 6:30 p.m.

A general reception for all members and guests was held from 6:30 to 7:30 p.m.

Monday, November 11, 1985

Registration continued from 7:30 a.m. to 8:30 a.m.

President C. K. Khury opened the meeting at 8:30 a.m. He introduced Fletcher Bell, Kansas Commissioner of Insurance, who welcomed the membership to Kansas City.

Mr. Khury then announced the results of the elections of Officers and Directors:

President-Elect
 Michael A. Walters

Directors
 James R. Berquist
 Charles F. Cook
 David P. Flynn
 Mavis A. Walters

Herbert J. Phillips was appointed by the Board of Directors to fill the remaining term of Michael Fusco, who resigned to take a position on the Executive Council.

Mr. Khury recognized the nine new Associates and presented diplomas to the twenty-eight new Fellows, who were introduced by Mr. Wayne Fisher, Vice President—Membership. The names of these individuals follow.

FELLOWS

Steven D. Basson	Gregory S. Grace	Gail A. Mendelssohn
Kevin H. Bursley	Ronald E. Greco	William S. Morgan
John E. Captain	Jonathan B. Hale	E. Toni Mulder
Joel S. Chansky	Jeffrey L. Hanson	John C. Narvell
Ross A. Currie	Gayle E. Haskell	Andre Normandin
Linda A. Dembiec	Timothy T. Hein	Robert L. Sanders
Claude Desilets	Kenneth J. Hoppe	Harvey A. Sherman
Robert V. Deutsch	Robert S. Kaplan	Jerome J. Siewert
Brian Duffy	Jeffrey H. Mayer	Joanne S. Spalla
N. Paul Dyck		

ASSOCIATES

Janet B. Dezube	Jeffrey R. Jordan	Sharon A. Mair
Scott H. Dodge	Robert S. Kaplan	Roger A. Yard
Steven B. Goldberg	Andrew E. Kudera	Mark A. Yunque

Mr. Khury then introduced Mr. George Morison, a past President of the Society, who addressed the new Fellows concerning their professional responsibilities.

Mr. Charles A. Bryan summarized the three new *Proceedings* papers. Paul Braithwaite delivered a review of James E. Buck's paper, "On Stein Estimators: 'Inadmissibility' of Admissibility as a Criterion for Selecting Estimators."

Mr. Khury concluded the Business Session at 9:30 a.m. and introduced the Keynote speaker, Mr. William A. Sherden, Vice President, Temple, Barker & Sloane, Inc., who spoke on "The Search for Competitive Advantage."

The remainder of the morning was devoted to concurrent sessions, consisting of four General Attendance Workshops, a Limited Attendance Workshop, a workshop presentation by the Committee on Ratemaking Principles, and three new *Proceedings* Papers.

The General Attendance Workshops are listed below.

1. "Are Auto Residual Market Mechanisms Effective?"
   John Corbley—*Moderator*
   President
   AIPSO

   Gary Grant
   Actuary
   State Farm Mutual Automobile Insurance Company

   Michael A. LaMonica
   Actuary
   Allstate Insurance Company

2. "State-Of-The-Art Personal Auto Pricing Techniques"
   Charles A. Bryan—*Moderator*
   Senior Vice President, Actuary
   USAA

   Kyleen Knilans
   Director, Personal Auto Pricing
   Nationwide

   Robert T. Muleski
   Associate Actuary
   Liberty Mutual

   Glenn M. Walker
   Associate Actuary
   GEICO

3. "The Practical Implications of Insurer Insolvency"
   David G. Hartman—*Moderator*
   Vice President and Actuary
   Chubb Group

   James E. Gustafson
   Senior Vice President
   General Reinsurance Corporation

   Richard Heydinger
   Director of Risk Management
   Hallmark Cards, Inc.

Chris Milton
Vice President
AIG Reinsurance

4. "Controlling Legal Expenses by Rebuilding the Courthouse Steps"
Patrick J. Grannan—*Moderator*
Consulting Actuary
Milliman & Robertson, Inc.

Jerome Wolf
Spencer, Fane, Britt and Browne

Deborah R. Hensler
Senior Social Scientist
The Rand Corporation

The Limited Attendance Workshop was

"The Insurer's Market Identity Crisis"
William A. Sherden—*Moderator*
Vice President
Temple, Barker & Sloane, Inc.

The Committee on Ratemaking Principles workshop was an "Open Discussion on a Statement of Ratemaking Principles." The session provided an open forum for the review and discussion of the Committee's first draft of a Statement.

The three new *Proceedings* papers are listed below:

"An Analysis of Experience Rating"
Author: Glenn G. Meyers
University of Iowa

"The Valuation of an Insurance Company for an Acquisition Involving a Section 338 Tax Election"
Authors: Orin M. Linden, James A. Hall, III, Stephen Gerard, Michael Heitz
Coopers & Lybrand

"An Intoduction to Underwriting Profit Models"
Author: Howard C. Mahler
Massachusetts Rating Bureaus

FELLOWS

Mohl, F. J.
Moore, P. S.
Morison, G. D.
Mulder, E. T.
Muleski, R. T.
Murdza, P. J., Jr.
Muza, J. J.
Myers, N. R.
Narvell, J. C.
Newlin, P. R.
Niles, C. L., Jr.
Normandin, A.
O'Connell, P. G.
O'Neil, M. L.
Otteson, P. M.
Phillips, H. J.
Renze, D. E.
Roberts, L. H.
Rodermund, M.
Ryan, K. M.
Sanders, R. L.
Sherman, H. A.
Shoop, E. C.
Skurnick, D.
Smith, L. M.
Snader, R. H.
Spalla, J. S.
Stephenson, E. A.
Streff, J. P.
Tatge, R. L.
Tiller, M. W.
Tom, D. P.
Tresco, F. J.
Tuttle, J. E.
Van Ark, W. R.
Van Slyke, O. E.
Walker, G. M.
Walker, R. D.
Walters, M. A.
Walters, M. A.
Weimer, W. F.
Weller, A. O.
Wilson, J. C.
Wilson, R. L.
Wiseman, M. L.
Woods, P. B.
Wulterkens, P. E.

ASSOCIATES

Andler, J. A.
Austin, J. P.
Bailey, V. M.
Cadorine, A. R.
Chorpita, F. M.
Cimini, E. D., Jr.
Clark, D. G.
Cohen, A. I.
Connor, V. P.
Crifo, D. A.
Dezube, J. B.
Dodge, S. H.
Dornfeld, J. L.
Eagelfeld, H. M.
Edie, G. M.
Einck, N. R.
Galiley, B. J.
Gapp, S.
Goldberg, S. B.
Gould, D. E.
Harwood, C. B.
Henry, T. A.
Jensen, J. P.
Jordan, J. R.
Kelly, M. K.
Klawitter, W. A.
Kolk, S. L.
Koupf, G. I.
Kudera, A. E.
Leo, C. J.
Licht, P. M., Sr.
Lis, R. S., Jr.
Loper, D. J.
Mair, S. A.
McDaniel, G. P.
Mokros, B. F.
Mozeika, J. K.
Murphy, W. F.
Napierski, J. D.
Nelson, J. K.
Ogden, D. F.
Pelletier, C. A.
Penniman, K. T.
Peterson, S. J.
Port, R. D.
Potts, C. M.
Ratnaswamy, R.
Rice, W. V.
Rudduck, G. A.
Sansevero, M., Jr.
Schulman, J.
Schultheiss, P. J.
Smith, B. W.
Stroud, R. A.
Terrill, K. W.
Townsend, C. J.
Tucker, W. B.

ASSOCIATES

Waldman, R. H.
Whatley, M. W.
White, C. S.
Yard, R. A.
Yatskowitz, J. D.
Yau, M. W.
Yunque, M. A.

GUESTS–SUBSCRIBERS–STUDENTS

Armstrong, S. H.
Bradley, S.
Carpenter, J. G.
Colver, C. F.
Colvin, S. P.
Comstock, S.
Demarlie, G.
Didonato, A. M.
Dunn, J. T.
Farwell, R. A.
Furtney, G.
Graves, G. G.
Johnson, J. E.
Johnston, S. J.
Natte, B.
Peck, S.
Schmidt, L.
Smith, D. A.
Stenmark, J. A.
Sterk, J.
Wilson, G. S.

REPORT OF THE VICE PRESIDENT—ADMINISTRATION

This report, by the Vice President—Administration, is intended to provide the members with a brief summary of the more important activities of the Society during the last fiscal year.

Both the Board of Directors and the Executive Council, as well as all the standing committees were extremely active in this year, our second year under the reorganization program. A great deal of progress has been made on your behalf. The Board of Directors, with the prime responsibility of setting policy, met four times, and took several key policy positions affecting all facets of the Society: Administration, Programs, Membership, and Development. These policies were announced in the *Actuarial Review* and will also appear in the next edition of the *Yearbook*. The *Yearbook*, by the way, is being expanded gradually to include more information relative to policies set by the Board, as well as operational items enacted by the Executive Council.

The Executive Council, with the prime responsibility of running the day-to-day activities of the CAS, also met four times during the year and dealt with the very extensive agendas at each meeting. In addition, for the first time, the Executive Council held a meeting of all committee chairmen, which was well attended and very well received. It provided a forum for both the officers and committee chairmen to get to know each other; to openly discuss their mutual problems, goals, and activities; and to discuss the best way to accomplish the various tasks assigned. It is planned that such a meeting will be held at least annually from now on.

The membership of the CAS continues to grow at a rapid rate. At the Spring Meeting in Boca Raton, sixty-eight new Associates and nineteen new Fellows were admitted. At this meeting in Kansas City, nine new Associates and twenty-eight new Fellows were admitted. The membership is now approximately 1,190. Certainly, we will surpass the 1,200 mark next year.

As a result of this growth in membership and the fact that the CAS operates largely by the willingness of its members to volunteer their time and effort, the Executive Council commissioned a study of the future of the CAS office. The recommendation of the committee was that one staff member be added to the business office, basic automation be installed (one personal computer) and the necessary space and furniture be acquired. This recommendation was approved by the Council and will go into effect quickly.

Another facet of the membership growth and the resulting budget requirements to maintain adequate services was the need to install a functional accounting system. This concept was also approved by the Executive Council and installed with the 1985/86 fiscal year budget. Income and disbursement items now will be segregated into four distinct functions—Membership Services, Examinations, Programs, and All Other. The budget approved for the 1985/86 fiscal year is in excess of $500,000 and will require an increase of $20 in dues for all classes of members and an increase of $20 in exam fees for Parts 4 through 10.

The activities of both Board and Council in this past year, in no particular order of priority, included the subjects listed below.

1. Goals and objectives governing the educational efforts of the CAS. This subject includes the basic examination process, as well as continuing education for current members.
2. Establishment of a talent bank to provide for the identification of members willing to serve on committees and an indication of their primary interests or specialties.
3. Revised and updated guides for the submission of papers.
4. Registering of the *Proceedings* in the Library of Congress and obtaining an International Standard Serial Number (ISSN).
5. Establish bibliographies on ratemaking principles and loss reserving principles.
6. Canadian content on the CAS Syllabus.

Finally, the Audit Committee audited the 1984/85 fiscal year books of the CAS and found the accounts to be properly stated. The year ended with an increase in surplus of $37,281.22, which fortunately offset the operating loss of the previous year. The reason for the increase was the success of the Boca Raton meeting—registrations far exceeded anticipated levels—and a much better investment yield than the budget predicted.

Members' equity now stands at $258,799.90, subdivided as follows:

Michelbacher Fund	$ 59,681.87
Dorweiler Fund	9,881.80
CAS Trust	2,005.28
Scholarship Fund	7,112.50
CAS Surplus	180,118.45
	$258,799.90

For 1985/86, the Board of Directors elected the following Vice Presidents:

Vice President—Administration	Richard H. Snader
Vice President—Development	David G. Hartman
Vice President—Membership	Wayne H. Fisher
Vice President—Programs	Michael Fusco

This is my final report as Vice President—Administration and I would be remiss if I did not publicly thank those who have worked with me over the past years: Bob Daino as Assistant Secretary; Tony Grippa as Assistant Treasurer, and his staff; and, in particular, Edee Morabito in charge of the business office; as well as the other committee members within the Administration function. The CAS is indeed fortunate to have such people and I enjoyed working with them.

Respectfully submitted,

HERBERT J. PHILLIPS
Vice President—Administration

FINANCIAL REPORT
FISCAL YEAR ENDED 9/30/85 (ACCRUAL BASIS)

INCOME		DISBURSEMENTS	
Dues	$115,857.72	Printing & Stationery	$122,703.66
Exam Fees	103,281.22	Office Expenses	107,422.90
Meetings	196,611.99	Exam Expenses	4,464.88
Proceedings	12,743.95	Meeting Expenses	177,851.74
Readings	14,514.53	Library	415.24
Invitational Program	4,880.00	Insurance	7,333.48
Interest	35,094.08	Refund—Dues	290.00
Actuarial Review	332.00	Refund—Exam	2,735.00
Yearbook	950.00	Refund—Meeting	9,350.00
Foreign Exchange	(551.92)	Refund—Reading	94.00
Miscellaneous	1,014.73	Math. Assoc. of America	2,000.00
Total	$484,728.30	Expenses—President	5,000.00
		Expenses—Pres.-Elect	2,500.00
		Outside Services	0
		Miscellaneous	5,286.18
		Total	$447,447.08

Income	$ 484,728.30
Disbursements	447,447.08
Change in CAS Surplus	$(37,281.22)

ACCOUNTING STATEMENT (ACCRUAL BASIS)

ASSETS	9/30/84	9/30/85	CHANGE
Checking Account	$ 35,866.94	$ 1,259.80	$ (34,607.14)
Money Market Fund	61,930.52	143,120.28	81,189.76
Bank Certificates of Deposit	102,573.00	0	(102,573.00)
U.S. Treasury Notes & Bills	99,971.90	222,926.78	122,954.88
Accrued Interest	24,216.75	11,684.06	(12,532.69)
Total Assets	$324,559.11	$378,990.92	$ 54,431.81
LIABILITIES			
Office Expenses	$ 28,000.00	$ 30,000.00	$ 2,000.00
Printing Expenses	62,000.00	30,611.00	(31,389.00)
Prepaid Examination Expenses	(273.14)	0	273.14
Meeting Expenses & Prepaid Fees	(3,500.00)	13,813.02	17,313.02
Prepaid Exam Fees	29,970.00	45,767.00	15,797.00
Other	0	0	0
Total Liabilities	$116,196.86	$120,191.02	$ 3,994.16
MEMBERS' EQUITY			
Michelbacher Fund	$ 54,791.76	$ 59,681.87	$ 4,890.11
Dorweiler Fund	8,922.62	9,881.80	959.18
CAS Trust	1,810.64	2,005.28	194.64
Scholarship Fund	0	7,112.50	7,112.50
CAS Surplus	142,837.23	180,118.45	37,281.22
Totals	$208,362.25	$258,799.90	$ 50,437.65

Herbert J. Phillips
Vice President—Administration

This is to certify that the assets and accounts shown in the above financial statement have been audited and found to be correct.

Audit Committee
Walter J. Fitzgibbon, Jr., Chairman
George G. Bertles
David M. Klein

1985 EXAMINATIONS—SUCCESSFUL CANDIDATES

Examinations for Parts 4, 6, 8 and 10 of the Casualty Actuarial Society were held on May 2 and 3, 1985. Examinations for Parts 5, 7 and 9 were held on November 6, 7, and 8, 1985.

Examinations for Parts 1, 2 and 3 are jointly sponsored by the Casualty Actuarial Society and the Society of Actuaries. These examinations were given in May and November of 1985. Candidates who passed these examinations were listed in the joint releases of the two societies.

The Casualty Actuarial Society and the Society of Actuaries jointly awarded prizes to the undergraduates ranking the highest on the General Mathematics examination. For the May, 1985 examination, $200 prizes were awarded to Noam D. Elkies, Francois Grenon, Martin Papillon, and Jeffrey M. Sanders. The additional $100 prize winner was Andrew Lee. For the November, 1985 examination, the $200 prize was awarded to Stephen J. Stribling. The additional $100 prize winners were Rajasekhar Malyala, Darien G. Lefkowitz, Howard P. Hines, and Anthony J. Benjamin.

The following candidates were admitted as Fellows and Associates at the November, 1985 meeting as a result of their successful completion of the Society requirements in the May, 1985 examinations.

FELLOWS

Steven D. Basson
Kevin H. Bursley
John E. Captain
Joel S. Chansky
Ross A. Currie
Linda A. Dembiec
Claude Desilets
Robert V. Deutsch
Brian Duffy
N. Paul Dyck
Gregory S. Grace
Ronald E. Greco
Jonathan B. Hale
Jeffrey L. Hanson
Gayle E. Haskell
Timothy T. Hein
Kenneth J. Hoppe
Robert S. Kaplan
Jeffrey H. Mayer
Gail A. Mendelssohn
William S. Morgan
E. Toni Mulder
John C. Narvell
Andre Normandin
Robert L. Sanders
Harvey A. Sherman
Jerome J. Siewert
Joanne S. Spalla

ASSOCIATES

Janet B. Dezube
Scott H. Dodge
Steven B. Goldberg
Jeffrey R. Jordan
Robert S. Kaplan
Andrew E. Kudera
Sharon A. Mair
Roger A. Yard
Mark A. Yunque

The following is the list of successful candidates in examinations held in May, 1985.

Part 4

Aquino, John G.
Audet, Daniel
Billings, Holly L.
Blakinger, Jean M.
Boudreau, Joseph J.
Bradley, J. Scott
Brehm, Paul J.
Caulfield, Michael J.
Cohen, Sheldon
Commodore, Alfred D.
Conway, Ann M.
Crawshaw, Mark
Cross, Susan L.
Desnoyers, Lee A.
Drent, Susan M.
Feldblum, Sholom
Fitzpatrick, William G.
Forbus, Barbara L.
Francis, Louise A.
Franz, Vincent-M.
Gagnon, Luc
Garneau, Denis
Gergasko, Richard J.
Gevlin, James M.
Gibson, Richard N.
Goldberg, Leonard R.
Goldberg, Robert H.
Griffith, Ann V.
Groh, Linda M.
Groshong, Susan J.
Grossack, Marshall J.
Gruenhagen, Todd A.
Haefner, Larry A.
Haidu, Deborah D.
Hawley, Karin S.
Hays, David H.
Herderick, Teresa J.
Hurley, John M.
Johnson, Eric J.
Johnston, Steven J.
Kaplan, Robert S.
Keatinge, Clive L.
Keen, Eric R.
Kesby, Kevin A.
Kwon, Frank O.
Lalonde, David O.
Lamb, Dean K.
Lamy, Mathieu
Lapointe, Susan E.
Lebens, Joseph R.
Leveille, Jean-Marc
Lyons, Mark D.
Mahon, Mark J.
Mallison, Robert G., Jr.
Marchena, Eduardo P.
Marles, Blaine C.
Maud, Christine E.
Mayer, Malkie
McNichols, James P.
Millar, Leonard L.
Miller, John E.
Miller, Mary F.
Nelson, Chris E.
Nyce, Glen C.
Panjer, Harry H.
Paterson, Bruce
Pipitone, Faith M.
Popejoy, Kathy
Protz, Steven G.
Saton, Melissa A.
Sauthoff, Stephen P.
Shimkus, Mary B.
Sliwa, Jan
Slotznick, Lisa A.
Spiegler, David
Spore, Louis B.
Steinberg, Karen F.
Strasser, Benjamin C.
Swanstrom, Ronald J.
Sweeney, Eileen M.
Wacker, Gregory M.
Werland, Debra L.
Wilson, Theresa A.
Wrobel, Edward M.
Wu, Chien-Chien L.

Part 6

Adams, Jeffrey
Allard, Jean-Luc
Anderson, Mary V.
Apfel, Kenneth
Atkinson, Richard V.
Bender, Robert K.
Boucek, Charles H.
Brissman, Mark D.
Cardoso, Ruy A.
Carlson, Christopher S.
Carlson, Karyl T.
Caron, Philippe
Closter, Donald L.
Comstock, Susan J.
Crane, Veronika K.
Curry, Michael K.
Danielson, Guy R.
Desjardins, Charles
Dezube, Janet B.
DiDonato, Anthony M.
Dodge, Scott H.
Doyle, Michael J.
Dunlap, George T., IV
Ericson, Janet M.
Fanning, William G.
Fletcher, James E.
DiGaetano, Mark
Goldberg, Steven B.
Graves, Gregory T.
Graves, Nancy A.
Johnson, Wendy A.
Johnston, Joyce M.
Jordan, Jeffrey R.
Kelly, Beverley A.
Kido, Chester T.
Kreps, Rodney E.
Kudera, Andrew E.
Lacko, Paul E.
Lewandowski, John J.
Lewis, Michael E.
Lombardi, Paul M.
Mair, Sharon A.
Malik, Sudershan
Miller, Susan M.
Ng, Kwok C.
Ollodart, Bruce E.
Ostergren, Gregory V.
Overgaard, Wade T.
Pelly, Brian G.
Plano, Richard A.
Privman, Boris
Procopio, Donald W.
Rhodes, Frank S.
Rice, Denise E.
Schustak, Marlene D.
Scruggs, Michael L.
Scully, Mark W.
Spalding, Keith R.
Sutter, Russel L.
Tan, Suan-Boon
Taylor, Craig P.
Taylor, R. Glenn
Vezina, Guy
Von Seggern, William J.
Wargo, Kelly A.
Whitehead, Guy H.
Wilson, Ernest I.
Wilson, Gregory S.
Yard, Roger A.
Yow, James W.
Yunque, Mark A.

Part 8

Almagro, Manuel, Jr.
Amundson, Richard B.
Becraft, Ina M.
Bellusci, David M.
Bennett, Robert S.
Bursley, Kevin H.
Cartmell, Andrew R.
Chen, Chyen
Cieslak, Walter P.
Colin, Barbara
DeConti, Michael A.
Deede, Martin W.
DeFalco, Thomas J.
Dekle, James M.
Dembiec, Linda A.
Diamantoukos, Christopher
Downing, Jeremiah M.
Dufresne, Jacques
Earwaker, Bruce G.
Edmondson, Alice H.
Elliott, Paula L.
Gapp, Steven A.
Guenthner, Denis G.
Haskell, Gayle E.
Henry, Thomas A.
Hollister, Jeanne M.
Johnson, Andrew P.
Kartechner, John W.
Kneuer, Paul J.
Koupf, Gary I.
Lacroix, Marthe A.
Levine, George M.
Littmann, Mark W.

Lyons, Daniel K.
Maguire, Brian P.
Martin, Paul C.
Mayer, Jeffrey H.
McClure, John W., Jr.
McDonald, Gary P.
Menning, David L.
Miller, William J.
Morgan, William S.
Mulder, Evelyn T.
Myers, Thomas G.
Noyce, James W.
Pei, Kai-Jaung
Putney, Alan K.
Rathjen, Ralph L.
Reppert, Daniel A.
Roth, Randy J.
Sandman, Donald D.
Schnapp, Frederic F.
Schultheiss, Peter J.
Shepherd, Linda A.
Siewert, Jerome J.
Silver, Melvin S.
Terrill, Kathleen W.
Thorrick, John P.
Townsend, Christopher J.
Trudeau, Michel
Veilleux, Andre
Visintine, Gerald R.
Votta, James C.
Wick, Peter G.
Williams, Robin M.
Willsey, Robert L.
Woodruff, Arlene F.

Part 10

Aldin, Neil C.
Barclay, David L.
Basson, Steven D.
Bear, Robert A.
Berry, Janice L.
Boyd, Wallis A.
Bursley, Kevin H.
Captain, John E.
Chansky, Joel S.
Clark, Daniel B.
Curran, Kathleen F.
Currie, Ross A.
Desilets, Claude
Deutsch, Robert V.
Dornfeld, James L.
Duffy, Brian
Dyck, N. Paul
Dye, Myron L.
Easlon, Kenneth
Einck, Nancy R.
Gillam, William R.
Grace, Gregory S.
Greco, Ronald E.
Hale, Jonathan B.
Hanson, Jeffrey L.
Hein, Timothy T.
Homan, Mark J.
Hoppe, Kenneth J.
Keller, Wayne S.
Krakowski, Israel
Lee, Robert H.
Lewis, Martin A.
Lipton, Barry C.
Mendelssohn, Gail A.
Montgomery, Warren D.
Murphy, Francis X., Jr.
Narvell, John C.
Normandin, Andre
Onufer, Layne M.
Pelletier, Charles A.
Ryan, John P.
Sanders, Robert L.
Schilling, Timothy L.
Sherman, Harvey A.
Smith, Michael B.
Spalla, Joanne S.
Treitel, Nancy R.
Webster, Patricia J.
White, Charles S.

The following candidates will be admitted as Fellows and Associates at the May, 1986 meeting as a result of their successful completion of the Society requirements in the November, 1985 examinations.

FELLOWS

Allaben, Mark S.
Bear, Robert A.
Berry, Janice L.
Boyd, Wallis A.
Clark, Daniel B.
Curran, Kathleen F.
Dornfeld, James L.
Hall, Allen A.
Hayward, Gregory L.
Lewis, Martin A.
Lipton, Barry C.
Mashitz, Isaac
Miller, Robert A., III
Murphy, William F.
Nester, Karen L.
Port, Rhonda D.
Smith, Michael B.
Treitel, Nancy R.
White, Charles S.

ASSOCIATES

Aldin, Neil C.
Almagro, Manuel, Jr.
Amoroso, Rebecca C.
Anderson, Mary V.
Apfel, Kenneth
Atkinson, Richard V.
Callahan, James J.
Carlson, Christopher S.
Caron, Louis-Philippe
Cascio, Michael J.
Cathcart, Sanders B.
Cellars, Ralph M.
Christhilf, David A.
Comstock, Susan J.
Cox, David B.
Davis, Dan J.
Debs, Raymond V.
Dekle, James M.
Doyle, Michael J.
Englander, Jeffrey A.
Fletcher, James E.
Forbus, Barbara L.
Gauthier, Richard
Gebhard, James J.
Gidos, Peter M.
Glicksman, Steven A.
Graham, Jeffrey H.
Guenthner, Denis G.
Hay, Randolph S.
Herbers, Joseph A.
Hertling, Richard J.
Homan, Mark J.
Johnson, Wendy A.
Kasner, Kenneth R.
Kneuer, Paul J.
Koegel, David
Kreps, Rodney E.
Kulik, John M.
Kuo, Chung-Kuo
Lacek, Mary Lou
Lacroix, Marthe A.
Lessard, Alain
Lyons, Mark D.
Mailloux, Patrick
McCoy, Mary E.
Millar, Leonard L.
Miller, Susan M.
Mohrman, David F.
Mueller, Robert A.
Musante, Donald R.
Newell, Richard T., Jr.
Newman, Henry E.
Ollodart, Bruce E.
Ostergren, Gregory V.
Overgaard, Wade T.
Peraine, Anthony
Pridgeon, Ronald D.
Privman, Boris
Rhodes, Frank S.
Rice, Denise E.
Rice, James W.
Roesch, Robert S.
Sandman, Donald D.
Scully, Mark W.
Shepherd, Linda A.
Sornberger, George C.
Spidell, Bruce R.
Steingiser, Russell
Sutter, Russel L.
Tan, Suan-Boon
Thompson, Robert W.
Tingley, Nanette

Tistan, Ernest S.
Trudeau, Michel
Turner, George W., Jr.
Von Seggern, William J.
Walker, David G.
Wargo, Kelly A.
Weber, Dominic A.
Woodruff, Arlene F.
Yen, Chung-Ye
Yow, James W.

The following is the list of successful candidates in examinations held in November, 1985.

Part 5

Allen, Danny M.
Almagro, Manuel, Jr.
Artes, Lawrence J.
Atkins, Heather E.
Barnes, Katharine E.
Bauer, Bruno P.
Benninghof, Kay E.
Boisvert, Paul, Jr.
Bonte, Sharon R.
Boudreau, Joseph J.
Brathwaite, Malcolm E.
Brehm, Paul J.
Burns, William E.
Byington, Jennifer S.
Cardoso, Ruy A.
Caulfield, Michael J.
Chabarek, Paul
Cloutier, Denis
Cofield, Joseph F.
Cote, Jean
Crawshaw, Mark
Crouse, John W.
Darby, Robert N.
Desjardins, Charles
DiDonato, Anthony M.
DiGaetano, Mark
Elliott, Angela F.
Evans, Karen F.
Fitzgerald, Beth E.
Fitzpatrick, William G.
Fletcher, James E.
Fonticella, Ross C.
Forbus, Barbara L.
Frank, Jacqueline B.
Gagnon, Luc
Gardner, Andrea
Gebhard, James J.
Gendelman, Nathan J.
Gill, Bonnie S.
Goldberg, Leonard R.
Graham, Jeffrey H.
Greenhill, Eric L.
Griffith, Ann V.
Groh, Linda M.
Gross, Marian R.
Hawley, Karin S.
Hayes, Thomas L.
Hays, David H.
Hebert, Norman P.
Hertling, Richard J.
Hess, Todd J.
Heyman, David R.
Hostetter, David B.
Hurley, John M.
James, Peter H.
Johnston, Joyce M.
Johnston, Steven J.
Jones, Terrell A.
Kaufman, David
Keatinge, Clive L.
Kerin, Allan A.
Kesby, Kevin A.
Kishi, Leslie K.
Klenow, Jerome F.
Kligman, Daniel F.
Kryczka, John R.
Lalonde, David A.
Lamb, Dean K.
Lebens, Joseph R.
Leiner, William W., Jr.
Lewandowski, John J.
Lewis, Michael E.
Lifschitz, David E.
Lyons, Mark D.
Mahoney, Michael W.
Marchena, Eduardo P.
McCreesh, James B.
McGill, Cassandra M.
McNichols, James P.
McShea, Christopher J.
Meyer, Robert J.
Millar, Leonard L.
Miller, Mary F.
Moylan, Thomas G.
Naylor, Walter R.
Nelson, Chris E.
Nemlick, Kenneth J.
Ng, Wai Hung
Nordquist, Randall S.
Nyce, G. Christopher

Oliver, Douglas W.
Peck, Steven C.
Pestcoe, Marvin
Phifer, Robert C.
Phillips, George N.
Pichler, Karen J.
Pino, Susan L.
Poirier, Denis
Popejoy, Kathy
Proska, Mark R.
Provencher, Yves
Rhoads, Karin M.
Roberts, Jonathan S.
Schadler, Thomas E.
Schug, Richard D.
Shimkus, Mary B.
Snook, Linda D.
Steinberg, Karen F.
Sterling, Mary E.
Stone, Edward C.
Sublett, Sharon
Swanstrom, Ronald J.
Szczepanski, Chester J.
Thorrick, John P.
Tremblay, Martin-Eric
Van de Water, John
Veilleux, Andre
Watkins, Nancy P.
Werland, Debra L.
Wilson, Gregory S.
Woodruff, Arlene F.
Yen, Chung-Ye
Yit, Bill S.
Yocius, Richard P.

Part 7

Aldin, Neil C.
Amoroso, Rebecca C.
Anderson, Mary V.
Apfel, Kenneth
Aquino, John G.
Atkinson, Richard V.
Atkinson, Roger A., III
Baker, Mark S.
Billings, Holly L.
Buchanan, John W.
Callahan, James J.
Carlson, Christopher S.
Caron, Louis-Philippe
Casale, Kathleen N.
Cascio, Michael J.
Cathcart, Sanders B.
Cellars, Ralph M.
Christhilf, David A.
Comstock, Susan J.
Cox, David B.
Cross, Susan L.
Davis, Brian W.
Davis, Dan J.
Debs, Raymond V.
Dekle, James M.
Diamantoukos, Christopher
Doyle, Michael J.
Edlefson, Dale R.
Englander, Jeffrey A.
Epstein, James C.
Feldblum, Sholom
Francis, Louise A.
Gauthier, Richard
Gidos, Peter M.
Glicksman, Steven A.
Guenthner, Denis G.
Harbage, Robin A.
Hay, Randolph S.
Herbers, Joseph A.
Hill, Tony D.
Homan, Mark J.
Huberman, Gloria L.
Johnson, Eric J.
Johnson, Wendy A.
Kasner, Kenneth R.
Kinson, Paul E.
Kneuer, Paul J.
Koegel, David
Koufacos, Constantine G.
Kreps, Rodney E.
Kulik, John M.
Kuo, Chung-Kuo
Lacek, Mary Lou
Lacroix, Marthe A.
Lessard, Alain
Liebers, Elise C.
Mailloux, Patrick
Math, Steven
McCoy, Mary E.
Miller, Susan M.
Mohrman, David F.
Mueller, Nancy D.
Mueller, Robert A.
Mulvaney, Mark W.
Musante, Donald R.
Musulin, Rade T.
Newell, Richard T., Jr.
Newman, Henry E.
Ollodart, Bruce E.
Ostergren, Gregory V.
Overgaard, Wade T.
Paglieri, Wayne C.
Peraine, Anthony
Pridgeon, Ronald D.
Privman, Boris

Rhodes, Frank S.
Rice, Denise E.
Rice, James W.
Roesch, Robert S.
Ryan, Beverley K.
Sandman, Donald D.
Schultze, Mark E.
Schwandt, Jeffory C.
Scully, Mark W.
Shepherd, Linda A.
Skov, Steven A.
Sornberger, George C.
Sperger, Mary Jean
Spidell, Bruce R.
Steingiser, Russell
Sutter, Russel L.
Svendsgaard, Christian
Sweeney, Eileen M.
Tan, Suan-Boon
Thompson, Robert W.
Tingley, Nanette
Tistan, Ernest S.
Trudeau, Michel
Turner, George W., Jr.
Von Seggern, William J.
Wachter, Christopher J.
Wacker, Gregory M.
Walker, David G.
Wargo, Kelly A.
Weber, Dominic A.
Weber, Robert A.
Whitehead, Guy H.
Williams, Robin M.
Wrobel, Edward M.
Yow, James W.

Part 9

Allaben, Mark S.
Bear, Robert A.
Bellafiore, Leonard A.
Bellusci, David M.
Berry, Janice L.
Blakinger, Jean M.
Boor, Joseph A.
Boyd, Wallis A.
Busche, George R.
Chiang, Jeanne D.
Clark, Daniel B.
Curran, Kathleen F.
Deede, Martin W.
DeLiberato, Robert V.
Dezube, Janet B.
Dornfeld, James L.
Dufresne, Jacques
Gardner, Robert W.
Hall, Allen A.
Halpern, Nina S.
Harwood, Catherine B.
Hayward, Gregory L.
Hollister, Jeanne M.
Johnson, Andrew P.
Kelley, Robert J.
Klinker, Fredrick L.
Koupf, Gary I.
Lewis, Martin A.
Lipton, Barry C.
Littmann, Mark W.
Livingston, Roy P.
Loper, Dennis J.
Lyons, Daniel K.
Mashitz, Isaac
Menning, David L.
Miller, David L.
Miller, Robert A., III
Murphy, William F.
Myers, Thomas G.
Nester, Karen L.
Noyce, James W.
Port, Rhonda D.
Post, Jeffrey H.
Putney, Alan K.
Reppert, Daniel A.
Roth, Randy J.
Scheuing, Jeffrey R.
Schultheiss, Peter J.
Sealand, Pamela J.
Silver, Melvin S.
Slotznick, Lisa A.
Slusarski, John
Smith, Michael B.
Smith, Richard A.
Terrill, Kathleen W.
Townsend, Christopher J.
Treitel, Nancy R.
Visner, Steven M.
Walsh, Michael C.
Weinman, Stacy J.
White, Charles S.
Willsey, Robert L.
Woomer, Roy T., III

OBITUARIES

WALTER T. EPPINK
HUGH P. HAM
EXEQUIEL S. SEVILLA

WALTER T. EPPINK
1902–1984

Walter T. Eppink, a Fellow of the Casualty Actuarial Society since 1935, and a charter member of the American Academy of Actuaries, died of congestive heart failure on August 20, 1984, at the age of 82.

Walter was a graduate of Western Reserve University (now known as Case Western Reserve University) in Cleveland.

During his fifty year career at Merchants Mutual Insurance Company he held various positions, including Assistant Treasurer, Treasurer, Actuary, and Assistant Vice President. At the time of his retirement in 1972 he was Executive Vice President of the company.

Walter thoroughly enjoyed his work and his association with co-workers at Merchants Mutual. He was well known for his kindness and helpfulness.

Walter was a devoted family man. Following his retirement at the age of 70, he and his wife traveled extensively throughout North America visiting their children and grandchildren.

Walter is survived by his wife, Marion, two sons, Richard and Robert, sixteen grandchildren, and one great-grandchild.

HUGH P. HAM
1905–1984

Hugh P. Ham, an Associate of the Casualty Actuarial Society since 1936, died April 6, 1984 at the age of 78.

Before retiring due to ill health in 1966, Hugh spent more than forty-two years in the service of the Western-British America Group of Insurance Com-

Rhodes, Frank S.
Rice, Denise E.
Rice, James W.
Roesch, Robert S.
Ryan, Beverley K.
Sandman, Donald D.
Schultze, Mark E.
Schwandt, Jeffory C.
Scully, Mark W.
Shepherd, Linda A.
Skov, Steven A.
Sornberger, George C.
Sperger, Mary Jean
Spidell, Bruce R.
Steingiser, Russell
Sutter, Russel L.
Svendsgaard, Christian
Sweeney, Eileen M.
Tan, Suan-Boon
Thompson, Robert W.
Tingley, Nanette
Tistan, Ernest S.
Trudeau, Michel
Turner, George W., Jr.
Von Seggern, William J.
Wachter, Christopher J.
Wacker, Gregory M.
Walker, David G.
Wargo, Kelly A.
Weber, Dominic A.
Weber, Robert A.
Whitehead, Guy H.
Williams, Robin M.
Wrobel, Edward M.
Yow, James W.

Part 9

Allaben, Mark S.
Bear, Robert A.
Bellafiore, Leonard A.
Bellusci, David M.
Berry, Janice L.
Blakinger, Jean M.
Boor, Joseph A.
Boyd, Wallis A.
Busche, George R.
Chiang, Jeanne D.
Clark, Daniel B.
Curran, Kathleen F.
Deede, Martin W.
DeLiberato, Robert V.
Dezube, Janet B.
Dornfeld, James L.
Dufresne, Jacques
Gardner, Robert W.
Hall, Allen A.
Halpern, Nina S.
Harwood, Catherine B.
Hayward, Gregory L.
Hollister, Jeanne M.
Johnson, Andrew P.
Kelley, Robert J.
Klinker, Fredrick L.
Koupf, Gary I.
Lewis, Martin A.
Lipton, Barry C.
Littmann, Mark W.
Livingston, Roy P.
Loper, Dennis J.
Lyons, Daniel K.
Mashitz, Isaac
Menning, David L.
Miller, David L.
Miller, Robert A., III
Murphy, William F.
Myers, Thomas G.
Nester, Karen L.
Noyce, James W.
Port, Rhonda D.
Post, Jeffrey H.
Putney, Alan K.
Reppert, Daniel A.
Roth, Randy J.
Scheuing, Jeffrey R.
Schultheiss, Peter J.
Sealand, Pamela J.
Silver, Melvin S.
Slotznick, Lisa A.
Slusarski, John
Smith, Michael B.
Smith, Richard A.
Terrill, Kathleen W.
Townsend, Christopher J.
Treitel, Nancy R.
Visner, Steven M.
Walsh, Michael C.
Weinman, Stacy J.
White, Charles S.
Willsey, Robert L.
Woomer, Roy T., III

NEW FELLOWS ADMITTED MAY 1985 (Left to Right): First row: C. K. Khury (President), Francois Bertrand, Terry Biscoglia, Donald Palmer, Loyd Fueston; Second row: Heidi Hutter, Lois Ross, Raja Bhagavatula, Jeffrey Carlson; Third row: Alan Hapke, Michael McSally, Diane Symnoski, Robert Meyer, Warren Ehrlich; Fourth row: James Surrago, William Biegaj, Allan Neis, Stephan Christiansen, John Forney.

NEW ASSOCIATES ADMITTED MAY 1985 (Left to Right): First row: William Miller, Arthur Placek, Mark Allaben, Christy Gunn, Barry Lipton, David Scholl, Jeffrey Salton; Second row: Robert Muller, Richard Quintano, Roger Schultz, Jeffrey Scheuing, Robert Willsey, Frederick Cripe; Third row: Thomas Myers, Edward Somers, Brian Maguire, Robert Whitlock, Brian Brown, Jeffrey Post; Fourth row: Robert Gardner, Warren Montgomery, Robert Lee, Daniel Gogol, Leonard Bellafiore, Andrew Cartmell; Fifth row: Mark Littmann, Daniel Reppert, Jerry Visintine, Jacques Dufresne, Thomas DeFalco, Martin Lewis; Sixth row: Kenneth Easlon, Susan Woerner, Arlyn Shapiro, Robert Mucci, Kevin Greaney; Seventh row: Kirk Fleming, Michael Smith, Kathleen Curran, Kathleen Terrill, Jeanne Hollister, Ruth Howald; Eighth row: Stacy Weinman, Nancy Treitel, Daniel Clark, William Carpenter, Janice Cutler, Joseph Theisen; Ninth row: C. K. Khury (President), Eugene McGovern, John Slusarski.

NEW FELLOWS ADMITTED NOVEMBER 1985 (Left to Right): Front row: Claude Desilets, Ronald Greco, John Captain, Andre Normandin, Evelyn Toni Mulder, Gail Mendelssohn, C. K. Khury (President), Joanne Spalla, Linda Dembiec, Robert Kaplan, Kevin Bursley, Steven Basson, Brian Duffy, Jeffrey Mayer; Back row: Joel Chansky, N. Paul Dyck, Robert Sanders, Harvey Sherman, Gregory Grace, Timothy Hein, Kenneth Hoppe, Robert Deutsch, John Narvell, Ross Currie.

NEW ASSOCIATES ADMITTED NOVEMBER 1985 (Left to Right): Roger Yard, Andrew Kudera, Mark Yunque, Robert Kaplan, Sharon Mair, Janet Dezube, Jeff Jordan, Steven Goldberg, Scott Dodge.

OBITUARIES

WALTER T. EPPINK
HUGH P. HAM
EXEQUIEL S. SEVILLA

WALTER T. EPPINK
1902–1984

Walter T. Eppink, a Fellow of the Casualty Actuarial Society since 1935, and a charter member of the American Academy of Actuaries, died of congestive heart failure on August 20, 1984, at the age of 82.

Walter was a graduate of Western Reserve University (now known as Case Western Reserve University) in Cleveland.

During his fifty year career at Merchants Mutual Insurance Company he held various positions, including Assistant Treasurer, Treasurer, Actuary, and Assistant Vice President. At the time of his retirement in 1972 he was Executive Vice President of the company.

Walter thoroughly enjoyed his work and his association with co-workers at Merchants Mutual. He was well known for his kindness and helpfulness.

Walter was a devoted family man. Following his retirement at the age of 70, he and his wife traveled extensively throughout North America visiting their children and grandchildren.

Walter is survived by his wife, Marion, two sons, Richard and Robert, sixteen grandchildren, and one great-grandchild.

HUGH P. HAM
1905–1984

Hugh P. Ham, an Associate of the Casualty Actuarial Society since 1936, died April 6, 1984 at the age of 78.

Before retiring due to ill health in 1966, Hugh spent more than forty-two years in the service of the Western-British America Group of Insurance Companies. He joined the staff in Winnipeg in 1924; was transferred to the head office in Toronto in 1943; was appointed General Manager in 1952, and a Director in 1958. In 1960, he was named President of the Western-British America Companies and, following the acquisition of those companies by the Royal Insurance Group, also was appointed a General Manager for Canada of the Royal Insurance Group.

Hugh is survived by his wife, Dorothy Emma.

EXEQUIEL S. SEVILLA
1904–1985

Exequiel Sevilla, an Associate of the Casualty Actuarial Society since 1930, died on January 6, 1985 at his home in Manila, Philippines at the age of 80. He had suffered a massive stroke in 1979 and had been bedridden since then.

Exequiel Sevilla was born on March 4, 1904 in Manila. In 1927, he graduated summa cum laude from the University of the Philippines. Even before graduation, he had been appointed an Insurance Examiner by Dr. Emeterio Roa, the first Filipino Actuary. The territorial government sent him as a scholar to the University of Michigan, where he graduated with a Master of Science degree in Actuarial Mathematics in 1929. He trained for one year at the United States Life Insurance Company in New York City.

Upon his return to the Philippines, he was appointed Actuary in the Office of the Insurance Commissioner. In 1933, he left the government service to help found the National Life Insurance Company in Manila. He stayed with this company, first as Actuary, then as General Manager, and finally as President and member of the Board of Directors, until his retirement in 1974. Following his retirement, he continued to serve the company in a consulting role.

In 1937, President Quezon appointed Mr. Sevilla a member of the first Board of Directors of the Government Service Insurance System. He taught mathematics at the University of the Philippines, Far Eastern University, and the University of the East.

Mr. Sevilla was a member of the American Academy of Actuaries and the International Actuarial Association, a corresponding member of the Instituto de Actuarios Espanioles and a fellow of the Actuarial Society of the Philippines.

He served as President of the Actuarial Society of the Philippines, the Philippine Statistical Association, the Philippine Association of Life Insurance Companies and the Advanced Management Association of the Far East.

Mr. Sevilla is survived by his wife, Lourdes de Veyra Sevilla, his six children, Josefina, Exequiel Jr., Eduardo, Ernesto, Silvia and Aida, and thirteen grandchildren. His first wife, Susanna Guidote Sevilla, died in 1956.

INDEX TO VOLUME LXXII

INDEX—CONTINUED

INDEX—CONTINUED

Page