VOLUME LXXXV NUMBERS 162 AND 163

BOOK 1 OF 2

PROCEEDINGS

OF THE

Casualty Actuarial Society

ORGANIZED 1914

1998

VOLUME LXXXV

Number 162—May 1998

Number 163—November 1998

Library of Congress Catalog No. HG9956.C3
ISSN 0893-2980

Printed for the Society by
United Book Press
Baltimore, Maryland

Typesetting Services by
Minnesota Technical Typography, Inc.
St. Paul, Minnesota

FOREWORD

Actuarial science originated in England in 1792 in the early days of life insurance. Because of 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. Eight years later, in Scotland, the Faculty of Actuaries was formed. In the United States, the Actuarial Society of America was formed in 1889 and the American Institute of Actuaries in 1909. These two American organizations merged in 1949 to become the Society of Actuaries.

In the early years of the 20th 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 Casualty Actuarial and Statistical Society of America in 1914. Dr. I. M. Rubinow, who was responsible for the Society's formation, became its first president. At the time of its formation, the Casualty Actuarial and Statistical Society of America had 97 charter members of the grade of Fellow. The Society adopted its present name, the Casualty Actuarial Society, on May 14, 1921.

The purposes of the Society are to advance the body of knowledge of actuarial science applied to property, casualty, and similar risk exposures, to establish and maintain standards of qualification for membership, to promote and maintain high standards of conduct and competence for the members, and to increase the awareness of actuarial science. The Society's activities in support of this purpose include 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 at the time of the Society's formation, 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, commercial multiple peril, and others. These contributions are found principally in original papers prepared by members of the Society and published annually in the *Proceedings of the Casualty Actuarial Society*. 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 industry over the years.

The membership of the Society includes actuaries employed by insurance companies, industry advisory 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 require successful completion of examinations, held in February, and in the spring and fall of each year in various cities of the United States, Canada, Bermuda, and selected overseas sites. In addition, Associateship requires completion of the CAS Course on Professionalism.

The publications of the Society and their respective prices are listed in the Society's *Yearbook*. The *Syllabus of Examinations* outlines the course of study recommended for the examinations. Both the *Yearbook*, at a charge of $40 (in U.S. funds), and the *Syllabus of Examinations*, without charge, may be obtained from the Casualty Actuarial Society, 1100 North Glebe Road, Suite 600, Arlington, Virginia 22201.

JANUARY 1, 1998

EXECUTIVE COUNCIL*

MAVIS A. WALTERS . . . .*President*
STEVEN G. LEHMANN . . . .*President-Elect*
CURTIS GARY DEAN . . . .*Vice President–Administration*
KEVIN B. THOMPSON . . . .*Vice President–Admissions*
SUSAN T. SZKODA . . . .*Vice President–Continuing Education*
PATRICK J. GRANNAN . .*Vice President–Programs & Communications*
ROBERT S. MICCOLIS . . . . .*Vice President–Research & Development*

THE BOARD OF DIRECTORS

*Officers**

MAVIS A. WALTERS . . . .*President*
STEVEN G. LEHMANN . . . .*President-Elect*

Immediate Past President†

ROBERT A. ANKER 1998

Elected Directors†

REGINA M. BERENS 1998
DAVID R. CHERNICK 1998
C. K. "STAN" KHURY 1998
DAVID L. MILLER 1998
SHOLOM FELDBLUM 1999
ALICE H. GANNON 1999
DAVID N. HAFLING 1999
RICHARD J. ROTH JR. 1999
PAUL BRAITHWAITE 2000
JEROME DEGERNESS 2000
MICHAEL FUSCO 2000
STEPHEN P. LOWE 2000

* Term expires at the 1998 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.

1998 PROCEEDINGS
CONTENTS OF VOLUME LXXXV

Page

1998 PROCEEDINGS
CONTENTS OF VOLUME LXXXV

Page

1998 PROCEEDINGS
CONTENTS OF VOLUME LXXXV

1998 PROCEEDINGS
CONTENTS OF VOLUME LXXXV

Page

NOTICE

Papers submitted to the *Proceedings of the Casualty Actuarial Society* are subject to review by the members of the Committee on Review of Papers and, where appropriate, additional individuals with expertise in the relevant topics. In order to qualify for publication, a paper must be relevant to casualty actuarial science, include original research ideas and/or techniques, or have special educational value, and must not have been previously copyrighted or published or be concurrently considered for publication elsewhere. Specific instructions for preparation and submission of papers are included in the *Yearbook* of the Casualty Actuarial Society.

The Society is not responsible for statements of opinion expressed in the articles, criticisms, and discussions published in these *Proceedings*.

Volume LXXXV, Part 1 **No. 162**

PROCEEDINGS
May 17, 18, 19, 20, 1998

A COMPARISON OF PROPERTY/CASUALTY INSURANCE FINANCIAL PRICING MODELS

STEPHEN P. D'ARCY AND RICHARD W. GORVETT

Abstract

A number of property/casualty insurance pricing models that attempt to integrate underwriting and investment performance considerations have been proposed, developed, and/or applied. Generally, empirical tests of these models have involved examining how well the models fit historical data at an industry level. This paper demonstrates how to apply a variety of property/casualty insurance financial pricing techniques to a single hypothetical, but representative, company. Both company and economic parameters are varied in order to examine the sensitivity of indicated underwriting profit margins from these techniques to different company situations and economic environments, and to highlight the differences between the techniques at a practical level. This paper *also serves as a practical guide for applying these models in order to encourage more widespread use of these approaches.*

1. INTRODUCTION

The determination of a "fair," or competitive, rate of return for property/casualty insurance underwriting operations has been the subject of increasing scrutiny over the last several decades among both academics and insurance practitioners. The five percent target underwriting profit margin promulgated by the National Convention of Insurance Commissioners in 1921 represented the first of many techniques that have been considered, and in some cases employed, to determine a fair rate of return. Although that first approach had little, if any, statistical or financial foundation, subsequent methods have attempted to determine insurance prices more rigorously, and with due consideration given to relevant insurance, economic, and financial market characteristics. An appropriate determination of fair insurance prices is important because capital will be attracted to—and retained by—the insurance industry only if its rates of return are comparable to those in other industries that are perceived to have similar levels of risk.

A variety of financial pricing models has now been proposed for property/casualty insurance, including the Target Total Rate of Return approach, the Capital Asset Pricing Model, several Discounted Cash Flow approaches, the Option Pricing Model, and the Arbitrage Pricing Model. In general, these models have been applied individually and without clearly showing how the necessary parameters can be determined from insurance financial statements. Several important studies do provide a degree of comparison among the different models. Myers and Cohn [25] compare the discounted cash flow model and the insurance CAPM, including sensitivity analysis of the various parameters. Cummins [9] provides a comparison of the discounted cash flow model and internal rate of return approach and illustrates the results of each method on one set of data. Doherty and Garven [15] contrast the insurance CAPM and the option pricing models over a range of values for each parameter. However, there has been no systematic comparison of all the financial pricing models or any documentation explaining how the relevant parameters should be determined for a particular insurer. This paper addresses those needs, first generating a financial statement for a hypothetical, but representative, insurer, and then applying each pricing model to this insurer to determine the appropriate premium level and underwriting profit margin. Finally, the models are examined over a range of parameter values that occur across insurers and over time to demonstrate which parameters need to be measured most accurately, and which models are most impacted by changes in different variables. This analysis illustrates potential strengths and weaknesses of each technique. By comparing the indications of fair underwriting profit margins under each of these pricing methods, their differences will be highlighted. This will allow both company management and regulators to better gauge the potential impact on prices of adopting one or another technique in various business environments.

The insurance pricing techniques applied to our representative insurance company include:

- target underwriting profit margin model,
- target total rate of return model,
- insurance capital asset pricing model,
- discounted cash flow (Myers–Cohn) model,
- internal rate of return model,
- option pricing model,
- arbitrage pricing model.

The company to which these techniques are applied is a fictitious entity, but quite representative of companies actively involved in the property/casualty insurance industry. As many of the techniques examined are best applied in a single line of business framework, we have chosen to model a company that writes only private passenger automobile (PPA) insurance. Representative financial values and ratios, as well as payout patterns, were selected based on an examination of both aggregate industry and individual company values. Values for other economic and insurance industry variables are derived from appropriate sources as described in Section 3 of the paper. The considerations involved in obtaining each of the parameters used in the models are shown, in order to illustrate how a company could use each technique.

2. REVIEW OF THE ALTERNATIVE PRICING MODELS

In 1921, the National Convention of Insurance Commissioners, by an overwhelming margin, approved the Majority report of the Committee on Fire Insurance. For two years, the Committee had been considering the issue of what was a reasonable underwriting profit margin. The report's conclusions included the following items:

- "Underwriting profit (or loss) is arrived at by deducting from earned premiums, all incurred losses and incurred expenses."
- "A reasonable underwriting profit is 5 percent, plus 3 percent for conflagrations … "

(See National Convention of Insurance Commissioners [28] and National Association of Insurance Commissioners [26] for more details.) A minority report recommended that investment income also be considered in determining a reasonable profit provision, but this recommendation was defeated (see Webb [35]). Thus, the position of the insurance regulatory community at that time was that only underwriting, and not investment, operations were relevant to the determination of a reasonable property/casualty profit level. Furthermore, the specific profit level recommended, five percent, was established apparently without meaningful statistical support.

Subsequent studies and reports began to question the appropriateness of ignoring investment income. This concern intensified in the 1960s and 1970s, as interest rates, and their volatility, increased. The National Association of Insurance Commissioners (NAIC) in 1970 [27] said that, "In determining profits, it is submitted that income from all sources should be considered." The NAIC, however, while criticizing the 1921 formula, did not recommend an alternative until its 1984 Investment Income Task Force Report, which recommended that the total rate of return on net worth should be used to measure insurance profitability.

In the meantime, actuaries started to develop (and sometimes use) several alternative pricing techniques that attempt to address both underwriting and investment considerations in pricing property/casualty insurance policies. Initially, something of a dichotomy existed among the techniques proposed: some concentrated on the underwriting side of the insurance process, with little consideration given to meaningful analysis of the investment process; others focused primarily on the investment side, without adequate understanding of the unique aspects of the insurance underwriting process. Recent research has attempted to give appropriate consideration to *both* aspects of the property/casualty insurance business.

This paper examines seven different pricing models, and applies each to a fictitious but representative insurance company. Each of the seven techniques is described below; additional details regarding specific calculations for each of the financial models are included in the Appendix. Variables used in the following formulas include:

P = premium

UPM = underwriting profit margin

L = losses and loss adjustment expenses

E = other expenses

S = equity (or adjusted statutory surplus)

IA = invested assets

IR = investment return

r_e = return on equity

r_f = risk-free rate of return

r_m = market rate of return

β_e = beta of the insurance company's stock

β_u = beta of the insurance underwriting process

k = funds-generating coefficient

t_i = tax rate on investment income

t_u = tax rate on underwriting income.

A. *Target Underwriting Profit Margin Model*

The Target Underwriting Profit Margin (Target UPM) Model determines an appropriate premium for a property/casualty insurance policy based upon a pre-selected underwriting profit margin. Thus, no consideration is given to the investment earnings produced by the insurance policy due to either the allocation of surplus in support of the policy or the delay between receipt of the premium and payment of the losses and expenses. The premium is determined strictly as a function of the expected losses, expenses, and the target underwriting profit margin as a percentage of premium. Historically, the target UPMs used have typically been 2.5 percent for workers compensation, and 5 percent for all other lines in most jurisdictions. However, in 1986, Florida adopted rule 4ER86-1 that established a formal procedure for including investment income in the ratemaking process by adjusting the target UPM downward to reflect the additional investment income attained in long-tailed lines over short-tailed lines.

For our representative PPA insurance company, the pricing and profit equations are:

$$UPM = 0.05 \tag{2.1}$$

$$P = \frac{L+E}{1-UPM} \qquad \text{or} \qquad UPM = 1 - \frac{L}{P} - \frac{E}{P}. \tag{2.2}$$

While this approach has been used in the property/casualty insurance industry for decades, and is relatively simple to apply, it is clearly the least "financially sophisticated" of the pricing models examined in this study, and in fact—efforts such as Florida's notwithstanding—the Target UPM model is not supported by financial considerations.

B. Target Total Rate of Return Model

A straightforward way of incorporating investment income into the ratemaking calculation is simply to target, rather than merely the underwriting margin, the *combined underwriting and investment* returns of an insurance policy. The total rate of return of a policy is viewed as having two components: investment and underwriting. If two of these three items are known, the third can be derived. Thus, in the Target Total Rate of Return (Target TRR) Model, the underwriting profit margin is determined based on a selected total rate of return and an estimate of the investment income on a policy. This is analogous to the process that has been historically used for the utility industry.

The target total rate of return can be calculated as

$$TRR = \frac{(IA \times IR) + (P \times UPM)}{S}. \tag{2.3}$$

The TRR reflects both investment income (the first term in Equation 2.3) and underwriting income (the second term) as a proportion of equity. Solving Equation 2.3 for the underwriting profit margin yields

$$UPM = \frac{(S \times TRR) - (IA \times IR)}{P}. \tag{2.4}$$

Now we need to specify an appropriate target total rate of return. As with utility regulation, this is the crux of the model. Although any number of methods might be viable, the Capital Asset Pricing Model (CAPM) has typically been used to select the TRR. This approach will be used in this paper. The CAPM formula is:

$$\mathrm{E}[r_e] = r_f + \beta_e(\mathrm{E}[r_m] - r_f), \tag{2.5}$$

where β_e is defined as

$$\beta_e = \frac{\mathrm{Cov}(r_e, r_m)}{\mathrm{Var}(r_m)}.$$

The TRR is then set equal to the expected return on equity, or the cost of equity capital, $\mathrm{E}[r_e]$. Thus, substituting Equation 2.5 into Equation 2.4 yields:

$$\begin{aligned} UPM &= \frac{(S \times [r_f + \beta_e(\mathrm{E}[r_m] - r_f)] - (IA \times IR)}{P} \qquad \text{or} \\ UPM &= \frac{S}{P}\left([r_f + \beta_e(\mathrm{E}[r_m] - r_f)] - \frac{IA \times IR}{S}\right) \end{aligned} \tag{2.6}$$

C. Insurance Capital Asset Pricing Model

The CAPM was first introduced into the finance literature in the mid-1960s by Sharpe [32], Lintner [22], and Mossin [24]. The CAPM, as described in Equation 2.5, expresses expected return on equity as consisting of two components: a risk-free component and a risk premium, which is essentially a reward for taking on risk. The degree of compensation for risk-taking is measured by the equity beta, which quantifies systematic, as opposed to nonsystematic (or diversifiable), risk. Diversifiable risk is not compensated by the market, since it can be eliminated through an appropriate investment diversification strategy.

The CAPM has been applied to insurance by several authors. Among the first were Biger and Kahane [5]. Fairley [16] developed the following underwriting profit margin formula based on the CAPM:

$$UPM = -kr_f + \beta_u(\mathrm{E}[r_m] - r_f). \tag{2.7}$$

Here, the appropriate underwriting profit margin is calculated as the risk premium associated with the systematic risk of the insurance underwriting process, offset by investment income, which is credited at the risk-free rate of return. The funds-generating coefficient reflects the fact that the insurance process produces investable assets generated by premium income prior to payout of expenses and claims. This coefficient is often estimated by a reserves-to-premium ratio. For a steady state insurer, this approach would be correct; if the company has changed premium or exposure volume, however, this calculation would need to be refined.

In addition to Fairley, Hill [18] and Hill and Modigliani [19] have also developed CAPM applications to property/casualty insurance. In particular, Hill and Modigliani have developed a model that considers the impact of taxes and in fact allows for differential tax rates. Letting t_i and t_u be tax rates as defined above, the Hill and Modigliani model can be expressed as:

$$UPM = -kr_f \frac{1 - t_i}{1 - t_u} + \beta_u(\mathrm{E}[r_m] - r_f) + \frac{S}{P} r_f \frac{t_i}{1 - t_u}. \tag{2.8}$$

It is Equation 2.8 that is modeled in this study.

D. Discounted Cash Flow Model (Myers–Cohn)

The Discounted Cash Flow (DCF) Model was developed for use in Massachusetts as a counterpart to the CAPM model that had been used there beginning in the 1970s. The model is described by Myers and Cohn [25] and takes the following general form:

$$P = \mathrm{PV}(L + E) + \mathrm{PV}(UWPT) + \mathrm{PV}(IBT), \tag{2.9}$$

where

PV = present value operator,

UWPT = tax generated on underwriting income, and

IBT = tax generated on income from the investment balance.

One of the keys to using the DCF model is to properly determine a method of discounting each component of the above equation. Those cash flows that are certain should be discounted at the risk-free rate, while risky cash flows must be discounted at an appropriate risk-adjusted rate.

D'Arcy and Garven [14] test the following DCF model, where all cash flows are discounted based on the risk-free rate (which is equivalent to assuming that $\beta_u = 0$ in the CAPM):

$$\begin{aligned} 1 = {} & \mathrm{PV}\left(\frac{E}{P}\right) + \mathrm{PV}\left(\frac{L}{P}\right) + \mathrm{PV}\left(t\left[1 - \frac{E}{P} - \frac{L}{P}\right]\right) \\ & + \mathrm{PV}\left(t\left[1 + \frac{S}{P}\right]\frac{L}{P}LPP\right), \end{aligned} \tag{2.10}$$

where LPP = the loss payout pattern. This equation is solved for L/P, the loss ratio. Then, the indicated UPM is calculated as $1 - (L/P) - (E/P)$.

In this study, we use this general approach, but refine it to reflect the discounting of risky cash flows at a risk-adjusted rate. Specifically, in order to determine an indicated DCF premium level, we have used Equation 6.2 of D'Arcy and Dyer [13], with the enhancement that different tax rates are allowed on underwriting versus investment operations. The UPM is then determined as $(P - E - L)/P$.

E. Internal Rate of Return Model

Whereas the Myers–Cohn discounted cash flow model described above considers flows between the insurer and the policyholder, the internal rate of return (IRR) model, for example as used by the National Council on Compensation Insurance (NCCI), looks at flows between the investor and the company. In particular, the flows under the IRR model include the commitment of surplus, the release of surplus, the investment income, and the underwriting profit (both of the last two being net of applicable taxes). The discount rate of these flows is solved for, so that the present value of the flows is zero; then, this discount rate is compared to the cost of capital. A financially fair premium is determined by setting the IRR equal to the cost of capital.

In this study, we have used the same approach as Cummins [9]. The cost of capital is determined by the CAPM. Exhibit 6, Part 2 displays the calculation of the IRR model fair premium for the base case.

F. Option Pricing Model

Recently, the option pricing model (OPM) has received increasing attention among both insurance academics and practitioners. The OPM is seen as having a great deal of promise as a property/casualty insurance pricing framework since an insurance policy can, essentially, be viewed as a package of contingent claims. The primary application of the OPM to property/casualty insurance to date is Doherty and Garven [15], who show that the present values of the claims held by the three claimholders to an insurance contract—shareholders, policyholders, and the tax authorities (government)—can be modeled as European call options. In order to actually value these claims, and then determine a competitive UPM and premium, Doherty and Garven assume two alternative valuation frameworks:

- asset returns are normally distributed, and investors exhibit constant absolute risk aversion (CARA) with regard to their preferences;
- asset returns are lognormally distributed, and investors exhibit constant relative risk aversion (CRRA) with respect to their preferences.

Although closed-form solutions are not derived, the premiums and UPMs can be found for both frameworks via a straightforward iteration process. The appropriate formulas in Doherty and Garven relating to these two valuation assumptions are their Equations 19 and 30, respectively. We apply the first of these two models in this study. (See the Appendix for further details.) A spreadsheet was created wherein the difference between the market value of the residual claim of the shareholders (V) and the initial paid-in equity (S) is "backsolved" to zero by varying the premium; the solving value represents the fair premium indication. This premium is net of expenses, which are then added in. The UPM is calculated as $(P^* - L - E)/P^*$, where P^* includes expenses.

G. *Arbitrage Pricing Model*

The Arbitrage Pricing Theory (APT), developed initially by Ross [31] and extended by Roll and Ross [30] and others, is, like the CAPM, an equilibrium model of security returns. However, the APT makes fewer assumptions than does the CAPM, and it also admits the possibility of more than one "factor" to which security returns are sensitive. The theory behind the APT specifies neither the number of such factors, nor their identity. Unlike the CAPM, the APT does not posit a special, or even necessarily *any*, role for a market return in determining individual security returns.

According to the Arbitrage Pricing Theory, security returns follow the process

$$r_i = \mathrm{E}[r_i] + \sum_{j=1}^{J} \beta_{ij} f_j + \epsilon_i, \tag{2.11}$$

where

r_i = return on the ith security,

β_{ij} = the sensitivity of the return on the ith security to the jth factor, and

f_j = a factor that influences security returns.

Then, the absence of arbitrage requires that the excess return on each security be a linear combination of the betas:

$$\mathrm{E}[r_i] - r_f = \sum_{j=1}^{J} \beta_{ij} \lambda_j, \tag{2.12}$$

where

λ_j = the risk premium corresponding to the factor f_j.

The APT has been applied to insurance by Kraus and Ross [21] and Urrutia [34]. Urrutia derives UPM formulae for an Arbitrage Pricing Model (APM), based on the above theoretical relationships. His differential-tax UPM equation takes the form

$$UPM = -\frac{1 - t_i}{1 - t_u} r_f k + \sum_{j=1}^{J} \beta_{UPM,j} \lambda_j + r_f \left(\frac{S}{P}\right) \frac{t_i}{1 - t_u}, \tag{2.13}$$

where

$$\beta_{UPM,j} = \frac{\mathrm{Cov}(UPM, f_j)}{\mathrm{Var}(f_j)}.$$

Generally, there are two approaches to testing the APT model. The first involves factor analysis, a statistical methodology that determines factors and betas that best "explain" the data (i.e., that minimize the covariance of residual returns). The second involves the pre-specification of variables that are hypothesized to influence returns, as in Chen, Roll, and Ross [7]. This second approach allows for economic intuition in the interpretation of results; it is this approach which we use in this paper. In particular, a number of macroeconomic variables were tested, with the inflation rate and the growth in industrial production being the two variables that appear most significant in explaining historical underwriting profit margins. Multivariate regression analysis is used to determine sensitivities of UPMs to these two variables. Selected parameter values are incorporated into Equation 2.13 to determine fair UPMs.

3. DATA AND METHODOLOGY

In addition to the selection of the pricing techniques and the identification of the appropriate formulas for each, as documented above, the following steps were involved in this study:

- development of the representative statutory company model,
- collection and development of information regarding company and economic variables,
- application of each of the pricing techniques to the representative company, and
- sensitivity tests of the models by varying certain company and economic parameters.

Each of these steps is discussed below.

A. Development of the Representative Statutory Company Model

It was decided, for the sake of simplicity and clarity of presentation, to concentrate on a fictitious but representative property/casualty insurance company that writes only one line of business in one state. Private passenger automobile insurance was selected due to its size and significance in the industry. Other lines can easily be modeled by the same techniques presented here, with appropriate changes in parameters.

The 1994 editions of several A. M. Best publications provided the basis for the development of the representative company model. The most recent statement year reflected in these editions is 1993. To begin, some basic financial values for the largest PPA companies in the industry were accumulated. Exhibit 9 summarizes the asset, liability, surplus, and net written premium values for the main PPA companies within each of the 20 largest PPA groups. The calculated ratios vary considerably, sometimes due to different operating philosophies, sometimes because a company writes a large amount of other business in addition to PPA. These ratios served as the basis for certain company parameter ranges, discussed later, to test model sensitivity.

Pages 2, 3, and 4 of the Property/Casualty Annual Statement are simulated for the representative insurance company in Exhibit 10. These simulated pages were developed from the consolidated data from A. M. Best. First, 0.1% of consolidated industry (all lines) earned premium was taken as the starting point for the fictitious company (Exhibit 10, Part 3). Then, company asset, liability, and surplus values were derived. Total assets for the company were calculated by applying an industry asset/earned premium ratio to the selected company earned premium. Consolidated industry total values for specific asset, liability, and income categories were compared to aggregate figures for companies in which private passenger automobile and homeowners predominate. Comparisons were generally made on the basis of percentages relative to the appropriate major item—e.g., each asset item as a percentage of total assets. These percentages are shown on the first three sheets of Exhibit 10. Generally, the percentages applying to PPA-and-homeowners-predominating companies were used as the basis for our selected company values.

The selected percentages—as well as the resulting asset, liability, and income items—for the representative company are shown on Exhibit 10, Parts 1 through 3.

PPA loss development patterns for the representative company were determined by analyzing the consolidated Schedule P data from A. M. Best, using standard actuarial techniques. PPA liability/medical and physical damage patterns were analyzed separately. The derivation of these patterns is included in Exhibit 11.

B. Collection and Development of Company and Economic Information

Exhibit 1 documents the information required by each of the pricing techniques in order to apply it to property/casualty insurance ratemaking. The initial or "base case" value assumed in this study for each variable, as well as a range of reasonable values, is included in the table. The variables are classified into three categories: Company Variables, Economic Variables, and Government Policy Variables, depending upon whether a particular variable is most influenced by company operating decisions, general economic conditions, or governmental policy decisions.

The Company Variables include equity, investment rate of return, standard deviation of investment returns, equity beta, and the funds-generating coefficient. The rationale for placing these variables in this category is that the company, through operating and investment decisions, determines the premium-to-surplus level, the investment policy (which affects both the investment rate of return and the standard deviation), internal factors which influence the beta of the firm's equity, and—to some extent—the claims payment patterns and philosophy.

The Economic Variables include the risk-free rate, market risk premium, risk adjusted discount rate, underwriting beta (and, analogously, the investment-claims correlation), the standard deviation of market returns, the standard deviation of losses, and annual growth rates and betas for inflation and industrial production. These parameters vary primarily due to effects exogenous to the company.

The final category of factors are considered Government Policy Variables, which include the tax rate, the ratio of the investment tax rate to the total tax rate, and the annual tax discount factor. Obviously, the government has sole control over the basic tax rate. The ratio of investment/total tax rate can be affected by the company, by changing the investment allocation, or by the government, by changing the rules about taxability of various investments. In light of the significant effect of such tax regulations as the Tax Reform Act of 1986, the government is assumed to have the greater influence over this variable. Similarly, although the tax discount factor is currently the 60-month moving average of mid-maturity Treasury issues, which would tend to make this an Economic Variable, the definition and calculation of this parameter could be changed by the government at any time.

The most critical step in applying any of the financial models is determining the values for the variables. No matter how accurate a particular model is felt to be, unless the correct parameters are included, the results will not be useful. Many of the prior applications of financial pricing models have either simply assumed particular values for certain variables or determined the values based on a one-time study of industry results. These approaches do not provide much guidance for someone who wants to apply these techniques to an individual company situation on an ongoing basis. In order to facilitate future applications of these models, the determination of parameter values is shown by basing the calculation, where possible, on the fictitious insurer's annual statement or supplementary financial reports, or on general economic information.

C. Company Variables

- *Equity*: Each of the financial pricing models requires either the premium-to-surplus ratio or a value for surplus itself. Although this may appear straightforward, it is not. The reason for the difficulty is the different definitions and uses of the surplus value. For example:

  — In the Target Total Rate of Return model, surplus relates to the amount of assets that an investor chooses to invest in any insurance operation, as opposed to deploying those assets in another investment.

  — In the Discounted Cash Flow model, surplus relates to the amount of invested funds that generate taxes that need to be covered by the premium.

  — In the Insurance CAPM, Internal Rate of Return, Option Pricing, and Arbitrage Pricing models, the surplus is both the amount of capital invested in the firm in support of writing a particular amount of business and the invested assets earning taxable investment income.

  Although each model terms this value "surplus," each model technically requires a slightly different definition of surplus. For consistency, the same value is used as the starting point for each method. As this parameter is extremely important, care should be taken in selecting the appropriate figure. In this study, the value for surplus is an adjusted statutory surplus value, or equity, that is determined as follows:

  Equity = (Statutory Surplus)
  + (Equity in the Unearned Premium Reserve)
  + (Difference Between Nominal and Risk-Free-Discounted Loss Reserves)
  + (Excess of Statutory Over Statement Reserves)

+ (Difference Between Market Value and Book Value for Bonds)

+ (Non-Admitted Assets)

– (Tax Liability on Equity in Unrealized Capital Gains).

For the fictitious company, equity equals $189,360 (dollar values in thousands). Premium-to-statutory surplus ratios for the top twenty private passenger auto insurers (Exhibit 9) range from 0.67 for ITT Hartford to 2.89 for American Premier Underwriters. This range, combined with the adjustments to statutory surplus, which were held constant, determined the range for equity values to be $122,132 to $399,692, as documented in Exhibit 10, Part 4.

Another way to determine the economic value of the insurer, which could be used for publicly traded firms, is to use market value, calculated by multiplying the number of shares outstanding times the current stock price. Our fictitious company is assumed to have a market value of $220,399, reflecting the average market-to-book ratio for stock property/casualty insurance companies at the end of 1993 of 1.46 multiplied by the statutory surplus. The market value could differ from the equity value for any number of reasons, including additional accounting conventions that cause a divergence between reported and economic value or other assets that are not reflected in an insurer's balance sheet (reputation, market niche, a book of business that will generate profits on renewal). For the models that consider surplus to be the investor's value in an insurer (all the models illustrated in this paper except the discounted cash flow), the market value represents the amount that the investor could obtain for giving up its investment in insurance. The market-to-book value ratio ranged from a low of 0.92 (market value of $138,881) to a high of 2.43 (market value of $366,828) for personal lines insurers. As these values all fall within the range obtained by varying only the premium-to-statutory surplus ratio, no further adjustments were made. However, the size of this range illustrates the importance of the selection of the appropriate equity value.

- *Investment rate of return*: The investment rate of return was calculated by summing net investment income, realized capital gains and unrealized capital gains and dividing the total by the average investable assets during the year. All investment income, realized or unrealized, was used to reflect the full effect of investment earnings. A base case value of eight percent was selected based on this calculation, with a range of plus or minus two percent. In practice, the average returns over a number of years should be taken to avoid distortions that could be caused by short-term fluctuations in investment results.

- *Standard deviation of investment returns*: The base case of 20 percent is the same value as used by Doherty and Garven [15]. This value could be calculated for a particular insurer by obtaining the standard deviation of the company's total investment rate of return (including both realized and unrealized capital gains). Our selected range is plus or minus ten percent around the base case.

- *Equity beta*: The base case is 1.0, which is the overall market beta. With regard specifically to insurance stocks, Hill [18] found equity betas averaging 0.61, and Hill and Modigliani [19] and Fairley [16] found betas of approximately 1.00. Fama and French [17] formed portfolios based on beta, and the portfolio betas ranged from .81 to 1.73. Thus, the selected range for insurance betas was .60 to 1.70, ranging from the value determined by Hill to approximately the 95th percentile based on Fama and French. Note that, rather than separately testing the sensitivity of UPMs to a selected range of internal rate of return values, we have assumed that the IRR values are determined via a CAPM approach and so have embedded UPM sensitivity to internal rates of return in our equity beta range.

- *Funds-generating coefficient*: k is the average length of time that the insurer holds (and invests) premiums before they are used to pay expenses and losses. The coefficient is calculated by multiplying the loss payments in the first year by 0.5, the loss payments in the second year by 1.5, and so forth. These values are then summed and divided by premiums. (Expenses are assumed to be paid when the premium is received, so they do not increase the total time-weighted sum of outgo.) For the base case, the ratio is 1.18. The sensitivity of UPM indications to the coefficient is examined by assuming that the company pays its losses and expenses, on average, one quarter of a year either faster, $k = .93$, or slower, $k = 1.43$, than the base case.

  The calculation of the coefficient for the Option Pricing Model is similar, but as this method calculates premiums net of expenses, the appropriate adjustment is to divide the weighted sum of loss payments by total losses, rather than premiums. This produces a k value of 1.5 for the base case, with a range of 1.18 to 1.81 to correspond with the insurance CAPM adjustment. The Discounted Cash Flow and Internal Rate of Return models both also depend on the loss payment pattern, but as the actual payments are used instead of a weighted average, the adjustment must be made to each payment. The calculation of the funds-generating coefficient for each model and the adjusted loss payout patterns are displayed in Exhibit 10, Part 4.

D. Economic Variables

- *Risk-free rate*: As is frequently done, we used the interest rate on U. S. Treasury bills as a proxy for the risk-free rate. As of February, 1996, both three- and six-month Treasury bills had a yield to maturity of approximately five percent. As the appropriate rate for this variable is the current, and not a past, rate, the base case value was set at five percent. This rate has ranged from 2.9 percent to 14.7 percent over the period 1974 to 1993 based on Ibbotson [20] data. The range was set at the twenty-year high and low values.

- *Market risk premium*: This is generally determined as the average excess return in the stock market over an investment in short-term Treasury bills. The time period 1926 through the most recent year is frequently used based on the Ibbotson Associates data series. For 1926 through 1994, the market risk premium is eight percent. Depending on the number of years included in the measurement and the selected years, the market risk premium fluctuates. The selected range is six to ten percent.

- *Risk-adjusted discount rate*: The RADR is used in the discounted cash flow model to discount risky cash flows, primarily losses. The consensus of research on the issue of discounting loss reserves indicates that the appropriate risk-adjusted discount rate is less than the risk-free rate. (See, for example, Butsic [6], Cummins [9], and D'Arcy [12].) This is because the RADR reflects a risky liability to the insurer, and the insurer may be viewed as a risk-averse entity. Conversely, the insurance policy can be viewed as a risky asset for the policyholder, an asset that has a negative beta since it increases in value when the value of the other assets of the policyholder decline due to a loss. Thus, the policyholder would expect to earn a rate of return below the risk-free rate based on either the Capital Asset Pricing Model or the Arbitrage Pricing Model. How *much* less, however, is an unsettled issue. If the CAPM is used to determine the risk-adjusted discount rate, then the differential will be a constant value, regardless of the level of interest rates. For extremely low interest rates, the RADR could even turn out to be negative. Conversely, if the RADR is proportional to the level of interest rates, then the differential will increase as interest rates increase and, regardless of how low interest rates were to fall, would be non-negative as long as interest rates were non-negative. This is an area requiring further research. For this paper, the RADR is assumed to be proportional to the level of interest rates. The risk-adjusted rate is held at 60 percent of the risk-free rate, with the RADR-to-risk-free ratio ranging from zero percent to 100 percent. (The zero percent is consistent with undiscounted loss reserves.) Given the base case risk-free rate assumption of five percent, the base case risk-adjusted discount rate is selected as three percent.

- *Underwriting beta*: There is no generally-accepted theoretical reason why underwriting results should be correlated with market returns, so measuring the value of the underwriting beta and the investment-claims correlation must be based on empirical results. Cummins and Harrington [10] test underwriting betas over the period 1970 to 1981 and find that values appear to range from −.20 to +.20, although the average is not significantly different from zero. D'Arcy and Garven [14] calculate a long-term correlation (1926 through 1985) of 0.0763. The base case is set at zero, with a range of −0.40 to +0.40.

- *Standard deviation of stock market returns*: This variable has historically been 22 percent (Doherty and Garven [15]), which is used here as the base case. The range of 12 to 40 percent was selected judgmentally.

- *Standard deviation of losses*: This should be measured by comparing actual losses with expected losses over a number of years. There is no information about initial expected losses in any financial statement of an insurer, although Schedule P shows loss development after the first accident year has occurred. Doherty and Garven [15] assume a value of 25 percent of losses for this parameter (i.e., a coefficient of variation, or Cov, of 25 percent), and that value is used here as the base case. The range of 12.5 percent of losses to 50 percent was determined judgmentally; these values correspond to assuming a Cov of one-half to twice the base case Cov.

- *CPI change*: Historical values were taken from Ibbotson Associates [20]. Based on recent inflation rates, an annual growth rate of three percent was chosen as the base case value, with a range from zero to six percent.

- *Industrial production growth rate*: Historical values were taken from Federal Reserve data—log-differences of index values were used. Based on recent growth rates, a base case value of two percent was chosen, with a range of zero to four percent.

- *CPI and industrial production betas*: The beta values were determined by running multivariate regressions of annual inflation and industrial production growth rates against the following dependent variable: historical auto UPMs (taken from A. M. Best *Aggregates and Averages*), plus the historical risk-free rate multiplied by the estimated historical funds generating coefficient. (See Urrutia [34], Equation 13.) Examination of the coefficients of the regressions, as well as covariance and variance calculations, over different periods of time led to the selection of base case values and ranges. (See Exhibit 8, Part 2.) While there is some evidence from the regressions for a base case inflation beta closer to 0.70 than to 0.50 based on much of the historical data period from 1948 to 1993, the last ten years of this period indicate an inflation beta much closer to zero. We have chosen a wide and symmetric range around our selected base case of 0.50 to indicate that these estimates require refinement; this could be the subject of future research.

E. Government Policy Variables

- *Tax rate*: The current corporate tax rates range from 15 to 39 percent, based on taxable income level. The base rates are 15, 25, 34 and 35 percent, with surcharges of three and five percent applying to segments of taxable income in order to equalize average and marginal tax rates. The appropriate tax rate is the projected marginal tax rate for the tax year in which the coverage will apply. This necessarily involves a projection. For the fictitious insurer, taxable income is projected to be between $335,000 and $10,000,000 for which a 34 percent marginal tax rate applies. The range is selected to be 28 to 40 percent, reflecting uncertainty over future, potentially retroactive, government tax policy.

- *Investment/total tax ratio*: Investment income is taxed at different rates depending on the source. Interest on federal and corporate bonds is fully taxable. Interest on municipal bonds purchased after August 7, 1986, is taxed at the 15 percent level. Seventy percent of dividends from non-controlled corporations are taxed at the 15 percent level, with the remainder fully taxed. Long-term capital gains are subject to a maximum tax rate of 28 percent. Based on the distribution of investment income earned by the fictitious insurer for the most recent year, the tax rate on investment income is 80 percent of the maximum level. This calculation is illustrated in Exhibit 10, Part 4. This value is allowed to range from 60 to 100 percent.
- *Tax discount factor*: The tax discount factor based on the 60-month moving average of mid-maturity Treasury issues ending February 1996 is 6.55 percent (Massachusetts Workers' Compensation Rate Filing [23]). The value for the period ending October 1994 was 7.03 percent. Thus, the base case value is seven percent. The range of four to ten percent was selected judgmentally.

F. Application of Techniques to the Representative Company

Exhibits 3 through 8 document the application of each of the six financial pricing techniques to the representative company described above and in Exhibit 10. Each exhibit shows, for the respective pricing model, the relevant parameter values and the indicated base case UPM. In addition, the sensitivity of the UPM indications to different premium/equity ratios is shown both numerically and graphically. For three of the models, an additional parameter is also varied in the graphs. For each model, UPM is an inverse function of premium/equity ratio.

G. Sensitivity Tests

For each pricing model, the sensitivity of UPM indications to each relevant parameter is determined by allowing the parameter to vary from the base case to the low and high ends of the reasonable range, keeping all other parameters constant at their base case values. The results of these sensitivity tests are summarized in Exhibit 1. In addition, Exhibit 2 summarizes the sensitivity of UPM indications to simultaneous changes in groups of variables. Relevant variables in each of the three categories (company, economic, and government policy) are varied while keeping the other groups constant; in addition, all relevant variables across all groups are varied simultaneously.

4. EMPIRICAL RESULTS AND ANALYSIS

A. Evaluation of Base Case Results

The six base case underwriting profit margin indications displayed on Exhibit 1 range from −4.9 percent to 1.7 percent. The Insurance CAPM, which produces the lowest value, is known to under-price, as it ignores insurance-specific risk (Ang and Lai [4] and Turner [33]). (This deficiency also applies to the Arbitrage Pricing model, at least in the form presented in this paper.) The next lowest value, the Target Total Rate of Return, does not consider the effect of taxes. The Discounted Cash Flow, Internal Rate of Return and Option Pricing models all cluster between 0.1 percent and 1.7 percent. A good case can be made for selecting the average of these three models, 0.7 percent, as the underwriting profit margin target for this insurer under base case economic and company conditions. Coincidentally, this is the same value that the Target Total Rate of Return model indicates for an equity beta of 1.7, which would produce a target rate of return, before taxes, of 18.6 percent. Note that this underwriting profit margin, and in fact the entire range of base case values, is significantly below the Target UPM model provision of five percent.

B. Sensitivity Analysis

B.1. Single Factor Variation

By examining the effect of changing each parameter over the range of reasonable values, the sensitivity of the pricing models to different conditions, as well as the importance of accurately measuring each variable, can be discerned. Examination of Exhibit 1 suggests that the variables most affecting the results are the equity, or premium-to-surplus ratio (for all except the Insurance CAPM and Arbitrage Pricing models), and the risk-free rate (for all except the Discounted Cash Flow model). For example, over the range of equity values examined, the underwriting profit margin changes by 14.6 percentage points for the Target Total Rate of Return model, 4.9 percentage points for the Discounted Cash Flow model, 12.7 percentage points for the Internal Rate of Return model, and 4.8 percentage points for the Option Pricing model. In each case, the higher the initial equity, the higher the indicated underwriting profit margin; the greater the amount that the insurance company has invested for a given volume of premium, the higher the price needs to be to provide an adequate return on capital. This volatility is a problem for the financial pricing models since it is difficult to measure true equity for a single-line, single-state insurer. The problem is vastly more complicated for a multiline, multistate insurer. The effect of varying the equity is illustrated in Exhibit 1, Part 2, which compares UPM indications across all models under a range of premium/equity ratio assumptions. In the graph, the low premium/equity ratio corresponds with a high equity value.

For an insurer that had the highest amount of equity, the indicated underwriting profit margins are 7.5 percent for the Target Total Rate of Return, −3.2 percent for the Insurance CAPM, 3.8 percent for the Discounted Cash Flow, 11.0 percent for the Internal Rate of Return, 3.2 percent for the Option Pricing model, and −1.2 percent for the Arbitrage Pricing model. Determining the appropriate underwriting profit margin target in this situation is difficult. Even ignoring the Insurance CAPM and Arbitrage Pricing models, the results vary by 7.8 percentage points, with an average of 6.4 percent. Selecting the appropriate value will be a difficult judgment call.

Similarly, the results are highly sensitive to the risk-free rate: across the range of reasonable parameter values, the results vary by 9.1 percentage points for the Target Total Rate of Return model, 11.6 percentage points for the Insurance CAPM and Arbitrage Pricing models, 0.8 percentage points for the Discounted Cash Flow model, 12.6 percentage points for the Internal Rate of Return model, and 11.0 percentage points for the Option Pricing model. Also, the effect of increasing the risk-free rate affects the results differently. For the Target Total Rate of Return and the Internal Rate of Return models, increasing the risk-free rate raises the indicated underwriting profit margin (given the model assumptions previously described). However, the indications for the Insurance CAPM, the Discounted Cash Flow, the Option Pricing, and the Arbitrage Pricing models all decline as the risk-free rate increases. Thus, if the risk-free rate were 14.7 percent, as it was in 1981, then the indicated underwriting profit margins vary from −14.5 percent for the Insurance CAPM model to 11.4 percent for the Internal Rate of Return model. Fortunately, the value of the risk-free rate is easy to obtain. Unfortunately, the models are very sensitive, in opposite directions, to this value.

On the opposite extreme, the results are not very sensitive, for example, to the tax discount factor. Since private passenger auto losses are paid relatively quickly, the effect of the tax discount factor is minimal. This would not be the case for a longer-tailed line. Also, neither the investment rate of return, the market risk premium nor the standard deviation of market returns affect the results of the Option Pricing model. This occurs because the base case underwriting beta is zero. In some respects, the choice of zero as a underwriting beta is not an ideal choice because the sensitivity analysis does not illustrate the effect of changing values that are multiplied by beta. However, all empirical analyses measuring the underwriting beta indicate that it is not significantly different from zero; taken in aggregate, zero is the best a priori estimate of the underwriting beta. (See, for example, Cummins and Harrington [10], Fairley [16], and Hill and Modigliani [19].) The effect of altering more than one variable at a time is discussed next.

B.2. Multiple Factor Variation

A point of critical importance in evaluating the results summarized on Exhibit 1 is that each of the high/low UPM indications reflects a change in only one parameter, while all of the others are kept fixed. In many cases, such a simplistic scenario will understate the potential sensitivity of a model to changes in a parameter. Often, this is because the selection of the base case value of another parameter minimizes the parameter's impact on a model's UPM calculation. For example, changes in the market risk premium do not influence the UPM indications for the Option Pricing Model, according to Exhibit 1—however, this is because we have selected an investment-claim correlation of zero for our base case assumption. This assumption, when incorporated into the option formulae, "zeroes out," or makes irrelevant, the market risk parameter.

Because of examples like this, in order to determine the true magnitude of potential impact of certain parameters on the various models, it is necessary and instructive to vary more than one of the parameters simultaneously. This is in keeping with reality, as it is entirely possible—and even probable—that multiple parameters will change concurrently.

The impact of selected multi-parameter changes on UPM indications for various models is demonstrated in Exhibit 2. The results vary markedly. Changing all the Company Variables from the lowest values to the highest values only changes the indicated underwriting profit margin by 9.1 percentage points for the Option Pricing model, 4.8 percentage points for the Discounted Cash Flow model and 5.1 percentage points for the Insurance CAPM and Arbitrage Pricing models. At the other extreme, the same changes shift the indicated underwriting profit margin by 34.5 percentage points for the Internal Rate of Return model and 46.8 percentage points for the Target Total Rate of Return model. Thus, the impact of different financial positions for companies will differ depending on the model.

The effect of changing the Economic Variables affects all the models significantly. The effect ranges from 12.1 percentage points for the Target Total Rate of Return model to 61.9 percentage points for the Option Pricing model. Changing economic conditions will affect all insurance companies and must be reflected in the parameters selected.

The impact of changing the Government Policy Variables, which all relate to taxation, is not as significant as the Company or Economic Variables. The indicated underwriting profit margins shift by only 2.6 percentage points for the Insurance CAPM and Arbitrage Pricing models, 5.6 percentage points for the Internal Rate of Return model, 5.7 percentage points for the Option Pricing model, and 6.1 percentage points for the Discounted Cash Flow model. They have no effect on the Target Total Rate of Return model, which ignores taxation.

Predictably, based on the above results, the impact of changing *all* variables simultaneously is extremely significant for each model.

5. PRACTICAL GUIDE TO USING FINANCIAL PRICING MODELS

When using a variety of financial pricing models to select an appropriate underwriting profit margin, the different models can be expected to generate different indications. Selecting the appropriate profit margin requires actuarial judgment, including a thorough understanding of the reliability of the inputs used in the models and the strengths and weaknesses of the different techniques. For example, if current interest rates are high, but there is considerable uncertainty about future levels, then the models that are least sensitive to changes in the risk-free interest rate (in this example the discounted cash flow model) should be given greater weight. If the value of the insurer's equity cannot be easily valued, such as in the case of a mutual insurer, then the models that are very sensitive to the initial equity value (in this case the Target Total Rate of Return and the Internal Rate of Return) should be given less weight than they would if the initial equity could be valued more accurately. When writing a line of business that is considered to have little insurance-specific risk (for example, fidelity), then models that ignore this factor, the Insurance CAPM or Arbitrage Pricing Model (as formatted in this study), would be more appropriate than they would be for lines with a high degree of insurance-specific risk (for example, homeowners).

For this study, all of the models are tested based on the same input data. These values were intentionally selected to be representative of the property/casualty insurance industry in general and of current economic conditions. Thus, the different indications result from differences in the basic structure of the individual models, and the effect of these differences for a representative insurance company can be quantified. Specifically, the Target Total Rate of Return model, which ignores taxes, produces an indication three to four percentage points below the level where other models tend to cluster. The Insurance CAPM, which ignores insurance-specific risk, produces a value approximately five percentage points below the cluster level, and the Arbitrage Pricing Model, which in the form used here also ignores insurance-specific risk, produces an indication about three percentage points lower than the cluster level. These differences can provide some guidance about adjustments that can be made to reflect omitted factors. However, the differences will need to be recalculated when applied in situations that differ from the example provided in this study, if, for example, the type of business or the financial parameters were to change.

Knowledge about the assumptions inherent in the models is also important for proper application. For example, the Internal Rate of Return model assumes that any underwriting losses are funded at the inception of the policy. Conversely, underwriting profits would also be reflected at the inception of the policy, reducing the initial surplus allocation for the policy. Although this assumption would have little effect when the indicated underwriting profit margins are approximately zero, it could introduce distortions in the case of sizeable underwriting profits or losses. Also, the Discounted Cash Flow model relies heavily on the appropriate risk-adjusted discount factor, a value that is difficult to measure given the lack of a public and liquid market trading insurance liabilities. Care must be taken, when using this model, that the risk-adjusted discount rate reasonably accounts for the risk involved in writing a particular line of business.

Selecting an appropriate underwriting profit margin is as much of an actuarial art as selecting the appropriate loss reserve level. The financial pricing models described in this study, although none is perfect, can be used to determine this value if properly applied and if used in conjunction with other models. Knowing the likely relationships among the respective indications and the sensitivity of the indications to specific parameters can help direct the user to select reasonable underwriting profit margins.

6. CONCLUSION

A. Implications for Insurance Companies and Regulators

The diverse results for the indicated underwriting profit margin, depending on the pricing model selected and the nature of the company and economic environment, should be convincing evidence that no one pricing model can be relied upon to provide the appropriate underwriting profit margin in all situations. Instead, insurers should apply a number of different pricing methods and evaluate the results in combination to select the target underwriting profit margin. The models discussed in this paper are possible techniques for companies to use. However, each model should be understood and its shortcomings noted in order to apply the techniques most appropriately.

Another conclusion that can be drawn from this research is that insurance is a very complex financial transaction. For people working in the insurance industry, this may seem to be an unusual statement. An insurance company collects premiums and pays losses and expenses: what is so complex about that? However, compared to stocks, bonds and options, for which the financial pricing models were originally developed, insurance is very complex. Owners of stock receive a periodic stream of dividends and the value of the stock when it is sold. Bondholders receive a fixed stream of income and a predetermined principal amount at maturity, subject only to default risk or an early call. Option holders receive at maturity the difference between the price of the underlying asset and the exercise price, if this is greater than zero.

Insurance, on the other hand, involves collecting a stream of premium payments over time in return for the promise to pay losses, if they occur, for which both the amount and timing are unknown. The mathematics for dealing with this degree of uncertainly have not been perfected. This complexity means that techniques that are readily applicable to other financial transactions are not necessarily going to provide reasonable results when applied to insurance. Thus, regulating insurance involves more than importing a financial technique that has been applied in another setting.

B. Future Research

Our goal in this paper has been to compare and contrast different asset pricing models in terms of their indicated fair rates of return for property/casualty insurance policies given various corporate and economic environments. We have focused on those pricing models that have been suggested and (at least somewhat) developed in the literature to date. Several of the models would benefit from more extensive development—perhaps the findings of this paper will help to suggest where such resources are best applied.

To date, the Arbitrage Pricing Model has received relatively little theoretical or empirical attention in terms of insurance applications. We have used a very basic pre-specified factors model for purposes of illustrative model comparisons in this paper. Two macroeconomic variables—inflation and industrial production growth rates—were found to be relatively significant in explaining adjusted (for investment income) UPMs over the 1948–1993 period (the other variables tested were a bond default premium, a bond horizon premium, and a New York Stock Exchange value-weighted stock return series). The positive relationship we found between inflation and adjusted UPM is interesting in light of the finding in Kraus and Ross [21] that the competitive premium should be affected by inflation only in so far as real rates of interest are impacted. This relationship should be analyzed further, perhaps by separately determining the sensitivity of UPMs to expected and unexpected inflation. Other insurance specific variables—e.g., catastrophe losses, leverage—should also be examined for significance. In addition, historical tax rates may well have an impact on historical UPM regressions with the pre-specified variables, and should be incorporated into the process of determining beta coefficients. Also, instead of assuming the relevancy of specific macroeconomic and insurance variables, a factor analysis approach might be worth investigating. We intend to examine these and other issues in a separate paper.

There are several other areas in which additional research might prove fruitful. For example, the Option Pricing Model requires distributional and risk preference assumptions in order to price the contingent claims. It would be instructive to examine the impact on OPM pricing indications of assuming return and loss distributions other than the normal and lognormal distributions. Another area involves surplus allocation: for practical applications of these models, a multiline and/or multistate insurer must be able to appropriately allocate surplus to its various business segments. Finally, additional research into appropriate parameter values for each model is certainly warranted before the models are actually used for insurance pricing purposes.

REFERENCES

[1] A. M. Best Company, *Best's Aggregates and Averages*, New Jersey, 1994 and prior.

[2] A. M. Best Company, *Best's Insurance Reports: Property-Casualty*, New Jersey, 1994.

[3] A. M. Best Company, *Best's Key Rating Guide: Property-Casualty*, New Jersey, 1994.

[4] Ang, James S., and Tsong-Yue Lai, "Insurance Premium Pricing and Ratemaking in Competitive Insurance and Capital Markets," *Journal of Risk and Insurance*, 54, 1987, pp. 767–779.

[5] Biger, Nahum, and Yehuda Kahane, "Risk Considerations in Insurance Ratemaking," *Journal of Risk and Insurance*, 45, 1978, pp. 121–132.

[6] Butsic, Robert P. "Determining the Proper Interest Rate for Loss Reserve Discounting: An Economic Approach," *Evaluating Insurance Company Liabilities*, 1988 Casualty Actuarial Society Discussion Paper Program, pp. 147–188.

[7] Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, "Economic Forces and the Stock Market," *Journal of Business*, 59, 1986, pp. 383–403.

[8] Cummins, J. David, "Risk-Based Premiums for Insurance Guaranty Funds," *Journal of Finance*, 43, 1988, pp. 823–839.

[9] Cummins, J. David, "Multi-Period Discounted Cash Flow Ratemaking Models in Property-Liability Insurance," *Journal of Risk and Insurance*, 57, 1990, pp. 79–109.

[10] Cummins, J. David, and Scott E. Harrington, "Property-Liability Insurance Rate Regulation: Estimation of Underwriting Betas Using Quarterly Profit Data," *Journal of Risk and Insurance*, 52, 1985, pp. 16–43.

[11] Cummins, J. David, and Scott E. Harrington, *Fair Rate of Return in Property-Liability Insurance*, Boston: Kluwer-Nijhoff Publishing Company, 1987.

[12] D'Arcy, Stephen P., "Use of the CAPM to Discount Property-Liability Loss Reserves," *Journal of Risk and Insurance*, 55, 1988, pp. 481–491.

[13] D'Arcy, Stephen P., and Michael A. Dyer, "Ratemaking: A Financial Economics Approach," *PCAS* LXXXIV, 1997, pp. 301–310.

[14] D'Arcy, Stephen P., and James R. Garven, "Property-Liability Insurance Pricing Models: An Empirical Evaluation," *Journal of Risk and Insurance*, 57, 1990, pp. 391–430.

[15] Doherty, Neil A., and James R. Garven, "Price Regulation in Property-Liability Insurance: A Contingent Claims Approach," *Journal of Finance*, 41, 1986, pp. 1031–1050.

[16] Fairley, William B., "Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results," *Bell Journal of Economics*, 10, 1979, pp. 192–210.

[17] Fama, Eugene F., and Kenneth R. French, "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 1992, pp. 427–465.

[18] Hill, Raymond D., "Profit Regulation in Property Liability Insurance," *Bell Journal of Economics*, 10, 1979, pp. 172–191.

[19] Hill, Raymond D., and Franco Modigliani, "The Massachusetts Model of Profit Regulation in Non-Life Insurance: An Appraisal and Extensions," Chapter 2 in Cummins and Harrington, 1987.

[20] Ibbotson Associates, *Stocks, Bonds, Bills, and Inflation: 1996 Yearbook*, Ibbotson Associates, 1996.

[21] Kraus, Alan, and Stephen A. Ross, "The Determination of Fair Profits for the Property-Liability Insurance Firm," *Journal of Finance*, 37, 1987, pp. 1015–1028.

[22] Lintner, John, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, 47, 1965, pp. 13–37.

[23] Massachusetts Workers Compensation Rate Filing, 1996.

[24] Mossin, Jan, "Equilibrium in a Capital Market," *Econometrica*, 34, 1966, pp. 768–783.

[25] Myers, Stewart C., and Richard Cohn, "A Discounted Cash Flow Approach to Property-Liability Insurance Rate Regulation," Chapter 3 in Cummins and Harrington, 1987.

[26] National Association of Insurance Commissioners, Report to the Special Subcommittee of the Fire and Marine Committee by Roy C. McCullough, *Proceedings of the National Association of Insurance Commissioners*, Seventy-Ninth Session, 1948, pp. 74–157.

[27] National Association of Insurance Commissioners, "Measurement of Profitability and Treatment of Investment Income in Property and Liability Insurance," *Proceedings of the National Association of Insurance Commissioners*, 1970 Annual Meeting, IIA, 1970, pp. 719–893.

[28] National Convention of Insurance Commissioners, Report of the Fire Insurance Committee, *Proceedings of the National Convention of Insurance Commissioners*, 1922, pp. 20–23.

[29] Roll, Richard, "A Critique of the Asset Pricing Theory's Tests," *Journal of Financial Economics*, 4, 1977, pp. 129–176

[30] Roll, Richard, and Stephen A. Ross, "An Empirical Investigation of the Arbitrage Pricing Theory," *Journal of Finance*, 35, 1980, pp. 1073–1103.

[31] Ross, Stephen A., "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 1976, pp. 341–360.

[32] Sharpe, William F., "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance*, 19, 1964, pp. 425–442.

[33] Turner, Andrew, "Insurance in an Equilibrium Asset-Pricing Model," Chapter 4 in Cummins and Harrington, 1987.

[34] Urrutia, Jorge L., "Determination of Competitive Underwriting Profit Margins Using the Arbitrage Pricing Model," *Journal of Insurance Issues and Practice*, 10, 1987, pp. 61–77.

[35] Webb, Bernard L., "Investment Income in Insurance Ratemaking," *Journal of Insurance Regulation*, 1, 1982, pp. 46–76.

EXHIBIT 1

Part 1

Indicated Underwriting Profit Margins Based on Range of Reasonableness for Each Variable

Variable	Base Case	Range	Target Total Rate of Return		Insurance CAPM		Discounted Cash Flow		Internal Rate of Return		Option Pricing Normal/CARA		Arbitrage Pricing	
Base Case Results			−3.6%		−4.9%		0.1%		1.7%		0.2%		−2.9%	
			Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
Equity	189,360	122,132 to 399,692	−7.1%	7.5%	−5.5%	−3.2%	−1.1%	3.8%	−1.7%	11.0%	−1.6%	3.2%	−3.5%	−1.2%
Investment Rate of Return	8.0%	6.0% to 10.0%	−0.2%	−7.0%					5.5%	−2.5%	0.2%	0.2%		
SD of Investment Returns	20.0%	10.0% to 30.0%									−1.2%	0.2%		
Equity Beta (also for IRR)	1.00	0.60 to 1.70	−6.0%	0.7%					−2.8%	7.8%				
Funds Generating Coefficient	1.18	0.93 to 1.43			−3.5%	−6.3%	0.3%	0.0%	1.5%	1.9%	1.1%	−0.8%	−1.5%	−4.3%
Risk-free Rate	5.0%	2.9% to 14.7%	−5.2%	3.9%	−2.9%	−14.5%	0.1%	−0.7%	−1.2%	11.4%	2.2%	−8.8%	−0.9%	−12.5%
Market Risk Premium	8.0%	6.0% to 10.0%	−5.1%	−2.0%	−4.9%	−4.9%			−1.0%	4.1%	0.2%	0.2%		
Risk-Adj/Risk-Free Rate Ratio	60.0%	0.0% to 100.0%					4.4%	−2.6%	6.2%	−7.3%				
Underwriting Beta & Inv-Cl Cor	0.00	−0.40 to +0.40			−8.1%	−1.7%					6.2%	−7.3%		
SD Market Returns	22.0%	12.0% to 40.0%									0.2%	0.2%		
SD of Losses	48,401	24,201 to 96,802									−0.3%	1.1%		
CPI Change	3.0%	0.0% to 6.0%											−4.4%	−1.4%
CPI Beta	0.50	0.00 to 1.00											−4.4%	−1.4%
Industrial Production Growth	2.0%	0.0% to 4.0%											−3.4%	−2.4%
Industrial Production Beta	0.25	0.00 to 0.50											−3.4%	−2.4%
Tax Rate (Total)	34.0%	28.0% to 40.0%			−5.2%	−4.6%	−0.8%	1.4%	0.5%	3.0%	−1.1%	1.6%	−3.2%	−2.6%
Investment/Total Tax Ratio	80.0%	60.0% to 100.0%			−5.9%	−3.9%	−1.5%	2.0%	0.2%	3.2%	−1.3%	1.7%	−3.9%	−1.9%
Tax Discount Factor	7.0%	4.0% to 10.0%					0.1%	0.2%	1.6%	1.8%				

Note: Each UPM indication represents a change in the relevant variable, with all other variables kept constant at base case values.

EXHIBIT 1

PART 2

SUMMARY OF MODEL INDICATIONS
SENSITIVITY OF UNDERWRITING PROFIT MARGINS TO PREMIUM/EQUITY RATIOS
(All other variables held constant at base case values)

	Premium / Equity Ratio									
Model	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30
Target UPM	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
Target Total Rate of Return	0.124	0.050	0.009	−0.017	−0.036	−0.049	−0.059	−0.067	−0.074	−0.079
Insurance CAPM	−0.024	−0.036	−0.042	−0.046	−0.049	−0.051	−0.053	−0.054	−0.055	−0.056
Discounted Cash Flow	0.061	0.034	0.019	0.009	0.002	−0.002	−0.006	−0.009	−0.011	−0.013
Internal Rate of Return	0.193	0.113	0.069	0.041	0.021	0.007	−0.004	−0.012	−0.019	−0.025
Option Pricing—Normal/CARA	0.048	0.028	0.017	0.009	0.003	−0.003	−0.008	−0.013	−0.018	−0.023
Arbitrage Pricing	−0.004	−0.016	−0.022	−0.026	−0.029	−0.031	−0.033	−0.034	−0.035	−0.036

EXHIBIT 1

PART 2—PAGE 2

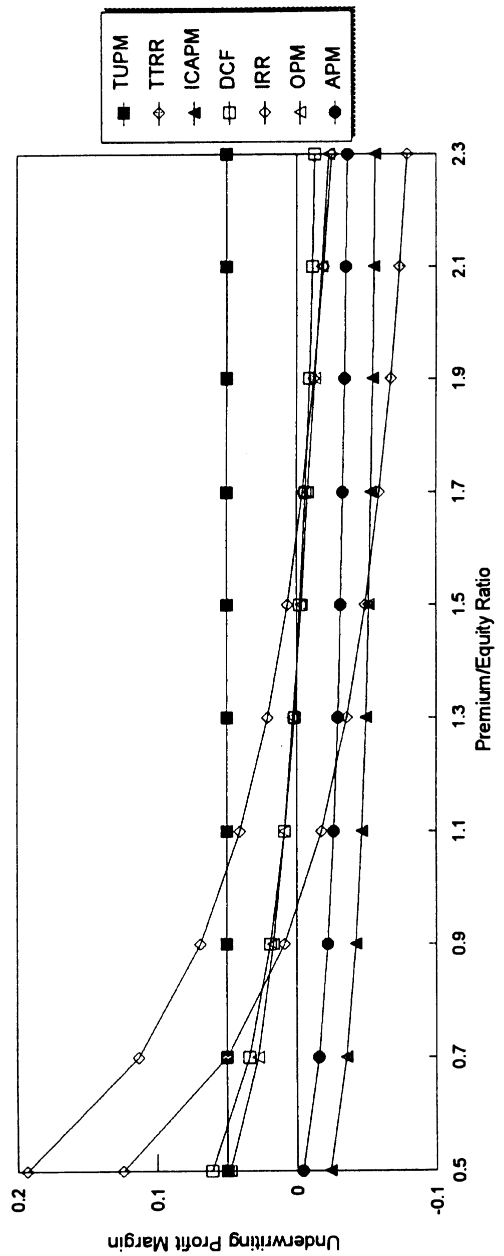


EXHIBIT 2

INDICATED UNDERWRITING PROFIT MARGINS BASED ON COMBINED RANGE EFFECTS OF GROUPS OF VARIABLES

Variable	Target Total Rate of Return		Insurance CAPM		Discounted Cash Flow		Internal Rate of Return		Option Pricing Normal/ CARA		Arbitrage Pricing	
Base Case Results	−3.6%		−4.9%		0.1%		1.7%		0.2%		−2.9%	
	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
Company Variables—Can be Influenced by the Company	−21.4%	25.4%	−6.9%	−1.8%	−1.4%	3.4%	−7.1%	27.4%	−2.9%	6.2%	−4.9%	0.2%
Economic Variables—Determined by Economic Conditions	−6.7%	5.4%	−18.5%	−0.5%	−7.9%	11.2%	−4.2%	12.9%	−26.4%	35.5%	−14.5%	5.1%
Government Policy Variables—Tax-Related			−5.9%	−3.3%	−2.0%	4.1%	−0.7%	4.9%	−2.2%	3.5%	−3.9%	−1.3%
All Variables Combined	−23.1%	46.6%	−26.6%	4.4%	−19.5%	31.9%	−12.1%	43.2%	−47.0%	40.4%	−22.6%	8.4%

Note: Each UPM indication represents a change in each of the relevant groups of variables, with all variables in the other groups kept constant at base case values.

Company Variables:	Equity Investment Rate of Return SD of Investment Returns Equity Beta (also for IRR) Funds Generating Coefficient	Economic Variables:	Risk-free Rate Market Risk Premium Risk-Adj./Risk-Free Rate Ratio Underwriting Beta & Inv-Cl Cor SD Market Returns SD of Losses CPI Change CPI Beta Industrial Production Growth Industrial Production Beta
Gov't. Policy Variables:	Investment/Total Tax Ratio Investment/Total Tax Ratio Tax Discount Factor		

EXHIBIT 3

Target Total Rate of Return Model
Sensitivity of Underwriting Profit Margins to Premium/Equity Ratios and Equity Betas

Risk-free rate	5.0%		Base Case UPM:		−0.036					
Expected market return	13.0%									
Investment rate of return	8.0%									
Invested assets	417,338									
Equity	189,360									
UPMs:										
Premium/equity ratio:	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30
Equity beta = 1.70	0.236	0.130	0.071	0.033	0.007	−0.012	−0.026	−0.038	−0.047	−0.055
Equity beta = 1.40	0.188	0.096	0.044	0.012	−0.011	−0.028	−0.040	−0.050	−0.058	−0.065
Equity beta = 1.20	0.156	0.073	0.027	−0.003	−0.023	−0.038	−0.050	−0.059	−0.066	−0.072
Equity beta = 1.00	0.124	0.050	0.009	−0.017	−0.036	−0.049	−0.059	−0.067	−0.074	−0.079
Equity beta = 0.80	0.092	0.027	−0.009	−0.032	−0.048	−0.060	−0.069	−0.076	−0.081	−0.086
Equity beta = 0.60	0.060	0.004	−0.027	−0.047	−0.060	−0.070	−0.078	−0.084	−0.089	−0.093

EXHIBIT 3

PART 2

TARGET TOTAL RATE OF RETURN MODEL

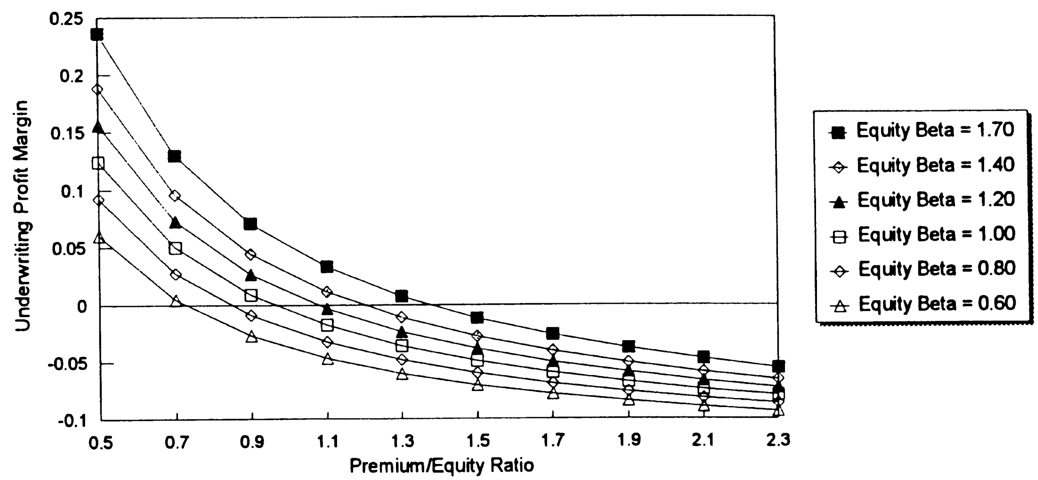


EXHIBIT 4

INSURANCE CAPITAL ASSET PRICING MODEL
SENSITIVITY OF UNDERWRITING PROFIT MARGINS TO PREMIUM/EQUITY RATIOS AND UNDERWRITING BETAS

Risk-free rate	5.0%		Base Case UPM:		−0.049					
Expected market return	13.0%									
Funds generating coeff.	1.18									
Tax rate on investment inc.	27.2%									
Tax rate on U/W income	34.0%									
UPMs:										
Premium/equity ratio:	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30
U/W beta = 0.40	0.008	−0.004	−0.010	−0.014	−0.017	−0.019	−0.021	−0.022	−0.023	−0.024
U/W beta = 0.20	−0.008	−0.020	−0.026	−0.030	−0.033	−0.035	−0.037	−0.038	−0.039	−0.040
U/W beta = 0	−0.024	−0.036	−0.042	−0.046	−0.049	−0.051	−0.053	−0.054	−0.055	−0.056
U/W beta = −0.20	−0.040	−0.052	−0.058	−0.062	−0.065	−0.067	−0.069	−0.070	−0.071	−0.072
U/W beta = −0.40	−0.056	−0.068	−0.074	−0.078	−0.081	−0.083	−0.085	−0.086	−0.087	−0.088

EXHIBIT 4

PART 2

INSURANCE CAPITAL ASSET PRICING MODEL

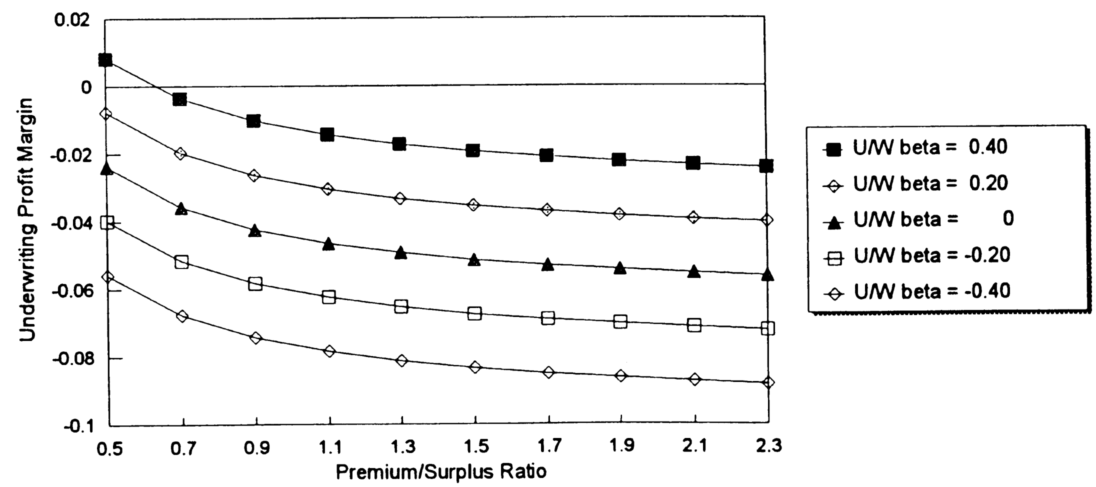


EXHIBIT 5

Discounted Cash Flow Model
Sensitivity of Underwriting Profit Margins to Premium/Equity Ratios and Risk-Adjusted Discount Rates

Risk-free rate	5.0%	Indicated Premium:	253,040
Risk-adj. discount rate	3.0%	Premium/Equity Ratio:	1.34
Tax discount rate	7.0%	Base Case UPM:	0.001
Tax rate on U/W income	34.0%		
Tax rate on investment inc.	27.2%		
Funds generating coeff.	1.18		
Expected losses	193,605		
Expenses	59,062		
Equity	189,360		

UPMs:

Premium/equity ratio:	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30
RADR = 0%	0.102	0.076	0.062	0.053	0.046	0.042	0.038	0.035	0.033	0.031
RADR = 3%	0.061	0.034	0.019	0.009	0.002	−0.002	−0.006	−0.009	−0.011	−0.013
RADR = 6%	0.021	−0.007	−0.023	−0.033	−0.039	−0.045	−0.048	−0.051	−0.054	−0.056

EXHIBIT 5

PART 2

DISCOUNTED CASH FLOW MODEL

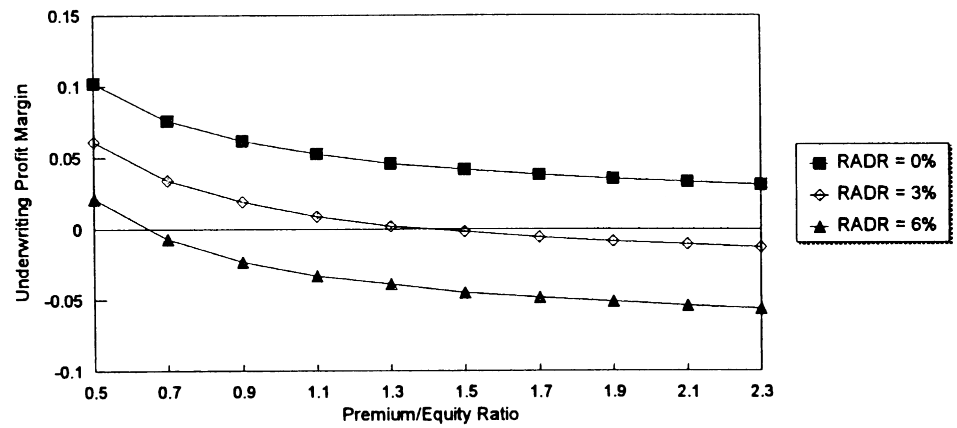


EXHIBIT 6

PART 1

INTERNAL RATE OF RETURN MODEL
(Based on Cummins)

SENSITIVITY OF UNDERWRITING PROFIT MARGIN TO PREMIUM/EQUITY RATIO

Tax Discount Rate:	0.07									
Investment Return Rate:	0.08		Base Case UPM:		0.017					
Tax Rate:	0.34									
Investment Tax Rate:	0.27									
Equity:	189,360									
Expected Losses:	193,605									
Equity/Reserves:	0.98									
IRR:	0.1300									
Premium/Equity Ratio:	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30
UPM:	0.193	0.113	0.069	0.041	0.021	0.007	−0.004	−0.012	−0.019	−0.025

EXHIBIT 6

PART 1—PAGE 2

INTERNAL RATE OF RETURN MODEL

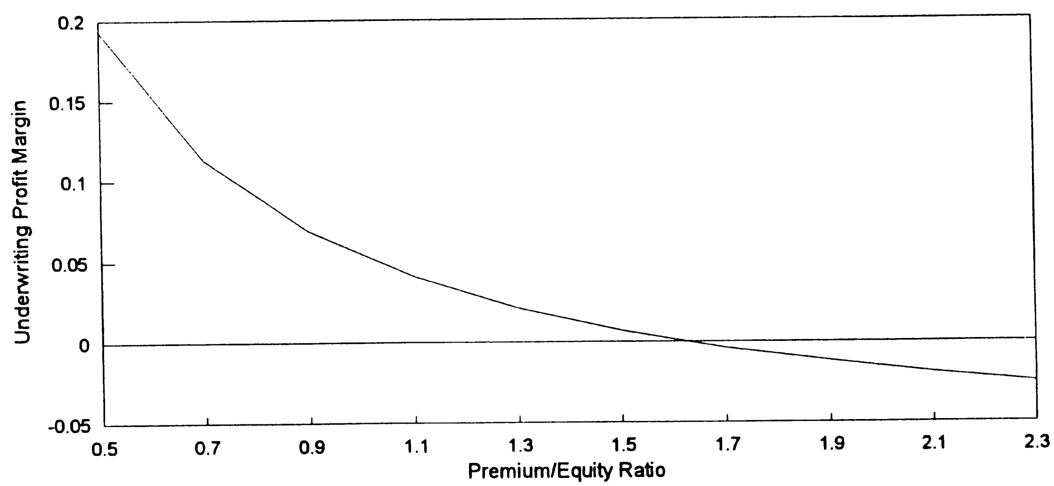


EXHIBIT 6

PART 2

INTERNAL RATE OF RETURN MODEL
(Based on Cummins)

CALCULATION OF DISCOUNTED NET CASH FLOWS FOR BASE CASE

QTR	Prems. (1)	Exps. (2)	Expected Loss (3)	Fed Tax Flow (4)	Total U/W Flow (5)	Accum U/W Acct (6)	Loss Reserve (7)	Supt. Surplus (8)	Avg. Surplus (9)	Invest. on Surp (10)	Income on U/W (11)	Surplus Flow (12)	Net Cash Flow (13)	IRR Discount Factor (14)	Disc. NCF (15)
					(2,866)			189,360				(186,494)	(186,494)	1.0000	(186,494)
1	257,009	(59,062)	(6,389)	(1,119)	190,439	92,354	42,012	183,111	186,236	3,725	1,847	6,249	10,305	0.9848	10,149
2	0	0	(19,361)	(1,119)	(20,480)	177,333	71,053	164,175	173,643	3,473	3,547	18,936	24,046	0.9552	22,969
3	0	0	(32,138)	(1,119)	(33,258)	150,464	87,316	132,741	148,458	2,969	3,009	31,434	35,786	0.9265	33,154
4	0	0	(44,916)	(1,119)	(46,036)	110,818	90,801	88,810	110,776	2,216	2,216	43,932	47,158	0.8986	42,375
5	0	0	(11,665)	351	(11,314)	82,143	79,136	77,401	83,105	1,662	1,643	11,409	13,815	0.8715	12,040
6	0	0	(11,665)	351	(11,314)	70,829	67,471	65,992	71,696	1,434	1,417	11,409	13,484	0.8453	11,398
7	0	0	(11,665)	351	(11,314)	59,515	55,807	54,583	60,287	1,206	1,190	11,409	13,153	0.8199	10,784
8	0	0	(11,665)	351	(11,314)	48,200	44,142	43,174	48,879	978	964	11,409	12,822	0.7952	10,196
12	0	0	(20,329)	701	(19,628)	32,729	23,813	23,291	33,233	2,659	2,618	19,883	23,724	0.7367	17,478
16	0	0	(10,842)	377	(10,465)	17,683	12,972	12,687	17,989	1,439	1,415	10,604	12,682	0.6520	8,268
20	0	0	(5,808)	204	(5,604)	9,649	7,163	7,006	9,847	788	772	5,681	6,816	0.5770	3,933
24	0	0	(2,904)	114	(2,790)	5,452	4,259	4,166	5,586	447	436	2,840	3,483	0.5106	1,779
28	0	0	(1,742)	67	(1,676)	3,220	2,517	2,462	3,314	265	258	1,704	2,085	0.4518	942

EXHIBIT 6

PART 2—PAGE 2

QTR	Prems. (1)	Exps. (2)	Expected Loss (3)	Fed Tax Flow (4)	Total U/W Flow (5)	Accum U/W Acct (6)	Loss Reserve (7)	Supt. Surplus (8)	Avg. Surplus (9)	Invest. on Surp (10)	Income on U/W (11)	Surplus Flow (12)	Net Cash Flow (13)	IRR Discount Factor (14)	Disc. NCF (15)
32	0	0	(968)	39	(929)	1,917	1,549	1,515	1,988	159	153	947	1,174	0.3999	470
36	0	0	(387)	26	(362)	1,272	1,162	1,136	1,326	106	102	379	530	0.3539	188
40	0	0	(194)	20	(173)	1,005	968	947	1,041	83	80	189	309	0.3132	97
44	0	0	(194)	17	(176)	830	774	757	852	68	66	189	287	0.2771	80
48	0	0	(194)	14	(180)	652	581	568	663	53	52	189	266	0.2452	65
52	0	0	(194)	10	(183)	470	387	379	473	38	38	189	244	0.2170	53
56	0	0	(194)	6	(187)	285	194	189	284	23	23	189	223	0.1921	43
60	0	0	(194)	2	(191)	96	0	0	95	8	8	189	200	0.1700	34
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1504	0
68	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1331	0
72	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1178	0
76	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1042	0
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0923	0
84	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0816	0
88	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0722	0
92	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0639	0
96	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0566	0
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0501	0
TOT	257,009	(59,062)	(193,605)	(1,476)	2,866					23,797	21,854	2,866	222,594		0

EXHIBIT 7

Option Pricing Model—Normal/Constant Absolute Risk Aversion (Based on Doherty–Garven)

Sensitivity of Underwriting Profit Margin to Premium/Equity Ratios

Initial Equity S(0)	189,360	Premium P(0):	194,060
Fund Generating Coefficient k	1.50		
Std. Dev. of Investment Returns s(i)	0.20	Equity V(0):	189,360
Expected Claims Costs E[L]	193,605		
Std. Dev. of Claims Costs s(L)	48,401	V(0)–S(0):	0.000
Inv.-Claims Correlation r(iL)	0.00		
Risk-free Interest Rate r(f)	0.05		
Statutory Tax Rate t	0.34		
Tax-Adjustment Parameter h	0.80		
Beta of Inv. Portfolio B(i)	0.38	Base Case UPM:	0.002
Expected Return on Market E[r(M)]	0.13		
Std. Dev. of Market Return s(M)	0.22		

UPMs:

Premium/Equity Ratio:	0.50	0.70	0.90	1.10	1.30	1.50	1.70	1.90	2.10	2.30
r(iL) = −0.40	0.114	0.095	0.083	0.074	0.067	0.060	0.053	0.046	0.040	0.034
r(iL) = −0.20	0.083	0.063	0.051	0.043	0.036	0.030	0.024	0.018	0.012	0.007
r(iL) = 0.00	0.048	0.028	0.017	0.009	0.003	−0.003	−0.008	−0.013	−0.018	−0.023
r(iL) = 0.20	0.011	−0.010	−0.021	−0.028	−0.034	−0.038	−0.043	−0.047	−0.051	−0.055
r(iL) = 0.40	−0.030	−0.051	−0.062	−0.069	−0.074	−0.078	−0.082	−0.085	−0.088	−0.091

EXHIBIT 7

PART 2

OPTION PRICING MODEL

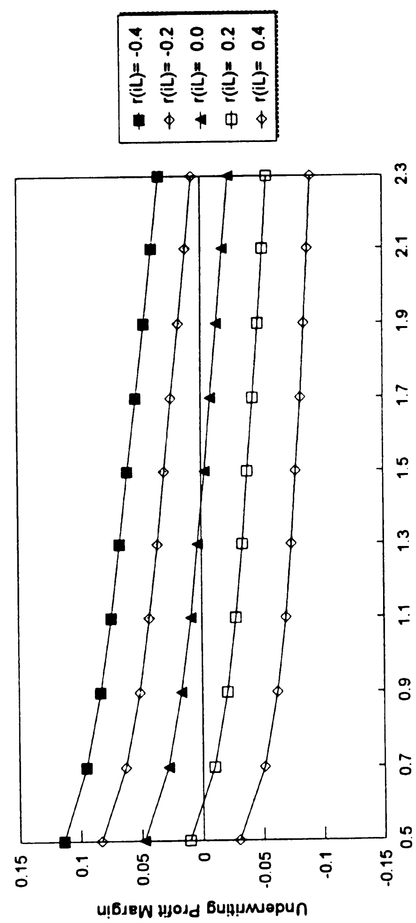


EXHIBIT 8

PART 1

ARBITRAGE PRICING MODEL
SENSITIVITY OF UNDERWRITING PROFIT MARGIN TO PREMIUM/EQUITY RATIO

Risk-Free Rate		5.0%	Base Case UPM:		−0.029					
Funds-Generating Coefficient		1.18								
Investment Tax Rate		27.2%								
Underwriting Tax Rate		34.0%								
Inflation Rate		3.0%								
Inflation Beta		0.500								
Industrial Prod. Growth Rate		2.0%								
Industrial Production Beta		0.250								
Premium/Equity Ratio:	0.5	0.7	0.9	1.1	1.3	1.5	1.7	1.9	2.1	2.3
UPM:	−0.004	−0.016	−0.022	−0.026	−0.029	−0.031	−0.033	−0.034	−0.035	−0.036

EXHIBIT 8

PART 1—PAGE 2

ARBITRAGE PRICING MODEL

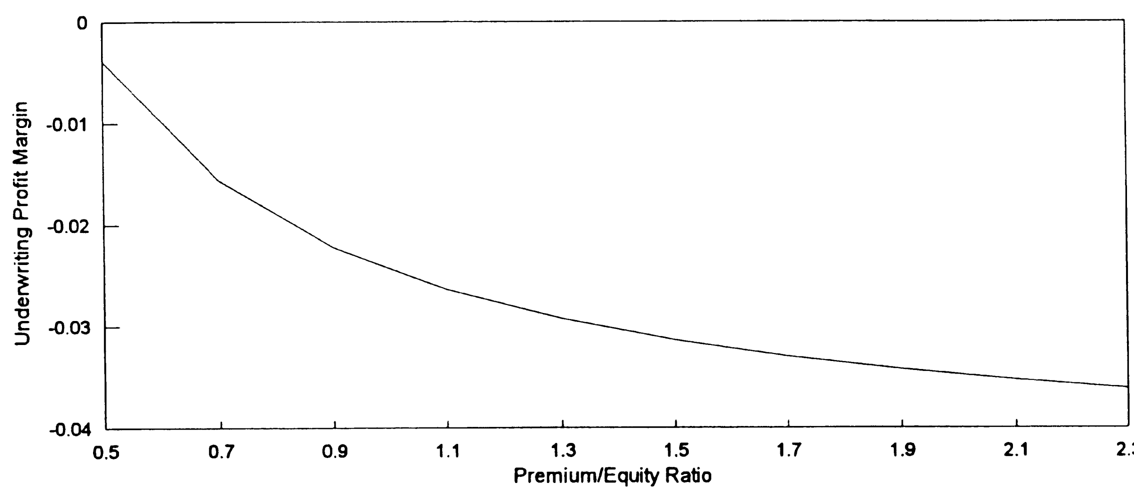


EXHIBIT 8

PART 2

ARBITRAGE PRICING MODEL
MULTIVARIATE REGRESSIONS OF PRE-SPECIFIED MACROECONOMIC FACTORS

This exhibit reports the results of multivariate regressions of pre-specified macroeconomic variables against the following dependent variable:

(Historical UPM + {(Historical Risk-Free Rate) × (Estimated Historical Funds-Generating Coefficient)}.

The regressions are performed with a specified constant of zero; a negative R-squared value indicates that the zero constant assumption is not appropriate. The time series for each of the five independent variables is taken from Ibbotson Associates [20] data, with the exception of the Change in Industrial Production, the source time series is A.M. Best's Aggregates & Averages data [1].

	CPI Change	Industrial Production Change	Bond Default Premium	Bond Horizon Premium	Value-Weighted NYSE	R-Squared
			1948–1993			
Coefficient	0.534	0.247	−0.414	−0.072	0.094	−0.000
Std. Error of Coeff.	0.152	0.117	0.263	0.082	0.039	
			1954–1993			
Coefficient	0.579	0.239	−0.283	−0.081	0.074	0.222
Std. Error of Coeff.	0.136	0.115	0.241	0.073	0.036	

EXHIBIT 8

PART 2—PAGE 2

	CPI Change	Industrial Production Change	Bond Default Premium	Bond Horizon Premium	Value-Weighted NYSE	R-Squared
			1964–1993			
Coefficient	0.715	0.236	−0.004	−0.045	−0.011	0.414
Std. Error of Coeff.	0.134	0.131	0.245	0.076	0.047	
			1974–1993			
Coefficient	0.673	0.258	−0.115	−0.090	−0.010	0.561
Std. Error of Coeff.	0.176	0.175	0.337	0.094	0.065	
			1948–1993			
Coefficient	0.683	0.242				−0.152
Std. Error of Coeff.	0.142	0.120				
			1964–1993			
Coefficient	0.710	0.245				0.395
Std. Error of Coeff.	0.109	0.125				
			1974–1993			
Coefficient	0.679	0.262				0.492
Std. Error of Coeff.	0.119	0.170				
			1984–1993			
Coefficient	0.191	0.231				0.116
Std. Error of Coeff.	0.176	0.175				

EXHIBIT 9

TOP 20 PRIVATE PASSENGER AUTO COMPANY FINANCIALS
For 1993 ($000)

Rank	Company Group	Assets	Liabilities	Surplus	NWP	Premium/ Surplus Ratio	Liabs./ Assets Ratio	Surplus/ Assets Ratio
1	State Farm	47,536,978	26,267,245	21,269,733	22,225,584	1.045	0.553	0.447
2	Allstate	27,698,530	20,553,474	7,145,056	15,801,617	2.212	0.742	0.258
3	Farmers	8,143,111	6,249,549	1,893,562	4,673,952	2.468	0.767	0.233
4	Nationwide	11,452,314	8,129,370	3,322,944	4,457,259	1.341	0.710	0.290
5	USAA	7,486,440	3,975,518	3,510,922	3,018,013	0.860	0.531	0.469
6	GEICO	3,307,701	2,477,676	830,025	1,882,237	2.268	0.749	0.251
7	Liberty Mutual	16,982,390	14,312,302	2,670,088	4,801,346	1.798	0.843	0.157
8	Prudential	2,690,053	2,045,524	644,529	1,519,302	2.357	0.760	0.240
9	Progressive	1,292,781	983,733	309,048	826,594	2.675	0.761	0.239
10	American Family	3,746,445	2,448,558	1,297,887	2,287,309	1.762	0.654	0.346
11	ITT Hartford	8,845,345	5,531,122	3,314,223	2,220,948	0.670	0.625	0.375
12	California State Auto Association	3,323,187	2,103,424	1,219,763	1,324,836	1.086	0.633	0.367
13	Travelers	10,368,761	8,122,916	2,245,845	2,238,146	0.997	0.783	0.217
14	Metropolitan	1,383,632	924,312	459,320	1,028,209	2.239	0.668	0.332
15	SAFECO	2,433,701	1,637,715	795,986	1,060,087	1.332	0.673	0.327
16	Aetna	2,809,288	2,138,362	670,926	1,097,364	1.636	0.761	0.239
17	20th Century	1,534,861	970,861	564,000	971,515	1.723	0.633	0.367
18	Erie	3,356,003	1,979,801	1,376,202	1,422,558	1.034	0.590	0.410
19	American Premier Underwriters	507,604	355,337	152,267	440,265	2.891	0.700	0.300
20	Allmerica	1,549,400	1,026,387	523,013	740,946	1.417	0.662	0.338
	TOTALS	166,448,525	112,233,186	54,215,339	74,038,087	1.366	0.674	0.326
	Average					1.690	0.690	0.310
	Standard Deviation					0.638	0.080	0.080
	Minimum					0.670	0.531	0.157
	Maximum					2.891	0.843	0.469
Per *1993 Aggregates and Averages* (1994) [1], PPA and HO Predominating:						1.601	0.681	0.319

Note: Per *A. M. Best Key Rating Guide* [3]; values for major PPA company within group.

EXHIBIT 10

PART 1

ASSETS
For 1993 ($000)

		Consolidated Industry		PPA and HO Predominating		Fictitious Company	
		Dollar Values	Proportion of Total Assets	Dollar Values	Proportion of Total Assets	Selected Proportion of Total Assets	Dollar Values
1	Bonds	417,776,558	0.622	135,480,072	0.627	0.627	296,679
2	Stocks:						
	2.1 Preferred Stocks	11,795,552	0.018	3,009,190	0.014	0.014	6,624
	2.2 Common Stocks	91,591,043	0.136	39,304,017	0.182	0.182	86,117
3	Mortgage Loans on Real Estate	4,472,774	0.007	272,741	0.001	0.001	473
4	Real Estate:						
	4.1 Properties Occupied by the Company	6,553,591	0.010	4,379,503	0.020	0.020	9,463
	4.2 Other Properties	2,152,292	0.003	451,660	0.002	0.002	946
5	Collateral Loans	76,435	0.000	21,608	0.000	0.000	0
6.1	Cash on Hand and on Deposit	5,189,814	0.008	1,859,610	0.009	0.009	4,259
6.2	Short-Term Investments	31,699,346	0.047	4,350,201	0.020	0.020	9,463
7	Other Invested Assets	7,321,187	0.011	1,506,706	0.007	0.007	3,312
8	Aggregate Write-Ins for Invested Assets	1,205,308	0.002			0.000	0
	8a Subtotals, Cash and Invested Assets	579,833,900	0.863	190,635,308	0.882	0.882	417,338
9	Agents' Balances or Uncollected Premiums:			14,482,101	0.067		
	9.1 In Course of Collection	14,930,076	0.022			0.020	9,463
	9.2 Booked but Deferred and Not Yet Due	27,724,017	0.041			0.038	17,981
	9.3 Accrued Retrospective Premiums	6,973,364	0.010			0.009	4,259

EXHIBIT 10

PART 1—PAGE 2

		Consolidated Industry		PPA and HO Predominating		Fictitious Company	
		Dollar Values	Proportion of Total Assets	Dollar Values	Proportion of Total Assets	Selected Proportion of Total Assets	Dollar Values
10	Funds Held By or Deposited With Reinsured Cos.	3,264,649	0.005	585,707	0.003	0.003	1,420
11	Bills Receivable, Taken for Premiums	1,065,926	0.002			0.000	0
12	Reinsurance Recoverable on LLAE Payments	10,579,195	0.016	1,200,334	0.006	0.006	2,839
13	Federal Income Tax Recoverable	1,959,091	0.003	1,027,140	0.005	0.005	2,366
14	Electronic Data Processing Equipment	2,083,622	0.003	1,347,599	0.006	0.006	2,839
15	Interest, Dividends and Real Estate Income D&A	7,926,145	0.012	2,761,278	0.013	0.013	6,151
16	Receivable from Parent, Subsidiaries, and Affils.	4,528,444	0.007	938,849	0.004	0.004	1,893
17	Equities & Deposits in Pools and Assocs.	2,491,321	0.004	730,527	0.003	0.003	1,420
18	A/R Relating to Uninsured A&H Plans	39,332	0.000			0.000	28
20	Aggregate Write-Ins for Other Than Invested Assets	8,139,150	0.012	2,485,871	0.011	0.011	5,205
21	TOTALS	671,538,233	1.000	216,194,714	1.000	1.000	473,172
	Selected Asset/Earned Premium Ratio:	2.008					

Note: Page 2, Annual Statement; per *Best's Aggregates & Averages—Property-Casualty* (1994) [1].

EXHIBIT 10

PART 2

LIABILITIES
For 1993 ($000)

		Consolidated Industry		PPA and HO Predominating		Fictitious Company	
		Dollar Values	Proportion of Total Liab + Surp	Dollar Values	Proportion of Total Liab + Surp	Selected Proportion of Total Liab + Surp	Dollar Values
1	Losses	280,522,065	0.418	73,866,581	0.342	0.342	161,825
1A	Reinsurance Payable on Paid Loss and LAE	1,820,414	0.003	407,265	0.002	0.002	946
2	Loss Adjustment Expenses	55,793,999	0.083	16,262,840	0.075	0.075	35,488
3	Contingent Commissions	1,942,406	0.003			0.003	1,420
4	Other Expenses	4,755,596	0.007	3,558,428	0.016	0.008	3,785
5	Taxes, Licenses, and Fees	2,647,299	0.004			0.005	2,366
6	Federal and Foreign Income Taxes	2,294,893	0.003	512,170	0.002	0.002	946
7	Borrowed Money	1,487,664	0.002	111,281	0.001	0.001	473
8	Interest	39,130	0.000	(286)	−0.000	0.000	0
9	Unearned Premiums	93,128,418	0.139	38,704,524	0.179	0.179	84,698
10	Dividends Declared and Unpaid:						
	(a) Stockholders	302,780	0.000	22,805	0.000	0.000	0
	(b) Policyholders	1,524,529	0.002	370,398	0.002	0.002	946
11	Funds Held by Company Under Reinsurance	6,941,513	0.010	1,333,242	0.006	0.006	2,839
12	Amounts Withheld or Retained	4,004,529	0.006	557,001	0.003	0.003	1,420
13	Provision for Reinsurance	3,258,238	0.005			0.000	0
14	Excess of Statutory Reserves	856,114	0.001			0.000	0
15	Net Foreign Exchange Adjustments	577,139	0.001	178,579	0.001	0.001	473

EXHIBIT 10

PART 2—PAGE 2

		Consolidated Industry		PPA and HO Predominating		Fictitious Company	
		Dollar Values	Proportion of Total Liab + Surp	Dollar Values	Proportion of Total Liab + Surp	Selected Proportion of Total Liab + Surp	Dollar Values
16	Drafts Outstanding	3,887,256	0.006	2,058,181	0.010	0.010	4,732
17	Payable to Parent, Subsidiaries, and Affiliates	3,486,585	0.005	881,546	0.004	0.004	1,893
18	Payable for Securities	2,010,972	0.003	396,826	0.002	0.002	946
19	Liability re Uninsured A&H Plans	22,112	0.000	170	0.000	0.000	0
20	Aggregates Write-Ins for Liabilities	17,959,263	0.027	7,916,048	0.037	0.036	17,034
21	Total Liabilities	489,262,914	0.729	147,137,599	0.681	0.681	322,230
22	Aggregate Write-Ins for Special Surplus Funds	15,356,274	0.023			0.000	0
23A	Common Capital Stock	6,539,022	0.010			0.000	0
23B	Preferred Capital Stock	736,974	0.001			0.000	0
23C	Aggregate Write-Ins for Other Than Special	2,347,697	0.003			0.000	0
24A	Gross Paid In and Contributed Surplus	73,893,613	0.110	29,817,617	0.138	0.137	64,825
24B	Unassigned Funds (Surplus)	83,803,704	0.125	39,239,497	0.182	0.182	86,117
24C	Less Treasury Stock, at Cost:						
	(1) Shares Common	378,920	0.001			0.000	0
	(2) Shares Preferred	23,045	0.000			0.000	16
25	Surplus as Regards Policyholders	182,275,319	0.271	69,057,114	0.319	0.319	150,958
26	TOTALS	671,538,233	1.000	216,194,713	1.000	1.000	473,172

Note: Page 3, Annual Statement; per *Best's Aggregates & Averages—Property-Casualty* (1994) [1].

EXHIBIT 10

PART 3

UNDERWRITING AND INVESTMENT EXHIBIT
For 1993 ($000)

		Consolidated Industry		PPA and HO Predominating		Fictitious Company	
	STATEMENT OF INCOME	Dollar Values	Proportion of Earned Premium	Dollar Values	Proportion of Earned Premium	Selected Proportion of Earned Premium	Dollar Values
	UNDERWRITING INCOME						
1	Premiums Earned	235,653,339	1.000	107,640,213	1.000	1.000	235,643
	DEDUCTIONS						
2	Losses Incurred	157,173,380	0.667	72,216,288	0.671	0.671	158,117
3	Loss Expenses Incurred	30,267,387	0.128	12,848,736	0.119	0.119	28,042
4	Other Underwriting Expenses Incurred	63,243,439	0.268	25,919,417	0.241	0.241	56,790
5	Aggregate Write-Ins for Underwriting Deductions	49,155	0.000	(560,465)	−0.005	−0.005	(1,178)
6	Total Underwriting Deductions	250,733,361	1.064	110,423,976	1.026	1.026	241,770
7	Net Underwriting Gain or (Loss)	(15,090,022)	−0.064	(2,783,763)	−0.026	−0.026	(6,127)
	INVESTMENT INCOME						
8	Net Investment Income Earned	32,645,415	0.139			0.078	18,380
9	Net Realized Capital Gains or (Losses)	9,817,573	0.042			0.023	5,420
9A	Net Investment Gain or (Loss)	42,462,988	0.180	10,883,286	0.101	0.101	23,800

EXHIBIT 10

PART 3—PAGE 2

		Consolidated Industry		PPA and HO Predominating		Fictitious Company	
	STATEMENT OF INCOME	Dollar Values	Proportion of Earned Premium	Dollar Values	Proportion of Earned Premium	Selected Proportion of Earned Premium	Dollar Values
	OTHER INCOME						
10	Net Balance Charge-Off Gain or (Loss)	(446,819)	−0.002			0.000	0
11	Finance and Service Charges	1,047,268	0.004			0.000	0
12	Aggregate Write-Ins for Miscellaneous Items	(895,740)	−0.004			0.000	0
13	Total Other Income	(295,291)	−0.001			0.000	0
14	Net Income Before Dividends and Income Taxes	27,077,675	0.115	8,099,523	0.075	0.075	17,673
14A	Dividends to Policyholders	2,709,126	0.011	1,124,630	0.010	0.010	2,356
14B	Net Income After Dividends, Before Income Taxes	24,368,549	0.103	6,974,893	0.065	0.065	15,317
15	Federal and Foreign Income Taxes Incurred	5,053,041	0.021	1,781,519	0.017	0.017	4,006
16	Net Income	19,315,508	0.082	5,193,374	0.048	0.048	11,311
	Selected Proportion of Company Earned Premium To That of Entire P/C Industry:	0.001					

Note: Page 4, Annual Statement; per *Best's Aggregates & Averages—Property-Casualty* (1994) [1].

EXHIBIT 10

PART 4

SUPPLEMENTAL FINANCIAL STATEMENT INFORMATION

Calculation of Adjusted Surplus	Base Case	Low	High
Statutory Surplus	150,958	83,730	361,290
Equity in the Unearned Premium Reserve	20,412	20,412	20,412
Nominal–Discounted Loss Reserves	8,289	8,289	8,289
Excess of Statutory Reserves	0	0	0
Market Value–Book Value of Bonds	13,928	13,928	13,928
Non-Admitted Assets	946	946	946
Tax Liability on Unrealized Capital Gains	(5,173)	(5,173)	(5,173)
Total Adjustments	38,402	38,402	38,402
Adjusted Surplus	189,360	122,132	399,692
Range: Premium to Surplus Ratio		2.891	0.670
Statutory Surplus		83,730	361,290

Funds-Generating Coefficient	
Prior Year's Losses	186,159
Time-Weighted Paid Losses	277,003
Prior Year's Earned Premium	235,643
Funds-Generating Coefficient	1.18

Investment Income Tax Rate	
Net Investment Income (Company)	18,380
Realized Capital Gaines (Company)	5,420
% Bonds Exempt from US Tax	0.292
% Dividends	0.086
Exclusion (70% of Dividends)	0.060
Tax Rate on Net Investment Income	0.273
Total Tax Rate on Investment Income	0.275
Ratio to Total Tax Rate	0.808

Investment Rate of Return	1992		1993
Total Invested Assets (PPA/HO)	177,103		193,397
Average Invested Assets		185,250	
Net Investment Income		10,883	
Realized Capital Gains		2,638	
Unrealized Capital Gains		1,786	
Total Investment Income		15,307	
Investment Rate of Return		0.083	

Payout Patterns

Year	Base	Fast	Slow
1	0.531	0.691	0.366
2	0.241	0.151	0.331
3	0.105	0.073	0.147
4	0.056	0.039	0.071
5	0.030	0.020	0.038
6	0.015	0.011	0.020
7	0.009	0.006	0.011
8	0.005	0.003	0.006

Payout Patterns

Year	Base	Fast	Slow
9	0.002	0.001	0.003
10	0.001	0.001	0.001
11	0.001	0.001	0.001
12	0.001	0.001	0.001
13	0.001	0.001	0.001
14	0.001	0.001	0.001
15	0.001	0.000	0.001

EXHIBIT 11

PART 1

PRIVATE PASSENGER AUTO LIABILITY/MEDICAL CONSOLIDATED INDUSTRY TOTALS

	Cumulative Net Paid Losses and ALAE ($000)									
Accident Year	12	24	36	48	60	72	84	96	108	120
1984	7,107,267	13,730,807	16,904,897	18,676,836	19,654,989	20,147,446	20,396,113	20,542,284	20,618,789	20,661,498
1985	7,797,650	15,395,558	19,096,053	21,216,278	22,337,968	22,903,322	23,207,201	23,358,240	23,433,206	
1986	8,680,599	17,211,774	21,458,533	23,840,126	25,120,488	25,765,753	26,075,742	26,243,078		
1987	9,674,040	19,363,480	24,098,471	26,757,800	28,171,290	28,854,812	29,247,314			
1988	10,902,007	21,694,958	26,909,462	29,735,250	31,204,226	31,949,941				
1989	12,016,958	24,023,854	29,718,129	32,761,578	34,382,265					
1990	13,250,396	25,970,550	32,000,373	35,193,034						
1991	13,305,353	25,822,406	31,810,715							
1992	14,372,292	27,676,795								
1993	15,656,520									

Note: Per *Best's Aggregates & Averages—Property-Casualty* (1994) [1].

EXHIBIT 11

PART 1—PAGE 2

Accident Year	12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	120–Ult.	
1984	1.932	1.231	1.105	1.052	1.025	1.012	1.007	1.004	1.002		
1985	1.974	1.240	1.111	1.053	1.025	1.013	1.007	1.003			
1986	1.983	1.247	1.111	1.054	1.026	1.012	1.006				
1987	2.002	1.245	1.110	1.053	1.024	1.014					
1988	1.990	1.240	1.105	1.049	1.024						
1989	1.999	1.237	1.102	1.049							
1990	1.960	1.232	1.100								
1991	1.941	1.232									
1992	1.926										
1993											
Average	1.967	1.238	1.106	1.052	1.025	1.013	1.007	1.003	1.002		
Wtd. Avg.	1.966	1.238	1.106	1.052	1.025	1.013	1.007	1.003	1.002		
3-Yr Wt Avg.	1.942	1.234	1.102	1.050	1.025	1.013	1.007	1.003	1.002		
Selected	1.942	1.234	1.102	1.050	1.025	1.013	1.007	1.003	1.002	1.008	
To-Ultimate	2.937	1.513	1.226	1.112	1.059	1.034	1.020	1.014	1.010	1.008	
% of Ult.	34.1%	66.1%	81.6%	89.9%	94.4%	96.8%	98.0%	98.7%	99.0%	99.2%	
Increm. %	34.1%	32.1%	15.4%	8.3%	4.5%	2.3%	1.3%	0.7%	0.3%	0.2%	0.8%
Avg. Pmt. Date	0.50	1.50	2.50	3.50	4.50	5.50	6.50	7.50	8.50	9.50	12.50
Avg. Time Until Pmt.	1.94										

EXHIBIT 11

PART 2

PRIVATE PASSENGER AUTO LIABILITY/MEDICAL CONSOLIDATED INDUSTRY TOTALS

	Incurred Losses and ALAE—Including Bulk and IBNR ($000)									
Accident Year	12	24	36	48	60	72	84	96	108	120
1984	19,881,411	20,458,822	20,723,492	20,788,321	20,793,344	20,796,050	20,803,403	20,818,151	20,819,122	20,826,114
1985	22,640,558	23,331,210	23,605,717	23,648,833	23,681,530	23,667,361	23,678,227	23,669,417	23,662,652	
1986	26,510,270	26,736,641	26,739,048	26,687,523	26,650,450	26,632,126	26,577,414	26,555,719		
1987	30,132,232	30,073,743	29,980,050	29,935,567	29,905,974	29,846,067	29,814,502			
1988	33,939,925	33,682,143	33,462,947	33,316,467	33,103,685	32,999,670				
1989	37,238,902	37,047,111	36,872,244	36,540,543	36,377,267					
1990	40,941,704	40,335,239	39,609,692	39,257,602						
1991	41,803,021	40,529,610	39,643,933							
1992	44,930,921	43,143,619								
1993	47,594,265									

Note: Per *Best's Aggregates & Averages—Property-Casualty* (1994) [1].

EXHIBIT 11

PART 2—PAGE 2

Accident Year	12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	120–Ult.	
1984	1.029	1.013	1.003	1.000	1.000	1.000	1.001	1.000	1.000		
1985	1.031	1.012	1.002	1.001	0.999	1.000	1.000	1.000			
1986	1.009	1.000	0.998	0.999	0.999	0.998	0.999				
1987	0.998	0.997	0.999	0.999	0.998	0.999					
1988	0.992	0.993	0.996	0.994	0.997						
1989	0.995	0.995	0.991	0.996							
1990	0.985	0.982	0.991								
1991	0.970	0.978									
1992	0.960										
1993											
Average	0.996	0.996	0.997	0.998	0.999	0.999	1.000	1.000	1.000		
Wtd. Avg.	0.991	0.994	0.996	0.998	0.999	0.999	1.000	1.000	1.000		
3-Yr Wt Avg.	0.971	0.985	0.992	0.996	0.998	0.999	1.000	1.000	1.000		
Selected	0.971	0.985	0.992	0.996	0.998	0.999	1.000	1.000	1.000	1.000	
To-Ultimate	0.943	0.971	0.985	0.993	0.997	0.999	1.000	1.000	1.000	1.000	
% of Ult.	106.1%	103.0%	101.5%	100.7%	100.3%	100.1%	100.0%	100.0%	100.0%	100.0%	
Increm. %	106.1%	−3.0%	−1.6%	−0.8%	−0.4%	−0.2%	−0.1%	−0.0%	−0.0%	0.0%	0.0%

EXHIBIT 11

PART 3

PRIVATE PASSENGER AUTO LIABILITY/MEDICAL CONSOLIDATED INDUSTRY TOTALS

	Case Incurred Losses and ALAE ($000)									
Accident Year	12	24	36	48	60	72	84	96	108	120
1984	14,971,043	18,740,436	19,858,672	20,351,846	20,561,437	20,657,368	20,711,038	20,743,177	20,758,601	20,777,998
1985	16,658,056	21,135,709	22,525,437	23,165,482	23,402,845	23,472,725	23,543,804	23,562,106	23,578,804	
1986	18,955,064	23,818,148	25,431,280	26,051,927	26,252,369	26,376,438	26,411,278	26,435,323		
1987	21,538,487	26,777,218	28,445,740	29,103,087	29,444,813	29,539,908	29,628,106			
1988	24,123,339	29,860,663	31,489,589	32,288,202	32,544,480	32,658,638				
1989	26,546,705	32,652,373	34,673,136	35,444,420	35,762,026					
1990	28,779,228	35,233,227	37,193,945	37,956,021						
1991	29,037,439	35,149,766	36,938,537							
1992	31,393,255	37,390,576								
1993	33,994,407									

Note: Per *Best's Aggregates & Averages—Property-Casualty* (1994) [1].

EXHIBIT 11

PART 3—PAGE 2

Accident Year	12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	120–Ult.	
1984	1.242	1.060	1.025	1.010	1.005	1.003	1.002	1.001	1.001		
1985	1.269	1.066	1.028	1.010	1.003	1.003	1.001	1.001			
1986	1.257	1.068	1.024	1.008	1.005	1.001	1.001				
1987	1.243	1.062	1.023	1.012	1.003	1.003					
1988	1.238	1.055	1.025	1.008	1.004						
1989	1.230	1.062	1.022	1.009							
1990	1.224	1.056	1.020								
1991	1.210	1.051									
1992	1.191										
1993											
Average	1.235	1.060	1.024	1.009	1.004	1.002	1.001	1.001	1.001		
Wtd. Avg.	1.230	1.059	1.024	1.009	1.004	1.002	1.001	1.001	1.001		
3-Yr Wt Avg.	1.208	1.056	1.023	1.009	1.004	1.002	1.001	1.001	1.001		
Selected	1.208	1.056	1.023	1.009	1.004	1.002	1.001	1.001	1.001	1.002	
To-Ultimate	1.332	1.102	1.044	1.021	1.011	1.007	1.005	1.004	1.003	1.002	
% of Ult.	75.1%	90.7%	95.8%	98.0%	98.9%	99.3%	99.5%	99.6%	99.7%	99.8%	
Increm. %	75.1%	15.6%	5.1%	2.2%	0.9%	0.4%	0.2%	0.1%	0.1%	0.1%	0.2%

EXHIBIT 11

PART 4

PRIVATE PASSENGER AUTO LIABILITY/MEDICAL CONSOLIDATED INDUSTRY TOTALS

Earned Premiums ($000)

Accident Year	Earned Premiums
1984	23,929,028
1985	26,545,102
1986	31,282,435
1987	35,721,716
1988	39,659,210
1989	42,942,782
1990	46,792,329
1991	49,990,387
1992	54,155,823
1993	57,846,167

Ratio of Paid Loss and ALAE to Earned Premiums

Accident Year	12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	120–Ult.
1984	0.297	0.574	0.706	0.781	0.821	0.842	0.852	0.858	0.862	0.863
1985	0.294	0.580	0.719	0.799	0.842	0.863	0.874	0.880	0.883	
1986	0.277	0.550	0.686	0.762	0.803	0.824	0.834	0.839		
1987	0.271	0.542	0.675	0.749	0.789	0.808	0.819			
1988	0.275	0.547	0.679	0.750	0.787	0.806				
1989	0.280	0.559	0.692	0.763	0.801					
1990	0.283	0.555	0.684	0.752						
1991	0.266	0.517	0.636							
1992	0.265	0.511								
1993	0.271									

EXHIBIT 11

PART 4—PAGE 2

Accident Year	Ratio of Case Incurred Loss and ALAE to Earned Premiums									
	12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	120–Ult.
1984	0.626	0.783	0.830	0.851	0.859	0.863	0.86	0.867	0.868	0.868
1985	0.628	0.796	0.849	0.873	0.882	0.884	0.887	0.888	0.888	
1986	0.606	0.761	0.813	0.833	0.839	0.843	0.844	0.845		
1987	0.603	0.750	0.796	0.815	0.824	0.827	0.829			
1988	0.608	0.753	0.794	0.814	0.821	0.823				
1989	0.618	0.760	0.807	0.825	0.833					
1990	0.615	0.753	0.795	0.811						
1991	0.581	0.703	0.739							
1992	0.580	0.690								
1993	0.588									

Note: Per *Best's Aggregates & Averages—Property-Casualty* (1994) [1].

EXHIBIT 11

PART 5

AUTO PHYSICAL DAMAGE
CONSOLIDATED INDUSTRY TOTALS

Cumulative Net Paid Losses and ALAE ($000)

Accident Year	12	24	12–24
1989		21,805,677	
1990	19,625,334	21,450,232	1.093
1991	19,037,412	20,637,002	1.084
1992	19,457,475	21,105,812	1.085
1993	20,529,226		

Cumulative Case Incurred Losses and ALAE ($000)

Accident Year	12	24	Ult. Est.	12–24	24–Ult.
1989		21,898,506	22,058,959		1.007
1990	21,415,626	21,618,021	21,741,587	1.009	1.006
1991	20,716,914	20,721,080	20,800,355	1.000	1.004
1992	21,075,920	21,148,591	21,236,170	1.003	1.004
1993	22,075,313				

Ratios of Loss and ALAE to Earned Premiums

Accident	Earned	Paid Loss and ALAE		Case Incurred Loss and ALAE		
Year	Premiums	12	24	12	24	Ult. Est.
1989	34,179,748		0.638		0.641	0.645
1990	35,015,802	0.560	0.613	0.612	0.617	0.621
1991	35,721,624	0.533	0.578	0.580	0.580	0.582
1992	36,737,466	0.530	0.575	0.574	0.576	0.578
1993	37,864,502	0.542		0.583		

Note: Per *Best's Aggregates & Averages—Property-Casualty* (1994) [1].

APPENDIX

In this appendix, we show examples of specific calculations for each of the six financial pricing models examined in this paper. It is hoped that these examples will provide further insight into the models, as well as encourage actuaries to implement some of these techniques themselves.

Target Total Rate of Return

Underwriting profit margins resulting from the Target Total Rate of Return model are shown in Exhibit 3. These UPMs are generated directly from Equation 2.6 in the text, assuming the parameters given on the exhibit. The Target TRR method equates the sum of an insurance company's underwriting and investment returns with a target total rate of return; this target, in our paper, is based on the Capital Asset Pricing Model. Thus, calculations of UPMs according to the Target TRR model require assumptions regarding values for the following parameters: the risk-free interest rate, the expected return on the equity market (which is equal to the risk-free rate plus an equity "risk premium"), the equity beta of the insurer, the company's invested assets and the rate of return on those investments, and the company's equity and its premium-to-equity ratio. Assumptions regarding five of these seven variables are documented at the top of Exhibit 3; alternative assumptions regarding the premium-to-equity ratio and the equity beta are shown below those, and on the graph on the exhibit.

As an example, the 0.236 UPM that is shown on the exhibit (assuming an equity beta of 1.70 and a premium-to-equity ratio of 0.50) is derived from Equation 2.6 as follows:

$$UPM = \left(\frac{1}{0.50}\right)\left([0.05 + 1.70(0.13 - 0.05)] - \frac{417{,}338 \times 0.08}{189{,}360 \times \dfrac{1.30}{0.50}}\right)$$

$$= 0.236,$$

where the (1.30/0.50) factor is an adjustment to equity to bring the premium-to-equity ratio to the assumed value of 0.50. (It is assumed that, for the base case corresponding to an equity value of 189,360, the premium-to-equity ratio is 1.30; thus, to test the sensitivity of the model to a different leverage and a different premium-to-equity ratio, the *equity* is adjusted in the UPM calculation—the premium level is held constant. Assuming that an equity value of 189,360 corresponds to a premium-to-equity ratio of 1.30, a ratio of 0.50 implies an equity of 492,336.)

Insurance Capital Asset Pricing Model

Underwriting profit margins resulting from the Insurance CAPM are shown in Exhibit 4. These UPMs are generated directly from Equation 2.8 in the text, assuming the parameters given on the exhibit. We use a differential tax version of the Insurance CAPM, and show the sensitivity of UPMs to changes in the underwriting beta and the premium-to-equity ratio on the exhibit and in the graph.

As an example, the 0.008 UPM on Exhibit 4, under underwriting beta and premium-to-equity ratio assumptions of 0.40 and 0.50, respectively, is calculated as follows:

$$UPM = -1.18 \times 0.05\frac{1-0.272}{1-0.340} + 0.40(0.13-0.05) + \frac{1}{0.50}0.05\frac{0.272}{1-0.340} = 0.008.$$

Discounted Cash Flow Model

Underwriting profit margins resulting from the DCF Model are shown in Exhibit 5. These UPMs are based on the concepts underlying Equation 2.9 in the text. In this framework, the present value of the premiums is set equal to the present value of all the cash flows emanating from the policy, including expenses, losses (and LAE), taxes on underwriting and taxes on investment income. The specific formula used (with one adjustment—see below) is as follows:

$$
\begin{aligned}
P\sum_{i=0}^{N}\frac{a_i}{(1+r_f)^i}
&= L\sum_{i=0}^{N}\frac{b_i}{(1+r_L)^i} + E\sum_{i=-M}^{N}\frac{c_i}{(1+r_f)^i} \\
&\quad + \frac{\left(P - E\sum_{i=-M}^{N}\frac{c_i}{(1+r_f)^i}\right)t}{1+r_f} \\
&\quad - Lt\left(\frac{\sum_{i=1}^{N}\frac{b_i}{(1+r_T)^{i-1}}}{1+r_L} + \sum_{j=2}^{N}\frac{\sum_{i=j}^{N}\frac{r_T b_i}{(1+r_T)^{i-j+1}}}{(1+r_L)^j}\right) \\
&\quad + r_f t\left(\sum_{j=1}^{N}\left[\frac{S\left(\sum_{i=j}^{N} b_i\right) + P - E - L\sum_{i=0}^{j-1} b_i}{(1+r_f)^j}\right]\right),
\end{aligned}
$$

where

a_i = fraction of premium received in time period i,

b_i = fraction of losses paid in time period i,

c_i = fraction of expenses paid in time period i,

S = owners' equity in insurer,

P = premiums,

L = losses and loss adjustment expenses,

E = underwriting expenses,

t = tax rate,

r_T = discount rate required for tax purposes,

r_f = risk free rate,

r_L = risk adjusted rate for losses,

M = number of time periods before policy effective date that the first prepaid expenses are paid, and

N = number of time periods after policy effective date that the last loss payment is made.

The above DCF formula—with an adjustment to allow for differential tax rates between underwriting and investment income—is implemented via a spreadsheet program in which future annual expected loss, expense, and tax cash flows are discounted to determine a fair premium. In order to solve this equation, we need to know the rates at which premium income is received and expenses, losses (including LAE) and taxes are paid, and the discount rates to use to calculate the present values. The cash flows are discounted at different rates. Premiums, expenses and investment income (which is assumed to be earned based on the risk-free rate) are discounted based on the risk-free rate. Losses (and LAE) and taxes on underwriting income based on losses, are discounted at a risk-adjusted rate. As indicated on Exhibit 5, for the base case, the risk-free interest rate is 5%, the risk-adjusted discount rate is 3% and the tax discount rate is 7%.

For the base case, it is assumed that the premium is received and expenses paid entirely when the policy is written. Losses (and LAE) are assumed to be paid in the middle of each year, with the loss payout pattern shown in Table 1 (based on Best's *Aggregates and Averages*, 1994). The discounted values of the loss payments are determined by dividing the percent of losses paid by $(1.03)^{(\text{Year}-0.5)}$.

Taxes are assumed to be paid at the end of each year. Taxes on underwriting income are determined based on the difference between premiums, and expenses and losses, with the loss reserves discounted based on the provisions of the Tax Reform Act of 1986. Thus, the incurred losses for tax purposes are the

TABLE 1

Year	% of Losses Paid	Discounted Value
1	0.531	0.523210
2	0.241	0.230548
3	0.105	0.097521
4	0.056	0.050496
5	0.030	0.026264
6	0.015	0.012749
7	0.009	0.007427
8	0.005	0.004006
9	0.002	0.001556
10	0.001	0.000765
11	0.001	0.000733
12	0.001	0.000712
13	0.001	0.000691
14	0.001	0.000671
15	0.001	0.000651
Total	1.000	0.957989

paid losses each year plus the ending reserves (discounted at the mandated seven percent rate) minus the beginning loss reserves (discounted at seven percent). Since losses are paid out over a 15-year period in this example, then the taxes are also paid out over the same 15-year period.

The calculation of taxes based on underwriting income is determined by a spreadsheet, which is available from the authors, that runs over the entire 15-year period. For the fifteenth (last) year, the incurred losses are the paid losses of 193.605 (.001 times the incurred losses of $193,605) minus the beginning reserve (discounted for half a year based on the mandated seven percent discount rate) of 187.165, or 6.44. This incurred loss is multiplied by the tax rate applicable to underwriting, 34 percent. This negative tax payment is then discounted back for 15 years, based on the risk-adjusted rate of three percent. Similar calculations are performed for every other year to determine the effect of the underwriting tax on losses. In addition, for the first year, the tax rate is multiplied by the difference between the premiums and expenses (both assumed to be paid at the inception of the policy). This tax payment is discounted back for one year (taxes are paid at the end of the year) at the risk-free rate of five percent.

The calculation of taxes based on investments is also determined by a spreadsheet that runs over the 15-year period during which losses are paid. Investment income is earned on the surplus allocated to the policy, which is released in proportion with loss payments, and the difference between the premiums received and expenses and cumulative losses paid out. Investment income is assumed to be earned based on the risk-free rate and discounted at the risk-free rate. For the first year, the entire surplus plus the premiums less expenses is invested. For the second year, 46.9 percent of the inital surplus (since 53.1 percent of the losses have been paid in the first year) plus the premiums less expenses and less 53.1 percent of the losses are invested. This pattern continues until all the losses are paid after 15 years.

For the base case, Equation (2.9) can be broken down as follows:

$\mathrm{PV}(P) = P$.

$\mathrm{PV}(L) = 193{,}605 \times (.957989)$.

$\mathrm{PV}(E) = 59{,}062$.

$\mathrm{PV}(UWPT)$ = PV(tax on premiums) − PV(tax on expenses)− PV (tax on losses).
PV(tax on premiums) $= 0.34 \times (1/1.05) \times P = 0.323810P$.
PV(tax on expenses) $= 0.34 \times (59{,}062) \times (1/1.05) = 19{,}125$.
PV(tax on losses) $= 0.34 \times (193{,}605) \times (.968011) = 63{,}720$.

$\mathrm{PV}(IBT)$ = PV(taxes on investment income from surplus)+ PV(taxes on investment income from premiums minus expenses minus paid losses).

PV(taxes on investment income from surplus) = $0.05 \times (0.272) \times (189{,}360) \times (1.887415)/(1.05)$.
PV(taxes on investment income from premiums minus expenses minus paid losses) = $0.05 \times (.272) \times (P - 59{,}062 -$ losses paid to date)/(1.05).

Solving for P yields 253,040. The underwriting profit margin associated with this premium is:

$$UPM = \frac{253{,}040 - 193{,}605 - 59{,}062}{253{,}040} = 0.001.$$

Internal Rate of Return Model

Underwriting profit margins resulting from the IRR model are shown in Exhibit 6, Part 1, along with the relevant parameter assumptions. The spreadsheet underlying the UPM calculations is shown in Exhibit 6, Part 2—the numbers on the sheet represent base case calculations, with an internal rate of return of 13% (equal to the expected market return of 5% risk-free plus an 8% market risk premium). The leftmost column indicates the timing of the cashflows; quarterly for the first two years and then annual. This spreadsheet can accomodate 25 years of payouts, but for the base case example all losses are settled within 15 years (60 quarters). The next column, labeled (1), is the premium income, which is calculated by the backsolver routine. Column 2 shows the expenses, an outgo, which are given and assumed to be all paid in the first quarter. Column 3 shows the loss payments. Column 4 shows the federal tax cash flow, calculated as $t(P - E -$ discounted losses). The taxes are calculated on an annual basis and, for the first two years, spread evenly over the four quarters.

Column 5 is the cash flow from underwriting, which is the sum of Columns 1 through 4. This cash flow totals \$2,866. The insurer is assumed, under this model, to receive this sum as soon as the policy is written. (This model was originally developed for workers compensation, which operated at an underwriting loss. The insurer was assumed to have to fund this underwriting loss at the inception of the policy. The same timing is assumed in this situation.) Column 6 is the accumulated value of the underwriting account. For the first quarter, the value is half of Column 5 plus the initial underwriting flow. For subsequent rows, the value is the average of Column 5 and Column 5 lagged one period plus Column 6 lagged one period.

Column 7 is the total required loss reserve, which is the total losses (193,605) minus paid to date. For the first year, losses are assumed to be incurred evenly over the year. Column 8 is the remaining surplus. The initial surplus allocated to the line is 189,360. Surplus is released as losses are paid. Column 9 is the average surplus, the average of Column 8 and Column 8 lagged one period. Column 10 is investment income on surplus; for the first eight quarters it is Column 9 times two percent; after that it is Column 9 times eight percent. Column 11 is the investment income on underwriting, which is two (or eight) percent of Column 6.

Column 12 shows the cash flows from surplus. For the top row, it is the sum of Columns 5 and 8. For each remaining row, it is Column 8 lagged by one period minus Column 8. Column 13, the net cash flow to capital providers, is the sum of the surplus cash flow (Column 12) and the after-tax investment returns (0.728 times the sum of Columns 10 and 11). Column 14 is the discount factor at the IRR rate; in this example, 13 percent annually. For the first quarter it is $(1/1.13)^{.125}$ (since the payment occurs midway through the first quarter). For the last row it is $(1/1.13)^{24.5}$. Column 15 is Column 13 times Column 14.

UPMs for non-base case parameter assumptions are derived by changing the parameter values, and then "backsolving" the Part 2 spreadsheet by changing the fair premium in Column 1 until the total discounted net cash flow in Column 15 is zero. (This can typically be done in spreadsheet programs by using a "backsolver" function.) As for the DCF model, the indicated

UPMs are then calculated based on the relationships between the premiums and expected losses and expenses. For example, the base case UPM of 0.017 on Exhibit 6, Part 1, is calculated as follows:

$$UPM = \frac{255{,}680 - 193{,}605 - 57{,}733}{255{,}680} = 0.017.$$

Option Pricing Model

Underwriting profit margins resulting from the OPM are shown in Exhibit 7. These UPMs are generated from a spreadsheet modeled after Equation 19 in Doherty and Garven [15]:

$$V_e = R_f^{-1}\left(\mathrm{E}^*(X)\mathrm{N}\left[\frac{\mathrm{E}^*(X)}{\sigma_x}\right] - \tau\mathrm{E}^*(W)\mathrm{N}\left[\frac{\mathrm{E}^*(W)}{\sigma_w}\right] + \sigma_x n\left[\frac{\mathrm{E}^*(X)}{\sigma_x}\right] - \tau\sigma_w \mathrm{n}\left[\frac{\mathrm{E}^*(W)}{\sigma_w}\right]\right),$$

where

V_e = the market value of the residual claim of the shareholders,

$R_f = 1 + r_f$,

E^* = the certainty-equivalent expectation operator,

τ = the corporate tax rate,

σ = standard deviation,

X = random variable representing the insurer's pre-tax end-of-period value after paying claims costs $\times$ (initial equity + premium income + investment income – claims costs),

W = random variable representing the insurer's taxable income (from both underwriting and investments), N[ ] = standard normal distribution value, and

n[ ] = standard normal density value.

This equation can be solved iteratively to determine the fair premium (in the spreadsheet, by using backsolver), since the premium level (net of expenses) is embedded in the X and W random variables above. The "fair" premium is defined where the market value of the residual claim of the shareholders is equal to the initial equity. For the base case, the following values apply:

$$V_e = 189{,}360.$$

$$R_f = 1.05.$$

$$\lambda = [\mathrm{E}(r_M) - r_f]/s^2(M) = 1.594388.$$

$$\mathrm{E}^*[L] = \mathrm{E}[L] - [\lambda/\beta(i)] \times [\rho(iL) \times s(i) \times s(L)] = 193{,}605.$$

$$\begin{aligned} \mathrm{E}^*[X] &= S(0) + [(S(0) + (k \times P(0))) \times r_f] + P(0) - \mathrm{E}^*[L] \\ &= 189{,}360 + [(189{,}360 + 1.5P(0))0.05] + P(0) - 193{,}605 \\ &= 213{,}837. \end{aligned}$$

$$\begin{aligned} s(X) &= \{[(S(0) + (k \times P(0)))^2 \times s^2(i)] + s^2(L) \\ &\quad - [2 \times (S(0) + (k \times P(0))) \times \rho(iL) \times s(i) \times s(L)]\}^{0.5} \\ &= 107{,}592. \end{aligned}$$

$$\begin{aligned} \mathrm{E}^*[W] &= [h \times (S(0) + (k \times P(0))) \times r_f] + P(0) - \mathrm{E}^*[L] \\ &= [0.8 \times (189{,}360 + 1.5P(0)) \times 0.05] + P(0) - 193{,}605 \\ &= 19{,}673. \end{aligned}$$

$$\begin{aligned} s(W) &= \{[(S(0) + (k \times P(0)))^2 \times s^2(i) \times h^2] + s^2(L) \\ &\quad - [2 \times (S(0) + (k \times P(0))) \times h \times \rho(iL) \times s(i) \times s(L)]\}^{0.5} \\ &= 90{,}840, \end{aligned}$$

where

$$M = \text{the market,}$$
$$i = \text{investment returns, and}$$
$$L = \text{claims costs.}$$

The system is then solved for $P(0)$, which turns out to be 194,060. To check that this value is correct, substitute the above values into the equation:

$$\begin{aligned}
189{,}360 = {} & (1/1.05) \\
& \times \{(213{,}837) \times \mathrm{N}[213{,}837/107{,}592] \\
& \quad - 0.34 \times (19{,}673) \times \mathrm{N}[19{,}673/90{,}840] + 107{,}592 \\
& \quad \times \mathrm{n}[213{,}837/107{,}592] - 0.34 \times (90{,}840) \\
& \quad \times \mathrm{n}[19{,}673/90{,}840]\} \\
= {} & 0.952381 \times \{[(213{,}837) \times (0.976566)] \\
& \quad - [0.34 \times (19{,}673) \times (0.585726)] \\
& \quad + [(107{,}592) \times (0.55354)] \\
& \quad - [0.34 \times (90{,}840) \times (.389696)]\} \\
= {} & 189{,}360.
\end{aligned}$$

Finally, expenses are added in to the premium, and the UPM is calculated in the usual way—for example, for the base case on Exhibit 7:

$$UPM = \frac{(194{,}060 + 59{,}062) - 193{,}605 - 59{,}062}{(194{,}060 + 59{,}062)} = 0.002.$$

Arbitrage Pricing Model

Underwriting profit margins resulting from the APT are shown in Exhibit 8, Part 1. These UPMs are generated directly from Equation 2.13 in the text, assuming the parameters given on the exhibit. We use a differential tax version of the Insurance APT and show the sensitivity of UPMs to changes in the premium-to-equity ratio on the exhibit and in the graph. Part 2 of Exhibit 8 shows the results of regressions of five different macroeconomic variables against historical underwriting profit margins. Based on these regressions, it was decided to use inflation and the growth in industrial production as the explanatory factors in the Arbitrage Pricing formula.

As an example, the -0.029 base case UPM on Part 1 is calculated as follows:

$$UPM = -\frac{1 - 0.272}{1 - 0.340} \times 0.05 \times 1.18 + (0.50 \times 0.03) + (0.25 \times 0.02) + 0.05\left(\frac{1}{1.3}\right)\left(\frac{0.272}{1 - 0.340}\right) = -0.029.$$

SMOOTHING WEATHER LOSSES: A TWO-SIDED PERCENTILE MODEL

CURTIS GARY DEAN, DAVID N. HAFLING, MARK S. WENGER, AND WILLIAM F. WILSON

Abstract

This paper presents a method for smoothing wind losses when calculating rate indications, but it can apply equally well to other weather events such as hail or freezing. It can be used in other situations such as smoothing out the effects of large losses or other large, but infrequent, events. The model is relatively easy to explain to non-actuaries, and it is not difficult to implement. The traditional approach applies a one-sided cap to losses. This paper presents a two-sided model that bounds losses on both the low and high sides.

1. INTRODUCTION

The current state-of-the-art in pricing for major loss-producing events lies in sophisticated computer models that combine both the damageability of risks and the damage causing potential for events such as hurricanes or earthquakes. Much effort and expense are being directed towards applying these models to insurance ratemaking. But the actuary, like any skilled craftsman, still needs simple, basic tools to handle tasks where more sophisticated methods cannot be readily applied.

2. STABILIZING RATEMAKING LOSS RATIOS

Although premiums and losses over a span of several years provide the basis for calculating a rate change indication, irregular events during that period can produce large swings in the rate indication. One weekend of tornadoes in a state can cause a large increase in a state's property rate indication even though it

TABLE 1

Accident Year	Earned Premium ($000)	Wind Loss Ratio	All Other Loss Ratio	Combined Loss Ratio
1992	$ 714	9.9%	45.0%	54.9%
1993	654	14.0	54.9	68.9
1994	750	3.0	43.4	46.4
1995	870	17.4	49.5	66.9
1996	907	40.0	61.0	101.0
Total	$3,895	17.9%	51.1%	69.0%

may be based on five years of data. Conversely, several years of exceptionally good weather in a state may drive the indication in the opposite direction. Table 1 contains an example of five years of ratemaking data for a line of insurance that includes coverage for wind losses. Partial loss ratios have been computed for wind losses and all other losses. The wind loss ratio ranges from a low of 3.0% in 1994 to a high of 40.0% in 1996.

The starting point for applying this smoothing procedure is to collect the wind loss data in the state for a longer time period. Seventeen years[1] of earned premium and wind loss ratios are displayed in Columns 1 and 2 of Table 2. From Column 2, two percentiles are computed and displayed: a 33rd percentile and a 67th percentile.[2] Normal wind loss ratios as shown in Column 3 are defined to be those loss ratios limited to the range of 5.5% to 14.0%, the 33rd and 67th percentiles. If a wind loss ratio falls below this range, then the 33rd percentile value is substituted. Correspondingly, if a wind loss ratio is above this range then the 67th percentile is used.

[1] Seventeen years of data was available for this line of insurance and state. The procedure does stabilize ratemaking loss ratios using data from this relatively short time period. But, this is *too short* a record to recognize the potential impact of catastrophic weather losses.

[2] Arranging the data from smallest to largest, $n_1, n_2 \ldots, n_m$, then n_1 is the 0th percentile and n_m is the 100th percentile. For data points in between, n_k is the $100 \times (k-1)/(m-1)$ percentile. Other percentiles are computed by interpolation.

TABLE 2

CALCULATION OF NORMAL WIND LOSSES AND ADDITIONAL WIND LOAD

Year	(1) Earned Premium ($000)	(2) Wind Loss Ratio	(3) Normal Wind Loss Ratio	(4) Difference = (2) − (3)	(5) Load = (1) × (4) ($000)	(6) Adjusted Wind Loss Ratio = (3)+ Wind Load
1980	402	0.0%	5.5%	−5.5%	− 22	7.6%
1981	462	9.6	9.6	0.0	0	11.7
1982	560	19.7	14.0	5.7	32	16.1
1983	601	1.4	5.5	−4.1	− 25	7.6
1984	664	13.7	13.7	0.0	0	15.8
1985	691	4.4	5.5	−1.1	− 8	7.6
1986	736	4.0	5.5	−1.5	− 11	7.6
1987	620	13.9	13.9	0.0	0	16.0
1988	669	0.5	5.5	−5.0	− 34	7.6
1989	673	8.4	8.4	0.0	0	10.5
1990	659	21.7	14.0	7.7	51	16.1
1991	710	14.8	14.0	0.8	6	16.1
1992	714	9.9	9.9	0.0	0	12.0
1993	654	14.0	14.0	0.0	0	16.1
1994	750	3.0	5.5	−2.5	− 19	7.6
1995	870	17.4	14.0	3.4	30	16.1
1996	907	40.0	14.0	26.0	236	16.1
Total	$11,342	12.4%			$237	12.4%

Calculation of Normal Range

33rd Percentile	5.5%
67th Percentile	14.0%

Calculation of Wind Load

Load = 237/11,342 = 2.1%

Normal Wind Loss Ratio

1. If "Wind Loss Ratio" < 33rd Percentile, then "Normal Wind Loss Ratio" = 33rd Percentile
2. If "Wind Loss Ratio" > 67th Percentile, then "Normal Wind Loss Ratio" = 67th Percentile
3. Otherwise, "Normal Wind Loss Ratio" = "Wind Loss Ratio"

The average off-balance of this bounding procedure is computed to determine a wind load. Column 4, the difference between Columns 2 and 3, represents how many loss ratio points to add or subtract to bring the wind loss ratio into the "normal" range. Column 5, the product of Columns 1 and 4, is the dollar impact of this bounding procedure. A load that recognizes the off-balance of the bounding is calculated by dividing the sum of Column 5 by the sum of Column 1.[3] An alternative calculation for the load would be to use the average of Column 4 rather than the earned premium weighted average. The last column, the Adjusted Wind Loss Ratio, is the sum of the Normal Wind Loss Ratio and the Wind Load.

Now the results of the calculation can be applied to the ratemaking data in Table 1. There are two steps to adjusting the wind loss ratios: (i) restrict each wind loss ratio to the normal range [5.5, 14.0], the 33rd and 67th percentiles, and (ii) add the balancing wind load factor to each loss ratio. The results are displayed in Table 3.

Compare Columns 1 and 2. Note that three out of five, or 60%, of the Unadjusted Wind Loss Ratios were capped by the bounding procedure. This outcome is consistent with the selection of the 33rd percentile to the 67th percentile as the normal range for the seventeen-year period. On average about 66% of the observed wind loss ratios would fall outside of the normal range. The total Adjusted Wind Loss Ratio (Column 4) is 13.7%, which lies between the unadjusted 17.9% in the five years of ratemaking data and 12.4%, the seventeen-year average wind loss ratio (Table 2, Column 2). The procedure blends both current and long-term experience.

---

[3]This method of calculating a wind load has the effect of allocating the excess wind losses based on earned premium. Sometimes excess losses are allocated using losses, for example by computing the ratio of excess losses to losses and then applying this factor to observed losses, but relating the excess losses to premium usually gives a more stable result.

TABLE 3

Accident Year	(1) Unadjusted Wind Loss Ratio	(2) Normal Wind Loss Ratio	(3) Wind Load	(4) Adjusted Wind Loss Ratio	(5) All Other Loss Ratio	(6) Combined Loss Ratio = (4) + (5)
1992	9.9%	9.9%	2.1%	12.0%	45.0%	57.0%
1993	14.0	14.0	2.1	16.1	54.9	71.0
1994	3.0	5.5	2.1	7.6	43.4	51.0
1995	17.4	14.0	2.1	16.1	49.5	65.6
1996	40.0	14.0	2.1	16.1	61.0	77.1
Total	17.9%	11.6%	2.1%	13.7%	51.1%	64.8%

3. SELECTION OF NORMAL LOSS RATIO RANGE

The selection of a normal loss ratio range from the 33rd percentile to the 67th percentile was based on judgment. Table 4 shows a sample of ranges.

The first row in the table results in no wind smoothing, whereas the last row is equivalent to uniformly spreading the average wind loss ratio to all years. Although the table displays minimum and maximums which are symmetric about the midpoint, a symmetric range is not necessary.

Typically, wind smoothing models only cap "upside" events that push loss ratios above some selected amount. The model presented here also adjusts for years that have very good weather because these years can drive ratemaking indications towards inadequate rates. For this reason, a two-sided capping model can be more effective at smoothing out loss ratios than a one-sided model. In fact, the one-sided model is just a special case of the two-sided where the lower cap is set at the 0th percentile. Because a two-sided model uses two parameters, it offers the opportunity for a better fit to the data than a one-sided, one-parameter model.

TABLE 4

Percentiles		Normal Loss Ratio Range		Wind
Lower	Upper	Lower	Upper	Load
0%	100%	0.0%	40.0%	0.0%
10	90	1.0	20.5	1.6
20	80	3.2	16.9	1.9
30	70	4.3	14.2	2.4
33	**67**	**5.5**	**14.0**	**2.1**
40	60	8.9	13.8	1.0
45	55	9.7	12.9	1.1
50	50	9.9	9.9	2.5

In selecting a percentile range, the actuary is confronted with the eternal tradeoff in ratemaking: stability versus responsiveness. A narrow percentile range will produce more stable loss ratios. But a narrow range may not be responsive to longer term changes in weather patterns, the geographic distribution of insureds, construction techniques, underwriting, or other factors that contribute to the level of risk. A wider range allows the ratemaking model to adjust more quickly for the changing level of risk, but at the cost of more year-to-year variability in the loss ratios.

4. FINDING A GOOD FIT TO THE HISTORICAL DATA

The selection of a normal loss ratio interval does not have to be left entirely to judgment. Quantitative measures can help eliminate weaker choices. To demonstrate this, the one-sided capping model mentioned above will be compared to the [33rd percentile, 67th percentile] bounding procedure. The last column of Table 2 shows the Adjusted Wind Loss Ratios after the bounding and load operations. The difference between the highest and lowest loss ratios in Column 6 of Table 2 is 8.5%. The width of the range can be considered a measure of the stability of the loss ratios resulting from application of the procedure with the selected percentiles.

When using one-sided capping for excess wind loss ratios, the lower bound is set at 0.0%. This forces the upper bound to be set at the 38th percentile in order to generate a range between the highest and lowest Adjusted Wind Loss Ratios as low as 8.5%! Any higher upper bound (with the lower bound fixed at the 0th percentile) will result in a wider range between the highest and lowest Adjusted Wind Loss Ratios, reducing stability. Table 5 compares the selected two-sided model and the one-sided model with the same level of year-to-year stability.

The difference between the high and low values of the Adjusted Wind Loss Ratios equals the width of the range of the Normal Wind Loss Ratios because the Wind Load is added to both of the endpoints of the normal range.

The range for each set of adjusted loss ratios is displayed at the bottom of Table 5. Below these ranges are two measures of how well the Adjusted Wind Loss Ratios fit the raw data in Column 2. The first measure is the sum of the squares of the differences between the adjusted and raw data, and the second measure is the sum of the absolute values of the differences. Under both of these measures the [33rd percentile, 67th percentile] bounding beats the one-sided capping for a given level of stability. For a selected level of stability represented by the range, the two-sided model produces a better fit to the raw data.

The average Wind Loss Ratio over the seventeen-year period is 12.4%. Year-to-year fluctuation in wind losses could be eliminated by taking all of the wind losses out of the ratemaking data and substituting this long-term average. Note that this 12.4% lies closer to the midpoint of the two-sided range [7.6%, 16.1%] than it does to the midpoint of the one-sided range [6.0%, 14.5%]. The two-sided model produces results which are better balanced about the long-term average for the sample data.

Frequently the actuary must rely on judgment to select ratemaking parameters. With this model the actuary does need to rely on judgment to select the desired degree of stability (i.e.,

TABLE 5

Two-Sided Versus One-Sided Capping

			Two-Sided Model			One-Sided Model		
Year	(1) Earned Premium ($000)	(2) Wind Loss Ratio		(3) Normal Wind Loss Ratio	(4) Adjusted Wind Loss Ratio = (3) + Load		(5) Normal Wind Loss Ratio	(6) Adjusted Wind Loss Ratio = (5) + Load
1980	402	0.0%		5.5%	7.6%		0.0%	6.0%
1981	462	9.6		9.6	11.7		8.5	14.5
1982	560	19.7		14.0	16.1		8.5	14.5
1983	601	1.4		5.5	7.6		1.4	7.4
1984	664	13.7		13.7	15.8		8.5	14.5
1985	691	4.4		5.5	7.6		4.4	10.4
1986	736	4.0		5.5	7.6		4.0	10.0
1987	620	13.9		13.9	16.0		8.5	14.5
1988	669	0.5		5.5	7.6		0.5	6.5
1989	673	8.4		8.4	10.5		8.4	14.4
1990	659	21.7		14.0	16.1		8.5	14.5
1991	710	14.8		14.0	16.1		8.5	14.5
1992	714	9.9		9.9	12.0		8.5	14.5
1993	654	14.0		14.0	16.1		8.5	14.5
1994	750	3.0		5.5	7.6		3.0	9.0
1995	870	17.4		14.0	16.1		8.5	14.5
1996	907	40.0		14.0	16.1		8.5	14.5
			Normal Range:			Normal Range:		
			33rd Percentile	5.5%		0th Percentile	0.0%	
			67th Percentile	14.0%		38th Percentile	8.5%	
			Wind Load =	2.1%		Wind Load =	6.0%	
			Range = (Maximum-Minimum)		8.5%	Range = (Maximum-Minimum)		8.5%
			Sum of squares of (4) – (2) =		839.0	Sum of squares of (6) – (2) =		1035.5
			Sum of absolute value of (4) – (2) =		80.7	Sum of absolute value of (6) – (2) =		94.3

the range in the Adjusted Wind Loss Ratios), but can also quantitatively search for minimum and maximum percentiles that fit the historical data well for the chosen level of stability. (It is possible that the sum-of-squares and the sum-of-absolute-values measures may give conflicting signals. The actuary will have to decide which measure is more meaningful for the situation.)

Table 6 shows various ranges for the Adjusted Wind Loss Ratios and corresponding percentiles that produce good fits to the data. The first column shows the stability constraint, how much variability is allowed in the Adjusted Wind Loss Ratios. Then the next columns display percentiles which satisfy the stability constraint and have low values for the sum-of-squares errors of the fit.[4] The last column shows the five-year Total Adjusted Wind Loss Ratio after application of the procedure to the ratemaking data in Table 1.

In the first row of Table 6 the Adjusted Wind Loss Ratio is a constant 12.4% = 9.9% + 2.5%; the long-term average wind loss ratio would be substituted for the actual wind loss ratios in the ratemaking experience period. In the last row, there is no smoothing on the data which itself has a 40 point range. Of course, as the stability constraint is loosened, the fit to the data improves.

5. CONCLUSION

The two-sided capping model presented here achieves the same end as the traditional "upside" capping model: the stabilization of loss ratios used in ratemaking. But, for the same degree of stabilization, two advantages of the two-sided model with the sample data were noted: (1) it fits the historical data

[4]The percentiles were computed by solving for the two percentiles that satisfied the stability constraint in the first column and that minimized the sum-of-squares error using the "Solver" routine in a Microsoft Excel spreadsheet. The iteration stopping point of the routine depended on the initial values. It was necessary to try a number of initial starting points and compare sum-of-squares errors for the resulting iteration stopping points and then pick the one with the lowest error.

TABLE 6

Wind Loss Ratio Range	"Good Fit" Normal Loss Ratio Bounds		Wind Load	Sum-of-Squares Error	1992–1996 Total Adjusted Wind Loss Ratio
	Lower	Upper			
0 points	50th percentile = 9.9%	50th percentile = 9.9%	2.5%	1608.3	12.4%
5 points	40th percentile = 8.9	64th percentile = 13.9	0.9	1078.5	13.1
10 points	37th percentile = 8.1	83rd percentile = 18.1	0.1	724.0	13.9
15 points	33rd percentile = 5.5	90th percentile = 20.5	0.6	469.1	14.5
20 points	11th percentile = 1.2	92nd percentile = 21.2	1.5	342.8	15.1
30 points	11th percentile = 1.2	97th percentile = 31.2	0.6	77.0	16.5
40 points	0th percentile = 0.0	100th percentile = 40.0	0.0	0.0	17.9

better, and (2) the range of resulting loss ratios is more evenly balanced around the long-term average loss ratio. Also, the two-sided model tempers the impact on the rate indication of unusually good weather during the ratemaking period.

This model offers the actuary considerable flexibility in stabilizing the effects of volatile losses on ratemaking. Choices range from a high degree of stabilization by choosing a [50th percentile, 50th percentile] range to complete responsiveness with a [0th percentile, 100th percentile] range, or anything in between.

Since percentiles involve ranking and counting, the concept is easier to explain to non-actuaries than a less intuitive concept such as standard deviation. Standard deviation has been used in some actuarial models to define the acceptable range of variability in weather losses, but when dealing with highly skewed distributions, a percentile is more meaningful and easier to understand.

REFERENCES

[1] Bradshaw, John, and Mark Homan, "Homeowners Excess Wind Loads: Augmenting the ISO Wind Procedure," Casualty Actuarial Society *Forum*, Summer 1993, pp. 339–350.

[2] Burger, George, Beth E. Fitzgerald, Jonathan White, and Patrick B. Woods, "Incorporating a Hurricane Model into Property Ratemaking," Casualty Actuarial Society *Forum*, Winter 1996, pp. 129–190.

[3] Clark, Karen M., "A Formal Approach to Catastrophe Risk Assessment and Management," *PCAS* LXXIII, 1986, pp. 69–92.

[4] Gogol, Daniel F., "An Actuarial Approach to Property Catastrophe Cover Rating," *PCAS* LXXXI, 1994, pp. 1–35.

[5] Hays, David H., and W. Scott Farris, "Pricing the Catastrophe Exposure in Property Insurance Ratemaking," Casualty Actuarial Society *Forum*, 1994, pp. 486–528.

[6] Hogg, Robert V., and Stuart A. Klugman, *Loss Distributions*, John Wiley & Sons, 1984.

[7] Homan, Mark J., "Reflecting Reinsurance Costs in Rate Indications for Homeowners Insurance," Casualty Actuarial Society *Forum*, Winter 1997, pp. 223–254.

[8] Wilkinson, Margaret E., "Estimating Probable Maximum Loss with Order Statistics," *PCAS* LXIX, 1982, pp. 195–209.

INVESTMENT-EQUIVALENT REINSURANCE PRICING

RODNEY E. KREPS

Abstract

Reinsurance pricing is usually described as market-driven. In order to have a more theoretical (and practical) basis for pricing, some description of the economic origin of reinsurance risk load should be given. A special-case algorithm is presented here that allows any investment criteria concerning return and risk to be applied to a combination of reinsurance contract terms and financial techniques. The inputs are the investment criteria, the loss distributions, and a criterion describing a reinsurer's underwriting conservatism. The outputs are the risk load and the time-zero assets allocated to the contract when it is priced as a stand-alone deal. Since most reinsurers already have a book of business and hence contracts mutually support each other, the risk load here can be regarded as a reasonable maximum. The algorithm predicts the existence of minimum premiums for rare event contracts, and generally suggests a reduction in risk load for pooling across contracts and/or years. Three major applications are: (1) pricing individual contracts, (2) packaging a reinsurance contract with financial techniques to create an investment vehicle, and (3) providing a tool for whole book management using risk and return to relate investment capital, underwriting, and pricing.

ACKNOWLEDGEMENT

Many thanks are due to Gary Venter, the reviewer, and Mike Steel for valuable input and discussions.

1. INTRODUCTION

There has been an evolution over the last few years toward looking at an insurance or reinsurance enterprise as a whole, rather than seeing underwriting, investments, dividend policy, and so forth as a set of disjoint pieces. Whereas in modern financial theory various approaches to the interaction of risk and reward are reasonably well developed, for insurance in general and reinsurance in particular the measurement of risk has been (and arguably still is) more of an art than a science. It is generally agreed that surplus creates capacity and writing business uses up surplus, but there is no agreement on how that happens.

This paper proposes a way of obtaining models for the special case where the contract is priced on a stand-alone basis; i.e., it is the reinsurer's only business. The risk loads (and hence pricing) derived here are maximal because reinsurers generally have an ongoing book of business. This book is mutually supporting in that usually it does not all go bad at the same time. Pricing on a stand-alone basis is equivalent to assuming that the whole book is fully correlated. Thus stand-alone pricing in general will result in larger risk loads than are actually needed.

Although the give and take of the market will ultimately determine what prices are actually charged for contracts, both insurers and reinsurers can use an economic pricing model in order to help decide whether to write a contract. For the insurer, the decision not to reinsure externally is a decision to self-reinsure. The intent of this paper is to present a paradigm that will allow the combination of a reinsurance arrangement and suitable financial techniques to be thought of as an investment alternative. This allows a firm's investment criteria to be applied to the decision.

What is actually done is to assume investment criteria in the form of a target mean return and risk measure thereon. From the paradigm, the necessary risk load and notionally allocated assets for the reinsurance arrangement are obtained.

The paradigm is as follows: When the reinsurer accepts a contract, it arranges to have available at every time of loss sufficient liquid assets to cover possible losses up to some safety level. These assets arise from premium and from surplus, both of which are invested in appropriate financial instruments. The reinsurer wishes to have at least as favorable return and risk over the period of the contract as it would target in other investments of the allocated assets.

Note that this is not, at least to the author's knowledge, how reinsurers currently do their pricing, nor is it advocated (except in special circumstances) as an operating procedure for reinsurers. It is a way of deriving risk loads by relating them to investment criteria. At the same time, it makes intuitive sense. Certainly reinsurers had better plan to have assets available to pay losses; otherwise they are planning for bankruptcy. This paradigm essentially looks at risk load as an opportunity cost and represents it as a (partially offset) cost of liquidity. This is not the only way to look at risk loads, but it is a simple and intuitive one.

The loss safety level is essentially a measure of reinsurer company conservatism. It is intuitive that some measure of company conservatism would be present in a risk load paradigm.[1] The more conservative the company, the higher the safety level and the less probable it is that the safety level will be exceeded. Higher safety levels will typically result in more expensive contracts.

A mundane example of a safety level occurs when a person decides to build a house in snow country. The question is, how strong to build the roof for snow load? If it is a cabin intended

---

[1]A financial economics point of view suggests that there is a market equilibrium price, and that it reflects a risk load independent of specific company attitudes. However, without wanting to get into a complex discussion, the author feels that many of the assumptions of an efficient market are not particularly well satisfied in the reinsurance arena. A reinsurer with a large portfolio of Florida Homeowners' policies may very well not take any more at all, much less at the price that some other reinsurer is willing to accept. What constitutes "large" depends upon the reinsurer management.

for use for only a few years, perhaps building to survive the ten year storm will be enough. If it is meant for the grandchildren, perhaps surviving the two hundred year storm is more appropriate. It is, of course, more expensive to build it stronger. In any case, some level is chosen depending on the builder's criteria.

The safety level used in the examples here will be the amount of loss associated with a previously chosen probability, such as the 99.9% level; i.e., the loss associated with a one thousand year return time. In some circumstances (see Section 2.3), the full amount of the contract may be the appropriate safety level. There are, of course, alternatives to a probability level. One would be to choose a loss safety level high enough so that the average value of the excess loss over that level is an acceptably small fraction of the mean loss. Another is when the average excess loss over the safety level times the probability of hitting the safety level is below some value. While it would be interesting to examine various choices in the context of different management styles, the essential point is that any quantifiable measure can be used.

Clearly, a risk load paradigm must involve the cost of capital and, more specifically, measures of investment return and risk for comparison to the capital markets. A *reductio ad absurdum* shows the argument: If capital were free and freely available, insurance, much less reinsurance, would be unnecessary. A firm in temporary trouble would simply borrow through difficulties.

The measure of investment risk used here will be the standard deviation (or variance) of the rate of return. Equally applicable would be one of the more sophisticated, strictly downside measures such as the semi-variance or the average value of the (negative) excess of return below some trigger point such as the risk-free rate. Especially where very large losses may generate negative results, such a downside risk measure may be desirable. These measures do not give pretty formulae but are easily used numerically. Again, any quantifiable measure is feasible.

There are two types of financial techniques that will be considered. Other techniques are possible; these are just two of the simplest. The first is where the reinsurer takes the capital that it would have put into the target investment (which could be, for example, corporate bonds) and puts it into a risk-free instrument such as government securities. This will be referred to as a ***switch***.[2] The cost associated with this is the loss of investment income, but there is also a gain in that investment default risk is reduced.

This technique results in simple formulae;[3] but it often results in a higher risk load and, hence, is more expensive (to the cedent) and therefore less competitive than the second technique: buying "put" options. These options are the right to sell the underlying target investment at a predetermined strike price at maturity. We only consider European options which can be exercised only at maturity. Here the strike price will be what an investment in risk-free securities would have brought, so that the reinsurer is buying the right to sell the target investment at a return not less than the risk-free rate.

The Black–Scholes[4] formula is used to price the option. The distribution of price for the reinsurer's target investment is assumed lognormal, so that this formula applies. The cost of these options will contribute to the risk load, but this is partly offset because the options both increase the return and decrease the variance of the target investment.

This treatment will not include the effects of reinsurer expenses, nor of taxes. However these could be included, especially in the simulation models in the latter part of the paper. For taxes, one would have to make some assumptions as to whether the

---

[2] To the author's knowledge, this is not a technical financial term. If it is, apologies are offered. The meaning intended is only what is stated here.

[3] For the variance measure of investment risk. As remarked earlier, other measures in general will not give simple formulae.

[4] See the discussion of Black–Scholes in the standard Part 5 reading [1, page 502 ff.]

contract would affect any possible Alternative Minimum Tax situation. Probably this could best be treated by looking at the reinsurer's whole underwriting book and investment structure with and without the contract of interest. This is beyond the scope of this paper.

In Section 2, the paper addresses the case of a single loss payment at the end of one year. In Section 2A, the switch is treated and in Section 2B, the option. These simple discussions will illustrate the general principles so that they will hopefully not be obfuscated by the details of the subsequent development. For readability, technical details are relegated to appendices. In Section 2C the limiting case of a high excess layer is presented, where it is shown that a minimum premium results. This is in accord with actual market behavior. In Section 3, the single payment case is extended to an arbitrary known time of loss. Section 3A is a numerical example, and Section 3B includes some general remarks on pooling and other subjects.

The multiple payment case is discussed in Section 4. In this case, there are no simple formulae available, and simulation modeling must be explicitly used. Section 5 contains concluding remarks on the whole paradigm.

2. SINGLE PAYMENT AT ONE YEAR

Let the principal determinants be:

S = the dollar safety level associated with the loss distribution;

L = the mean value of the loss;

σ_L = the standard deviation of the loss;

r_f = the risk-free rate;

y = the yield rate of the target investment;

σ_y = the standard deviation of the investment yield rate; and

P = the premium net to the reinsurer after expenses.

Quantities derived from the above are:

A = the assets allocated by the reinsurer;

F = the funds initially invested = premium and assets less option cost, if applicable; and

R = the risk load.

The premium in all cases is the risk load plus the expected loss discounted at the risk-free rate. Note that this premium does not include any reinsurer expenses. For a single payment at one year,

$$P = R + \frac{L}{1 + r_f}. \tag{2.1}$$

The constraints of the paradigm may now be stated as: (1) the investment result from F as input must be at least S, and (2) the standard deviation of the overall result must be no larger than σ_y.

Although the fundamental cash flow relations are stochastic, it is possible in this Section to get explicit formulae for the mean and variances involved, and hence get explicit forms for the risk load. In Section 4, the mean is easily obtained, but the variance of the final outcome of the stochastic cash flows has to be determined by simulation.

A. Switch Case

At time zero, the reinsurer has an inflow of P and an outflow of

$$F = (P + A). \tag{2.2}$$

Since the investment is in risk-free securities, at the end of the year the reinsurer has an inflow of $(1 + r_f)F$ and an outflow of the loss. The internal rate of return (IRR) on these cash flows is defined by the fundamental stochastic relation

$$(1 + \text{IRR})A = (1 + r_f)F - \text{loss}, \tag{2.3}$$

where both the loss and the IRR are stochastic variables. Taking the mean value of this equation and asking that the mean value of the IRR be the yield rate y gives

$$(1+y)A = (1+r_f)F - L, \tag{2.4}$$

which may be expressed as[5]

$$R = \frac{(y-r_f)}{(1+r_f)}A. \tag{2.5}$$

Another equation is needed to solve the system, and there are two constraints that must be satisfied—a loss safety constraint and an investment variance constraint. In general, it is clear that by making the asset base large enough, the fractional variability of results can be made as small as desired and the funds available as large as desired. Hence there is always a solution. Both constraints may be phrased as placing lower limits on the allocated assets, so satisfying the more restrictive will satisfy both.

For the safety constraint, requiring the funds available at the year end to be greater than or equal to the safety level gives

$$(1+r_f)F \geq S. \tag{2.6}$$

Combining Equations 2.4 and 2.6 to eliminate F,

$$A \geq \frac{(S-L)}{1+y}, \tag{2.7}$$

and consequently, from Equations 2.5 and 2.7, the risk load at the equality is

$$R = \frac{(y-r_f)(S-L)}{(1+r_f)(1+y)}. \tag{2.8}$$

Having the risk load, Equation 2.1 gives the premium before expenses.

This is the result for the safety constraint. For the variance constraint, since there is no variability in the investment return

[5]For readability, derivations of more than one line are done in Appendix C.

(because it is risk-free), the standard deviation of the IRR is given by Equation 2.3 as

$$A\sigma_{\text{IRR}} = \sigma_L. \tag{2.9}$$

The investment constraint is that the IRR should have a variance less than or equal to that of the target investment, which gives

$$A \geq \sigma_L/\sigma_y \tag{2.10}$$

and, using Equation 2.5 again,

$$R = \frac{(y - r_f)}{(1 + r_f)}(\sigma_L/\sigma_y). \tag{2.11}$$

Given typical values for the loss distribution and the target investment, the latter is likely to be the more stringent constraint. This will be true when

$$(S - L)/\sigma_L < (1 + y)/\sigma_y. \tag{2.12}$$

For a one-in-a-thousand safety level and a normal distribution, the number on the left is around 3. For more positively skewed distributions, it will be larger; but, in the experience of the author, it is seldom as large as five for typical reinsurance layers. However, in the example used later of an unlimited cover with a lognormal distribution with coefficient of variation two, the ratio on the left is over ten. The unlimited cover is a mathematical convenience for illustration rather than a realistic contract, at least since pollution losses began to develop extremely adversely. Plausible values for the ratio on the right are easily around twelve for bonds and higher than five for equities. For example, for a bond with an 8% yield and an 8% standard deviation, the ratio is $1.08/0.08 = 13.5$; for a stock with a 12% yield and a 20% standard deviation, the ratio is $1.12/0.20 = 5.6$.

B. Option Case

At time zero, the reinsurer will receive the premium but keep the initial assets invested in the target investment. It will also buy an option to sell the target investment at the end of the year for the value that a risk-free investment would have reached. By doing so it has an instrument that eliminates that portion of the investment return distribution that lies below the risk-free rate. This will have the effect both of increasing the mean return from the investment and decreasing its standard deviation.

The rate of the put option (cost per dollar of investment protected) is here denoted c, and depends upon the underlying investment parameter σ, which is determined by y and σ_y and defined in Appendix A. For small values of the ratio of σ_y to $(1 + y)$, it is approximately true[6] that

$$\sigma = \sigma_y/(1 + y), \qquad \text{and} \tag{2.13}$$

$$c = \frac{1}{\sqrt{2\pi}}\sigma(1 - \sigma^2/24). \tag{2.14}$$

However, the examples below use the exact formula from Appendix A. At time zero, the reinsurer has an inflow of P and an outflow of $(P + A)$. The funds available for investment have decreased by the cost of the option. Specifically, Equation 2.2 becomes

$$F = P + A - cF, \tag{2.15}$$

$$\text{so} \qquad F = (P + A)/(1 + c). \tag{2.16}$$

Since the investment is now in risky securities (hedged at the bottom end to stay above or equal to the risk-free rate), at the end of the year the reinsurer has an inflow of $(1 + \text{invest})F$ and an outflow of the loss. The internal rate of return on these cash flows is defined by a fundamental stochastic relation similar to Equation 2.3:

$$(1 + \text{IRR})A = (1 + \text{invest})F - \text{loss}. \tag{2.17}$$

Again, requiring that the mean value of IRR be the target yield rate gives

$$(1 + y)A = (1 + i)F - L, \tag{2.18}$$

[6]See Appendix A.

where i is the mean investment return (determined in Appendix B). This does not simplify easily, but fundamentally we have two unknowns, R and A, and this is one equation relating them. The other equation will come from whichever is the more restrictive constraint, as before.

The loss safety constraint on the funds available is again Equation 2.6:

$$(1 + r_f)F \geq S. \tag{2.6}$$

It should be noted that the actual funds available are likely to be larger than this, since r_f represents the *minimum* value of the realizable investment return, thanks to the option. Combining Equations 2.6 and 2.18 to eliminate F, the allocated assets are

$$A \geq \frac{1}{1+y}\left\{\frac{1+i}{1+r_f}S - L\right\}. \tag{2.19}$$

This is larger than in the switch case since $i > y > r_f$. The expression for the risk load at equality is[7]

$$R = \frac{1}{(1+r_f)(1+y)}\{S[(1+y)(1+c) - (1+i)] - L[y - r_f]\}. \tag{2.20}$$

For $i = r_f$ and $r = 0$, the results of the previous section are, of course, obtained in the above two formulae.

In order to express the investment variance constraint, it is necessary to decide the correlation between the loss and the investment return. The linkage by inflation suggests that there may be a negative correlation. If inflation rises, typically claims costs rise and bond values fall. In the interest of simplicity the assumption will be made that the correlation is zero, although there is no essential complication introduced by taking a non-zero value. The standard deviation of the investment return is derived in Appendix B and written as σ_i. When the variance of the IRR is

[7] See Appendix C.

required to be that of the target investment, Equation 2.17 (with zero correlation) implies that

$$(A\sigma_y)^2 = (F\sigma_i)^2 + (\sigma_L)^2. \tag{2.21}$$

The value of the initial fund F from Equation 2.18 may be substituted into this, resulting[8] in a quadratic equation for A of the form

$$-aA^2 + 2bA + c = 0, \tag{2.22}$$

with

$$a = (\sigma_y)^2(1+i)^2 - (\sigma_i)^2(1+y)^2,$$

$$b = L(1+y)(\sigma_i)^2, \qquad \text{and}$$

$$c = L^2(\sigma_i)^2 + (\sigma_L)^2(1+i)^2.$$

All three coefficients are positive; the last two because of their explicit construction, and the first because the option both decreases the variance and increases the mean of the investment return compared to the target values.

The positive solution is

$$A = \frac{b + \sqrt{b^2 + ac}}{a} \tag{2.23}$$

and[9]

$$R = A\frac{(1+c)(1+y) - (1+i)}{1+i} + L\left[\frac{1+c}{1+i} - \frac{1}{1+r_f}\right]. \tag{2.24}$$

As $\sigma_i \to 0$, the solution for A goes back to the ratio of standard deviations. With $i = r_f$ and $c = 0$, the risk load returns to the earlier form found in the switch case, as it should.

---

[8]See Appendix C. The forms corresponding to a non-zero correlation are also given there.

[9]See Appendix C.

C. High Excess Layer and Minimum Premium

An interesting application of these formulae is in the case of a high excess layer or any similar finite rare event cover. A non-zero rate on line (ratio of premium to limit) is predicted even for cases where the loss probability goes to zero.

For simplicity, take the loss distribution to be binomial: there is a probability p of hitting the layer, and if it does get hit, it is a total loss. Note that the 99.9% level is not an appropriate way to get the safety level (especially for $p < 0.001$). There is still in fact an intuitive value: the safety level S is taken to be the limit (total amount payable) of the layer.

The mean loss L is pS and the variance of the loss is $p(1-p)S^2$. As the probability p gets smaller, corresponding to higher and higher layers, in both the switch and option cases the variance constraint forces A and R both to zero as $\sqrt{p}$. However, the safety constraint in both cases is linear in p with a non-zero intercept. In the option case, the rate on line (ROL) in the limit as p goes to zero is

$$\mathrm{ROL} = \frac{(1+y)(1+c)-(1+i)}{(1+r_f)(1+y)}. \tag{2.25}$$

This is obtained by setting $L = 0$ in Equation 2.20 and recognizing ROL as the ratio of R to S. As usual, the switch version may be obtained by letting $c = 0$ and $i = r_f$, which results in

$$\mathrm{ROL} = \frac{1}{1+r_f} - \frac{1}{1+y} = \frac{y-r_f}{(1+r_f)(1+y)}. \tag{2.26}$$

The latter form suggests that the minimum ROL is of the order of the real target return; i.e., the excess of the return over the risk-free rate. However, often the option form Equation 2.25 will produce a smaller number. For the investment values used below it is typically half as large. As the investment standard deviation gets small, the switch ROL stays the same (of course) and the option ROL gets small because the option cost gets small and the mean investment return approaches the target yield. It is important to remember that Equations 2.25 and 2.26 and this discussion are all at the limit where $p = 0$. For this value, the standard deviation of the loss distribution is zero, which implies that the variance constraint is always satisfied. However, for a small but fixed probability, say in the range from 2% to 0.1%, which is typical of catastrophe contracts, as the target standard deviation of investment is made small, the variance constraint will eventually become dominant.

In the market, a minimum ROL is generally justified by underwriters as a charge for using surplus. This approach is consistent with that view and also allows for quantification of the charge.

3. SINGLE PAYMENT AT VARIABLE TIME

If all the returns in the preceding are interpreted as total return up to time t, then the formulae hold without modification. When we wish to express the returns in terms of the equivalent annualized returns, the results hold after the following replacements are made:

$$(1+i) \rightarrow (1+i)^t,$$

$$(1+y) \rightarrow (1+y)^t, \qquad \text{and}$$

$$(1+r_f) \rightarrow (1+r_f)^t.$$

The forms for the option rate and the standard deviations given in Appendix B contain the time dependence.

A. Numerical Example

For any one-payment situation, the recommended procedure is as follows: (1) calculate the four risk loads and allocated assets under the safety and variance constraints for the option and switch cases; (2) find for each financial technique the constraint with the larger allocated assets—this is the dominant one;

(3) compare the dominant risk loads for different techniques and choose the smaller—this is the preferred[10] solution.

This calculation is easily put on a spreadsheet. For a specific example, the following annualized values are used:[11] yield rate $y = 5.3\%$; standard deviation of the yield rate $\sigma_y = 8.4\%$; risk-free rate $r_f = 3.6\%$. The loss distribution is assumed to be lognormal with mean of \$1M (million) and standard deviation of \$2M. The loss safety level is taken as the 99.9% level, \$22,548,702. For a one-year interval this makes the left-hand side of Equation 2.12 equal to 10.8, while the right-hand side is 12.5, suggesting that variance will be the dominant constraint for the switch. For a two-year interval, the right-hand side changes to 8.9 and safety is dominant in the switch. The large value of the left-hand side is due to the fact that this is an unlimited contract.

As an example of the recommended procedure, the following results can be derived from the formulae in the preceding sections for a time of two years.

	SWITCH		OPTION	
constraint	variance	safety	variance	safety
assets	\$15,963,111	\$19,434,097	\$23,024,033	\$20,737,421
risk load	\$528,184	\$643,031	\$316,332	\$283,248

The results are incorporated in Table 1.

For the switch, the safety constraint is dominant; for the option the variance constraint is dominant. Of the two, the option risk load is smaller, and hence preferred.

[10]Preferred from the point of view of the cedent, and preferred from the point of view of offering competitive advantage to the reinsurer—less charge for the same return and risk. On the other hand, the reinsurer may prefer to charge more if the market will bear it. Of course, a higher market rate can always be recast as a more profitable target investment return.

[11]These values are long-term rates from [1, Table 7-1, page 131] except for the standard deviation, which is a private estimate. The return rates are clearly too small to represent current (January 1997) conditions where returns are high and deviations apparently small. Anyone using the pricing technique will use current values appropriate to their own targets.

TABLE 1

VALUES FOR THE OPTION TECHNIQUE ON A SINGLE PAYMENT

Time (years)	1	2	3	4
Option rate	3.18%	4.49%	5.50%	6.35%
Risk load	$235,225	$316,332	$399,548	$502,444
Risk-loaded premium	1,200,476	1,248,042	1,298,882	1,370,526
Total premium	1,379,857	1,434,531	1,492,967	1,575,317
Allocated assets	32,522,839	23,024,033	20,095,065	19,446,192
Initial investment	32,685,050	23,228,830	20,278,801	19,574,132
Determining constraint	variance	variance	safety	safety
Safety value	3,087 years	1,309 years	1,000 years	1,000 years
Annualized Std/target std	100%	100%	97%	93%

In order to get, for example, the second column of Table 1, time is taken as two years. Following the formulae and notation of the appendices for the investment, $2\mu = 9.69\%$ and $\sigma\sqrt{2} = 11.26\%$ at two years. The target investment mean and standard deviation are 10.88% and 12.53% as calculated from the lognormal formulae. The option rate is 4.49%. The mean and standard deviation of the option-protected investment are 14.21% and 8.95%, respectively higher and lower than the target, as previously advertised. The investment minimum value is 7.33%, the risk-free cumulative return.

The calculated risk loads and asset values are given above for both the option and the switch, and the option-variance technique-constraint combination is chosen.

Please note again that any form of loss distribution could have been used, including underwriter's intuition or simulation result. All that is needed for this choice of risk load is the mean, standard deviation, and safety level. Reinsurer expenses, needed to calculate total premium from risk loaded premium, are taken as 13% of the total.

The table also lists the safety level implied by the chosen asset allocation, and the ratio of the standard deviations of the annualized yield to the target standard deviation. Whichever is not the determining constraint is, of course, more than satisfied. It is noteworthy that as the contract period becomes longer, the safety constraint becomes the more restrictive. In numerical explorations this often seems to be true.

B. Pooling and Other Remarks

It is an intuitive expectation that the total risk load may be reduced by pooling. Pooling over contracts will be considered here; pooling over years will be considered later. The one-year contract from Table 1 has a risk load of $235,225. If two contracts are combined into a single contract, then the safety level on the combined contract is generally less than the sum of the individual safety levels, unless the contracts are fully correlated.[12] Specifically, taking the approximation that the sum of two uncorrelated lognormals may, for these purposes, be represented by a lognormal, the safety level for the combined contract is $29,455,245, which is only 65.3% of the sum.[13]

The risk load for the combined contract over one year is $331,156, which is 70.4% of the sum of the individual risk loads. This risk load results from the option-variance technique and constraint. However, one may question whether some other investment risk measure might have given a different result. The author knows of no general theorem, but experimentation has given consistent reduction in risk load from pooling.

More intuitively, both the safety levels and investment risk measures will be primarily sensitive to the tail of the loss distribution. When two contracts are imperfectly correlated, the bulk

---

[12]Or effectively taken as such, as in the high excess example.

[13]The sum is represented by a lognormal with mean and variance equal to twice the mean and variance of the individual contract. Since the individual contract has mean $1,000,000 and standard deviation $2,000,000, the sum has mean is $2,000,000 and standard deviation is $2,828,427, which implies the 99.9% level mentioned. The 99.9% level on the individual contract is $22,548,701, so twice that level (which would correspond to perfect correlation) is $45,097,403. The ratio is 65.3%.

of the tail results from one or the other of the contracts going bad, not both. The effect generally is to shorten the tail relative to the mean, making measures that depend on extreme values take on less dangerous significance.

A glance at the values in Table 1 shows that it is possible that if the loss is very bad, say at the 0.001% level, then the ending value will be negative. That is, the reinsurer will lose all the premium and allocated assets, and still have to put in more money to fulfill the contract. At the very least, this result cannot be from a lognormal distribution, which never becomes negative.

Nevertheless, it is convenient to express the mean and standard deviation of the distribution of the ending values in terms of the annualized parameters of a geometric Brownian motion investment with the same mean and standard deviation at the time horizon. This allows a direct comparison with the original investment possibility. It is in this sense that the combination of reinsurance contract and switch/option can be thought of as an equivalent investment, if the return and standard deviation are the same as the target.

To the extent that the investment risk measure is valid for general distributions, a comparison can always be made.

Should a reinsurer actually follow through on the indicated financial technique for each contract? Almost surely not, unless the reinsurer is very conservative or this is the only contract. The latter could be the case for a specialty reinsurer set up for a single contract—for example for a large catastrophe contract. In general, a method relating investment criteria to reinsurance contracts could be useful when specifically engineered deals are made to connect reinsureds and investors looking for new opportunities. Considering the hunger of capital for uncorrelated risks, this kind of bundling would seem natural.

This procedure takes as input the financial targets and safety criterion and produces as output the risk load and the allocated assets. It is also possible to take the financial targets and allocated assets as input (more the financial point of view). The two constraints then become requirements on the loss distribution. The corresponding risk loads will emerge. Knowing the desired loss characteristics and the necessary risk loads, market knowledge can be used to do selective underwriting and keep the overall distribution within acceptable risk levels at the target rate of return. This point of view is really more applicable to the book as a whole and requires a treatment of multiple payments.

4. MULTIPLE PAYMENTS

When there are multiple loss payments possible, the same basic paradigm is used but needs a more complex formulation. In the single payment case, simultaneously enforcing the safety constraint and the rate of return through the mean value of the stochastic equation gave an easy solution for the risk load. The risk load appropriate to a particular safety constraint is almost as easy to find in the multiple payment case. However in contrast to the single payment case, there is no simple formula from the variance constraint, but the constraint can be evaluated by simulation for any given level of allocated assets.

The main complication lies in the construction of the safety constraint: in the definitions appropriate to safety levels at different times and in different circumstances. For example, if the first year has a very large loss, do the safety levels for subsequent years change? There are many different formulations possible, all of which will lead to risk loads. Competitive efficiency would suggest looking for a formulation with the smallest possible risk load. This usually will involve using the least possible capital for the shortest period of time, and may depend upon the specifics of the payout pattern expected.

The suggested general procedure is:

1. Express the fundamental stochastic process on a spreadsheet. It is now more complex than a simple equation be-

cause of the interaction of the fund, loss, and investment levels at different times but it is still easily expressed. The complications come from the fact that there are separate simulations of the loss and investment variables at each time. Further, there may have to be other cash flows in either direction between the reinsurer's general assets and the fund set up for this contract, depending upon how the safety levels are defined.

2. Define the safety levels. Since the whole point of this exercise is to use notionally allocated funds to obtain the opportunity cost of capital, the definitions need to be fixed at time zero so the pricing can be done. The simplest version would be to define a single safety level, say 99.9% on the distribution of ending cumulative loss values or the largest 99.9% loss level encountered at any time. The problem with this formulation is that much of the time there will be unnecessary liquidity available which will add to the risk load. A more sophisticated version would be to define levels dependent on the loss distributions at each time. The definitions need not necessarily even result in fixed amounts; the amounts could be conditional on how the losses manifest over time during each particular simulation. A general rule of thumb suggested for safety level definitions is that, if the losses are almost entirely at one time, then the outcome of whatever definitions are used should closely approximate the single payment case for that time.

3. Use the definitions of the safety levels to determine what funds need to be available at various points in time and what options need to be bought (notionally) to protect the values of those funds. Funds not totally consumed at a given time can be carried forward and should be option protected for the expected carryforward. For example, if it is decided to use the 99.9% level at year one as a safety level, in almost all simulations the loss in year one

will be considerably smaller than this safety level. If so, the net can be carried forward to form part of, all of, or more than the safety level for year two. In the latter case it may be that funds flow back to the reinsurer's general account. The option cost on the carryforward will depend on the actual amount of funds carried and the time period when they exist. The time zero present value of the projected average cost is probably not a bad prescription for the initial funds necessary for these options. The switch case is an easier problem because of the lack of option costs, but it does not give the advantage of reducing the standard deviation of the investment and increasing its average return. Hence it will usually give larger risk loads.

4. Find the risk load corresponding to the target return for the safety constraint chosen. This can be done by using a trial risk load and running simulations to ascertain the value of the average final cash result of all the transactions. If this value does not correspond to the desired average return, then try another risk load until the desired target is attained. A faster, simpler, and usually almost as accurate procedure begins by putting all the stochastic variables at their mean values. Then the value of the final cash result of all the transactions is deterministic and can be adjusted to the desired value by varying the risk load.[14] This latter procedure can also be used as a starting point for the former.

5. Simulate to see if the variance constraint is satisfied. If it is and the return is acceptable, stop. It is convenient to represent the variability of the final cash result of all the transactions in terms of the annualized standard deviation of a lognormal investment with the same return and variance at the horizon.

[14]Using Goal Seek in Excel, for example.

6. If the variance constraint is not satisfied, then add more initial capital and simulate again. Clearly, the addition of enough capital beyond that required by the safety constraint can reduce the variability to any desired point. If during the course of iterations the variance constraint is more than satisfied, then take away some capital.

7. Repeat step 6 until both constraints are satisfied. The whole process can be treated numerically like a root-finding procedure, but it is necessary to be careful of the simulation uncertainty in the mean and standard deviation in creating the estimates of the next value of initial capital to be tried.

If it is decided to work with cumulative loss values in establishing definitions of safety levels, the time value of money for the loss in year one must be accounted for with an appropriate rate in order that that loss be economically comparable to a loss in year two. Since the reinsurer can think of this as borrowing from itself, the rate used is the risk-free rate. In the switch case, this is obvious, since the securities held are risk-free. In the option case, this still seems appropriate, since the lower limit which will be realized is the risk-free rate.

5. CONCLUDING REMARKS

The usefulness of safety levels is that they make explicit the minimum funds to be allocated. Unless the safety level is 100%, there is always the possibility in a particular realization of loss and investment that some safety level(s) will be breached. In this case the general account of the reinsurer will have to contribute to the cedent. This does not affect the validity of the original pricing, but the reinsurer's attitude toward this possibility will influence how the safety level was set and hence the price.

For convenience the losses are assumed to occur at the end of each year, although there is no great difficulty in generalizing to arbitrary times. Since simulations are being used, any measures of risk and return that can be defined on individual results can be used. Also, in real-world scenarios the individual years of multi-year contracts may well have some correlation simply because they are from the same firm or exposures. In the simulation environment, the overall contract can still be evaluated if one is willing to quantify the correlation.

A simplification used here is to ignore the fact that the spot rates for risk-free investment depend upon the length of time invested, usually rising with time. For example, incremental losses could be discounted back to time zero using the different spot rates. Here only one single risk-free rate is assumed to apply, for all times of a contract. However, if desired the calculations can be reformulated to include the current spot rates and the view of what the future values of the spot rates are likely to be over the contract period.

In the single payment case, the IRR is used because it is unequivocally defined and provides a natural way of talking about returns. It is not actually necessary to look at the IRR, and only the end result need be considered. In the multiple payment case, the IRR may not even be definable as a real number. This is particularly obvious when the final value is negative because of large losses, but can also happen otherwise. In order to consider the end value (future value of the cash flows), it is necessary to set up some description of the investment policy on the released funds. The target investment is the obvious choice.

It is intuitive that there should be a reduction in risk load from pooling over years, even allowing for the increased cost of liquidity of the later contract. Numerical experimentation seems to indicate that the benefits of pooling over time are usually present for uncorrelated contracts.

The pricing here is extremal pricing, in that each contract is priced as a stand-alone entity. In reality, each contract written is supported by the whole surplus of the reinsurer. A more accurate treatment of the actual risk load needed to satisfy investment criteria would be to consider the whole book with and without the proposed contract. Perhaps a satisfactory compromise would be to scale the extremal risk load contemplated here by the ratio of the overall portfolio risk charge to the sum of the extremal risk loads.

If this paradigm is to be used in connection with a complete book of business, both the general unavailability of options for periods of more than one year and the changing nature of the ongoing book suggest something like looking at the distribution of the one-year forward value of the discounted payment streams, and re-evaluating the necessary risk load annually.

REFERENCES

[1] Brealey, Richard and Stewart Myers, *Principles of Corporate Finance*, 4th Edition, 1991, McGraw-Hill, Inc.

[2] Malliaris, A. G. and W. A. Brock, Stochastic Methods in Economics and Finance, North-Holland, 1982.

APPENDIX A

The form of the Black–Scholes formula for the price of a European call option on a security is[15]

$$\text{call price} = \Phi(\Delta_1)P_0 - \Phi(\Delta_2)PV(E),$$

where

$PV(E)$ = present value of the exercise price discounted at the risk-free rate,

P_0 = price of the security at time zero,

$$\Delta_1 = \frac{\ln\left(\dfrac{P_0}{PV(E)}\right)}{\sigma\sqrt{t}} + \sigma\sqrt{t}/2, \qquad \text{and}$$

$$\Delta_2 = \Delta_1 - \sigma\sqrt{t},$$

σ is a parameter of the distribution of the underlying security, and

$\Phi(x)$ is the cumulative distribution function for the normal distribution; that is

$$\Phi(x) = \int_{-\infty}^{x} \frac{\exp\left(-\dfrac{z^2}{2}\right)}{\sqrt{2\pi}} dz.$$

This function is available in at least one standard spreadsheet program.

A call option is the right to buy an underlying security at an exercise price at time t. The logarithm of the value of the security is assumed to follow a normal distribution with parameters μt and $\sigma\sqrt{t}$ for the mean and standard deviation, respectively.[16] Given

[15] See Brealey and Myers [1, page 502].

[16] This is known as a geometric Wiener process or geometric Brownian motion process. See the development of Black–Scholes in [2] pages 220–223, and the discussion of the Brownian motion on pages 36-38, especially equation (7.13) and the development leading to it.

the expected annual yield rate y and its standard deviation σ_y, then

$$\sigma^2 = \ln\{1 + [\sigma_y/(1+y)]^2\},$$

and

$$\mu = \ln(1+y) - \sigma^2/2.$$

These equations are simply the inversion of the results for the mean and standard deviation of a log-normal distribution for $1 + y$. The approximation in Equation 2.13 comes from the first-order Taylor expansion of the relation for σ^2:

$$\ln(1+x) \approx x,$$

$$\text{so} \qquad \sigma^2 \approx [\sigma_y/(1+y)]^2.$$

The price for a put option, which is actually the contract of interest here, is given by put-call parity as

$$\text{put price} = \text{call price} + PV(E) - P_0.$$

Here, $PV(E)$ equals P_0 since we want the exercise price to be the same as the value which would result from growth at the risk-free rate. Hence the put price equals the call price, and for either option the

$$\text{option cost} = P_0\Phi(\sigma\sqrt{t}/2) - P_0\Phi(-\sigma\sqrt{t}/2)$$

so the

$$\text{option rate} = \Phi(\sigma\sqrt{t}/2) - \Phi(-\sigma\sqrt{t}/2)$$

$$= \sqrt{\frac{2}{\pi}}\int_0^{\sigma\sqrt{t}/2} \exp\left(-\frac{z^2}{2}\right) dz.$$

The exponential may be expanded as $(1 - z^2/2)$ and integrated to get the approximation of Equation 2.14 for t equal one. For the order of magnitude of numbers used here this approximation is actually rather good.

APPENDIX B

As stated in Appendix A, the probability density function for the investment value (which is 1 + return) is lognormal with parameters μt and $\sigma\sqrt{t}$. That is,

$$f(x) = \frac{1}{\sigma x\sqrt{2\pi t}}\exp\left\{\frac{-(\ln(x)-\mu t)^2}{\sigma^2 t}\right\}.$$

The investment hedged with the option to time t has the characteristics (r_f is the risk-free rate):

$$\begin{aligned}\text{investment} &= x \qquad \text{for} \quad x \geq (1+r_f)^t\\ &= (1+r_f)^t \qquad \text{for} \quad x < (1+r_f)^t.\end{aligned}$$

What is needed are the moments of the investment, in particular its mean and standard deviation.

Define

$$\begin{aligned}Fn &= \int_0^{(1+r_f)^t} x^n f(x)\,dx\\ &= \Phi(\zeta - n\sigma\sqrt{t})\exp\{n\mu t + n^2\sigma^2 t/2\},\end{aligned}$$

where $\zeta = \sqrt{t}(\ln(1+r_f)-\mu)/\sigma$. In general,

$$\begin{aligned}\text{moment}(n) &= \int_0^{\infty} \text{investment}^n f(x)\,dx\\ &= (1+r_f)^{nt}\int_0^{(1+r_f)^t} f(x)\,dx + \int_{(1+r_f)^t}^{\infty} x^n f(x)\,dx.\end{aligned}$$

Using the results for Fn above, the moment of order n of the investment is

$$\begin{aligned}\text{moment}(n) &= (1+r_f)^{tn}F_0 + \exp\{n\mu t + n^2\sigma^2 t/2\} - Fn\\ &= (1+r_f)^{tn}\Phi(\zeta) + \exp\{n\mu t + n^2\sigma^2 t/2\}\\ &\quad \times [1-\Phi(\zeta - n\sigma\sqrt{t})].\end{aligned}$$

The mean value is just moment(1) and the variance of the investment is $\{\text{moment}(2) - \text{moment}(1)^2\}$. The standard deviation σ_i of the investment is, of course, the square root of the variance.

APPENDIX C

Derivation of Equation 2.5: Substitute for L and F in Equation 2.4:

Equation 2.1 may be solved for L as

$$L = (1 + r_f)(P - R). \tag{C.1}$$

Substitute F from Equation 2.2 and L from Equation C.1 into Equation 2.4:

$$\begin{aligned}(1 + y)A &= (1 + r_f)(P + A) - (1 + r_f)(P - R)\\ &= (1 + r_f)A + (1 + r_f)R.\end{aligned}$$

Solving for R gives Equation 2.5.

Derivation of Equation 2.20: Equation 2.16 can be written

$$(1 + r)F = P + A = A + L/(1 + r_f) + R,$$

from Equation 2.1. Rearranging to solve for R, and subsequently using Equations 2.6 for F and 2.19 for A,

$$\begin{aligned}R &= (1 + r)F - A - L/(1 + r_f)\\ &= (1 + r)\frac{S}{1 + r_f} - \frac{1}{1 + y}\left\{\frac{1 + i}{1 + r_f}S - L\right\} - L/(1 + r_f)\\ &= \frac{S}{1 + r_f}\left[(1 + r) - \frac{1 + i}{1 + y}\right] + L\left[\frac{1}{1 + y} - \frac{1}{1 + r_f}\right]\\ &= \frac{1}{(1 + r_f)(1 + y)}\{S[(1 + y)(1 + r) - (1 + i)] - L[y - r_f]\}.\end{aligned}$$

Derivation of Equation 2.22: Equation 2.18 can be written

$$F = \frac{(1 + y)A + L}{1 + i}.$$

Substituting for F in Equation 2.21 gives

$$A^2\sigma_y^2 = [(1 + y)^2A^2 + 2AL(1 + y) + L^2]\frac{\sigma_i^2}{(1 + i)^2} + \sigma_L^2.$$

Multiplying through by the denominator and collecting terms,

$$0 = A^2[(1+y)^2\sigma_i^2 - \sigma_y^2(1+i)^2] \\ + 2AL(1+y)\sigma_i^2 + L^2\sigma_i^2 + \sigma_L^2(1+i)^2.$$

This is Equation 2.22. If there is a correlation ρ_{iL} between investment and loss, then this equation becomes

$$0 = A^2[(1+y)^2\sigma_i^2 - \sigma_y^2(1+i)^2] \\ + 2A(1+y)\sigma_i[L\sigma_i - \sigma_L\rho_{iL}(1+i)] \\ + L^2\sigma_i^2 + \sigma_L^2(1+i)^2 - 2L\sigma_i\sigma_L(1+i)\rho_{IL}.$$

Derivation of Equation 2.24: By substituting for F from Equation 2.16 into Equation 2.18, we get

$$(1+y)A = (1+i)\frac{P+A}{1+r} - L.$$

Multiplying through by the denominator and using Equation 2.1 for P,

$$A(1+y)(1+r) = (1+i)\left(R + \frac{L}{1+r_f} + A\right) - L(1+r).$$

Rearranging terms,

$$(1+i)R = A[(1+y)(1+r) - (1+i)] \\ + L\left[(1+r) - \frac{1+i}{1+r_f}\right].$$

Equation 2.24 for R results immediately.

WORKERS COMPENSATION EXCESS RATIOS: AN ALTERNATIVE METHOD OF ESTIMATION

HOWARD C. MAHLER

Abstract

This paper presents an alternative method of calculating excess ratios for workers compensation insurance. While the method shares many similarities with that presented by Gillam [1], there are important differences in approach. The (adjusted) data is relied upon directly for lower limits. For higher limits this is supplemented by a mixed Pareto-exponential distribution fitted to the (adjusted) data.

1. INTRODUCTION

The excess ratio for a limit L is defined as the ratio of losses excess of L to the total ground-up losses. If $f(x)$ is the probability density function for the size of loss distribution, then the excess ratio is defined as:

$$R(L) = \frac{\int_L^\infty (x-L)f(x)\,dx}{\int_0^\infty xf(x)\,dx}.$$

The excess ratio can also be written in terms of the limited expected value $\mathrm{E}[X;L]$ and the mean $\mathrm{E}[X]$: $R(L) = 1 - \mathrm{E}[X;L]/\mathrm{E}[X]$. See for example Hogg and Klugman [2].

The excess ratio is an important statistic with many applications. For example, it can be used to calculate excess loss factors for workers compensation insurance, as discussed in Gillam [1]. Generally, the excess loss factor for a limit is the product of an excess ratio and a permissible loss ratio with the possible addition of a risk load.[1]

[1] The similar excess loss and allocated expense factors (ELAFs) are for use in the ALAE option for retrospective rating. Let $1 + a$ be the factor to load losses for allocated loss adjustment expense. Then for an accident limit L, one computes $R(L/(1 + a))$ and multi-

Since excess loss factors are typically calculated by hazard group and accident limit, excess ratios need to be estimated by hazard group and by accident limit. This paper will show one method of estimating such excess ratios, with an emphasis on general principles rather than on the important details that may affect the estimate in specific situations.

2. DATA

As always, the first step is to collect the appropriate data. As in Gillam [1], Unit Statistical Plan data is used at third, fourth, and fifth report. All medical only losses are assumed to be below any accident limit. For lost time claims, the Unit Statistical Plan data has the individual claim size[2] for those claims greater than $2,000. (Those claims of size less than $2,000 may be reported on a grouped basis; all of their losses are below any accident limit.)

The reported class codes can be used to divide the data into hazard groups.[3] Using the reported information the proportion of loss dollars excess of any accident limit can be calculated.[4]

In order to illustrate the method in this paper, it will be applied to Massachusetts workers compensation data from composite policy years 1988/1989 at 5th report,[5] 1989/1990 at 4th report and 1990/1991 at 3rd report. In practice it is often appropriate to examine indications using data from several evaluation dates.

---

plies by a permissible loss and allocated LAE ratio, in order to get an ELAF. This method of calculating ELAFs assumes that the expected ALAE ratio is approximately the same for all claim sizes. The effects of variations in this assumption are beyond the scope of this paper. (ALAE data by claim has only recently started to be collected by workers compensation rating bureaus.)

[2]Paid losses plus case reserves, divided between medical and indemnity.

[3]There are four hazard groups, with hazard group 4 having the highest expected claim severity.

[4]Provided the accident limit is greater than $2,000 (times any adjustment factors).

[5]Composite policy year 1988/1989 includes all data from policies with effective dates from July 1, 1988 to June 30, 1989. Fifth report is evaluated 66 months from policy inception.

3. ADJUSTMENTS TO DATA

The claim severity is adjusted from that observed in the data to that expected in the policy effective period at the appropriate report.[6] Each claim is multiplied by an appropriate trend, law amendment and development factor. The product of these adjustment factors for a particular example is shown in Exhibit 1. These adjustment factors should be calculated in a manner consistent with the procedures that produced the rates.[7] In other words, whatever procedure was used to project past losses to ultimate in the effective period in order to estimate rates, should also be applied to the data in order to estimate excess ratios on a consistent basis.

In the example presented in Exhibit 1, the adjustment factors vary by injury kind and between medical and indemnity. To the extent that expected severity trend, development, or law changes differ significantly by claim size within injury kind, the adjustment factors could be varied by claim size interval as well. This refinement is beyond the scope of this paper.

The excess loss factors are used in pricing excess coverage on a per occurrence basis, as discussed in Gillam [1]. Therefore, we are interested in a distribution of loss values for accidents rather than claims. After the above adjustments, all claims are grouped into accidents except medical only claims.[8] A computer program was used which groups data by hazard group, accident date, and policy number, on the assumption that a single policy will not incur two or more accidents on one particular date.[9]

---

[6]The factors in Exhibit 1 include an estimate of average development to ultimate. However, the impact of the dispersion of claim sizes due to development beyond fifth report has not been taken into account.

[7]While the adjustment factors are an important part of the process, they do not represent a difference in the method presented. Therefore, the details are beyond the scope of this paper. Gillam [1] gives an example. See Feldblum [3] for a general discussion of workers compensation ratemaking.

[8]Medical only losses are much smaller than the accident limits purchased, and thus none of them will exceed a relevant loss limit. All medical only losses are assumed to be primary.

[9]Claims without a reported accident date are grouped by hazard group and claim number.

Exhibit 2 shows a comparison of excess ratios computed using ungrouped claim data and data grouped into accidents. For ungrouped data, the excess ratios obtained after adjusting the limit by dividing by a factor of 1.1 were also examined.[10]

The differences between excess ratios calculated from the ungrouped and grouped data were relatively small. At lower limits the 1.1 factor seemed to produce too much of an adjustment, but at higher limits it approximated the effect of the grouping of the data into accidents.

It should be noted that while these results may be interesting, they are far from conclusive. They represent the results for one state for one point in time. At the higher limits random fluctuations are expected to produce differing results over time. Even more importantly, the method used to group claims into accidents is far from perfect. Thus, it is inappropriate to assume the difference represents an error in either method of accounting for multi-claim accidents.

The accident data resulting from the grouping process forms the basis of the analysis.[11] The excess ratios computed from this data are shown in Exhibit 3.

4. CURVE FITTING PROCEDURE

The mean residual life statistic provides a convenient way to examine the tails of loss distributions.[12] The mean residual life at a limit x is defined as $e(x) = (\text{dollars excess of } x)/(\text{number of accidents larger than } x)$. Figure 1 displays the mean residual lives for each of the four hazard groups. As expected, the higher the hazard group the larger the mean residual life. However, as we reach higher limits the data in the two smaller hazard groups,

---

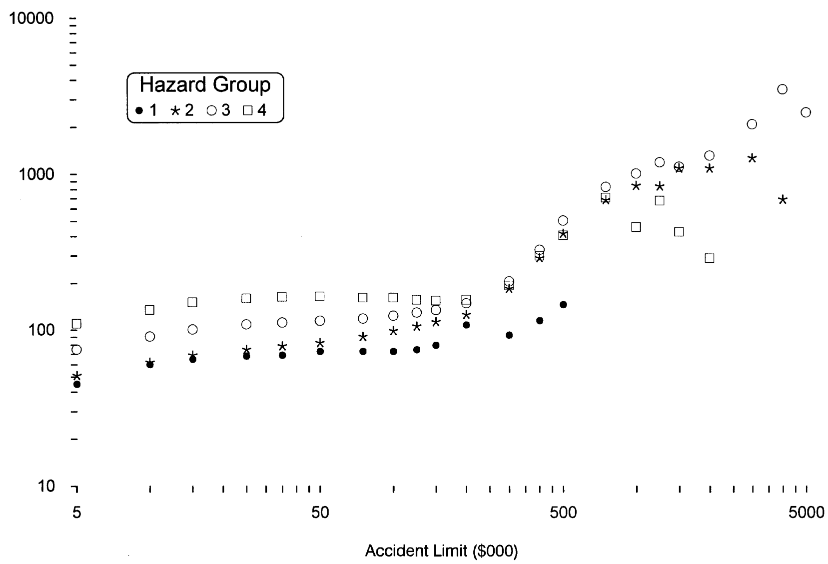
[10] This is similar to the method in Gillam [1].

[11] For the three composite policy years combined there were a total of 157,726 lost-time accidents of which 13,699 had adjusted values greater than $100,000.

[12] See for example Hogg and Klugman [2]. The mean residual life is the average excess cost of a claim that exceeds a given limit.

FIGURE 1

OBSERVED MEAN RESIDUAL LIVES ($000) BY HAZARD GROUP



Three Composite Policy Years of Data, Adjusted for Trend, Development, & Law Changes.
Based on the following number of lost-time claims: HG1 1391, HG2 97876, HG3 57307, HG4 1114.

1 and 4, becomes sparse. The chance of a very large claim appearing in the data for these hazard groups is too small[13] to get a reliable estimate of the mean residual life at high limits.

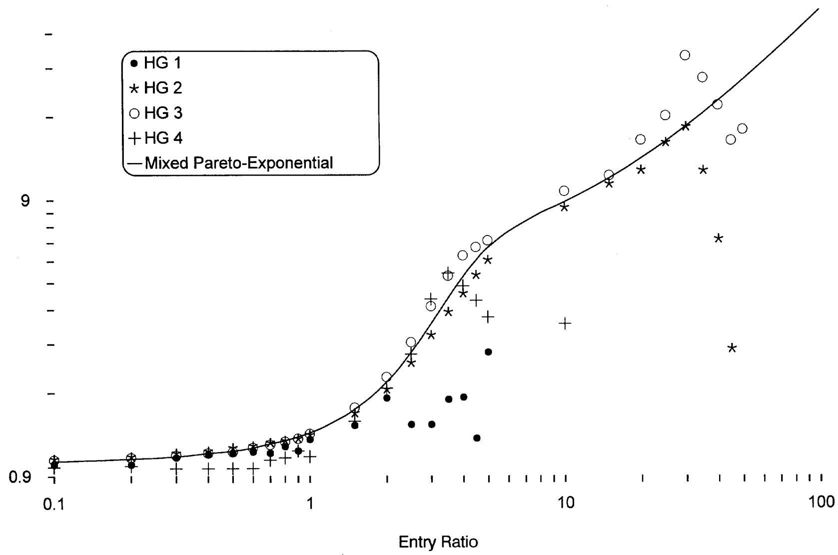
The hazard groups seem to have a similar pattern, with the mean residual life increasing, at least up to an accident limit of several million dollars. A number of adjustments are made to the accident data in order to fit a distribution to it.

---

[13]For example, we can estimate that on average we expect about 0.4 accidents greater than $1 million for hazard group 1. This is based on 96 accidents greater than $100,000 in the data set for hazard group 1 and a tail probability of the fitted mixed Pareto-exponential distribution of .0038 at an entry ratio of 12.4. $(.0038)(96) \approx .4$. In the reported data there were 39 accidents greater than $1 million of which none were in hazard group 1, 11 were in hazard group 2, 27 were in hazard group 3, and 1 was in hazard group 4.

FIGURE 2

MEAN RESIDUAL LIVES BY HAZARD GROUP
Data Truncated and Shifted at $100,000 and then Normalized to a Mean of Unity



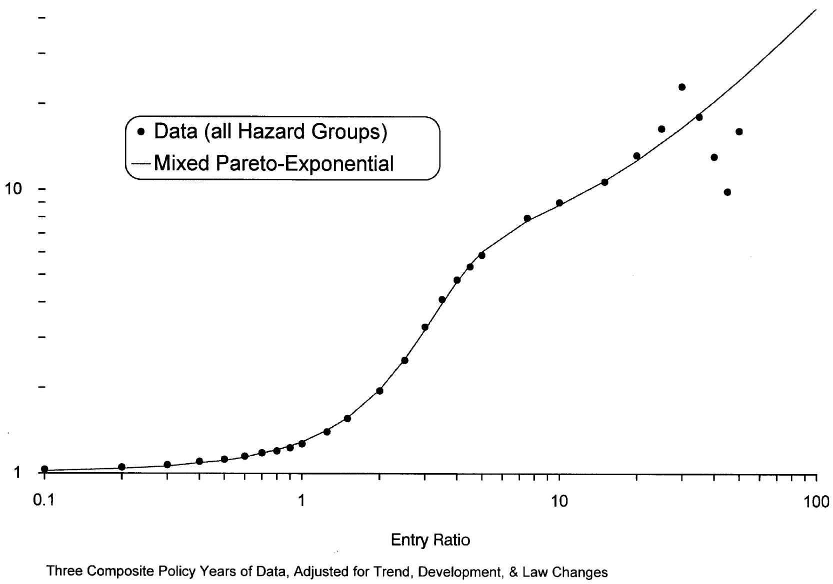
Three Composite Policy Years of Data, Adjusted for Trend, Development, & Law Changes.
Based on the following number of lost-time claims: HG1 1391, HG2 97876, HG3 57307, HG4 1114.

First, the accident data for third, fourth, and fifth report are combined. Next, for each of the four hazard groups, the data are truncated and shifted at $100,000.[14] Finally, each of these four sets of data is normalized to a mean of unity. Figure 2 shows the mean residual lives for the resulting truncated, shifted and then normalized data by hazard group. Bearing in mind the limited data for hazard groups 1 and 4, it is plausible that the normalized

[14] Accidents with losses less than or equal to $100,000 are eliminated from consideration (for now). Those of size $x > \$100{,}000$ have $100,000 subtracted from them and appear in the truncated and shifted data as $x - \$100{,}000$.

FIGURE 3

MEAN RESIDUAL LIVES, OBSERVED VS. FITTED
Data Truncated and Shifted at $100,000 and then Normalized to a Mean of Unity



hazard group data might all come from approximately the same distribution. These four sets of normalized data are combined, as displayed in Figure 3.

A mixture[15] of Pareto and exponential distributions is fit to this combined data[16] using the method of maximum likeli-

---

[15]See for example, Hogg and Klugman [2] for a discussion of the mixture of loss distribution models. The probability density function is $f(x) = pg(x) + (1-p)h(x)$, where g and h are each probability density function.

[16]The data has been combined across reports, injury kinds, and hazard groups, representing over 13,000 accidents over $100,000 in size.

hood.[17] The Pareto and exponential curves are standard size of loss distributions, described in Exhibits 5 and 6. The mixed distribution is (p) (Pareto distribution) + ($1-p$) (exponential distribution) where p is a fitted parameter with a value between zero and one. Together with the two Pareto parameters (shape and scale) and the single exponential parameter, the mixed distribution has a total of four parameters.

The fitted parameters are displayed in Exhibit 4. Figure 3 compares the mean residual lives for the fitted curve and the observed data. Figure 4 shows the probability density functions for the mixed Pareto-exponential as well as the Pareto and exponential distributions. For small entry ratios the mixed curve behaves as the short-tailed exponential, while for larger entry ratios it behaves as the long-tailed Pareto.

Figure 5 compares the excess ratios for the mixed distribution to that of the exponential and the Pareto. As derived in the Appendix, the excess ratio for the mixed distribution is a weighted average of the excess ratios of the individual distributions, with the weights being (p) (mean of Pareto) and ($1-p$) (mean of the exponential). In this case, the weights are .2132 and .7868.[18] Thus, for lower entry ratios the excess ratio of the mixed distribution is close to that for the exponential. At higher limits, the excess ratio for the short-tailed exponential is too small to contribute significantly. Therefore, the excess ratio of the mixed distribution for higher entry ratios is approximately 21% of that for the Pareto.

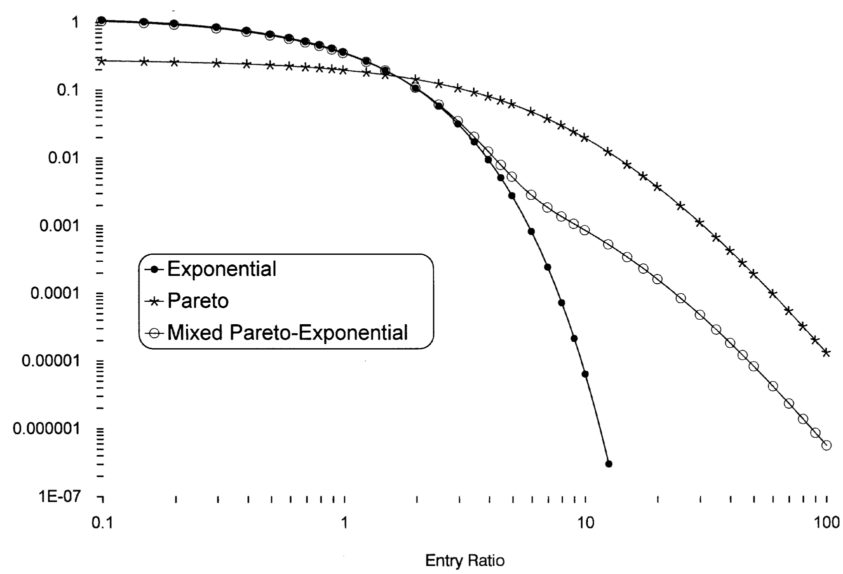
---

[17]The result of the maximum likelihood method has a mean slightly different from unity, so the scale parameters of the Pareto and exponential have been adjusted so as to have the desired mean of unity. The method of maximum likelihood is a commonly used method for fitting size of loss distributions to either grouped or ungrouped data, as discussed in Hogg and Klugman [2]. In this case, the method was applied to the individual data points rather than data grouped into intervals.

[18]This is based on $p = .04294$, a mean of the Pareto of $12.83704/(3.58490-1) = 4.9662$ and a mean of the exponential of .82205, as shown in Exhibit 4. $.2132 = (.04294)(4.9662)/\{(.04294)(4.9662) + (.95706)(.82205)\}$.

FIGURE 4

PROBABILITY DENSITY FUNCTIONS
Data Truncated and Shifted at $100,000 and then Normalized to a Mean of Unity



Pareto parameters: 3.58490 and 12.83704. Exponential parameter: .82205. Weight to the Pareto: .04294.

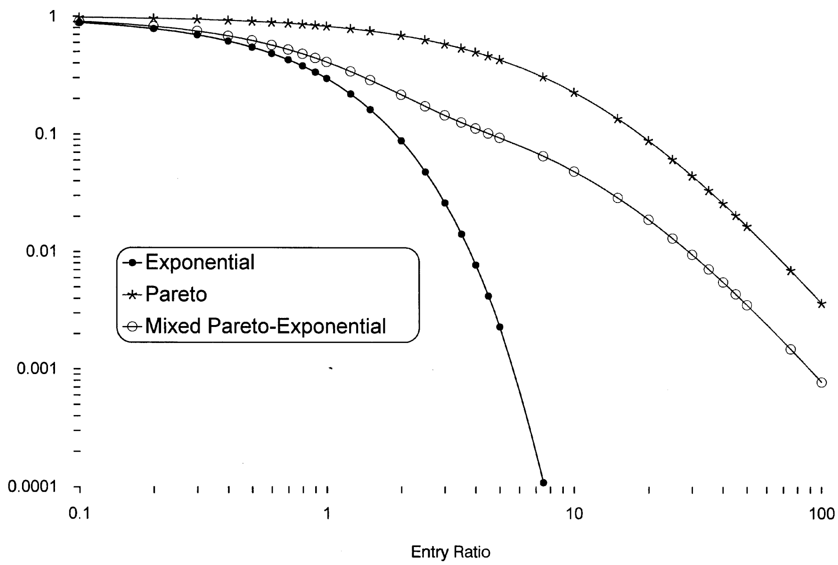
5. ESTIMATION OF EXCESS RATIOS

For each hazard group this fitted curve, scaled to the observed mean, is used in Exhibit 7 to estimate the excess ratios for the data truncated and shifted at $100,000.

The excess ratios for accident limits less than or equal to $100,000 are determined directly from the data. For accident limits L above $100,000, the excess ratio is estimated from the product of (empirical excess ratio at $100,000) $\times$ (excess ratio estimated from mixed Pareto-exponential curve for $L - \$100,000$). See the Appendix. The former is shown in Exhibit 2, the latter in

FIGURE 5

EXCESS RATIOS
Data Truncated and Shifted at $100,000 and then Normalized to a Mean of Unity



Pareto parameters: 3.58490 and 12.83704. Exponential parameter: .82205. Weight to the Pareto: .04294.

Exhibit 7, while the product is in Exhibit 8.[19] Figure 6 compares the estimated and observed excess ratios.

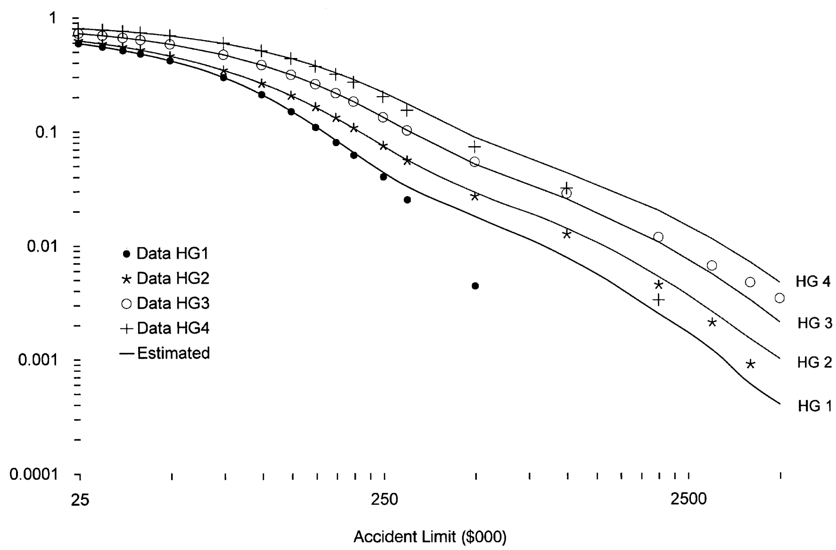
This method provides a smooth transition from relying on data for lower accident limits to relying on a fitted curve to provide some of the information at higher accident limits. It is important to note that even at higher accident limits an important contribution to the excess ratio is $R(100{,}000)$ which is calculated directly from the data.

---

[19]It should be noted that for a limit of $100,000 the two methods automatically give the same answer since the excess ratio estimated from the curve at 0 is always unity.

FIGURE 6

EXCESS RATIOS BY HAZARD GROUP
Observed versus Estimated



Three Composite Policy Years of Data, Adjusted for Trend, Development, & Law Changes.
For accident limits of $100,000 or less, the estimated excess ratio is equal to the observed excess ratio.

6. SELECTION OF A TRUNCATION POINT

The $100,000 truncation point was selected to permit the maximum reliance on reported data while still retaining enough data above the truncation point to permit the reasonable fitting of a loss distribution. For this technique and data set, a truncation point around $100,000 achieves the desired balance. Other values such as $50,000 or $150,000 could also have been used without substantially altering the estimated excess ratios.

In general, the truncation point should be a round number prior to the "thinning out" of the data. In this data set there are over 13,000 accidents with values greater than $100,000, with the two smallest hazard groups having about 100 or 200 accidents.[20] For the two larger hazard groups, a higher truncation point could have been selected, but for hazard groups 1 and 4 a higher truncation point would make it difficult to get a reliable average value to use to normalize the data.[21]

7. FEATURES OF THE PROCEDURE

This procedure allows us to rely on the actual data for the lower layers where there is a larger volume of data less subject to random fluctuation. The task of fitting curves to the smaller accidents is avoided totally.

Fitting curves to the combined data regardless of injury kind allows claims to be grouped into accidents.[22] It also avoids relying on the sometimes arbitrary or judgmental division of claims between injury kind.[23] The mixed Pareto-exponential distribution fit to the truncated and shifted data assigns the preponderance of weight to the short-tailed exponential distribution.[24] The long-tailed Pareto distribution models the behavior of the extreme tail of the accident distribution and has a very large effect on the estimated excess ratios for limits over $500,000.

Thus, the estimation procedure can be viewed in terms of three layers. The layer of losses below $100,000 is estimated without curve fitting. The layer from $100,000 to about $500,000 is

[20] There are 96 accidents from hazard group 1 and 228 accidents from hazard group 4.

[21] These average values are used in Exhibit 7 in order to calculate excess ratios by hazard group.

[22] An accident may consist of claims of several different injury kinds. For the calculation of the effect of accident limits it is not inherently necessary to divide dollars between injury kinds.

[23] Note, however, that prior to grouping by accident, claims of differing injury kinds have somewhat different adjustment factors applied to them, as shown in Exhibit 1.

[24] As is common in the use of mixed distribution, a mixture of a longer and shorter tailed distribution was selected. Originally, the short-tailed distribution was a Weibull. However, the fitted Weibull portions of the mixed distribution were very close to an exponential. Therefore, the one parameter exponential was substituted for the two parameter Weibull of which the exponential is a special case.

modeled largely by the exponential distribution. The layer above about \$500,000 is modeled largely by the Pareto distribution.[25]

8. CONCLUSION

Actuaries should be familiar with the Pareto distribution, the exponential distribution, and truncated and shifted data. These basic concepts have been employed together in a procedure with a powerful ability to fit the observed data. This procedure of estimating excess ratios is likely to be useful in various practical applications.

[25]The parameters of the fitted Pareto-exponential determine the approximate layers above \$100,000. Although it may be conceptually useful to think of it that way, there is no actual division into layers above \$100,000.

REFERENCES

[1] Gillam, William R., "Retrospective Rating: Excess Loss Factors," *PCAS* LXXVIII, 1991, pp. 1–40.

[2] Hogg, Robert V. and Stuart A. Klugman, *Loss Distributions*, New York, Wiley, 1984.

[3] Feldblum, Sholom, "Workers' Compensation Ratemaking," *Casualty Actuarial Society Forum*, Special Edition including 1993 Ratemaking Call Papers.

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

EXHIBIT 1

COMBINED TREND, LAW, AND DEVELOPMENT FACTORS

INDEMNITY

Composite Pol. Yr.	Injury Kind 1	Injury Kind 2	Injury Kind 3	Injury Kind 4	Injury Kind 5
88/89	1.79	1.82	1.41	1.28	1.04
89/90	1.53	1.87	1.38	1.26	0.95
90/91	1.42	1.56	1.44	1.31	0.91

MEDICAL

Composite Pol. Yr.	Injury Kind 1	Injury Kind 2	Injury Kind 3	Injury Kind 4	Injury Kind 5	Injury Kind 6
88/89	2.29	2.29	2.29	1.85	1.85	1.85
89/90	3.93	2.06	2.15	1.74	1.61	1.67
90/91	3.30	1.96	2.00	1.62	1.38	1.50

Notes: Product of separate factors calculated to bring all losses to ultimate and a common level, consistent with a 10/1/96 effective date. Injury Kind 1 = Fatal, Injury Kind 2 = Permanent Total, Injury Kind 3 = Major Permanent Partial, Injury Kind 4 = Minor Permanent Partial, Injury Kind 5 = Temporary Total, Injury Kind 6 = Medical Only.

EXHIBIT 2

OBSERVED EXCESS RATIOS FOR UNADJUSTED DATA[1]

Hazard Group 2			
		Claim Data	
Limit ($000)	Accident Data[2]	Using Limit	Using Limit ÷ 1.1
25	.5230	.5130	.5373
100	.1417	.1342	.1553
500	.0167	.0157	.0171
1,000	.0087	.0078	.0087
2,000	.0042	.0039	.0043
Hazard Group 3			
25	.6335	.6259	.6465
100	.2369	.2276	.2560
500	.0311	.0295	.0324
1,000	.0128	.0118	.0138
2,000	.0042	.0041	.0047

[1]The data for three separate reports, 88/89 at 3rd, 87/88 at 4th, 86/87 at 5th have been combined and then an excess ratio has been calculated. The data have *not* been adjusted for trend, law amendments, or development.

[2]Claims with the same hazard group, accident date, and policy number are grouped into the same accident.

EXHIBIT 3

EXCESS RATIOS BASED ON ADJUSTED DATA[1]

Accident Limit ($000)	Hazard Group 1[2]	Hazard Group 2	Hazard Group 3	Hazard Group 4[2]
25	0.5950	0.6288	0.7283	0.8064
30	0.5530	0.5888	0.6960	0.7817
35	0.5142	0.5521	0.6655	0.7581
40	0.4791	0.5184	0.6366	0.7353
50	0.4177	0.4586	0.5831	0.6918
75	0.2974	0.3441	0.4709	0.5935
100	0.2106	0.2643	0.3832	0.5098
125	0.1494	0.2072	0.3146	0.4353
150	0.1086	0.1647	0.2604	0.3715
175	0.0804	0.1327	0.2171	0.3165
200	0.0622	0.1081	0.1827	0.2699
250	0.0400	0.0754	0.1333	0.2011
300	0.0252	0.0559	0.1021	0.1526
500	0.0044	0.0271	0.0541	0.0730
1,000	0.0000	0.0126	0.0286	0.0317
2,000	0.0000	0.0045	0.0118	0.0033
3,000	0.0000	0.0021	0.0066	0.0000
4,000	0.0000	0.0009	0.0047	0.0000
5,000	0.0000	0.0000	0.0034	0.0000

[1] Massachusetts Workers Compensation, Composite Policy Years 88/89 at 5th, 89/90 at 4th, 90/91 at 3rd.
[2] Note there is relatively little data for hazard groups 1 and 4 since they each represent only between 1 and 2 percent of total premiums. In this reported data there were no accidents greater than $1,000,000 for hazard group 1 and only one for hazard group 4. Thus the empirical excess ratios at higher limits for these hazard groups are poor estimates of future expected excess ratios.

EXHIBIT 4

MIXED PARETO-EXPONENTIAL DISTRIBUTION

Parameters:

Pareto Shape = s	3.58490
Pareto Scale = b	12.83704
Exponential Scale = c	0.82205
Weight to Pareto = p	0.04294
Mean = 1	Coef. of Var. = 1.94
Variance = 3.75	Skewness = 30

Excess Ratios

Entry Ratio	Excess Ratio	Entry Ratio	Excess Ratio
0.1	0.9057	11	0.0431
0.2	0.8217	12	0.0387
0.3	0.7470	13	0.0350
0.4	0.6806	14	0.0317
0.5	0.6214	15	0.0288
0.6	0.5687	20	0.0188
0.7	0.5217	25	0.0131
0.8	0.4797	30	0.0095
0.9	0.4422	35	0.0071
1.0	0.4088	40	0.0055
1.25	0.3397	45	0.0044
1.5	0.2872	50	0.0035
1.75	0.2469	55	0.0029
2.0	0.2157	60	0.0024
2.5	0.1722	65	0.0020
3.0	0.1444	70	0.0017
3.5	0.1255	75	0.0015
4.0	0.1118	80	0.0013
4.5	0.1014	85	0.0011
5.0	0.0929	90	0.0010
6.0	0.0797	95	0.0009
7.0	0.0694	100	0.0008
8.0	0.0610		
9.0	0.0540		
10.0	0.0481		

Note: See the Appendix for a sample calculation of an excess ratio.

EXHIBIT 5

PARETO DISTRIBUTION

$$F(x;s,b) = 1 - \left(1 + \frac{x}{b}\right)^{-s}$$

$$f(x;s,b) = \frac{s}{b}\left(1 + \frac{x}{b}\right)^{-(s+1)}$$

$$\mathrm{E}(X^y) = \frac{b^y\Gamma(y+1)\Gamma(s-y)}{\Gamma(s)}, \qquad -1 < y < s$$

If y is an integer N,

$$\mathrm{E}(X^N) = \frac{b^N N!}{\prod_{i=1}^{N}(s-i)} \qquad N < s$$

$$\text{Mean} = \frac{b}{s-1} \qquad \text{Variance} = \frac{b^2 s}{(s-1)^2(s-2)}$$

$$\text{Coefficient of Variation} = \sqrt{\frac{s}{s-2}} \qquad s > 2$$

$$\text{Skewness} = \frac{2(s+1)}{s-3}\sqrt{\frac{s-2}{s}} \qquad s > 3$$

$$\text{Excess Ratio} = R(x) = \left(1 + \frac{x}{b}\right)^{1-s}$$

$$\text{Mean Residual Life} = e(x) = \frac{b+x}{s-1}$$

Note: s is the shape parameter, b is the scale parameter.

EXHIBIT 6

EXPONENTIAL DISTRIBUTION

$$F(x;c) = 1 - e^{-x/c}$$

$$f(x;c) = \frac{1}{c}e^{-x/c}$$

$$\mathrm{E}(X^y) = c^y\Gamma(1+y) \qquad y > -1$$

If y is an integer N,

$$\mathrm{E}(X^N) = c^N N! \qquad N > -1$$

$$\text{Mean} = c$$

$$\text{Variance} = c^2$$

$$\text{Coefficient of Variation} = \text{Standard Deviation} \div \text{Mean} = 1$$

$$\text{Skewness} = 2$$

$$\text{Excess Ratio} = R(x) = e^{-x/c}$$

$$\text{Mean Residual Life} = e(x) = c$$

Note: c is the scale parameter.

EXHIBIT 7

EXCESS RATIOS TRUNCATED AND SHIFTED TO $100,000

	Hazard Group 1		Hazard Group 2		Hazard Group 3		Hazard Group 4	
Accident Limit ($000)	Entry Ratio	Excess Ratio	Entry Ratio	Excess Ratio	Entry Ratio	Excess Ratio	Entry Ratio	Excess Ratio
125	0.3446	0.7165	0.2531	0.7810	0.2020	0.8201	0.1541	0.8590
150	0.6891	0.5265	0.5062	0.6180	0.4041	0.6781	0.3083	0.7413
175	1.0337	0.3983	0.7592	0.4962	0.6061	0.5657	0.4624	0.6429
200	1.3783	0.3110	1.0123	0.4049	0.8081	0.4765	0.6165	0.5606
250	2.0674	0.2086	1.5185	0.2838	1.2122	0.3490	0.9248	0.4335
300	2.7566	0.1565	2.0247	0.2131	1.6162	0.2671	1.2331	0.3438
500	5.5131	0.0856	4.0493	0.1107	3.2325	0.1348	2.4662	0.1746
1,000	12.4046	0.0371	9.1110	0.0533	7.2731	0.0669	5.5489	0.0852
2,000	26.1874	0.0120	19.2343	0.0200	15.3543	0.0279	11.7144	0.0399
3,000	39.9702	0.0055	29.3576	0.0098	23.4355	0.0145	17.8798	0.0223
4,000	53.7531	0.0030	39.4809	0.0057	31.5167	0.0087	24.0453	0.0139
5,000	67.5359	0.0019	49.6042	0.0036	39.5979	0.0056	30.2107	0.0093

Entry Ratio = (LIMIT – $100,000 Truncation Point)/Average Size for Data Truncated and Shifted to $100,000.
Average Size of Loss Data Truncated and Shifted to $100,000 by Hazard Group:

HG 1	72,554
HG 2	98,782
HG 3	123,744
HG 4	162,194

Excess ratio is computed for a Pareto-exponential distribution with parameters:

Pareto Shape	Pareto Scale	Exponential	Weight to Pareto
3.58490	12.83704	0.82205	0.04294

EXHIBIT 8

ESTIMATED EXCESS RATIOS BASED ON ADJUSTED DATA AND CURVES FIT TO DATA TRUNCATED AND SHIFTED AT $100,000

Accident Limit ($000)	Hazard Group 1	Hazard Group 2	Hazard Group 3	Hazard Group 4
25	0.5950	0.6288	0.7283	0.8064
30	0.5530	0.5888	0.6960	0.7817
35	0.5142	0.5521	0.6655	0.7581
40	0.4791	0.5184	0.6366	0.7353
50	0.4177	0.4586	0.5831	0.6918
75	0.2974	0.3441	0.4709	0.5935
100	0.2106	0.2643	0.3832	0.5098
125	0.1509	0.2064	0.3143	0.4379
150	0.1109	0.1633	0.2598	0.3779
175	0.0839	0.1311	0.2168	0.3278
200	0.0655	0.1070	0.1826	0.2857
250	0.0439	0.0750	0.1337	0.2210
300	0.0330	0.0563	0.1024	0.1753
500	0.0180	0.0293	0.0517	0.0890
1,000	0.0078	0.0141	0.0256	0.0434
2,000	0.0025	0.0053	0.0107	0.0203
3,000	0.0012	0.0026	0.0056	0.0114
4,000	0.0006	0.0015	0.0033	0.0071
5,000	0.0004	0.0010	0.0021	0.0047

Note: For accident limits of $100,000 or less the excess ratio is taken directly from Exhibit 3. For accident limits larger than $100,000, the excess ratio is a product of that for $100,000 in Exhibit 3 and the excess ratio shown in Exhibit 7. For example, for hazard group 3 at a limit of $1 million, (.3832)(.0669) = .0256.

APPENDIX

Excess Ratios, Truncated and Shifted Data

Let $f(x)$ be the size of loss probability density function. Then the excess ratio for limit L is given by:

$$R(L) = \frac{\int_L^\infty (x-L)f(x)\,dx}{\int_0^\infty xf(x)\,dx} = \frac{\text{average dollars of loss excess of } L}{\text{average size of loss}} = \frac{\text{total dollars of loss excess of } L}{\text{total dollars of loss}}.$$

Assume we have a truncation point of T. Assume we look at the size of loss distribution for the data truncated and shifted at T. So for a loss $x > T$, we instead look at $x - T$. Then the excess ratio for the truncated and shifted data for ground up limit $L > T$ can be written as

$$\hat{R}(L-T).$$

Assume we were computing the (observed) excess ratio for a \$500,000 accident limit, for hazard group 3 data[26]

$$R(\$500{,}000) = \frac{\text{HG3 Losses Excess of \$500,000}}{\text{Total HG3 Losses (including Medical Only)}}.$$

We can also express this in terms of the data truncated and shifted at \$100,000 as follows:

$$R(\$500{,}000) = \frac{\text{HG3 Losses Excess of \$500,000}}{\text{HG3 Losses Excess of \$100,000}} \times \frac{\text{HG3 Losses Excess of \$100,000}}{\text{Total HG3 Losses (including Medical Only)}}.$$

[26]For 3rd, 4th, and 5th report combined, adjusted for trend, law changes, and development.

However, the second term is the excess ratio at \$100,000, $R(\$100{,}000)$, while the first term is $\hat{R}(\$400{,}000)$ = excess ratio at \$400,000 for the data truncated and shifted at \$100,000. Thus

$$R(\$500{,}000) = \hat{R}(\$400{,}000) \times R(\$100{,}000).$$

In general, for limits $L > \$100{,}000$

$$R(L) = \hat{R}(L - \$100{,}000) \times R(100{,}000).$$

In the methodology used here, $\hat{R}(L - \$100{,}000)$ is estimated via a curve fit to the data truncated and shifted at \$100,000, while $R(100{,}000)$ is estimated from the data.

Excess Ratios, Mixed Distributions

Let a (mixed) distribution be a weighted average of two other distributions:

$$f(x) = pg(x) + (1-p)h(x).$$

Then the mean is a weighted average of the two means:

$$\begin{aligned} m_f &= \int_0^\infty x f(x)\,dx = \int_0^\infty x\{pg(x) + (1-p)h(x)\}\,dx \\ &= p\int_0^\infty x g(x)\,dx + (1-p)\int_0^\infty x h(x)\,dx \\ &= pm_g + (1-p)m_h. \end{aligned}$$

The excess ratio for limit L is given by:

$$\begin{aligned} R_f(L) &= \frac{\int_L^\infty (x-L) f(x)\,dx}{\int_0^\infty x f(x)\,dx} \\ &= \frac{p\int_L^\infty (x-L)g(x)\,dx + (1-p)\int_L^\infty (x-L)h(x)\,dx}{pm_g + (1-p)m_h} \\ &= \frac{pm_g R_g(L) + (1-p)m_h R_h(L)}{pm_g + (1-p)m_h}. \end{aligned}$$

So the excess ratio for a mixed distribution is a weighted average of the excess ratios for the individual distributions, with weights equal to the product of the mean of each distribution times the weight in the mixture of each distribution.[27]

For example, for the mixed Pareto-exponential distribution with parameters: 3.5849, 12.83704, .82205, .04294, at an entry ratio of 2, the excess ratio is computed as follows:

excess ratio for Pareto (3.5849, 12.83704) at entry ratio 2 (of the mixed distribution)

$$= \left(1 + \frac{2}{12.83704}\right)^{1-3.5849} = .6878$$

excess ratio for exponential (.82205) at entry ratio 2 (of the mixed distribution)

$$= e^{-2/.82205} = .0878$$

mean for Pareto (3.5849, 12, 83704)

$$= \frac{12.83704}{3.5849 - 1} = 4.9662$$

mean for exponential (.82205) = .82205

excess ratio for Pareto-exponential at entry ratio 2

$$= \frac{(.04294)(4.9662)(.6878) + (1 - .04294)(.82205)(.0878)}{(.04294)(4.9662) + (1 - .04294)(.82205)}$$

$$= .2157/1 = .2157.$$

This matches the value shown on Exhibit 4.

Moments of Mixed Models

Moments of a mixed model are a weighted average of the moments of the individual distributions. For example, for the mixed Pareto-exponential distribution with parameters: 3.5849,

[27] This is closely related to the similar result for increased limit factors discussed in Venter [4].

12.83704, .82205, .04294, the moments are a weighted average using weights of .04294 applied to the moments of the Pareto and 1 – .04294 = .95706 applied to the moments of the exponential.

Moment	Pareto	Exponential	Pareto-exponential
1	4.9662	.82205	1
2	80.4478	1.35153	4.7479
3	5296.86	3.3331	230.64

Then the variance of the Pareto-exponential is $4.7479 - 1^2 = 3.7479$. Note that the variance of the mixed distribution is not the weighted average of the individual variances. The skewness of the Pareto-exponential is

$$\{230.64 - (3)(4.7479)(1^2) + 2(1^3)\}/3.7479^{1.5} = 30.1.$$

The coefficient of variation is $(\sqrt{3.7479})/1 = 1.94$. These match the values shown on Exhibit 4.

AN APPLICATION OF GAME THEORY: PROPERTY CATASTROPHE RISK LOAD

DONALD F. MANGO

Abstract

Two well known methods for calculating risk load—Marginal Surplus and Marginal Variance—are applied to output from catastrophe modeling software. Risk loads for these marginal methods are calculated for sample new and renewal accounts. Differences between new and renewal pricing are examined. For new situations, both current methods allocate the full marginal impact of the addition of a new account to that new account. For renewal situations, a new concept is introduced which we call "renewal additivity."

Neither marginal method is renewal additive. A new method is introduced, inspired by game theory, which splits the mutual covariance between any two accounts evenly between those accounts. The new method is extended and generalized to a proportional sharing of mutual covariance between any two accounts. Both new approaches are tested in new and renewal situations.

ACKNOWLDGEMENTS

The author would like to thank Sean Ringsted, Eric Lemieux and Paul Kneuer for their reviews of earlier versions of the paper, and the Committee on Review of Papers for their helpful commentary, all of which greatly improved the paper.

1. INTRODUCTION

The calculation of risk load continues to be a topic of interest in the actuarial community—see Bault [1] for a recent survey of well known alternatives. One area of great need, where the

CAS literature is somewhat scarce, is calculation of risk loads for property catastrophe insurance.

Many of the new catastrophe modeling products produce occurrence size-of-loss distributions for a series of simulated events. These output files might contain an event identifier, event probability, and modeled loss amount for that event for the selected portfolio of exposures. Given such output files for a portfolio before and after the addition of a new account, one could calculate the before and after portfolio variance and standard deviation. The difference will be called the *marginal impact* of that new account on the portfolio variance or standard deviation.[1]

Two of the more well known risk load methods from the CAS literature—what shall be called "Marginal Surplus" (MS) from Kreps [3] and "Marginal Variance" (MV) from Meyers [6]—use the marginal change in portfolio standard deviation (variance) due to the addition of a new account to calculate the risk load for that new account. However, problems arise when these marginal methods are used to calculate risk loads for the renewal of accounts in a portfolio. These problems can be traced to the *order dependency* of the marginal risk load methods.

Order dependency is a perplexing issue. Many marginal risk load methods—whether based on variance, standard deviation, or even a selected percentile of the loss distribution—suffer from it. It is also not just an actuarial issue; even the financial community struggles with it. "Value at Risk" (VAR) is an attempt by investment firms to capture their risk in a single number. It is a selected percentile of the return distribution (e.g., 95th) for a portfolio of financial instruments over a selected time frame (e.g., 30 days). VAR can be calculated for the entire portfolio or for a desired subset (e.g., asset class). But so-called "marginal"

[1]The variance and standard deviation are "between account" and "between event," and ignore any parameter uncertainty associated with the modeled loss amount for a given event and account.

or "component" VAR has, to this point, eluded satisfactory solution in the finance community precisely because of what will be termed *renewal additivity*. Finance professionals charged with assessing how much VAR a certain financial instrument or asset class contributes to the total VAR are dealing with the same unresolved order dependency issue. As the finance and insurance worlds blend more and more, perhaps actuaries will combine forces with quantitative analysts and Certified Financial Analysts (CFAs) to determine a solution.

The remainder of this paper is organized as follows. Section 2 describes the basic characteristics of a catastrophe occurrence size-of-loss distribution. Sections 3 and 4 describe the application of the MV and MS methods to a simplified occurrence size-of-loss distribution. Sections 5 and 6 calculate risk loads both in assembling or building up a portfolio of risks and in subsequently renewing that portfolio. Section 7 discusses the differences between build-up and renewal results.

Section 8 introduces a new concept to the theory of property catastrophe risk loads—renewal additivity. However, the concept is not new to the field of game theory. Section 9 introduces game theory concepts underlying a new approach. Section 10 extends and generalizes the effect of the new approach to sharing of covariance between accounts. Section 11 concludes by applying the new approach to some examples.

2. THE CATASTROPHE OCCURRENCE SIZE-OF-LOSS DISTRIBUTION

For demonstration purposes throughout the paper, a simplified version of an occurrence size-of-loss distribution will be used. It captures the essence of typical catastrophe modeling software output, while keeping the examples understandable.[2]

[2]In particular, only single event or occurrence size-of-loss distributions will be considered. Many models also produce multi-event or aggregate loss distributions. Occurrence

A modeled event denoted by identifier i is considered an independent Poisson process with occurrence rate[3] λ_i. To simplify the mathematics, following Meyers [6], the binomial approximation with probability of occurrence p_i [where $\lambda_i = -\ln(1-p_i)$] will be employed. This is a satisfactory approximation for small[4] λ_i.

For an individual account or portfolio of accounts, the model produces a modeled loss for each event L_i. A table containing the event identifiers i, the event probabilities p_i and modeled losses L_i will be referred to as an "occurrence size-of-loss distribution."

From Meyers [6], the formulas for expected loss and variance are

$$E(L) = \Sigma_i[L_i \times p_i], \qquad \text{and} \tag{2.1}$$

$$\text{Var}(L) = \Sigma_i[L_i^2 \times p_i \times (1-p_i)], \tag{2.2}$$

where Σ_i = sum over all events.

The formula for covariance of an existing portfolio L (with losses L_i) and a new account n (with losses n_i) is

$$\text{Cov}(L,n) = \Sigma_i[L_i \times n_i \times p_i \times (1-p_i)]. \tag{2.3}$$

Note that $\text{Cov}(L,n)$ is always greater than zero when each of L_i, n_i, p_i, and $(1-p_i)$ are greater than zero.

The total variance of the combined portfolio $(L+n)$ is then

$$\text{Var}(L+n) = \text{Var}(L) + \text{Var}(n) + 2\,\text{Cov}(L,n). \tag{2.4}$$

---

size-of-loss distributions reflect only the *largest* event that occurs in a given year. Aggregate loss distributions reflect the sum of losses for all events in a given year. Clearly, the aggregate distribution would provide a more complete picture, but for purposes of the exposition here, the occurrence distribution works well and the formulas are substantially less complex.

[3]This implies that the loss for a given event and account is fixed and known.

[4]An event with a probability of 0.001 (typical of the more severe modeled events) would have $\lambda = 0.0010005$.

3. THE MARGINAL SURPLUS (MS) METHOD

This is a translation of the method described in Rodney Kreps's paper, "Reinsurer Risk Loads from Marginal Surplus Requirements" [3] to property catastrophe calculations.

Consider:

L_0 = losses from a portfolio before a new account is added,

L_1 = losses from a portfolio after a new account is added,

S_0 = Standard deviation of L_0,

S_1 = Standard deviation of L_1,

R_0 = Return on the portfolio before new account is added, and

R_1 = Return on the portfolio after new account is added.

Borrowing from Mr. Kreps, assume that needed surplus, V, is given by[5]

$$V = z \times \text{standard deviation of loss} - \text{expected return}, \quad (3.1)$$

where z is, to cite Mr. Kreps [3, p. 197], "a distribution percentage point corresponding to the acceptable probability that the actual result will require even more surplus than allocated." Then

$$\begin{aligned} V_0 &= z \times S_0 - R_0, \quad \text{and} \\ V_1 &= z \times S_1 - R_1. \end{aligned} \quad (3.2)$$

The difference in returns $R_1 - R_0 = r$, the risk load charged to the new account. The marginal surplus requirement is then

$$V_1 - V_0 = z \times [S_1 - S_0] - r. \quad (3.3)$$

Based on the required return, y, on that marginal surplus (which is based on management goals, market forces and risk appetite),

[5]Mr. Kreps sets needed surplus equal to $z \times$ standard deviation of return − expected return. If one assumes premiums and expenses are invariant, then $\text{Var(Return)} = \text{Var}(P - E - L) = \text{Var}(L)$.

the MS risk load would be

$$r = [yz/(1+y)][S_1 - S_0]. \tag{3.4}$$

4. THE MARGINAL VARIANCE (MV) METHOD

The Marginal Variance Method is based on Glenn Meyers's paper, "The Competitive Market Equilibrium Risk Load Formula for Catastrophe Ratemaking" [6].

For an existing portfolio L and a new account n, the MV risk load r would be

$$\begin{aligned} r &= \lambda \times \text{Marginal Variance of adding } n \text{ to } L \\ &= \lambda \times [\text{Var}(n) + 2\,\text{Cov}(L,n)], \end{aligned} \tag{4.1}$$

where λ is a multiplier similar to $yz/(1+y)$ from the MS method although dimensioned to apply to variance rather than standard deviation.[6]

5. BUILDING UP A PORTFOLIO OF TWO ACCOUNTS

Exhibit 1 shows the occurrence size-of-loss distribution and risk load calculations for building up (assembling) a portfolio of two accounts, X and Y. It is assumed X is written first and is the only risk in the portfolio until Y is written.

5.1. MS Method

Pertinent values from Exhibit 1 for the Marginal Surplus method are summarized in Table 1.

Item 1 is the change in portfolio standard deviation from adding each account, or *marginal* standard deviation.

---

[6]Mr. Meyers develops a variance-based risk load multiplier by converting a standard deviation-based multiplier using the following formula [6, p. 573]: λ = (Rate of Return $\times$ Std Dev Mult^2)/(2 $\times$ Avg Capital of Competitors).

TABLE 1

BUILDING UP X AND Y: MARGINAL SURPLUS METHOD

	Account X	Account Y	Account X +Account Y	Portfolio $(X+Y)$
(1) Change in Standard Deviation	\$4,429.00	\$356.00	\$4,785.00	\$4,785.00
(2) Risk Load Multiplier	0.33	0.33	—	0.33
(3) Risk Load = (1) × (2)	\$1,461.71	\$117.43	\$1,579.14	\$1,579.14

Item 2 is the Risk Load multiplier of 0.33. Using Kreps's formula, a return on marginal surplus y of 20% and a standard normal multiplier z of 2.0 (2 standard deviations, corresponding to a cumulative non-exceedance probability of 97.725%) would produce a risk load multiplier of

$$yz/(1+y) = 0.20 \times 2/1.20 = 0.33 \text{ (rounded).} \qquad (5.1)$$

Item 3 is the Risk Load, the product of Items 1 and 2.

Since X is the first account, the marginal standard deviation from adding X equals the standard deviation of X, Std Dev (X) = \$4,429. This gives a risk load of \$1,461.71.

The marginal standard deviation from writing Y equals Std Dev $(X+Y)$ − Std Dev (X) or \$356, implying a risk load of \$117.43.

The sum of these two risk loads $X+Y$ is \$1,461.71 + \$117.43 = \$1,579.14. This equals the risk load that this method would calculate for the combined account $(X+Y)$.

5.2. MV Method

Pertinent values from Exhibit 1 for the Marginal Variance method are summarized in Table 2.

TABLE 2

BUILDING UP X AND Y: MARGINAL VARIANCE METHOD

	Account X	Account Y	Account X +Account Y	Portfolio $(X+Y)$
(1) Change in Variance	19,619,900	3,279,059	22,898,959	22,898,959
(2) Risk Load Multiplier	0.000069	0.000069	—	0.000069
(3) Risk Load = (1) × (2)	\$1,353.02	\$226.13	\$1,579.14	\$1,579.14

Item 1 is the change in portfolio variance from adding each account, or *marginal* variance.

Item 2 is the Variance Risk Load multiplier λ of 0.000069. To simplify comparisons between the two methods (recognizing the difficulty of selecting a MV-based multiplier[7]), the MS multiplier was converted to a MV basis by dividing by Std Dev $(X+Y)$:

$$\lambda = 0.33/4{,}785 = 0.000069. \tag{5.2}$$

This means the total risk load calculated for the portfolio by the two methods will be the same, although the individual risk loads for X and Y will differ between the methods.

Item 3 is the Risk Load, the product of Items 1 and 2.

Since X is the first account, the marginal variance from adding X equals the variance of X, Var(X) = \$19,619,900. This gives a risk load of \$1,353.02.

The marginal variance from writing Y equals Var($X+Y$) – Var(X), or \$3,279,059, implying a risk load of \$226.13.

The sum of these two risk loads is \$1,353.02 + \$226.13 = \$1,579.14. This equals the risk load which this method would calculate for the combined account $(X+Y)$.

[7] Mr. Meyers [6, p. 572] admits that in practice "it might be difficult for an insurer to obtain the (lambdas) of each of its competitors." He goes on to suggest an approximate method to arrive at a usable lambda based on required capital being "Z standard deviations of the total loss distribution" [6, p. 574].

TABLE 3

RENEWING X AND Y: MARGINAL SURPLUS METHOD

	Account X	Account Y	Account X +Account Y	Portfolio ($X + Y$)
(1) Change in Standard Deviation	\$4,171.00	\$356.00	\$4,526.00	\$4,785.00
(2) Risk Load Multiplier	0.33	0.33	—	0.33
(3) Risk Load = (1) × (2)	\$1,376.27	\$117.43	\$1,493.70	\$1,579.14
(4) Build-Up Risk Load	\$1,461.71	\$117.43	\$1,579.14	\$1,579.14
(5) Difference	(\$85.45)	\$0.00	(\$85.45)	\$0.00

6. RENEWING THE PORTFOLIO OF TWO ACCOUNTS

Exhibit 2 shows the natural extension of the build-up scenario—renewal of the two accounts, in what could be termed a "static" or "steady state" portfolio (one with no new entrants).

As for applying these methods in the renewal scenario, renewing policy X is assumed equivalent to adding X to a portfolio of Y; renewing Y is assumed equivalent to adding Y to a portfolio of X.

6.1. MS Method

Pertinent values from Exhibit 2 for the Marginal Surplus method are summarized in Table 3.

The marginal standard deviation for adding Y to X is \$356.00, same as it was during build-up—see Section 5.1. The risk load of \$117.43 is also the same.

However, adding X to Y gives a marginal standard deviation of Std Dev $(X + Y)$ – Std Dev (Y) = \$4,171.00. This gives a risk load for X of \$1,376.27, which is \$85.45 less than \$1,461.71, the risk load for X calculated in Section 5.1.

TABLE 4

RENEWING X AND Y: MARGINAL VARIANCE METHOD

	Account X	Account Y	Account X +Account Y	Portfolio $(X+Y)$
(1) Change in Variance	22,521,000	3,279,059	25,800,059	22,898,959
(2) Risk Load Multiplier	0.000069	0.000069	—	0.000069
(3) Risk Load = (1) × (2)	\$1,553.08	\$226.13	\$1,779.21	\$1,579.14
(4) Build-Up Risk Load	\$1,353.02	\$226.13	\$1,579.14	\$1,579.14
(5) Difference	\$200.06	\$0.00	\$200.06	\$0.00

The sum of these two risk loads in Table 3 is \$1,376.27 + \$117.43 = \$1,493.70. This is also \$85.45 less than the total risk load from Section 5.1.

6.2. MV Method

Pertinent values from Exhibit 2 for the Marginal Variance method are summarized in Table 4.

The marginal variance for adding Y to X is 3,279,059, same as it was during build-up—see Section 5.2. The risk load of \$226.13 is also the same.

However, adding X to Y gives a marginal variance of $\mathrm{Var}(X+Y) - \mathrm{Var}(Y)$, or 22,521,000. The risk load is now \$1,553.08, which is \$200.06 more than the \$1,353.02 calculated in Section 5.2.

The sum of these two risk loads is \$1,553.08 + \$226.13 = \$1,779.21. This is also \$200.06 more than the total risk load from Section 5.2.

7. EXPLORING THE DIFFERENCES BETWEEN NEW AND RENEWAL

Why are the total Renewal risk loads different from the total Build-Up risk loads?

In Section 5.1 (Build-Up), the marginal standard deviation for X, $\Delta\,\text{Std Dev}(X)$, was

$$\Delta\,\text{Std Dev}(X) = \text{Std Dev}\ (X) = \sqrt{\Sigma_i[X_i^2 \times p_i \times (1-p_i)]}, \qquad (7.1)$$

where X_i = modeled losses for X for event i, while in Section 6.1 (Renewal), the marginal standard deviation was

$$\begin{aligned}\Delta\,\text{Std Dev}(X) &= \text{Std Dev}(X+Y) - \text{Std Dev}(Y)\\ &= \sqrt{\Sigma_i[(X_i+Y_i)^2 \times p_i \times (1-p_i)]}\\ &\quad - \sqrt{\Sigma_i[Y_i^2 \times p_i \times (1-p_i)]}. \qquad (7.2)\end{aligned}$$

For positive Y_i, this value is less than Std Dev (X). Therefore, one *would* expect the Renewal risk load to be less than the Build-Up.[8]

Unfortunately, when the MS method is applied in the renewal of all the accounts in a portfolio, the sum of the individual risk loads will be less than the total portfolio standard deviation times the multiplier. This is because the sum of the marginal standard deviations (found by taking the difference in portfolio standard deviation with and without each account in the portfolio) is less than the total portfolio standard deviation.[9] Please recall that the square root operator is *sub-additive*: the square root of a sum is less than the sum of the square roots.[10]

---

[8]For example, assume $\text{Var}(X) = 9$, $\text{Var}(Y) = 4$, $\text{Cov}(X,Y) = 1.5$; then

$$\Delta\,\text{Std Dev}(X) = \sqrt{\text{Var}(X)} = \sqrt{9} = 3, \qquad \text{for}\quad X \text{ alone},$$

$$\Delta\,\text{Std Dev}(X) = \sqrt{9+4+2\times 1.5} - \sqrt{4} = 4-2 = 2 < 3, \qquad \text{for}\quad X \text{ added to } Y.$$

[9]The same issue is raised in Mr. Gogol's discussion. He suggests correcting for this sub-additivity by using a weighted average of the contract's own standard deviation and its last-in marginal standard deviation. The weight is chosen so the sum of these redefined marginal impacts equals the total portfolio standard deviation [2, p. 363].

[10]For example, $\sqrt{9+16} < \sqrt{9} + \sqrt{16}$.

What about marginal variance? In Section 5.2 (Build-Up), the marginal variance $\Delta\,\mathrm{Var}(X)$ was

$$\begin{aligned}\Delta\,\mathrm{Var}(X) &= \mathrm{Var}(X)\\ &= \Sigma_i[X_i^2 \times p_i \times (1 - p_i)],\end{aligned} \tag{7.3}$$

while in Section 6.2 (Renewal) the marginal variance was

$$\begin{aligned}\Delta\,\mathrm{Var}(X) &= \mathrm{Var}(X+Y) - \mathrm{Var}(Y)\\ &= [\mathrm{Var}(X) + 2\,\mathrm{Cov}(X,Y) + \mathrm{Var}(Y)] - \mathrm{Var}(Y)\\ &= \mathrm{Var}(X) + 2\,\mathrm{Cov}(X,Y)\\ &> \mathrm{Var}(X).\end{aligned} \tag{7.4}$$

Since $2\,\mathrm{Cov}(X,Y)$ is greater than zero, one would expect the Renewal risk load to be greater than the Build-Up.

However, when the MV method is applied in the renewal of all the accounts in a portfolio, the sum of the individual risk loads will be more than the total portfolio variance times the multiplier. This is because the sum of the marginal variances (found by taking the difference in portfolio variance with and without each account in the portfolio) is greater than the total portfolio variance. The covariance between any two risks in the portfolio is double counted: when each account renews, it is allocated the full amount of its shared covariance with all the other accounts.

8. A NEW CONCEPT: RENEWAL ADDITIVITY

The renewal scenarios point out that these two methods are not what I call *renewal additive*, defined as follows:

> For a given portfolio of accounts, a risk load method is *renewal additive* if the sum of the renewal risk loads calculated for each account equals the risk load calculated when the entire portfolio is treated as a single account.

Neither the MS nor the MV method is renewal additive: MS because the square root operator is sub-additive; MV because the covariance is double counted. So why should renewal additivity matter? Consider what happens when either of these non-renewal -additive methods is used to renew the portfolio. The MV method would result in quoted renewal premiums the sum of whose risk loads would be greater than the required total risk load of ($\lambda \times$ total portfolio variance). One would in essence overcharge every account. The opposite is true for the MS case, where one would undercharge every account.

In order for the MS or MV methods to be renewal additive, one must assume an entry order for the accounts. Since the marginal impacts depend on the size of the existing portfolio, the entry order selected for an account could determine whether it is written or declined.

Renewal additivity reduces the renewal risk load calculation to an allocation of the total portfolio amount back to the individual accounts. An objective, systematic allocation methodology for renewals would be desirable. Examples of many such allocation methodologies can be found in the field of game theory.

9. A NEW APPROACH FROM GAME THEORY

Two ASTIN papers by Jean Lemaire—"An Application of Game Theory: Cost Allocation" [4], and "Cooperative Game Theory and Its Insurance Applications" [5]—focus on general insurance applications of game theory. Lemaire also provides an extensive list of real world applications of game theory [4, p. 77], including tax allocation among operating divisions of McDonnell-Douglas, maintenance costs of the Houston medical library, financing of large water resource development projects in Tennessee, construction costs of multi-purpose reservoirs in the U.S., and landing fees at Birmingham airport. Consider this example from [5]:

> "The Treasurer of ASTIN (player 1) wishes to invest the amount of 1,800,000 Belgian francs on a short term

(3 months) basis. In Belgium, the annual interest rate is a function of the sum invested.

Deposit (in Belgian Francs)	Annual Interest Rate
0–1,000,000	7.75%
1,000,000–3,000,000	10.25%
3,000,000–5,000,000	12.00%

The ASTIN Treasurer contacts the Treasurers of the International Actuarial Association (I.A.A.–player 2) and of the Brussels Association of Actuaries (A.A.Br.–player 3). I.A.A. agrees to deposit 900,000 francs in the common fund, A.A.Br. 300,000 francs. Hence the 3-million mark is reached and the interest rate will be 12%. How should the interests be split among the three associations?" [5, p. 18]

Games such as this are referred to as "cooperative games with transferable utilities." They typically feature:

1. participants (players) that have some benefits (or costs) to share (political power, savings, or money),
2. the opportunity to share benefits (costs) results from cooperation of all participants or a sub-group of participants,
3. freedom for players to engage in negotiations, bargaining, and coalition formation, and
4. conflicting player objectives: each wants to secure the largest part of the benefits (smallest share of the costs) for himself. (See [5, p. 20].)

Cooperative games can be used as models for situations where participants must share or allocate an amount of money. Players may want to maximize or minimize their allocation depending on the nature of the problem. If the group is deciding who pays what share of the total tax bill, players will want to minimize their share. If on the other hand the group is deciding how to split a pot of bonus money, players will want to maximize their share.

The total amount to be allocated is determined by the *characteristic function*, which associates a real number $v(S)$ to each coalition (group) S of players. It can be either sub-additive or super-additive, defined as follows:

Sub-Additive

$$v(S) + v(T) > v(S \text{ union } T) \quad \text{for} \quad \text{every disjoint } S \text{ and } T.$$

Super-Additive

$$v(S) + v(T) < v(S \text{ union } T) \quad \text{for} \quad \text{every disjoint } S \text{ and } T.$$

In the actuarial association example above, the characteristic function would be the money earned by each coalition (combination) of associations. It is an example of a super-additive characteristic function where the players seek to maximize their allocation. An example of a sub-additive characteristic function would be the insurance premium for a risk purchasing group: the sum of the individual members' insurance premiums is more than the insurance premium for the risk purchasing group as a whole. These players would seek to minimize their allocations, since they want to be charged the lowest premium. (Equivalently, these players want to maximize their savings as a result of joining the group—savings being the difference between their allocation from the group and their stand-alone premium.)

Allocation Rules

A player's marginal impact depends on its entry order. In the example, the "allocation [to the three associations] of course depends on the order of formation of the grand coalition" [5, p. 27]. In the interests of fairness and stability, a new member should probably receive an allocation amount somewhere between its stand-alone value and its full marginal impact on the coalition characteristic function—but where in between? How much is fair? These questions must be answered simultaneously for all the players, balancing questions of stability, incentives to split from the group, bargaining power, and marginal impact to the coalition characteristic function value.

To help answer the allocation question, game theory has developed a set of standards or rules for allocations. First, legitimate allocation methods must be additive—the sum of the players' allocations must equal the total amount to be allocated. The MV and MS methods are not (renewal) additive: they either allocate too much (MV) or too little (MS) in the renewal case.

Second, a coalition should be stable, which roughly translates to fair. There must not be incentives for either a single player or a sub-group of players to split from the group and form a faction. These "rules of fairness" are referred to as the conditions of individual and collective rationality (see [4, p. 66–68]). Individual rationality means a player is no worse off for having joined the coalition. Collective rationality means no subgroup would be better off on its own.

These rules can be formalized into a set of acceptable ranges of allocations for each player. This set defines what is known as the *core* of the game. It consists of all allocations satisfying these fairness and stability conditions.

Consider the Brussels Association of Actuaries (A.A.Br.–player 3) from the example. They have 300,000 francs, and on their own could earn 7.75%. If they join as the third player, they will push the coalition rate of return from 10.25% to 12.00%. How much should they earn? Certainly not less than 7.75%—it would not be individually rational for them to join. Conversely, they should not earn so much that players 1 and 2 end up earning less than 10.25%—that would not be collectively rational for them. In that case, players 1 and 2 would be better off forming their own faction. Similar exercises can be performed from the perspective of players 1 and 2. The resulting set of acceptable allocations defines the boundaries of the core (see [5, p. 26]).

TABLE 5

TRANSLATION FROM GAME THEORY TO PROPERTY CAT RISK LOAD

Game Theory	Property Cat Risk Load
Player	Account
Coalition	Portfolio
Characteristic Function	Portfolio Variance or Standard Deviation

Translating to Property Cat Risk Load

Given this brief introduction, a reasonable first attempt at translating from the game theory context might be as shown in Table 5.

Because of the covariance component, portfolio variance is a super-additive characteristic function: the variance of a portfolio is greater than the sum of the individual account variances. Standard deviation, on the other hand, is a sub-additive characteristic function because of the sub-additivity of the square root operator: the standard deviation of a portfolio is less than the sum of the individual account standard deviations.

This means, from the game theory perspective at least, that the choice between variance and standard deviation is material. It determines whether the characteristic function is sub-additive or super-additive. This is a fundamental paradox of the game theory translation of the risk load problem, and will require further research to resolve.

Setting aside this paradox for the moment, however, the risk load problem fits remarkably well into the game theory framework. The "players" want to minimize their allocations of the portfolio total risk load. The allocation should fairly and objectively assign risk loads to accounts in proportion to their contribution to the total. Using the current definition of marginal impact of a renewal account, however, an entry order would have to be assumed in order to make the allocation additive. The results of that allocation would be heavily dependent on the selected order, however.

How can one choose the entry order of a renewal? A well known allocation method from game theory may provide the answer.

The Shapley Value

The Shapley value (named for Lloyd Shapley, one of the early leaders of the game theory field) is an allocation method that is:

1. additive,
2. at the centroid of the core, and
3. order independent.

It equals the average of the marginal impacts taken over all possible entrance permutations—the different orders in which a new member could have been added to the coalition[11] (i.e., a new account could have been added to a portfolio).

For example, consider a portfolio of accounts A and B to which a new account C is added. Shown in Table 6 are the marginal variances for adding C in the 6 possible entrance permutations ("ABC" in Column 1 below means A enters first, then B, then C).

[11] Lemaire provides this more complete definition of the Shapley value [5, p. 29]: "The Shapley value can be *interpreted* as the mathematical expectation of the admission value, when all orders of formation of the grand coalition are equiprobable. In computing the value, one can assume, for convenience, that all players enter the grand coalition one by one, each of them receiving the entire benefits he brings to the coalition formed just before him. All orders of formation of N are considered and intervene with the same weight $1/n!$ in the computation. The combinatorial coefficient results from the fact that there are $(s-1)!(n-s)!$ ways for a player to be the last to enter coalition S: the $(s-1)$ other players of S and the $(n-s)$ players of $N \backslash S$ (those players in N which are not in S) can be permuted without affecting i's position."

TABLE 6

ENTRY PERMUTATIONS FOR ACCOUNT C

(1) Permutation	(2) C Enters ...	(3) Marginal Variance
ABC	After A & B	$\mathrm{Var}(C) + 2 \times \mathrm{Cov}(C,A) + 2 \times \mathrm{Cov}(C,B)$
ACB	After A	$\mathrm{Var}(C) + 2 \times \mathrm{Cov}(C,A)$
BAC	After B & A	$\mathrm{Var}(C) + 2 \times \mathrm{Cov}(C,A) + 2 \times \mathrm{Cov}(C,B)$
BCA	After B	$\mathrm{Var}(C) + 2 \times \mathrm{Cov}(C,B)$
CAB	First	$\mathrm{Var}(C)$
CBA	First	$\mathrm{Var}(C)$

The Shapley value is the straight average of Column 3, Marginal Variance, over the six permutations:

$$\begin{aligned}\text{Shapley Value} &= [\text{Sum(Column 3)}]/6\\ &= [6\,\mathrm{Var}(C) + 6\,\mathrm{Cov}(C,A) + 6\,\mathrm{Cov}(C,B)]/6\\ &= \mathrm{Var}(C) + \mathrm{Cov}(C,A) + \mathrm{Cov}(C,B).\end{aligned} \tag{9.1}$$

Or, to generalize, given

$$\begin{aligned}L &= \text{existing portfolio} \qquad \text{and}\\ n &= \text{new account},\end{aligned} \tag{9.2}$$

$$\text{Shapley Value} = \mathrm{Var}(n) + \mathrm{Cov}(L,n).$$

Before seeing this result, there might have been concerns about the practicality of this approach—how much computational time might be required to calculate all the possible entrance permutations for a portfolio of thousands of accounts? This simple reduction formula eliminates those concerns. The Shapley value is as simple to calculate as the marginal variance.

Comparing the Shapley value to the marginal variance formula from Section 4:

$$\text{Marginal Variance} = \mathrm{Var}(n) + 2\,\mathrm{Cov}(L,n), \tag{9.3}$$

whereas the Shapley value only takes 1 times the covariance of the new account and the existing portfolio.

One can also calculate the Shapley value under the marginal standard deviation method. However, due to the complex nature of the mathematics—differences of square roots of sums of products—no simplifying reduction formula was immediately apparent.[12]

Therefore, the remainder of the paper will focus on the MV method and the variance-based Shapley value. Life will be much easier (mathematically) working with the variances, and very little is lost by choosing variance. Citing Mr. Bault [1, p. 82], from a risk load perspective, "both (variance and standard deviation) are simply special cases of a unifying covariance framework." In fact, Bault goes on to suggest "in most cases, the 'correct' answer is a marginal risk approach that incorporates covariance."[13]

10. SHARING THE COVARIANCE

The risk load question, framed in a game-theoretical light, has now become:

> How do accounts share their mutual covariance for purposes of calculating risk load?

The Shapley method answers, "Accounts split their mutual covariance equally." At first glance this appears reasonable, but consider the following example.

Assume two accounts, E and F. F has 100 times the losses of E for each event. Their total shared covariance is

$$\begin{aligned} 2\mathrm{Cov}(E,F) &= 2\Sigma_i[E_i \times F_i \times p_i \times (1-p_i)] \\ &= 2\Sigma_i[E_i \times 100E_i \times p_i \times (1-p_i)]. \qquad (10.1) \end{aligned}$$

[12] Those wishing to employ standard deviation can use approximate methods to calculate the Shapley value. Two approaches suggested by John Major are (i) taking the average of marginal value if first in and last in; and (ii) employing Monte Carlo simulation to sample a subset of the possible entrance permutations, presumably large enough to achieve satisfactory convergence while being much more computationally efficient.

[13] Kreps also incorporates covariance in his "Reluctance" R [3, p. 198], which has the formula $R = [yz/(1+y)]/(2SC+\sigma)/(S'+S)$, where C is the correlation of the contract with the existing book. The Risk Load is then equal to $R\sigma$.

The Shapley value would equally divide this total covariance between E and F, even though their relative contributions to the total are clearly not equal. There is no question that E should be assessed some share of the covariance. The issue is whether there is a more equitable share than simply half.

One could develop a generalized covariance sharing (GCS) method which uses a weight $W_i^X(X,Y)$ to determine an account X's share of the mutual covariance between itself and another account Y for event i:

$$\text{CovShare}_i^X(X,Y) = W_i^X(X,Y) \times 2 \times X_i \times Y_i \times p_i \times (1 - p_i). \tag{10.2}$$

Then Y's share of that mutual covariance would simply be

$$\text{CovShare}_i^Y(X,Y) = [1 - W_i^X(X,Y)] \times 2 \times X_i \times Y_i \times p_i \times (1 - p_i). \tag{10.3}$$

The total covariance share allocation for account X over all events would be:

$$\text{CovShare}_{\text{Tot}}^X = \Sigma_Y \Sigma_i [\text{CovShare}_i^X(X,Y)], \tag{10.4}$$

where Σ_Y = sum over every other account Y in the portfolio.

The Shapley method is in fact an example of the generalized covariance sharing method with $W_i^X(X,Y) = 50\%$ for all X, Y and i.

Returning to the example with E and F, one could develop an example of a weighting scheme that assigns the shared covariance by event to each account in proportion to their loss for that event. $W_i^E(E,F)$, account E's share of the mutual covariance between itself and account F for event i, equals

$$\begin{aligned} W_i^E(E,F) &= [E_i/(E_i + F_i)] \\ &= [E_i/(E_i + 100E_i)] = (1/101) \\ &= \text{roughly } 1\% \text{ of their mutual covariance for event } i. \end{aligned}$$

This shall be called the "Covariance Share" (CS) method.

TABLE 7

BUILDING UP X AND Y: SHAPLEY VALUE METHOD

	Account X	Account Y	Account X +Account Y	Portfolio ($X+Y$)
(1) Change in Variance	19,619,900	1,828,509	21,448,409	22,898,959
(2) Risk Load Multiplier	0.000069	0.000069	—	0.000069
(3) Risk Load = (1) × (2)	\$1,353.02	\$126.10	\$1,479.11	\$1,579.14

TABLE 8

BUILDING UP X AND Y: COVARIANCE SHARE METHOD

	Account X	Account Y	Account X +Account Y	Portfolio ($X+Y$)
(1) Change in Variance	19,619,000	950,658	20,570,558	22,898,959
(2) Risk Load Multiplier	0.000069	0.000069	—	0.000069
(3) Risk Load = (1) × (2)	\$1,353.02	\$65.56	\$1,418.57	\$1,579.14

11. APPLYING THE SHAPLEY AND CS METHODS TO THE EXAMPLE

Consider the Shapley and CS methods applied to the two Account example for both Build-Up and Renewal.

11.1. Portfolio Build-up

Exhibit 3 shows the Build-Up of accounts X and Y from Section 5, but for the Shapley and CS methods. Pertinent values for the Shapley value are summarized in Table 7.

Pertinent values for the Covariance Share are summarized in Table 8.

Both Shapley and CS produce the same risk load for X as the MV method on Build-Up: \$1,353.02. This is because there is no covariance to share: X is the entire portfolio at this point. How-

TABLE 9

COMPARISON OF BUILD-UP RISK LOADS FOR ACCOUNT Y

Marginal Variance (MV)—Section 5.2	$226.13
Shapley Value	$126.10
Difference from MV	**$100.03**
Covariance Share (CS)	$ 65.56
Difference from MV	**$160.57**

TABLE 10

RENEWING X AND Y: SHAPLEY VALUE METHOD

	Account X	Account Y	Account X +Account Y	Portfolio $(X+Y)$
(1) Change in Variance	21,070,450	1,828,509	22,898,959	22,898,959
(2) Risk Load Multiplier	0.000069	0.000069	—	0.000069
(3) Risk Load = (1) × (2)	$1,453.05	$126.10	$1,579.14	$1,579.14
(4) Build-Up Risk Load	$1,353.02	$126.10	$1,479.11	$1,579.14
(5) Difference	$100.03	$0	$100.03	$0

ever, compare the results of the three variance-based methods for account Y (see Table 9).

Compared to MV, which charges account Y for the full increase in variance $\text{Var}(Y) + 2\text{Cov}(X,Y)$, the Shapley method only charges Y for $\text{Var}(Y) + \text{Cov}(X,Y)$. The same can be said for the CS method, although the share of the mutual covariance depends on each account's relative contribution by event, weighted and summed over all events. Now consider what happens to that *difference from MV* upon renewal.

11.2. Renewal

Exhibit 4 shows the renewal of X and Y for the Shapley and CS methods. Pertinent values for the Shapley method are summarized in Table 10.

TABLE 11

RENEWING X AND Y: COVARIANCE SHARE METHOD

	Account X	Account Y	Account X +Account Y	Portfolio $(X+Y)$
(1) Change in Variance	21,948,301	950,658	22,898,959	22,898,959
(2) Risk Load Multiplier	0.000069	0.000069	—	0.000069
(3) Risk Load = (1) × (2)	\$1,513.59	\$65.56	\$1,579.14	\$1,579.14
(4) Build-Up Risk Load	\$1,353.02	\$65.56	\$1,418.57	\$1,579.14
(5) Difference	\$160.57	\$0	\$160.57	\$0

TABLE 12

COMPARISON OF BUILD-UP AND RENEWAL RISK LOADS FOR ACCOUNT X

	Shapley	**Cov Share**
Renewal	\$1,453.05	\$1,513.59
Build-Up	\$1,353.02	\$1,353.02
Additional Renewal Risk Load over Build-Up	**\$100.03**	**\$160.57**
Difference from MV	**\$100.03**	**\$160.57**

Pertinent values for the Covariance Share method are summarized in Table 11.

With both the Shapley and CS methods, the sum of the risk loads for Account X and Account Y equals the risk load for Account $(X+Y)$, namely \$1,579.14. This means that both new methods are renewal additive.

To see what happened to the *difference from MV*, compare the risk loads calculated at Renewal for X with those at Build-Up (see Table 12).

The difference from MV during Build-Up is simply the portion of X's risk load attributable to its share of covariance with Y. It was missed during Build-Up because it was unknown—account Y had not been written.

12. CONCLUSION

This paper introduces two new approaches to determination of renewal risk load that address concerns with renewal additivity and point out the issue of covariance sharing between accounts. The ideal solution in practice might involve using a marginal method for the pricing of new accounts, and a renewal additive method for renewals.

This paper also represents a first step in addressing the perplexing question of order dependency. As mentioned in the introduction, order dependency stretches beyond the confines of actuarial pricing to the finance community at large. It will likely take a joint effort between finance professionals and actuaries to reach a satisfactory solution.

Finally, this paper brings important information from game theory to the *Proceedings*. Game theory is a rich field for actuaries to find new ideas on cost allocation, fairness, and order dependency. Many sticky social issues (taxation, voting rights, utility costs) have been resolved using ideas from game theory. Further research can be done on several questions raised during the review of this paper, including the relative bargaining power of accounts, portfolio departure rules, lack of account information, and the unresolved paradox of the sub-additive MS characteristic function versus the super-additive MV characteristic function.

REFERENCES

[1] Bault, Todd, Discussion of Feldblum: "Risk Loads for Insurers," *PCAS* LXXXII, 1995, pp. 78–96.

[2] Gogol, Daniel F., Discussion of Kreps: "Reinsurer Risk Loads from Marginal Surplus Requirements," *PCAS* LXXIX, 1992, pp. 362–366.

[3] Kreps, Rodney, "Reinsurer Risk Loads from Marginal Surplus Requirements," *PCAS* LXXVII, 1990, pp. 196–203.

[4] Lemaire, Jean, "An Application of Game Theory: Cost Allocation," *ASTIN Bulletin* 14, 1, 1984, pp. 61–81.

[5] Lemaire, Jean, "Cooperative Game Theory and its Insurance Applications," *ASTIN Bulletin* 21, 1, 1991, pp. 17–40.

[6] Meyers, Glenn, "The Competitive Market Equilibrium Risk Load Formula for Catastrophe Ratemaking," *PCAS* LXXXIII, 1996, pp. 563–600.

EXHIBIT 1

BUILD UP PORTFOLIO OF TWO ACCOUNTS

			Loss for Account		
Event *i*	**Prob *p*(*i*)**	**1-*p*(*i*)**	***X***	***Y***	**Portfolio (*X* + *Y*)**
1	2.0%	98.0%	25,000	200	25,200
2	1.0%	99.0%	15,000	500	15,500
3	3.0%	97.0%	10,000	3,000	13,000
4	3.0%	97.0%	8,000	1,000	9,000
5	1.0%	99.0%	5,000	2,000	7,000
6	2.0%	98.0%	2,500	1,500	4,000

E (*L*)	1,290	179	1,469
Var (*L*)	**19,619,900**	**377,959**	**22,898,959**
Std Dev (*L*)	4,429	615	4,785

Covariances	***X***	***Y***
X	**19,619,900**	1,450,550
Y	1,450,550	**377,959**

		X	***Y***	***X* + *Y***
Change in Std Deviation		4,429	356	4,785
Risk Load (Std Dev)		1,461.71	117.43	1,579.14
Multiplier :	**0.33**	**Risk Load for *X* + *Y* :**		1,579.14

Change in Variance		19,619,900	**3,279,059**	22,898,959
Risk Load (Variance)		1,353.02	226.13	1,579.14
Multiplier :	**0.000069**	**Risk Load for *X* + *Y* :**		1,579.14

EXHIBIT 2

RENEW THE PORTFOLIO OF TWO ACCOUNTS

			Loss for Account		
Event *i*	**Prob *p*(*i*)**	**1-*p*(*i*)**	***X***	***Y***	**Portfolio (*X* + *Y*)**
1	2.0%	98.0%	25,000	200	25,200
2	1.0%	99.0%	15,000	500	15,500
3	3.0%	97.0%	10,000	3,000	13,000
4	3.0%	97.0%	8,000	1,000	9,000
5	1.0%	99.0%	5,000	2,000	7,000
6	2.0%	98.0%	2,500	1,500	4,000

E (*L*)	1,290	179	1,469
Var (*L*)	**19,619,900**	**377,959**	**22,898,959**
Std Dev (*L*)	4,429	615	4,785

Covariances	***X***	***Y***
X	**19,619,900**	1,450,550
Y	1,450,550	**377,959**

		X	***Y***	***X* + *Y***
Change in Std Deviation		4,171	356	4,526
Risk Load (Std Dev)		**1,376.27**	117.43	1,493.70
0.33	*Build Up Risk Load*	1,461.71	117.43	1,579.14
	Difference	**(85.45)**		**(85.45)**

Change in Variance		22,521,000	3,279,059	25,800,059
Risk Load (Variance)		**1,553.08**	226.13	1,779.21
0.000069	*Build Up Risk Load*	1,353.02	226.13	1,579.14
	Difference	**200.06**		**200.06**

EXHIBIT 3

BUILD UP A PORTFOLIO OF TWO ACCOUNTS—ALTERNATIVES

			Covariance Share $	
Event *i*	**Prob *p*(*i*)**	**1-*p*(*i*)**	***X***	***Y***
1	2.0%	98.0%	9,920,635	79,365
2	1.0%	99.0%	14,516,129	483,871
3	3.0%	97.0%	46,153,846	13,846,154
4	3.0%	97.0%	14,222,222	1,777,778
5	1.0%	99.0%	14,285,714	5,714,286
6	2.0%	98.0%	4,687,500	2,812,500

		Total
2,328,401	572,699	2,901,100

Chg in Variance		***X***	***Y***
If added 1st		19,619,900	377,959
If added 2nd	after 1		3,279,059
	after 2	22,521,000	
Average	*(Shapley Value)*	21,070,450	1,828,509

Shapley Value	19,619,900	1,828,509	21,448,409
Risk Load (Shapley)	**1,353.02**	**126.10**	**1,479.11**
0.000069			**1,579.14**
	Deferred Risk Load		**100.03**

Covariance Share	19,619,900	950,658	20,570,558
Risk Load (Cov Share)	**1,353.02**	**65.56**	**1,418.57**
0.000069			**1,579.14**
	Deferred Risk Load		**160.57**

EXHIBIT 4

RENEW THE PORTFOLIO OF TWO ACCOUNTS—ALTERNATIVES

			Covariance Share $	
Event *i*	**Prob *p*(*i*)**	**1-*p*(*i*)**	***X***	***Y***
1	2.0%	98.0%	9,920,635	79,365
2	1.0%	99.0%	14,516,129	483,871
3	3.0%	97.0%	46,153,846	13,846,154
4	3.0%	97.0%	14,222,222	1,777,778
5	1.0%	99.0%	14,285,714	5,714,286
6	2.0%	98.0%	4,687,500	2,812,500

		Total
2,328,401	572,699	2,901,100

Chg in Variance		***X***	***Y***
If added 1st		19,619,900	377,959
If added 2nd	after 1		3,279,059
	after 2	22,521,000	
Average	*(Shapley Value)*	21,070,450	1,828,509

Shapley Value	21,070,450	1,828,509	22,898,959
Risk Load (Shapley)	1,453.05	126.10	**1,579.14**
0.000069	**Risk Load for Portfolio (*X* + *Y*)**		**1,579.14**

Covariance Share	21,948,301	950,658	22,898,959
Risk Load (Cov Share)	1,513.59	65.56	**1,579.14**
0.000069	**Risk Load for Portfolio (*X* + *Y*)**		**1,579.14**

A BUYER'S GUIDE FOR OPTIONS ON A CATASTROPHE INDEX

GLENN MEYERS

Abstract

In the wake of recent catastrophes, a new way of transferring insurance risk was born. In December 1992, the Chicago Board of Trade began trading contracts on an index sensitive to insurer catastrophe experience. Such indices provide an insurer a means to transfer a portion of its catastrophe risk to the capital markets by buying future and option contracts.

The cost of using these contracts to transfer catastrophe risk is compared to the cost of raising sufficient capital to retain the risk and the cost of conventional reinsurance. We derive equations that give the optimal participation in the future and option contracts, and in reinsurance. The cost of using these contracts can be compared to the cost of the capital that they replace.

1. INTRODUCTION

In the wake of recent catastrophes, a new way of transferring insurance risk was born. In December 1992, the Chicago Board of Trade began trading contracts on an index sensitive to insurer catastrophe experience.

These contracts gave insurers an additional financial strategy for handling catastrophe risk. Two other common strategies are:

1. buying reinsurance; and
2. raising sufficient capital to maintain solvency while retaining the risk.

Another innovation that has gained popularity in the wake of recent catastrophes is the use of catastrophe models in insurance ratemaking and underwriting. These models combine meteorological and geological science with engineering damageability studies and insurance exposure information to estimate potential losses for an insurance portfolio.

The purpose of this paper is to show how to use catastrophe models to estimate the costs and benefits of contracts on a catastrophe index relative to other means of managing the catastrophe risk.

2. MOTIVATION FOR TRADING CONTRACTS ON A CATASTROPHE INDEX

Risk of loss is usually transferred from one with insufficient capital to absorb a loss to one(s) who can absorb it. The size of an insurance catastrophe, which at its worst is measured in billions, is small compared to the money invested in the capital markets, which is measured in trillions. There are insurers with a demand for risk transfer, and there are investors who can meet this demand. However, one needs to find a contract that meets the institutional needs of the investors and the insurers.

Investors are ill-equipped to deal with counterparty risk, i.e., the risk that the insurer knows something about the transfer that will be to the investors disadvantage. One way to reduce this risk is to base the contract on the combined results of several insurers, i.e., a catastrophe index.

Trading contracts on an index introduces additional risk for the insurer in that the money it recovers from a catastrophe index contract may differ substantially from its own catastrophe losses. In investment language, this is referred to as basis risk.[1]

---

[1]For a more complete explanation of basis risk, see Hull [5, p. 32].

The insurer would like its losses to be highly correlated with the index, as is the case for reinsurance,[2] so that its basis risk is small.

The investor seeks to maximize profit while adding the least amount of risk to its total investment portfolio. Usually the returns on available investments tend to be positively correlated over time. For example, the returns on stocks tend to be correlated with the general economy. If the value of the index is uncorrelated with the seller's other investments, the investor will take on less risk by selling contracts on the index than he would if he took on an otherwise equivalent investment on the stock market.

Both the insurer and the investor want their risk to be quantified. As this paper will illustrate, both risks can be quantified with the use of a catastrophe model and a tabulation of the underlying exposures.

3. A STATISTICAL DESCRIPTION OF THE CONTRACTS

This paper will focus on catastrophe index contracts as they are traded on an exchange such as the Chicago Board of Trade. The form of the contracts that are traded is explained below. The scale of the index is arbitrary. In this paper we set the scale so that the expected value of the index at expiration is \$1.00.

A **call option contract** gives the buyer the right to buy the index at an agreed upon price at a specified date. The agreed upon price is called the strike price.

As an example, suppose an investor sells a one year option contract with a strike price of \$1.00 for a premium of \$0.20 to an insurer. If there are no catastrophes during the year and the value of the index is zero on December 31, the insurer would

[2]The coefficient of correlation between losses and reinsurance recoveries will be 1.00 for quota share reinsurance agreements. If there is a reinsurance limit, the coefficient of correlation will be less than 1.00.

not want to buy the index for $1.00, so it would not exercise its option. The investor would keep the $0.20. However, if the index is valued at $3.00 on December 31, the insurer would buy the index for $1.00 and the investor would lose $1.80.

A **call option spread** is a package of two option contracts where one buys an option at one strike price and simultaneously sells another option at a higher strike price. The difference between the two strike prices is called the covered layer of the spread.

To continue our example, suppose the investor sells a call option spread to an insurer for the $1.00 to $2.00 layer for a net premium of $0.10. This means that the insurer is buying insurance on the index for the $1.00 to $2.00 layer for $0.10.

In terms of the transaction details, this means that the investor sells the insurer an option with a strike price of $1.00 for a premium of $0.20, and insurer sells the investor an option with a strike price of $2.00 for a premium of $0.10. If the final value of the index is zero, neither party exercises its option and the investor keeps its $0.10. If the final value of the index is $3.00, the investor exercises its option to buy the index from the insurer for $2.00 and the insurer exercises its option to buy the index from the investor for $1.00. The net effect is that the investor gives the insurer (and loses) $0.90. This is the most that the investor can lose on this contract.

If the final value of the index is $1.50, the insurer exercises its option and the investor does not. The investor pays the insurer $0.50 and ends up losing $0.40.

The purpose of the call option spread is to limit the liability of the seller, in much the same way that reinsurers limit their liability on catastrophe reinsurance contracts. If an insurer wants the full coverage, it can buy a series of call option spreads from different sellers, with the cost of the coverage being the sum of the premiums for the call option spreads.

When an insurer buys these contracts, it reduces the overall variability of its financial results and, at least in principle, it will need less capital to support its business.

We illustrate these points with a statistical argument. Let:

- X be a random variable for the insurer's losses prior to buying contracts on a catastrophe index;
- Y be a random variable for the final contract value;
- ρ be the coefficient of correlation between X and Y; and
- σ_Z be the standard deviation of any random variable, Z.

If an insurer buys n contracts on the index, the random variable for its net loss is $X - nY$, and a quantification of its risk is given by:

$$\sigma_{X-nY} = \sqrt{\sigma_X^2 - 2n\rho\sigma_X\sigma_Y + n^2\sigma_Y^2}. \qquad (3.1)$$

Note that the insurer will reduce its risk if $2\rho\sigma_X > n\sigma_Y$. There may be motivation to buy an options contract if ρ is positive and n is not too large. Exactly how many contracts will be bought depends upon the price. More on this below.

Let:

- U be a random variable for the investor's gain on its current portfolio;
- V be a random variable for the investor's gain on a prospective investment; and
- ν be the coefficient of correlation between U and V.

Further suppose that $\sigma_V^2 = n^2\sigma_Y^2$ and that U and Y are uncorrelated.

If the investor buys the prospective investment, a quantification of its risk is given by:

$$\sigma_{U+V} = \sqrt{\sigma_U^2 + 2\nu\sigma_U\sigma_V + \sigma_V^2}. \qquad (3.2)$$

If an investor sells n contracts on the index, the random variable for its net return is $U + nY$, and the equivalent quantification of its risk is given by:

$$\sigma_{U+nY} = \sqrt{\sigma_U^2 + n^2\sigma_Y^2}. \tag{3.3}$$

Since $\sigma_V^2 = n^2\sigma_Y^2$, the investor will face less risk by selling the catastrophe contracts when $\nu > 0$. Thus the investor should have a preference for selling the catastrophe contracts.[3] Again, it depends upon the price.

4. THE COST OF CAPITAL

The ultimate reason an insurer would want to purchase contracts on a catastrophe index is to reduce its cost of doing business. One of the key costs of the insurance business is the cost of capital. In this paper, we assume that the amount of capital needed for an insurer to adequately support the risks it writes is given by:

$$C = T\sigma_X. \tag{4.1}$$

Our choice of Equation 4.1 deserves some discussion since there is no universal agreement on a capitalization formula. For example, the NAIC risk based capital formula might be one possible alternative, but it does not recognize the catastrophe risk. Another alternative is the "expected policyholder deficit," which is the expected payment by the policyholders (or guaranty fund) in case the insurer goes insolvent (see AAA Report [1]). This formula is sensitive only to the tail of the loss distribution.

We offer the following two arguments in favor of Equation 4.1. First, we feel that most insurers are worried about losing even a small portion of their capital. Equation 4.1 is sensitive to the entire range of losses. Second, the mathematics needed to im-

[3]This is often called the "zero beta" argument. This is in reference to the Capital Asset Pricing Model. See Chapter 8 of Brealy and Myers [3].

plement this formula are relatively simple. However, many of the ideas in this paper can be implemented with other capitalization formulas.

Continuing, if the insurer buys n contracts on the catastrophe index, the needed capital becomes:

$$C(n) \equiv T\sigma_{X-nY} = T\sqrt{\sigma_X^2 - 2n\rho\sigma_X\sigma_Y + n^2\sigma_Y^2}. \qquad (4.2)$$

To obtain the reduction of capital indicated by the difference between Equations 4.1 and 4.2, the insurer must buy n contracts at a price determined by the market forces of supply and demand. Let P be equal to the price of a single contract less the expected return on the contract, i.e., the net cost of the contract. Then nP is the net cost of the contracts being substituted for capital.

Let K denote the rate of return the insurer pays to secure the needed capital. K will depend on the riskiness of the insurer's enterprise and the cost of competing investments.

When the insurer buys n contracts, its cost of capital plus its capital substitute is:

$$R(n) \equiv KT\sqrt{\sigma_X^2 - 2n\rho\sigma_X\sigma_Y + n^2\sigma_Y^2} + nP. \qquad (4.3)$$

To minimize its cost of providing insurance, the insurer will choose the value of n that minimizes $R(n)$. To determine this n, we find:

$$R'(n) = \frac{KT(n\sigma_Y^2 - \rho\sigma_X\sigma_Y)}{\sqrt{\sigma_X^2 - 2\rho n\sigma_X\sigma_Y + n^2\sigma_Y^2}} + P. \qquad (4.4)$$

Setting $R'(n) = 0$, and then solving for n yields:[4]

$$n = \frac{\rho\sigma_X}{\sigma_Y} - \frac{\sigma_X}{\sigma_Y}\sqrt{\frac{P^2(1-\rho^2)}{K^2T^2\sigma_Y^2 - P^2}}. \qquad (4.5)$$

[4]The details of this derivation are provided in the Appendix.

Here we see that the number of contracts needed to minimize the cost of providing insurance decreases:

1. as the price of the contracts, quantified by P, increases;
2. as ρ decreases, i.e. as the basis risk, quantified by ρ, increases;
3. as the cost of capital, quantified by K and T, decreases; and
4. as the scale of the index, quantified by σ_Y, increases.

If you set $P = 0$, Equation 4.5 reduces to a familiar expression for the "optimal hedge postion" otherwise known as the "hedge ratio".[5]

The quantities P and K depend upon market conditions. K also depends on the overall risk of the insurer. T depends upon the risk aversion of the insurer. To obtain the quantities σ_X, σ_Y and ρ you need a catastrophe model. It is to this we now turn.

5. AN ILLUSTRATIVE CATASTROPHE MODEL

The following information can be provided by a catastrophe model:

1. h—the natural event causing the catastrophe, numbered from 1 to s;
2. p_h—the probability of event h;
3. i—the location, e.g. county or ZIP code, numbered from 1 to m;
4. E_i—the number of exposure units at location i for all the insurers in the index, appropriately scaled so that the expected value of the index at expiration is \$1.00;

[5] See, for example, Hentschel and Smith [4]. There are several articles that may be of interest in this volume of the *JRI*, which is titled *Symposium on Financial Risk Management in Insurance Firms.*

5. e_i—the number of exposure units for the insurer at location i; and
6. L_{ih}—the damage caused to a unit of exposure at location i by event h.

For the examples in this paper, we will assume only one class of property. In practice one should add another subscript to allow for different classes each with different L_{ih}s.

The assembling of this information is a formidable task, and those who have done so regard the results of their efforts as proprietary. In this paper we use an illustrative catastrophe model published by Glenn Meyers [6]. Meyers' model has the following properties.

1. The covered area consists of a state with 50 counties. The east coast is exposed to the ocean and therefore to hurricanes.
2. Hurricanes travel only from east to west. They come in various strengths and affect either five or ten counties.
3. For the inland counties, the damage per exposure unit is 70% of the damage per unit in the county immediately to the east.

Table 1 provides a schematic map of the state along with the index exposures, E_i.

Tables 2A and 2B provide the probability, p_h, of each event h, and the loss per unit of exposure, L_{ih}, by each landfall county for each event. L_{ih} decreases by 70% as each event moves inland by one county. The index loss for each event h is given by:

$$\text{Index}(h) = \frac{\sum_{i=1}^{50} p_h e_i L_{ih}}{\text{Average Annual Hurricane Loss}}. \qquad (5.1)$$

TABLE 1

INDEX EXPOSURES BY COUNTY

i	E_i	i	E_i	i	E_i	i	E_i	i	E_i	Ocean
1	0.010	2	0.030	3	0.030	4	0.010	5	0.010	∿∿∿
6	0.010	7	0.030	8	0.030	9	0.010	10	0.010	∿∿∿
11	0.010	12	0.010	13	0.010	14	0.010	15	0.010	∿∿∿
16	0.010	17	0.010	18	0.010	19	0.010	20	0.010	∿∿∿
21	0.010	22	0.010	23	0.010	24	0.090	25	0.090	∿∿∿
26	0.010	27	0.010	28	0.010	29	0.010	30	0.010	∿∿∿
31	0.010	32	0.010	33	0.010	34	0.010	35	0.010	∿∿∿
36	0.050	37	0.010	38	0.050	39	0.050	40	0.010	∿∿∿
41	0.050	42	0.010	43	0.050	44	0.050	45	0.010	∿∿∿
46	0.010	47	0.030	48	0.010	49	0.010	50	0.010	∿∿∿

In this example we assume that only one hurricane can happen in a given year. To allow for multiple hurricanes in a year, one could create synthetic "events" by randomly selecting hurricanes that can happen in a single year, and simulating a very large version of Table 2.

The probability of a hurricane happening is 0.5000.

We also give the probability distribution of the final index values in Table 2. We consider this information to be valuable to potential investors who want to estimate the risk they are taking. This probability distribution is also shown graphically in Figure 1.

6. CALCULATING σ_X, σ_Y, AND ρ

Given the information from the previous section, we calculate:

$$\sigma_Y = \sqrt{\sum_{h=1}^{s}\left(\sum_{i=1}^{m} E_i L_{ih}\right)^2 p_h - \left(\sum_{h=1}^{s}\sum_{i=1}^{m} E_i L_{ih} p_h\right)^2}. \quad (6.1)$$

It is possible for a large multiline insurer to have the same catastrophe exposure as a small monoline property insurer. The capital

TABLE 2A

SMALL HURRICANES

h	i, at Landfall	L_{ih}	p_h	Index Loss for h
1	5	41.46	0.016181	0.4601
2	5	82.91	0.012945	0.9201
3	5	124.37	0.004854	1.3802
4	10	41.46	0.016181	0.4601
5	10	82.91	0.012945	0.9201
6	10	124.37	0.004854	1.3802
7	15	41.46	0.016181	0.2874
8	15	82.91	0.012945	0.5748
9	15	124.37	0.004854	0.8622
10	20	41.46	0.016181	0.2874
11	20	82.91	0.012945	0.5748
12	20	124.37	0.004854	0.8622
13	25	41.46	0.016181	1.6969
14	25	82.91	0.012945	3.3938
15	25	124.37	0.004854	5.0907
16	30	41.46	0.016181	0.2874
17	30	82.91	0.012945	0.5748
18	30	124.37	0.004854	0.8622
19	35	41.46	0.016181	0.2874
20	35	82.91	0.012945	0.5748
21	35	124.37	0.004854	0.8622
22	40	41.46	0.016181	0.8803
23	40	82.91	0.012945	1.7605
24	40	124.37	0.004854	2.6408
25	45	41.46	0.016181	0.8803
26	45	82.91	0.012945	1.7605
27	45	124.37	0.004854	2.6408
28	50	41.46	0.016181	0.3585
29	50	82.91	0.012945	0.7170
30	50	124.37	0.004854	1.0755

carried by each insurer depends on its entire book of business and should be taken into account when calculating the coefficient of correlation of its losses with the catastrophe index. To do this let:

$$X = X_1 + X_2 \tag{6.2}$$

TABLE 2B

LARGE HURRICANES

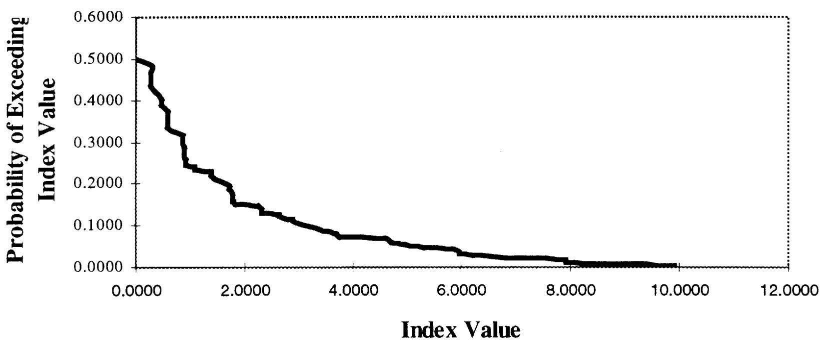
h	i, at 1st Landfall	i, at 2nd Landfall	L_{ih} at 1st and 2nd Landfall	p_h	Index Loss for h
31	5	10	124.37	0.004854	2.7604
32	5	10	165.82	0.006472	3.6806
33	5	10	207.28	0.003236	4.6007
34	10	15	124.37	0.004854	2.2424
35	10	15	165.82	0.006472	2.9899
36	10	15	207.28	0.003236	3.7374
37	15	20	124.37	0.004854	1.7244
38	15	20	165.82	0.006472	2.2992
39	15	20	207.28	0.003236	2.8740
40	20	25	124.37	0.004854	5.9530
41	20	25	165.82	0.006472	7.9373
42	20	25	207.28	0.003236	9.9216
43	25	30	124.37	0.004854	5.9530
44	25	30	165.82	0.006472	7.9373
45	25	30	207.28	0.003236	9.9216
46	30	35	124.37	0.004854	1.7244
47	30	35	165.82	0.006472	2.2992
48	30	35	207.28	0.003236	2.8740
49	35	40	124.37	0.004854	3.5030
50	35	40	165.82	0.006472	4.6707
51	35	40	207.28	0.003236	5.8384
52	40	45	124.37	0.004854	5.2816
53	40	45	165.82	0.006472	7.0422
54	40	45	207.28	0.003236	8.8027
55	45	50	124.37	0.004854	3.7163
56	45	50	165.82	0.006472	4.9551
57	45	50	207.28	0.003236	6.1939
58	5		124.37	0.004854	1.3802
59	5		165.82	0.006472	1.8403
60	5		207.28	0.003236	2.3003
61	50		124.37	0.004854	1.0755
62	50		165.82	0.006472	1.4340
63	50		207.28	0.003236	1.7925

where:

- X_1 represents the catastrophe losses that are estimated with a catastrophe model; and

FIGURE 1

INDEX LOSS EXCEEDING PROBABILITY



- X_2 represents the other insurer losses, which are assumed to be uncorrelated with X_1.

Then:

$$\sigma_{X_1} = \sqrt{\sum_{h=1}^{s}\left(\sum_{i}^{m} e_i L_{ih}\right)^2 - \left(\sum_{h=1}^{s}\sum_{i=1}^{m} e_i L_{ih} p_h\right)^2}. \qquad (6.3)$$

σ_{X_2} must be obtained from an analysis of the insurer's other business.

Let ρ_k be the coefficient of correlation of X_k with the index. We assume $\rho_2 = 0$.

Then:

$$\rho_1 = \frac{\displaystyle\sum_{h=1}^{s}\left(\sum_{i=1}^{m}(e_i L_{ij})\right)\left(\sum_{i=1}^{m}(E_i L_{ih})\right)p_h - \left(\sum_{h=1}^{s}\sum_{i=1}^{m} e_i L_{ih} p_h\right)\cdot\left(\sum_{h=1}^{s}\sum_{i=1}^{m} E_i L_{ih} p_h\right)}{\sigma_{X_1}\sigma_Y} \qquad (6.4)$$

and

$$\rho = \frac{\rho_1\sigma_{X_1}\sigma_Y + \rho_2\sigma_{X_2}\sigma_Y}{\sigma_{X_1+X_2}\sigma_Y} = \frac{\rho_1\sigma_{X_1}}{\sqrt{\sigma_{X_1}^2 + \sigma_{X_2}^2}}. \qquad (6.5)$$

7. EXAMPLES USING THE ILLUSTRATIVE MODEL

The examples given in this section will be based on an option with a zero strike price contract as described in Section 2. We chose this contract because it offers the insurer the maximum amount of protection and can be replicated by a series of the more popular call option spreads.

Using Table 1 as a reference, we create six sample insurers. Each insurer's book of business has a different geographical distribution.

1. All County Insurance Company has exposure in all counties in proportion to the industry as charted in Table 1.
2. Uni-County Insurance Company has the same exposure in all counties.
3. Northern Counties Insurance Company has exposure in counties 1–25 in proportion to the industry as charted in Table 1. It has no exposures in counties 26–50.
4. Big County Insurance Company has all its exposure in county 25.
5. Southern Counties Insurance Company has exposure in counties 26–50 in proportion to the industry as charted in Table 1. It has no exposures in counties 1–25.
6. Small County Insurance Company has all its exposure in county 1.

To facilitate comparisons among the six insurers, we have scaled the exposure of each so that σ_{X_i} is the same for each insurer.

TABLE 3
INSURER PARAMETERS

Parameter	Value
K	0.20
T	10
σ_{X_1}	30,000,000
σ_{X_2}	40,000,000
σ_Y	1.819

TABLE 4
INSURER PARAMETERS

Insurer #	Expected Loss	ρ_1	ρ
1	16,496,571	1.000	0.600
2	19,404,690	0.867	0.520
3	11,246,179	0.743	0.446
4	6,942,082	0.693	0.416
5	11,255,277	0.609	0.365
6	6,942,082	0.147	0.088

Table 2 lists the parameters, both selected and calculated from the model, common to each insurer.

The parameters in Table 3 are sufficient to describe the cost of providing coverage without buying any contracts on the catastrophe index. The needed insurer capital is:

$$C(0) = T\sigma_X = 10\sqrt{30{,}000{,}000^2 + 40{,}000{,}000^2} = 500{,}000{,}000.$$

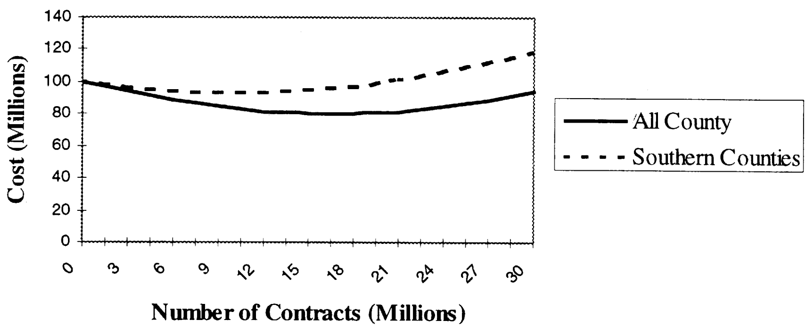
The cost of providing this capital is:

$$R(0) = KC(0) = 100{,}000{,}000.$$

We now introduce option contracts on the catastrophe index. Table 4 gives the expected loss for each insurer resulting from scaling the exposure, along with ρ_1 and ρ calculated from the illustrative model using Equations 6.4 and 6.5.

FIGURE 2

COST OF CAPITAL+NET COST OF CONTRACTS



As discussed in Section 4, the insurer wants to choose n so as to minimize its cost of capital, $KC(n)$, plus the net cost of the n contracts, nP. Figure 2 shows the cost for selected insurers as a function of n for $P = 0$.

As Figure 2 illustrates, there is an optimal number, n, of contracts that will minimize the cost of writing insurance subject to catastrophes. The number n can be calculated using Equation 4.5. Tables 5 and 6 show the ns calculated from Equation 4.5 for each of the insurers in our example. The cost of insuring is then given by Equation 4.3 for these ns.

Table 5 is sorted in order of P to illustrate the effect of the contract price. As the price increases, the optimal number of contracts decreases and the cost of insuring increases.

Table 6 is sorted in order of insurer to illustrate the effect of the insurer's correlation with the catastrophe index. As the correlation increases, the optimal number of contracts increases, and the cost of insuring decreases.

Without the catastrophe contracts, All County must raise an additional \$20,000,000 in capital. This provides a yardstick for measuring the efficiency of the contracts. For example, if $P = 0.6$, the cost of insuring catastrophes for All County is only an additional \$8,801,889 if it buys the optimal number of contracts. All County reduces its cost of insuring its catastrophe exposure by 56%. At the same time, Big County Insurance's additional cost of insuring its catastrophe exposure is reduced by only 17%.

It is possible for n to be negative. This simply indicates that if the price of the contract is sufficiently high, it is better to be a seller than a buyer of the catastrophe contracts.

8. CONTRACTS ON A CATASTROPHE INDEX VS. REINSURANCE

The examples given show that contracts on a catastrophe index can reduce the cost of providing insurance, even if the correlation between the insurer's catastrophe losses are not highly correlated with the index. However, it is possible that conventional reinsurance may be an even lower cost of providing insurance. In this section we show how to investigate this possibility.

Reinsurance can be viewed as an option contract on a catastrophe index, with the index being the insurer's own experience. We take this view here. Properly interpreted, Equations 4.3 and 4.5 provide the means of finding out how much reinsurance to buy, and the expected benefit of buying it.

We will use the examples in the preceding section to show that reinsurance can give a lower cost of providing insurance.

A quota-share reinsurance contract corresponds to the option contract with $\rho_1 = 1$. We find a net cost of reinsurance, denoted by P_R, that provides the same cost of insurance as the corresponding contract on the catastrophe index. If reinsurance can be obtained for a lower net cost, we conclude that insurance can be provided at a lower cost.

TABLE 5

THE EFFECT OF THE CONTRACT PRICE

Insurer #	Number of Contracts	Cost of Insuring	P
1	16,496,571	80,000,000	0.0
1	15,285,243	83,178,275	0.2
1	14,062,815	86,113,360	0.4
1	12,817,677	88,801,889	0.6
1	11,537,127	91,238,074	0.8
2	14,306,818	85,394,944	0.0
2	13,013,800	88,127,104	0.2
2	11,708,935	90,599,676	0.4
2	10,379,829	92,809,065	0.6
2	9,012,923	94,749,092	0.8
3	12,264,212	89,500,107	0.0
3	10,909,035	91,817,535	0.2
3	9,541,442	93,862,895	0.4
3	8,148,442	95,632,421	0.6
3	6,715,825	97,119,635	0.8
4	11,428,496	90,951,642	0.0
4	10,051,340	93,099,730	0.2
4	8,661,567	94,971,339	0.4
4	7,245,975	96,562,639	0.6
4	5,790,124	97,867,049	0.8
5	10,048,063	93,082,705	0.0
5	8,638,639	94,951,482	0.2
5	7,216,303	96,537,301	0.4
5	5,767,543	97,836,244	0.6
5	4,277,580	98,841,576	0.8
6	2,425,986	99,609,960	0.0
6	917,729	99,944,446	0.2
6	– 604,346	99,976,132	0.4
6	– 2,154,698	99,700,825	0.6
6	– 3,749,142	99,111,318	0.8

The P_Rs were calculated by trial and error as follows:

1. Select a P_R.
2. Find n_R using Equation 4.5.

TABLE 6

THE EFFECT OF INSURER CORRELATION WITH THE INDEX

Insurer #	Number of Contracts	Cost of Insuring	P
1	16,496,571	80,000,000	0.0
2	14,306,818	85,394,944	0.0
3	12,264,212	89,500,107	0.0
4	11,428,496	90,951,642	0.0
5	10,048,063	93,082,705	0.0
6	2,425,986	99,609,960	0.0
1	15,285,243	83,178,275	0.2
2	13,013,800	88,127,104	0.2
3	10,909,035	91,817,535	0.2
4	10,051,340	93,099,730	0.2
5	8,638,639	94,951,482	0.2
6	917,729	99,944,446	0.2
1	14,062,815	86,113,360	0.4
2	11,708,935	90,599,676	0.4
3	9,541,442	93,862,895	0.4
4	8,661,567	94,971,339	0.4
5	7,216,303	96,537,301	0.4
6	−604,346	99,976,132	0.4
1	12,817,677	88,801,889	0.6
2	10,379,829	92,809,065	0.6
3	8,148,442	95,632,421	0.6
4	7,245,975	96,562,639	0.6
5	5,767,543	97,836,244	0.6
6	−2,154,698	99,700,825	0.6
1	11,537,127	91,238,074	0.8
2	9,012,923	94,749,092	0.8
3	6,715,825	97,119,635	0.8
4	5,790,124	97,867,049	0.8
5	4,277,580	98,841,576	0.8
6	−3,749,142	99,111,318	0.8

3. Find the cost of insurance using Equation 4.3 with $P = P_R$ and $n = n_R$.
4. If the cost of insurance is not equal to the target cost, try another P_R.

TABLE 7

OPTIONS VS. REINSURANCE

Insurer #	Cost of Insuring	P	P_R
1	88,801,889	0.6000	0.6000
2	92,809,065	0.6000	0.7820
3	95,632,421	0.6000	2.0073
4	96,562,639	0.6000	5.6528
5	97,836,244	0.6000	2.0165
6	99,700,825	0.6000	8.1631

We use the option contract from Table 6 with $P = 0.6$. The P_Rs that provide the same cost of providing insurance are given in Table 7.

For Insurer 1, All County Insurance Company, there is no difference because its losses correlate perfectly with the index losses. If the net cost for reinsurance to Insurer 2, Uni-County Insurance Company, is between 0.6000 and 0.7820, reinsurance is less expensive. There is more leeway for reinsurance for the regional insurers, Insurers 3 and 5, and considerably more leeway for reinsurance with the single-county insurers, Insurers 4 and 6.

9. SUMMARY

The cost of capital and its substitutes is determined by a variety of market conditions that are beyond the control of the insurer. To efficiently use its capital, the insurer has to constantly analyze the opportunities that are presented to it. This paper shows how a catastrophe model can be used to evaluate the costs and benefits of alternative catastrophe risk management tools for insurers. The alternatives include:

1. raising sufficient capital to contain the catastrophe risk;
2. buying options on a catastrophe index; and
3. buying reinsurance.

These alternatives are quantified by the cost of providing insurance, which depends upon:

1. the price of the contracts and/or reinsurance, as quantified by P and P_R;
2. the basis risk, as quantified by ρ; and
3. the cost of capital, as quantified by K, T and σ_X.

The quantities P and K depend upon market conditions, and T depends upon the risk aversion of the insurer. The quantities σ_X, σ_Y, and ρ are obtained from the catastrophe model.

With these quantities one can calculate the optimal number of contracts (or the optimal amount of reinsurance) to buy with Equation 4.5 and then quantify the cost of providing insurance with Equation 4.3. The cost of the various alternatives can be compared to provide the best insurance value.

REFERENCES

[1] American Academy of Actuaries Property/Casualty Risk Based Capital Task Force, Report on Reserve and Underwriting Risk Factors, Casualty Actuarial Society *Forum*, Summer 1993 Edition, pp. 28–71.

[2] Brealy, Richard A. and Stewart C. Myers, *Principles of Corporate Finance*, McGraw-Hill, 1991.

[3] Cole, Joseph B. and Richard L. Sandor, "Calculating the Hedging Layer for a Synthetic Reinsurance Contract," *Journal of Reinsurance* 1, 2, Winter 1993, pp. 29–40.

[4] Hentschel, Ludger and Clifford W. Smith, "Risk in Derivitives Markets: Implications for the Insurance Markets," *Journal of Risk and Insurance* 64, 2, June 1997, pp. 34–56.

[5] Hull, John C., *Options, Futures and Other Derivatives*, Third Edition, Prentice Hall, 1997.

[6] Meyers, Glenn, "The Competitive Market Equilibrium Risk Load Formula for Catastrophe Ratemaking," *PCAS* LXXXIV, 1996, Casualty Actuarial Society, pp. 563–600.

APPENDIX

DERIVATION OF EQUATION 4.5

We seek to solve the equation:

$$\frac{K \cdot T \cdot (n\sigma_Y^2 - \rho\sigma_X\sigma_Y)}{\sqrt{\sigma_X^2 - 2n\rho\sigma_X\sigma_Y + n^2\sigma_Y^2}} + P = 0.$$

Moving the P and the denominator to the other side of the equation, and squaring yields:

$$\begin{aligned} P^2(\sigma_X^2 - 2n\rho\sigma_X\sigma_Y + n^2\sigma_Y^2) &= K^2T^2(n\sigma_Y^2 - \rho\sigma_X\sigma_Y)^2 \\ &= K^2T^2(n^2\sigma_Y^2 - 2n\rho\sigma_X\sigma_Y^3 + \rho^2\sigma_X^2\sigma_Y^2) \end{aligned}$$

The above equation can be put into the form: $an^2 + bn + c = 0$ with

$$\begin{aligned} a &= \sigma_Y^2(K^2T^2\sigma_Y^2 - P^2); \\ b &= -2\rho\sigma_X\sigma_Y(K^2T^2\sigma_Y^2 - P^2); \qquad \text{and} \\ c &= \sigma_X^2(K^2T^2\sigma_Y^2\rho^2 - P^2). \end{aligned}$$

The solution for n is of the form

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

with:

$$\frac{-b}{2a} = \frac{2\rho\sigma_X\sigma_Y(K^2T^2\sigma_Y^2 - P^2)}{2\sigma_Y^2(K^2T^2\sigma_Y^2 - P^2)} = \frac{\rho\sigma_X}{\sigma_Y}; \qquad \text{and}$$

$$\begin{aligned} \frac{b^2 - 4ac}{4a^2} &= \frac{4\sigma_X^2\sigma_Y^2\rho^2(K^2T^2\sigma_Y^2 - P^2)^2 - 4\sigma_Y^2(K^2T^2\sigma_Y^2 - P^2)\sigma_X^2(K^2T^2\sigma_Y^2\rho^2 - P^2)}{4\sigma_Y^4(K^2T^2\sigma_Y^2 - P^2)^2} \\ &= \frac{\sigma_X^2\rho^2(K^2T^2\sigma_Y^2 - P^2) - \sigma_X^2(K^2T^2\sigma_Y^2\rho^2 - P^2)}{\sigma_Y^2(K^2T^2\sigma_Y^2 - P^2)} \\ &= \frac{\sigma_X^2}{\sigma_Y^2} \cdot \frac{P^2(1 - \rho^2)}{K^2T^2\sigma_Y^2 - P^2}. \end{aligned}$$

Then:

$$n = \frac{-b - \sqrt{b^2 - 4ac}}{2a} = \frac{\rho\sigma_X}{\sigma_Y} - \frac{\sigma_X}{\sigma_Y}\sqrt{\frac{P^2(1-\rho^2)}{K^2T^2\sigma_Y^2 - P^2}}.$$

Squaring the equation in the first step introduces an extraneous root. The solution with the positive square root is the extraneous root since it indicates one should buy more contracts when $P > 0$ than when $P = 0$.

THE IMPACT OF INVESTMENT STRATEGY ON THE MARKET VALUE AND PRICING DECISIONS OF A PROPERTY/CASUALTY INSURER

TRENT R. VAUGHN

Abstract

This paper examines the impact of investment strategy on the market value and pricing decisions of a property/casualty insurance company. Section 2 utilizes classic financial theory to demonstrate the irrelevance of investment policy in perfect capital and product markets. Sections 3 through 6 illustrate four possible sources of investment policy relevance: imperfect information in property/casualty (P/C) insurance product markets, guaranty funds, conflicts of interest between shareholders and current policyholders, and taxes.

Lastly, Section 7 will discuss the implications of the optimal investment strategy on an insurer's pricing decisions. This section will close with a discussion of three commonly posed questions: (1) Is insurance a negative-net present value (NPV) transaction to the policyholder? (2) Does excess capital depress insurance prices? and (3) Does diversification create value?

1. INTRODUCTION

In 1995 and 1996, the bullish stock market produced large investment earnings for the property/casualty insurance industry. In fact, the increase in the industry's net income in 1996 was driven largely by continued growth in realized capital gains [12]. Not all insurers, however, benefited equally from the booming equity market. P/C insurance companies vary considerably in the proportion and composition of funds invested in the equity market.

Historically, long-term investment returns on common stocks have outperformed returns on bond portfolios [11, page 33]. Likewise, riskier common stock portfolios have generally produced better returns over the long run. In general, investors require some payoff for accepting greater investment risks, and this payoff comes in the form of higher expected returns. Given this relationship, many insurance managers may adopt a riskier investment strategy in order to increase earnings, return on equity (ROE), and shareholder value.

However, will adopting a riskier investment strategy, such as a higher proportion of equity investments, really increase shareholders' wealth?[1] In other words, can insurance management change the *total* value of the company by changing its asset allocation?

Section 2 of this paper utilizes classic financial theory to demonstrate the irrelevance of investment policy in perfect capital and product markets. In perfect markets, the only decision capable of creating or destroying value is the firm's underwriting decisions. Asset allocation does not matter.

In reality, however, insurers do not operate in perfect markets, and the insurer's investment choices can affect value. Sections 3 through 6 illustrate four possible sources of investment policy relevance: imperfect information in P/C insurance product markets, guaranty funds, conflicts of interest between shareholders and current policyholders, and taxes.

Given these relevant market imperfections, a value-maximizing asset allocation is possible. But what impact does this optimal investment portfolio have on competitive insurance prices? And how should this impact be reflected in insurance pricing models? Section 7 will discuss the implications of the optimal investment strategy on an insurer's pricing decisions.

---

[1]Shareholders' wealth is a measure of the total market value of the shareholders' investment in the firm.

Lastly, Section 7 will close with a discussion of three commonly posed questions: (1) Is insurance a negative-NPV transaction to the policyholder? (2) Does excess capital depress insurance prices? and (3) Does diversification create value?

2. INVESTMENT POLICY IRRELEVANCE IN PERFECT MARKETS

Nonfinancial Firms vs. Insurance Companies

Most of modern financial management theory focuses on decisions by nonfinancial firms, such as manufacturing and retailing concerns. Nonfinancial firms differ significantly from insurance firms in their investments, operations, and financing.

Nonfinancial firms make investments in product markets, the markets that bring together the buyers and sellers of goods and services. The structure of the various product markets ranges from pure monopoly to perfect competition. Those companies that acquire assets and capabilities in attractive product markets will earn superior profits [4]. Conversely, nonfinancial firms obtain financing in the capital markets, where competition is intense and profits are difficult to achieve.

Insurance companies operate in the reverse manner. Insurers make investments in the intensely competitive capital markets, where economic profits are elusive. On the right-hand side of the balance sheet, they obtain financing partially from insurance product markets. These markets may not be perfectly competitive in all niches at all times, allowing the possibility for superior profits.

A classic problem in finance considers the optimal capital structure, or financing decisions, for a nonfinancial firm. Given highly efficient capital markets, can the nonfinancial firm's financing decisions create value? In order to isolate the effect of the firm's financing decisions on value, its current assets and operations are usually considered fixed.

Given the inherent differences in their operating environment, it is logical to modify this problem for insurance companies. As mentioned, insurance companies deal with capital markets on the asset side. For insurers, we must consider the optimal investment strategy, given fixed financing from policyholders and shareholders.

Modigliani and Miller's Propositions I and II

In their famous Proposition I, Modigliani and Miller (abbreviated as "MM" throughout this paper) proved that the value of the nonfinancial firm is determined entirely on the left side of the balance sheet by the assets it owns [16]. The firm's capital structure, the mix of different securities it has issued, does not impact firm value. MM's proof assumes (1) perfect capital markets, including no taxes, and (2) the firm's financing decisions have no impact on the firm's investment decisions.

The implications of Proposition I are shown graphically in Figure 1. In perfect capital markets, the firm's debt ratio has no impact on its operating income (since investment strategy is fixed) or on the total firm value. As such, the expected rate of return on the firm's assets (r_A) is independent of the firm's debt ratio and is displayed as a horizontal line.

For low debt levels, the expected return on the debt (r_D) equals the risk-free rate of interest. As the firm borrows past a certain point, the firm's debtholders demand a higher interest rate; and the r_D curve slopes upward.

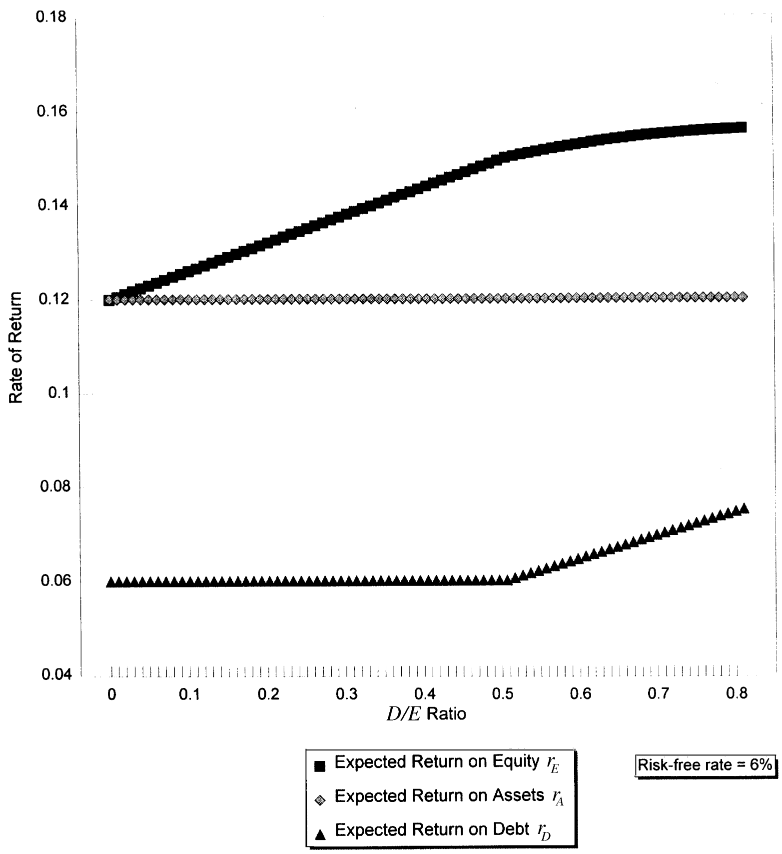
The expected return on the levered equity (r_E) is shown as the top curve on the graph. For low debt levels, r_E increases linearly with the debt ratio. Eventually, the slope of r_E decreases, as debtholders bear more of the business risk of the firm. The exact formula for the r_E curve is given by MM's Proposition II:

$$r_E = r_A + (D/E) \times (r_A - r_D).$$

Here, D/E represents the debt-to-equity ratio, expressed in market values.

FIGURE 1

MODIGLIANI AND MILLER'S PROPOSITION I



The Insurer as a Levered Equity Trust

In a 1968 *Proceedings* paper, J. Robert Ferrari proposed viewing the P/C insurer as a levered equity trust [9]. In other words, Ferrari visualized the insurer as borrowing funds from policyholders, then investing the combined policyholder and shareholder funds in financial assets.

The interest rate on the funds borrowed from policyholders is reflected in the premium charged. In the absence of taxes, the premium equals the present value of the expected losses and expenses under the policy.[2] In this sense, the "debt" issued by insurers is comparable to zero-coupon debt issued by corporations. That is, the return to the insurance "debtholders" is not provided by a regular interest payment, but by discounting the expected loss payment.[3]

Assuming a tax-free, perfectly competitive economy, the expected rate of return on the insurer's levered equity is given by the MM Proposition II formula. In this context, r_A represents the expected return on the insurer's asset portfolio, r_D represents the expected return on the insurer's liabilities, and the "D" term represents the present value of the insurer's liabilities.

Figure 2 displays the Proposition II formula graphically for a hypothetical insurance company. The company depicted in this figure holds a very conservative investment portfolio, as evidenced by the close proximity of the expected investment return (r_A) to the risk-free rate. Furthermore, the company's liabilities are risk-free at the current debt ratio. In other words, r_D is equal to the risk-free rate.[4]

Now assume that the company in Figure 2 decides to implement a more aggressive investment stance, perhaps investing a larger proportion of the portfolio in blue-chip stocks.

Figure 3 displays the consequences of this new investment policy. As shown, the riskier investment strategy results in an increased expected investment return (r_A). As r_A increases, the expected return to shareholders (r_E) increases according to the Proposition II formula.

---

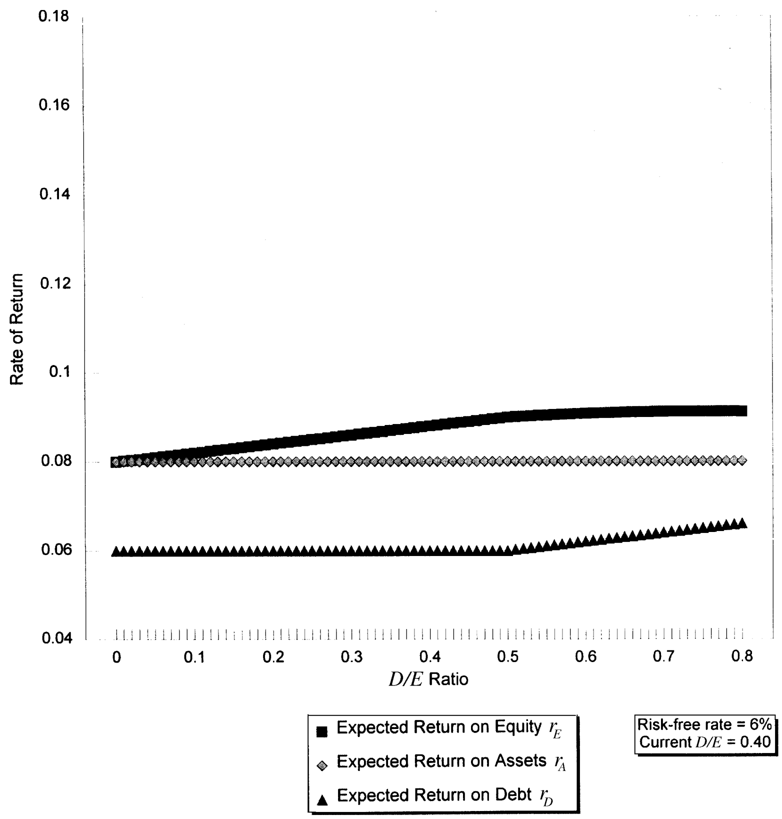
[2]This fact is demonstrated in the Myers and Cohn [18] article discussed in Section 7.

[3]See page 685 of [2] for a comparison of interest-paying bonds and pure-discount bonds.

[4]The covariance of insurance losses with the capital market return is often very low. In terms of the capital asset pricing model (CAPM), the beta of insurance losses is often close to zero, implying an expected return equal to the risk-free rate. However, this is not a necessary assumption for the proof.

FIGURE 2

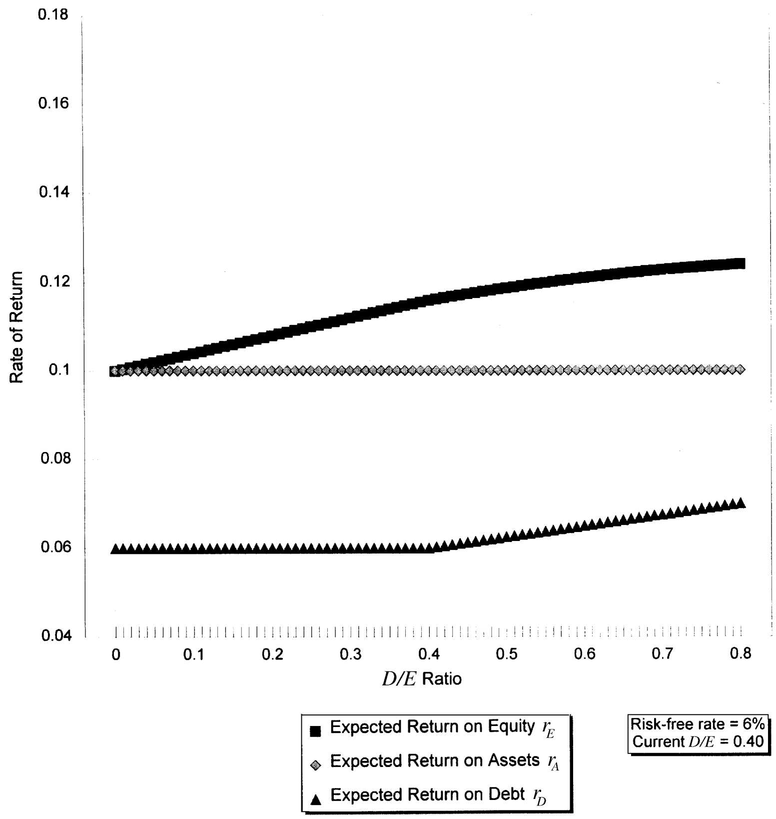
MODIGLIANI AND MILLER'S PROPOSITION II
(Conservative Investment Portfolio)



Does this higher expected return make shareholders better off? Unfortunately, the MM theory implies that the higher *expected* return will be exactly offset by a higher *required* return by shareholders. Specifically, the beta of the levered equity is expressed by an equation very similar in form to Proposition II: $B_E = B_A + (D/E) \times (B_A - B_D)$. Therefore, the systematic risk and required

FIGURE 3

MODIGLIANI AND MILLER'S PROPOSITION II
(Riskier Investment Strategy)

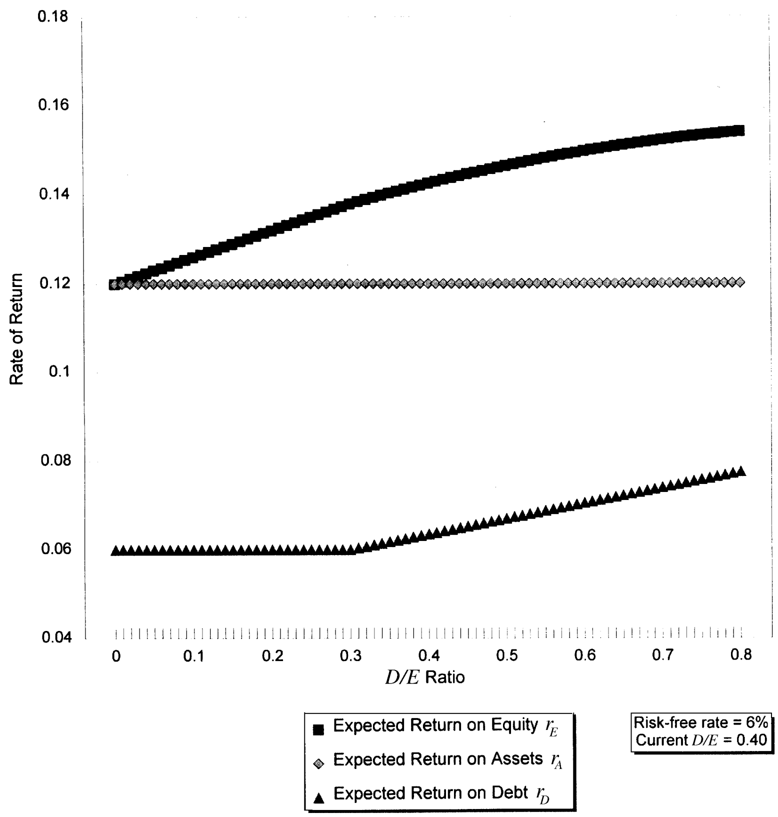


return on the levered equity increase exactly in lockstep with the expected return, leaving shareholder wealth unchanged [2, p. 456].

While the riskier investment policy displayed in Figure 3 increased r_A and r_E, the insurer's liabilities remained risk-free at the current debt ratio. Assume that our hypothetical insurer now

FIGURE 4

MODIGLIANI AND MILLER'S PROPOSITION II
(Riskiest Investment Portfolio)



decides to go for broke, investing entirely in risky stocks, collateralized mortgage obligations (CMOs), and derivatives.[5] Figure 4 shows the results of this new investment policy. Not surprisingly, the r_A line and the r_E curve each notch further up.

[5]In the U.S., investment regulations would most likely preclude such a risky asset allocation.

But from the policyholders' standpoint, the new investment policy results in a much greater risk that the insurer will default on its promises. Accordingly, the policyholders will mark down the value of the insurer's promise and demand a lower premium. Equivalently, the insurer will now have to assume more expected losses and expenses to maintain the same level of premium—and the same ongoing "debt" ratio. That is, the insurer is now paying a higher interest rate on the funds borrowed from policyholders. This is reflected in Figure 4, as r_D now exceeds the risk-free rate.

The systematic risk of the levered equity in Figure 4 is again given by the formula: $B_E = B_A + (D/E) \times (B_A - B_D)$. The B_D term adjusts to reflect the additional risk assumed by policyholders. Once again, required return increases in accordance with expected return, leaving shareholders' wealth unchanged. The critical assumption is that policyholders correctly identify and adjust for the investment change.

Thus, assuming perfectly competitive capital markets and insurance product markets, an insurance company's investment policy has no impact on the value of the company and the wealth of its shareholders. In reality, perfect MM conditions rarely exist, and investment policy can affect value. This paper explores the implications of several possible market imperfections on investment strategy, beginning with imperfect information in insurance markets.

3. IMPERFECT INFORMATION IN INSURANCE PRODUCT MARKETS

Section 2 demonstrated that a riskier investment strategy will increase the expected and required rate of return on levered equity. Moreover, a dramatic investment change may also increase the default risk and systematic risk of the insurer's liabilities, thereby eliciting an increase[6] in the insurer's r_D; new policyholders will then demand lower premiums.

[6]Any investment strategy that increases r_A will also increase r_E (Figure 3). However, we did not discuss the *degree* of investment risk required to induce an increase in r_D

The Section 2 proof assumes perfect competition as an insurance product market model. In a perfectly competitive market, every buyer possesses perfect information regarding the price and quality of the insurance promise. If the riskier investment policy forces an increase in r_D, new policyholders will recognize this.

In real world insurance markets, buyers rarely have perfect information. For instance, how many drivers are aware of the investment policies of competing personal auto carriers? And could a P/C insurer modify its investment strategy enough to increase the risk of its liabilities without policyholders noticing?

Fortunately, most market participants would agree that P/C insurance in the U.S. is a tremendously competitive business. Several thousand U.S. and foreign insurance companies compete for the business of millions of domestic customers. Consumer groups, state insurance departments, rating agencies, agents and brokers all work to ensure that sufficient information is promulgated to insurance buyers. These market conditions ensure that any investment change dramatic enough to impact the risk of the insurer's liabilities would be fully appreciated by buyers.

4. GUARANTY FUNDS

In a competitive insurance market, policyholders will recognize an investment shift that increases the risk of the insurer's liabilities. By shifting to a very risky investment policy, the insurance company will force the policyholders to share in the investment risk. While the policyholders are now sharing in these risks, they are also getting paid for it, by paying a lower insurance premium for the same policy. As long as these premium changes are conducted on *fair* terms, shareholders' wealth will not increase.

(Figure 4). For an insurer with a strong surplus position and a reasonably diversified underwriting portfolio, it may take a hair-curling investment change to increase r_D.

Under the U.S. guaranty fund system, however, policyholders will not bear all of the risks and costs of insolvency. A large part of these risks and costs will be absorbed by the guaranty fund.[7] In this case, *shareholders* will gain[8] from an investment change that increases r_D.

For example, assume that policyholders of insolvent insurance companies will be reimbursed on a full and timely basis by the guaranty fund. The riskiness of an insurance company's investment strategy is then irrelevant to the policyholder. Policyholders will discount expected losses and expenses at the risk-free rate regardless of the company's investment risk.

Thus, if an investment change increases r_D, the value of the shareholders' stake in the firm is increased. In effect, the guaranty fund's promise allows the company to obtain a subsidized loan from new policyholders. The value of the company is increased by the NPV of this loan. Provided that all parties are aware of the loan guarantee prior to the transaction, the entire increase in value will fall to the shareholders. The riskier the investment policy, the more valuable this loan guarantee becomes.

5. CONFLICTS OF INTEREST BETWEEN SHAREHOLDERS AND CURRENT POLICYHOLDERS

An investment policy which increases the insurer's r_D also creates a transfer of value from *current* policyholders to shareholders—even without the assumption of imperfect information or guaranty funds.

How does this transfer of value work? At the time the current policyholders purchased their policies, they discounted the insurer's promises at the lower risk-free rate. But after the invest-

[7]Guaranty funds do not protect all policyholders. For instance, guaranty fund protection does not apply to policyholders of non-admitted insurance companies. Guaranty funds also do not provide "full" coverage in terms of amounts or certain lines of business.

[8]Remember: While every riskier investment strategy will increase r_A (see Figure 3), not every investment change will increase r_D (see Figure 4).

ment change, r_D exceeds this risk-free rate. If these policyholders were free to renegotiate the outstanding portion of their policies, the premium they would be willing to pay would be lower.[9] Yet, these contracts are not typically renegotiable—many, in fact, have already expired.

The total value of the insurance company's assets, however, is not changed by the switch into riskier investments. Assuming efficient capital markets, all financial assets are bought and sold at their fair value. With the value of the assets unchanged, the current policyholders' loss is the shareholders' gain.

Clearly, only a company that is already in financial trouble would adopt a riskier investment strategy for the purpose of creating a transfer of value from current policyholders to shareholders. However, the example serves to illustrate the general rule that a shift in the risk of the firm's assets benefits shareholders at the expense of debtholders.[10]

In sum, the imperfect information, guaranty funds, and transfer of value effects all encourage the insurer to invest in riskier securities.[11] Section 6 discusses the possibility that taxes may have an opposite effect, encouraging the insurer to invest significant amounts in bonds.

6. TAXES AND INVESTMENT POLICY

The proof of investment policy irrelevance in Section 2 relies on the assumption of a perfectly competitive, *tax-free* economy.

---

[9] Assuming no guaranty fund applies.

[10] See Brealey and Myers [2, p. 492] for a general discussion, and Galai and Masulis [10, pp. 62–64] for a rigorous proof.

[11] In practice, increasing the riskiness of an insurance company's assets may actually have the perverse effect of decreasing shareholder wealth for two other reasons: (1) insurance buyers generally prefer not to share in the investment risk of the company, and (2) since insurance companies are subject to regulatory solvency constraints, increasing the volatility of its investments will increase the likelihood of losing an important intangible asset—franchise value. For insurers with valuable growth opportunities, the reduction in franchise value may dominate the guaranty fund and transfer of value effects. See [14, pp. 450–457 and 644].

The impact of taxes on the optimum capital structure and value of the corporation has been a source of debate in financial theory. Three competing theories have emerged: MM's original theory corrected for taxes, Miller's equilibrium theory, and a compromise theory. This section will briefly describe the three competing theories and discuss the implications of each theory on an insurer's choice between taxable bonds and common equity. The section will conclude with a brief discussion of municipal securities and dividend-paying stocks.

MM—The "Corrected" Theory

The original MM theory described in Section 2 concludes that a company's decision to borrow or lend money does not impact its market value. In 1963, MM modified the original theory to accommodate corporate taxes [17]. This modified theory stresses the corporate tax advantage of borrowing: debt interest is deductible at the corporate level, whereas dividends and retained earnings are not.

The "corrected" MM theory does not explicitly address investor taxes; only corporate taxes are relevant. This simplification implies that investor taxes are independent of the firm's debt policy; in other words, the effective *personal* tax rate on corporate debt equals the effective personal tax rate on corporate equity.

Under MM's revised theory, there is a clear tax disadvantage to corporate lending. A firm's decision to invest in taxable bonds decreases the value of the firm; this decrease in value is equal to the present value of *corporate* taxes paid on the investment.

For instance, suppose a hypothetical firm decides to issue \$1,000 of new equity and invest the money in a perpetual, risk-free, taxable bond with a 10% coupon. Assume the corporate tax rate is 35%. The firm's value is then reduced by the present value of a perpetual tax payment of $0.10 \times \$1,000 \times .35 = \$35$. The correct discount rate for these tax payments is generally assumed to be the interest rate on the bond; thus, the value of the company is reduced by $\$35/0.10 = \$350$.

However, suppose the same hypothetical firm were to invest the newly raised capital in risk-free (zero beta) common equity.[12] Given that the effective *personal* tax rates on equity income and bond income are equivalent, the expected return on the common stock equals the risk-free rate of interest. Yet, the firm is taxed at a lower rate on the common stock at the corporate level. This lower tax rate results from two provisions of the tax code: (1) corporations are only taxed on 30% of the dividends received from other corporations, and (2) equity income in the form of unrealized capital gains escapes taxation entirely. The result is a higher after-tax return to the firm on the common stock than the taxable bond.

Of course, to the extent that taxable interest income is offset by underwriting losses, investing in corporate bonds creates no tax disadvantage to the insurer. Yet, since the insurer cannot be certain of the actual underwriting losses, the safest strategy in this simplified MM world would be to invest solely in common stocks. Common stocks will offer the firm a higher after-tax return than taxable bonds of equivalent risk.

Varying Personal Tax Rates—Miller's Debt and Taxes

MM's corrected theory implies that an insurer's optimal investment strategy is to invest solely in common stock. Of course, we don't see any insurance companies doing this in practice.

The MM theory also leads to unrealistic implications for the optimal behavior of nonfinancial firms. At its extreme, the MM theory implies that industrial firms should employ entirely debt-financed capital structures. This extreme prediction, however, ignores an important cost of higher debt levels—the increased cost of bankruptcy and financial distress.

---

[12]Zero beta common stock is not "risk-free" in the same sense as government debt; in this context, risk-free merely implies that the stock possesses no systematic, or undiversifiable, risk.

Specifically, MM's corrected theory asserts that a borrowing firm creates value through the corporate tax shield of debt. In practice, the value of this corporate tax shield will be partially offset by the expected cost of bankruptcy and financial distress.

Yet, many financial economists still worried about the implications of the theory. Compared to the enormous value of corporate tax shields, the expected costs of financial distress were generally low—implying that most firms should operate at extremely high debt levels. Merton Miller compared the situation to making horse-and-rabbit stew: mix in one horse and one rabbit, and it tastes an awful lot like horse stew.

Miller resolved the horse-and-rabbit stew dilemma by specifically introducing *investor* taxes into the mix [15]. Miller's revised theory assumed that the effective tax rate on equity is zero (due to the deferring of capital gains), but that individual tax rates on interest income varied from zero (for example, investments in pension funds) to rates that exceeded the corporate tax rate (for high-income individuals).

In Miller's world, the total amount of corporate debt would adjust to minimize the sum of corporate and personal taxes. All investors in tax brackets less than or equal to the corporate tax rate would hold corporate debt. Investors in higher tax brackets would hold equity or municipal (tax-free) bonds. The expected rate of return on risk-free common stocks would equal $r_f \times (1 - T_c)$, the risk-free interest rate times the complement of the *corporate* tax rate.

Under Miller's theory, there is no tax disadvantage to a firm investing in taxable bonds. For example, consider the hypothetical firm discussed in the previous subsection. The bond purchased earns interest of 10% before corporate taxes, but 6.5% after corporate taxes. The required return to the firm's *shareholders* on a risk-free investment is also 6.5% (10% × (1 – 0.35)). Thus, the investment has an NPV of zero and firm value is unchanged.

Miller's theory implies that an insurance company should invest solely in taxable bonds. Specifically, risk-free taxable bonds will offer investors a (pre-tax) return of r_f; risk-free common stock will offer a (pre-tax) return of $r_f \times (1 - T_c)$. Hence, tax-free investors gain value by investing in corporate bonds, which offer a higher return than common stocks of equivalent risk. To the extent that interest income is shielded by underwriting losses, the company is a tax-free investor. To the extent that interest is not shielded by underwriting losses, the company earns the same *after-tax* return on bond income as it would on risk-free equity income ($r_f \times (1 - T_c)$).

Miller coined the term "bondholders' surplus" to describe the extra (pre-tax) return on taxable bonds. According to the Miller model, this extra return is largest in the case of risk-free bonds [15, p. 271]. Therefore, the theory implies that insurers should invest in super-safe government debt to maximize bondholders' surplus. This results in a fortunate counterweight to the Sections 4 and 5 argument, which implied that insurers should invest in risky assets at the expense of the guaranty fund and current policyholders.

A Compromise Theory

DeAngelo and Masulis [6] and others have described a compromise theory that avoids the extreme assumptions and implications of the MM and Miller theories. The adherents of this view contend that Miller's assumptions are somewhat extreme and were not intended to be a realistic description of the tax code. Instead, most economists would agree that there is a moderate tax advantage to corporate borrowing (and corresponding tax *disadvantage* to corporate lending). However, this tax effect is less than the MM corrected theory predicts [2, p. 484].

Specifically, under this compromise theory, the tax code's effect on common stock returns is described by the factor T^*, which is between zero (MM) and the full corporate tax rate (Miller).

That is, the expected return on risk-free common stock is given by $r_f \times (1 - T^*)$.

To the extent that bond interest is shielded by underwriting losses, the company is a tax-free investor and benefits by investing in taxable bonds.[13] To the extent that interest income is not shielded by underwriting losses, the company generally earns a higher after-tax return on equity income than on bond income. For instance, if we assume that equity income escapes taxation at the corporate level,[14] the expected after-tax return on risk-free common stock is given by $r_f \times (1 - T^*)$; the expected after-tax return on risk-free bonds is $r_f \times (1 - T_c)$, where T_c is greater than T^*. Thus, the company should attempt to invest in taxable bonds up to the point where bond interest equals expected underwriting losses, with the balance invested in common stocks.

Fortunately, this optimal investment strategy under the compromise theory also works well in both the MM and Miller worlds. It satisfies Miller's mandate that the portion of investment returns shielded by underwriting losses should be derived from taxable bonds. It also satisfies the MM belief that common stocks offer a higher after-tax return than bonds for corporate investors in a tax-paying position.

While the theory to this point indicates that insurers should invest significant amounts in common stocks, it does not suggest which stocks are most appealing to insurers. For example, should insurers invest in high-dividend stocks to capitalize on the corporate dividend exclusion? Or should they invest in low-dividend stocks and benefit from the tax exclusion on unrealized capital gains?

---

[13] Under the compromise theory, risk-free taxable bonds will offer the company a (pre-tax) return of r_f; risk-free common stock will offer a (pre-tax) return of $r_f \times (1 - T^*)$.

[14] In Section 7, Table 1 will demonstrate how an insurer can structure its asset portfolio so that equity returns escape taxation at the corporate level.

Common Stocks and Dividend Policy

Under the current (1996) tax law, many individual investors pay higher tax rates on dividends than capital gains. Today's tax code prescribes a maximum marginal tax rate of 39.6% on dividends and 28% on realized capital gains.[15] Furthermore, since capital gains taxes are deferred until the stock is sold, the effective tax rate on capital gains could be much less than 28%.

Many economists contend that the higher effective tax rate on dividends implies that high-dividend stocks must be priced to offer a higher pre-tax rate of return than low-dividend stocks of equivalent risk.[16] This differential compensates for the tax disadvantage of dividends and provides that both types of stocks offer identical after-tax returns.

In this case, the implication to corporate stock purchasers is clear. The higher pre-tax return on high-dividend stocks allows the benefits of the corporate dividend exclusion to overwhelm the tax deferral on capital gains. Insurers should respond by selecting stocks with high-dividend payouts—for example, utilities, real estate investment trusts (REITs), and oil companies.

Another group of economists, led by the late Fischer Black [1], offers a different view. This group agrees that a large class of high-income investors would prefer to invest in companies with low-dividend payouts. Other investors, such as corporate investors with short investment horizons, would pay lower taxes on dividends than capital gains and would prefer high-dividend stocks. And tax-free investors would remain neutral, paying no taxes on either dividends or capital gains.

The proponents of the alternative theory argue that a wide enough variety of stocks already exists to satisfy investors of

---

[15] At the time of writing, the Taxpayers Relief Act of 1997 had not yet been finalized. The Act promises to lower the capital gains tax rate for certain long-term investments.

[16] See [2, pp. 430–431], for a detailed explanation. Also, Table 16-2, on p. 433, summarizes the research findings on the effect of dividend yield on returns.

any tax position. No company can increase (or decrease) its share price by modifying its dividend strategy; consequently, high-dividend stocks offer the same expected pre-tax return as low-dividend stocks of equivalent risk.

This alternative theory provides a simple prescription for all common stock investors, whether businesses or individuals: *given the equality of pre-tax returns on all stocks of the same risk, the equity investor should pick those stocks that minimize his or her taxes*.

For the corporate investor, the tax-minimizing stock selection depends on the investment time horizon. Suppose the corporation is merely looking for a short-term parking spot for some extra cash that will be needed in one year. The effective tax rate on dividends will be 10.5% (assuming 30% of dividends are taxed at 35%), while the effective tax rate on capital gains is 35%. This corporation should choose a high-dividend stock.

But the longer the corporation's investment time horizon, the lower the effective tax rate on capital gains becomes. For long-term corporate investors, low-dividend stocks become the investment vehicle of choice.

This suggests an optimal common stock strategy for the P/C insurer. The common stock portion of the investment portfolio is intended as a relatively permanent capital base for supporting current and future underwriting. In this sense, the insurer should select zero-dividend growth stocks, selling only as required to pay larger-than-expected insurance losses. If the insurer is forced to realize capital gains to pay insurance losses, these realized gains will be offset by underwriting losses, and still escape taxation.

Of course, an insurer that followed this strategy precisely would have no extra cash to pay out as shareholder dividends. Would this zero-payout strategy affect the value of the insurer's shares? Provided that the "dividends-are-irrelevant" school is right, the market value and pre-tax return of the insurance company would not be impacted by its dividend policy.

A Comment on Municipal (Tax-Exempt) Bonds and Preferred Stock

The discussion above concentrates on the broad asset classes of taxable bonds and common equity. Historically, insurers have also purchased large amounts of preferred stock and tax-exempt bonds. The standard investment approach has emphasized tax-exempt bonds and the dividends-received deduction to minimize federal income taxes and maximize net income [13, pages 4: 66–67].

Prior to the Tax Reform Act (TRA) of 1986, tax-exempt bonds and preferred dividends offered special tax advantages to insurers. The TRA of 1986 eliminated most of these tax advantages, rendering these investments inferior to other alternatives. For example, tax-exempt bonds should offer similar pre-tax returns as growth stocks of identical risk.[17] But the insurer will now be taxed on the tax-exempt bond income according to the proration and Alternative Minimum Tax (AMT) provisions of the tax code, while unrealized capital gains on the growth stocks remain tax-exempt.

Likewise, many insurance company portfolio managers in the past espoused the view that high-dividend stocks offered higher pre-tax yields. The TRA of 1986 equalized personal tax rates on dividends and realized capital gains, converting many investment managers to the Fischer Black school. The new mantra became minimizing taxes on stock purchases, and preferred stocks lost much of their original appeal.

---

[17] For the individual investor, municipal bonds are taxed at a lower rate than equity returns: municipal bond interest is tax-free, while investors still pay taxes on dividends and realized gains from stocks. Moreover, the vast majority of municipal bonds are held by individuals [3, p. 43]. Due to the personal tax advantages of municipals, these securities should actually offer a slightly *lower* pre-tax yield than common stock of equivalent risk. For a detailed proof, see [6, p. 26].

7. INTERACTION OF INVESTMENT POLICY AND INSURANCE PRICING

The Traditional DCF Pricing Model Under MM Assumptions

Myers and Cohn [18] describe a Discounted Cash Flow (DCF) insurance pricing model that is widely used for internal profitability studies and regulatory purposes. The model derives the fair insurance premium, which provides shareholders with an adequate expected rate of return. This fair premium is partly a function of the present value of expected losses and expenses resulting from the policy.

In theory, the discount rate used to capitalize expected losses and expenses should vary by line of business. Certain lines of business, such as credit or unemployment insurance, possess a high degree of systematic risk and deserve a correspondingly high discount rate. Most of the risk of other lines, such as crop/hail or earthquake, is diversifiable—the risk-free rate may be appropriate for these coverages.

Moreover, the theory presented in this paper demonstrates that the correct discount rate for insurance liabilities also depends on the insurer's asset allocation. Insurers with very risky investment strategies merit higher hurdle rates.

The DCF model also includes two adjustments to the fair premium. First, the present value of taxes on investment income from both policyholder-supplied and shareholder-supplied funds must be included in the fair premium. Second, the fair premium is reduced by the present value (PV) of the corporate tax shield from underwriting losses. This tax shield is calculated as 35% of expected underwriting losses.

In sum, the traditional DCF pricing formula calculates the fair premium as follows:

fair insurance premium

= PV of expected losses and expenses

+ PV of expected tax on investment income from policyholder and shareholder supplied funds

− PV of corporate tax shield from underwriting losses

Most DCF pricing models in practical use assume the MM (corrected) theory of debt and taxes. These models typically *assume* that (1) the MM model correctly describes the tax disadvantage of corporate lending, and (2) insurers invest solely in taxable bonds. As such, the common equity value of the insurer is reduced by the tax disadvantage of corporate investment.

But given the tax disadvantage of corporate lending in the MM world, why would the insurance company invest solely in taxable bonds? The discussion in Section 6 indicates that the safest strategy in this scenario would be to invest entirely in risk-free (zero beta) common stock, selling only enough shares at the end of the period to pay actual indemnity losses. The insurer will pay taxes in good years—for example when indemnity losses are less than expected (perhaps even leading to an underwriting profit), and investment returns are positive (shares sold to pay the indemnity losses are sold at a capital gain). In bad years, the insurer will earn tax carry-overs—for example, when indemnity losses are high (a large underwriting loss) and investment returns are negative (shares sold to pay indemnity losses generate a capital loss).

On average, expected underwriting losses will offset expected realized capital gains. Provided that all equity returns come as capital gains, and tax credits can be carried forward or back, the insurer's expected tax bill will be zero.

Thus, the present value of expected insurance company tax is zero. The DCF insurance premium is given by the present value of expected losses and expenses—there is no need to adjust the DCF premium for insurance company tax or tax shields.

The DCF Model Under Miller and Compromise Theories

Under the Miller theory of debt and taxes, the optimal investment strategy comprises 100% taxable bonds. Since Miller's model implies that there is no tax disadvantage to corporate lending, there is no need to include the present value of corporate taxes on investment income as part of the insurance premium. Likewise, corporate tax shields on debt create no value in Miller's world and the third term in the DCF equation drops off as well. The DCF pricing model then indicates that the market insurance premium equals the present value of expected losses and expenses.

Under the compromise theory, the optimal investment portfolio includes the proper proportion of both taxable bonds and growth stocks. As in the optimal MM strategy of all-equity investing, the insurer will pay taxes in some years and earn tax credits in other years. Provided the insurer has correctly estimated expected underwriting losses, the expected tax amount will be zero. As such, the market premium will also be given by the discounted value of expected losses and expenses.

An Illustrative Example

Table 1 provides an illustrative one-period example to demonstrate the impact of the optimal investment decision on the insurer's pricing decision. The model assumes that expected losses and expenses of $500 will be paid at the end of the period. The appropriate discount rate for the expected losses equals the risk-free rate of 6%. Surplus of $500 has been allocated to support the business. As noted in the previous two subsections, the fair premium is equal to the present value of expected losses and expenses, or $500/1.06 = $471.70.

The insurer will invest in some combination of risk-free common stock and taxable bonds. Since the insurer's assets and li-

TABLE 1

IMPACT OF INVESTMENT STRATEGY ON PRICING DECISIONS

		(1) MM	(2) Miller	(3) Compromise
a	Risk-free rate	6.0%	6.0%	6.0%
b	Expected losses and expenses at end of period	500	500	500
c	Surplus allocated to support policy	500	500	500
d	Premium collected at beginning of period = b/(1.0 + a)	471.70	471.70	471.70
e	Total assets at beginning of period = c + d	971.70	971.70	971.70
f	Taxable bonds at beginning of period	0.00	971.70	471.70
g	Expected return on taxable bonds	6.0%	6.0%	6.0%
h	Common stock at beginning of period = e − f	971.70	0.00	500.00
i	Expected return on common stock	6.0%	3.9%	4.5%
j	Common stock price per share at beginning of period	1.00	1.00	1.00
k	Common shares purchased = h/j	971.70	0.00	500.00
l	Expected share price at end of period = j × (1.0 + i)	1.06	1.04	1.05
m	Expected number of shares sold to pay losses	471.70	0.00	0.00
n	Expected number of shares not sold = k − m	500.00	0.00	500.00
o	Expected underwriting income = d − b	(28.30)	(28.30)	(28.30)
p	Expected taxable bond interest = f × g	0.00	58.30	28.30
q	Expected realized capital gains = m × i	28.30	0.00	0.00
r	Expected unrealized capital gains = n × i	30.00	0.00	22.50
s	Expected taxable income = o + p + q	0.00	30.00	0.00
t	Expected tax payment = 0.35 × s	0.00	10.50	0.00
u	Expected dollar return to shareholders = o + p + q + r − t	30.00	19.50	22.50
v	Expected percentage return to shareholders = u/c	6.0%	3.9%	4.5%
w	Required percentage return to shareholders	6.0%	3.9%	4.5%

abilities are *both* risk-free, the insurer's shareholders require a total return equal to the expected return that they could achieve from *other* risk-free common stocks. The marginal corporate tax rate is 35%.

Column 1 displays the expected tax and ROE in an MM world. As noted in Section 6, the insurer will invest entirely in risk-free growth stocks, selling only as necessary to pay losses and expenses. As shown in the table, expected realized capital gains are offset by expected underwriting losses and the $471.70 premium provides the insurer's shareholders with the required 6% return.

Next, recall that in the Miller world the expected return on risk-free common stock equals $r_f \times (1 - T_c)$. Therefore, the insurer's shareholders require an expected return of $6\% \times (1 - 0.35) = 3.9\%$. Also, recall that the optimum strategy holds only taxable bonds. As shown, this strategy provides shareholders with the required return of 3.9%.

Finally, we must specify T^* for the compromise world, which lies somewhere between 0% and 35%. Let's assume $T^* = 25\%$. Thus, risk-free common stock offers an expected return of $r_f \times (1 - 0.25) = 4.5\%$. As described in Section 6, the insurer's optimum strategy equates taxable bond interest and expected underwriting losses, with the balance invested in risk-free growth stocks. Again, this strategy will provide the required return to shareholders.

Is Insurance a Negative-NPV Transaction for the Insured?

In the popular version of the DCF model (that is, an MM world with all insurers investing 100% in taxable bonds) the PV of tax on investment income outweighs the PV of the corporate tax shield from underwriting losses. Therefore, the fair insurance premium exceeds the discounted value of expected indemnity losses and insurance expenses. From the policyholder's standpoint, the insurance purchase is a negative-NPV transaction.

But provided that the insurer follows an optimal investment strategy, the DCF model implies that the fair premium is given by the discounted value of expected losses and expenses under each of the three theories of debt and taxes.

Since each prospective policyholder pays a premium equal to the discounted value of expected losses and expenses, insurance is a zero-NPV transaction to the policyholder—assuming problems such as adverse selection and moral hazard are not significant, and that the insured could not handle the risks at a lower expense level than the insurance company.

Does Excess Capital Depress Insurance Prices?

In the traditional DCF model, the present value of taxes on the investment income from shareholder-supplied funds is included in the fair premium. The greater the marginal surplus required for a given insurer to write a policy, the higher this tax amount will be. This logic is the foundation of the popular idea that excess surplus contributes to lower pricing in our industry.

For example, assume two insurance companies are competing for the same account. Insurer A has excess capital and requires no additional marginal capital to write the account. Insurer B is already operating at a high premium-to-surplus ratio and requires $500 of additional capital. The traditional DCF model implies that the fair premium for insurer B is higher than the fair premium for Insurer A[18] by the discounted value of the investment income tax on the $500.

Yet, provided that Insurer B follows an optimal investment strategy, the additional capital required creates no tax disadvantage. Under all three theories of debt and taxes, the *expected* return on the additional capital equals the shareholders' *required* return.

[18]Assuming Insurer B has ready access to capital markets and ignoring issue costs.

For instance, assume that Insurer B decides to invest the marginal capital in risk-free securities. Table 2 summarizes the investment choice, expected return and required return under each of the theories:

TABLE 2

SUMMARY OF RETURNS

Theory	Investment Choice	Pre-tax Expected Return	After-tax Expected Return	Shareholders' Required Return
MM	Common stock	r_f	r_f	r_f
Miller	Taxable bonds	r_f	$r_f \times (1 - T_c)$	$r_f \times (1 - T_c)$
Compromise	Common stock	$r_f \times (1 - T^*)$	$r_f \times (1 - T^*)$	$r_f \times (1 - T^*)$

Does Diversification Create Value?

The traditional view in the insurance industry has held that a diversified underwriting portfolio reduces risk and creates value. As in the previous subsection, this view is based on the notion that there is a tax disadvantage associated with excess surplus. The greater the marginal surplus required to write a given policy, the higher the premium.

For instance, the loss experience for a given policy may be negatively correlated with the current loss exposures in an insurer's book. The marginal surplus required to support the policy may be very low, perhaps even negative. Under the traditional DCF assumptions, such a policy would be very attractive to the insurer: the insurer can offer a lower premium and still meet its financial goals.

According to this view, present values do not add up. The insurer must evaluate every policy as a potential addition to its current book of business. Underwriting decisions become extremely complex.

If we instead assume that every insurer follows an optimal investment strategy, then value additivity is restored. Insurance pricing is independent of marginal surplus, and every policy can be evaluated on its own merits. Diversification for its own sake does not increase value.

Optimal Asset Allocation—The Theory Versus Reality

Do insurers really follow an optimal investment strategy? If so, fair premium equals the discounted value of expected losses and expenses. Furthermore, the insurer's expected tax bill in both the MM and compromise worlds is zero, and the insurer pays no shareholder dividends.

However, insurance companies on average do pay taxes. Insurers also pay shareholder dividends. One possible explanation for this discrepancy may lie in investment laws and regulations imposed on the industry. For example, investment laws may preclude the insurer from holding stocks without established dividend records. Also certain laws and regulations may limit common stock holdings to a certain percentage of assets or surplus. This maximum amount may be below the theoretically optimal amount.

If these investment restrictions do indeed preclude insurers from holding optimal investment portfolios, competitive insurance premiums will adjust until insurance shareholders earn their required return. In this case, insurance premiums will exceed the discounted value of expected losses and expenses. Insurance would be a negative-NPV transaction to the insured, even in the absence of adverse selection and moral hazard problems. There would also be a moderate tax disadvantage to holding excess surplus, and insurers with excess surplus would price at a lower level.

Still, many insurers have the capacity to increase common stock holdings and enjoy increased tax advantages. Moreover, part of the industry's tax bill may result from excessive trading and unnecessary realized capital gains. Insurers are often motivated to realize capital gains unnecessarily in an effort to "dress up" the income statement. Statutory accounting rules allow realized capital gains to contribute to earnings, whereas unrealized capital gains are direct contributions to surplus. In an efficient market, investors see through transparent accounting conventions to real value. In this case, efforts to boost earnings through premature asset sales offer no benefit, while only resulting in higher taxes.

8. CONCLUSION

Actuaries are becoming more involved in the insurer's asset allocation decision. Recently, dynamic financial analysis (DFA) models have been utilized to maximize investment income and earnings subject to certain solvency constraints. But in an efficient market, the asset allocation decision is irrelevant. Indeed, any investment change that increases earnings will simultaneously increase the riskiness of those earnings, leaving share price unchanged.

When it comes to asset allocation decisions, a little financial theory may be much better than a thousand simulations. Specifically, one must specify the *source* of value from changing the asset mix. For instance, a riskier investment strategy may increase value by creating a loan subsidy from the guaranty fund.[19]

As noted earlier, only a financially troubled company would attempt to prop up share price at the expense of the guaranty fund or current policyholders. A better approach to the problem focuses on the impact of government taxation on the insurer's

---

[19]The guaranty fund mechanism, of course, was not intended to subsidize riskier investment strategies. To this end, regulators could take a page out of the Pension Benefit Guarantee Corporation's (PBGC) book, varying the guaranty fund assessment according to the riskiness of the insurer's asset portfolio.

optimum investment choice. This requires an understanding of the theories of debt and taxes, as well as the relationship between dividend yield and common stock returns. Section 6 of this paper analyzed these issues and recommended an optimum investment approach for each view of debt and taxes.

Moreover, under an optimum investment strategy, the fair premium for the property/liability policy will be given by the discounted value of expected losses and expenses—no adjustment will be required for the tax disadvantage of corporate investments.[20] This implies that the fair insurance premium is independent of the amount of surplus allocated to the policy.

Here we have an apparent contradiction with current actuarial practice. Actuaries expend a great deal of time and energy allocating surplus to line of business, profit center, etc., as part of the normal ratemaking process [5, pp. 541–547]. Furthermore, many pricing models take this exercise a step further and require the actuary to specify the release of this surplus over time [8, pp. 20–25]. Instead of directing so much energy to these endeavors, one may be better served to ask, "How can we modify our asset allocation to make surplus less tax-inefficient?"

Of course, more research still remains to determine the optimum asset portfolio for an insurer and the impacts of this portfolio on the actuary's financial pricing models. In closing, a quote from Myers and Cohn's classic DCF paper still remains relevant today [18, p. 65]:

> There is little in the insurance literature regarding the optimal asset portfolio, given taxes, for an insurance company. Are insurance companies' common-stock values reduced by the seeming tax disadvantage associated with corporate purchases of taxable marketable

---

[20]The traditional view has held that the fair insurance premium must include a provision for the tax disadvantage of corporate investments, even assuming the insurer has adopted an optimum asset allocation. See, for instance, Derrig's recent paper. [7]

securities? … The present-value approach as it is employed in this report … [is] probably not exactly correct in specifying fair insurance premiums, and it is not clear just how the approach should be modified so as to take corporate taxes properly into account.

REFERENCES

[1] Black, Fischer and Myron S. Scholes, "The Effects of Dividend Yield and Dividend Policy on Common Stock Prices and Returns," *Journal of Financial Economics*, May 1974, pp. 1–22.

[2] Brealey, Richard A., and Stewart C. Myers, *Principles of Corporate Finance*, Fifth Edition, McGraw-Hill, 1996.

[3] Cummins, J. David, and Elizabeth Grace, "Tax Management and Investment Strategies of Property-Liability Insurers," *Journal of Banking and Finance*, 18, 1994, pp. 43–73.

[4] Collis, David J., and Cynthia A. Montgomery, "Competing on Resources: Strategy in the 1990s," *Harvard Business Review*, July–August 1995, pp. 118–128.

[5] D'Arcy, Steven P., "Special Issues," *Foundations of Casualty Actuarial Science*, Third Edition, pp. 541–576.

[6] DeAngelo, Harry, and Ronald W. Masulis, "Optimal Capital Structure Under Corporate and Personal Taxation," *Journal of Financial Economics*, March 1980, pp. 5–29.

[7] Derrig, Richard A., "Theoretical Considerations of the Effect of Federal Income Taxes on Investment Income in Property-Liability Ratemaking," *The Journal of Risk and Insurance*, 61, 1994, pp. 691–709.

[8] Feldblum, Sholom, "Pricing Insurance Policies: The Internal Rate of Return Model," CAS Part 10 Study Note, May 1992, pp. 1–63.

[9] Ferrari, J. Robert, "The Relationship of Underwriting, Investments, Leverage, and Exposure to Total Return on Owners' Equity," *PCAS* LV, 1968, pp. 295–302.

[10] Galai, Dan, and Ronald W. Masulis, "The Option Pricing Model and the Risk Factor of Stock," *Journal of Financial Economics*, 3, 1976, 53–81.

[11] Ibbotson Associates, Inc., *Stocks, Bonds, Bills and Inflation, 1997 Yearbook*, Ibbotson Associates, 1997.

[12] "It Was a Very Mixed Year," *Best's Review*, *Property/Casualty Edition*, May 1997, pp. 36–38.

[13] Maginn, John L., and Donald L. Tuttle, *Managing Investment Portfolios: A Dynamic Process*, Second Edition, Warren, Gorham & Lamont, 1990.

[14] Merton, Robert C., *Continuous-Time Finance*, Revised Edition, Blackwell Publishers Inc., 1992.

[15] Miller, Merton H., "Debt and Taxes," *Journal of Finance*, May 1977, pp. 261–276.

[16] Modigliani, Franco, and Merton H. Miller, "The Cost of Capital, Corporate Finance and the Theory of Investment," *American Economic Review*, June 1958, pp. 261–297.

[17] Modigliani, Franco, and Merton H. Miller, "Corporate Income Taxes and the Cost of Capital: A Correction," *American Economic Review*, June 1963, pp. 433–443.

[18] Myers, Stewart C., and Richard A. Cohn, "A Discounted Cash Flow Approach to Property-Liability Insurance Rate Regulation," *Fair Rate of Return in Property-Liability Insurance*, Klower-Nijhoff, 1987, pp. 55–78.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXVII

REINSURER RISK LOADS FROM MARGINAL SURPLUS REQUIREMENTS

RODNEY E. KREPS

DISCUSSION BY PAUL J. KNEUER

1. INTRODUCTION

Writing an insurance risk increases the variability of an insurer's results. This has direct economic costs to the insurer, such as not being able to write other attractive risks or to comfortably maintain the desired degree of risk in its asset portfolio. Extra risk also reduces the value of its future profits in the capital market, that is, its stock price or similar valuation. Insurers require premiums that allow enough expected profit to overcome these costs.

In "Reinsurer Risk Loads from Marginal Surplus Requirements" and "Investment-Equivalent Reinsurance Pricing," published in this volume, Rodney Kreps has increased our understanding of how a reinsurer (actually, any insurer) commits its capital to risks. Based on simple microeconomic assumptions and consequential expressions, Kreps has developed powerful models relating an insurer's capitalization and a prospective contract's risk profile to develop a minimum acceptable premium for the contract. Premiums below this minimum cause an insurer to dilute its earnings and should be rejected. Even though reinsurers do not appear to manage their capital on an individual contract basis as Kreps' calculations assume, the model has received the highest actuarial compliment: it is actively used to price business. This is notable at several of the newly established catastrophe reinsurance markets.

2. THE KREPS MODEL

Fill your bowl to the brim and it will spill.
Keep sharpening your knife and it will blunt.

— The *Tao*[1]

Kreps begins by assuming that every insurance contract can be uniquely associated with a marginal amount of an insurer's surplus. This amount is computed assuming that each insurer selects and maintains a certain probability of ruin and then finds the amount by which its surplus must increase to maintain that probability if a proposed contract is written. For proposed contracts that are small relative to the insurer's existing business, this is equivalent to requiring that the ratio of the insurer's surplus to the standard deviation of its results does not change after the contract is written. A proposed contract must have an expected profit that adequately rewards the required marginal amount of surplus, or else it would dilute the insurer's return and is thus declined.

To see how Kreps' results are used, recall his Equation 2.4,

$$\text{Minimum premium for a contract} = \mu + \sigma\mathscr{R} + E - yB/(1+y),$$

where μ and σ are the mean and standard deviation of the losses on the proposed contract and y is the insurer's target return on equity. E is the insurer's marginal expense. B is the "bank," if any, that the insured has "built up." Kreps defines $\mathscr{R}$ as the insurer's "reluctance" to assume additional degrees of risk.

---

[1]See Mitchell [11, Chapter 9]. The *Tao*, literally the "Way," is a short collection of Chinese philosophical writing that was probably first gathered in the sixth century B.C., but which is still influential for its simple, natural, and organic way of describing human perception and behavior.

This excerpt and others from Mitchell's readable translation help illustrate the forces that insurers must seek to balance. Readers may want to consider that critical thinking about focus, competition, success, and control is much older and deeper than our current microeconomic analysis of insurers.

For high-level catastrophe coverage on national ceding companies, each proposed contract has a correlation close to 1, since these treaties are only exposed by a very few physical hazards that already expose the reinsurers' other business. Kreps argues that for contracts that are relatively small additions to the reinsurers' portfolio, the average ratio of marginal surplus to marginal standard deviation is equal to the average ratio. With a correlation of 1, and using z to represent the ratio of the insurer's current surplus to the standard deviation of the losses on its existing portfolio, Kreps' reluctance becomes:

$$\mathscr{R} = yz/(1+y). \tag{2.1}$$

At this writing, few reinsureds claim large positive "banks" (to be kind) and few reinsurers see "banks" as economic rather than rhetorical obligations; so the B term is ignored as well, and Kreps' conclusion (with a simple substitution) shows that the minimum premium must be:

$$\mu + (yz/(1+y))\sigma + E. \tag{2.2}$$

This minimum premium has two contract-specific terms: the expected losses plus a charge for the marginal contribution to the insurer's standard deviation. The latter term can be thought of as an interest rate, $y/(1+y)$, applied to a marginal amount of surplus, $z\sigma$. The two terms are generally independent. This conclusion is in sharp contrast to earlier actuarial theory and practice, which based risk charges either directly on the expected losses (incurred or unpaid) or, indirectly, through the premiums on notional allocations of surplus.

The values of $\mathscr{R} = yz/(1+y)$ are similar at many catastrophe reinsurance markets and share derivations based on similar views of acceptable ruin scenarios and required returns to capital. While it is inappropriate to detail specific market pricing in an industry forum, I can also note that prices often show similar variations in the relative contribution of the μ and $\mathscr{R}\sigma$ terms for different contracts. The following four examples are several brokers' consensus estimates of prices at a recent date for different property catastrophe reinsurance layers for a hypothetical U.S. nationally-exposed cedant, expressed as annual rates-on-line (ROL):[2]

Annual Layer and Retention ($ Millions)	Layer Penetration Recurrence Time (Years)	Estimated Price (ROL)	Pure Premium (ROL)	Standard Deviation (ROL)	Implied Loss Ratio
10 xs 10	10	18.00%	10.0%	30.0%	55.5%
10 xs 20	40	9.50%	2.5%	15.6%	26.3%
30 xs 30	100	5.00%	1.0%	9.9%	20.0%
40 xs 60	1,000	2.25%	0.1%	3.1%	4.4%

For many cedants, the expected loss ratios vary across the different layers of their programs by this factor of ten or more, decreasing in the higher layers as the risk charge contribution takes on more importance compared to the expected losses. Kreps' formula easily explains this and other surprising[3] variations in risk loads visible in the current reinsurance market.

[2]The annual losses to each layer are approximated as a binomial process. Pure Premium = 1/Recurrence Time. Standard deviation is the square root of the product of pure premium and the complement of pure premium.

[3]Another excellent reinsurance example where Kreps' approach improves our understanding of current pricing is "second event" covers. Reinsurance actuaries frequently treat prices expressed as rate-on-line as if they were probabilities. This common shorthand cannot explain second-event cover prices.

For the hypothetical cedant reviewed earlier, a $10 million excess $10 million second-event layer (i.e., pays up to $10 million for a second loss during the year in excess of $10 million) has a pure premium of approximately 1/2% (as a rate-on-line). Using Poisson assumptions, the pure premium for the original layer, when expressed as a rate-on-line, is actually the probability of one or more losses; so the complement of the pure premium is the probability of no losses, here 90%. This produces a Poisson frequency of $-\ln(90\%) = 10.54\%$. It follows that the probability of exactly one loss is $90\% \times 10.54\%$ or 9.48%. (The Poisson probability of exactly one loss is the probability of no losses times the frequency.) The probability of no more than one loss is 99.48%. The probability of two or more losses is 0.52%.

Many actuaries are tempted to perform a similar calculation on the price for the original layer, which is an 18% rate on line. If 18% is a fair compensation for assuming the risk of one or more losses to the layer, including the value of assuming the variability in the layer, then 82.0% is the consistent price for a "no losses" cover. Continuing with this common logic, and treating the price as a risk-adjusted probability, produces a Poisson "risk-adjusted" frequency of 19.8%; and the risk-adjusted price for coverage of exactly one loss is 16.3%. Thus the price for a second-event coverage, under this logic, would be

The Kreps minimum premium formulation is clear, understandable and powerful. It also avoids the problematic assumptions[4] needed to allocate an insurer's total surplus to product. However, the model ignores important considerations:

- Other things equal, an insurer prefers to reduce its probability of ruin below the current level.
- If marginal results are very attractive, an insurer may choose to grow and increase its probability of ruin beyond the current target.
- The market capitalization rate applied to an insurer's future profits must depend on the kind and amount of business that the insurer assumes.
- Insurers identify and separately manage distinct risk categories, such as lines of business and exposure zones. They do not directly examine the covariance between a proposed contract and their entire existing portfolios. This is particularly true for catastrophe reinsurers that analyze contracts using modern event-modeling software.
- Insurers do not always calculate unlimited means of the losses for their contracts. They generally evaluate expectations only over scenarios with realistic probabilities. For example, current event-modeling software includes only foreseeable events with

---

1.7% (100%, less 82.0% for the value of "covering" the no loss case, less 16.3% for the value of exactly one loss.) In financial economics, this probability-like measure is called a *martingale*, and some recent research has developed utility functions that produce risk loaded prices that are martingales.

Unfortunately, brokers agree that in today's market this second-event cover would actually cost something above a 5% rate-on-line. Real-world prices are not martingales and the common arithmetic of treating a reinsurance rate-on-line as a risk-adjusted frequency is empirically wrong. Kreps' approach correctly indicates the higher price by noting that the standard deviation of the second-event layer is above that of the third excess layer and is even more highly correlated with reinsurers' results. The risk load must be a significantly greater part of the limit than the 4% in that higher layer (5% price less 1% pure premium).

Second-event cover pricing, as well as the up-front discount for a mandatory 100% reinstatement premium, is strong empirical evidence for a formulation like Kreps'.

[4] See Kneuer [6], Miller and Rapp [10], Roth [12], and Bass and Khury [1].

estimated annual recurrence probabilities above 10^{-4} to 10^{-7}. Less frequent (or apparent) events are omitted, so the reported means are understated. There is no theoretical reason why the unlimited mean even needs to be finite.

- Kreps' process is circular.[5] Insurers evaluate proposed contracts based on their expected return on marginal surplus. But the marginal surplus requirements depend on the order in which proposed contracts are evaluated. In Kreps' calculation, an insurer compares each proposed contract's contribution to the variance of a portfolio consisting of every other current contract. This is equivalent to assuming, a priori, that each contract is equally desirable. That may not be the case because some contracts may be selected before others. A different amount of imputed marginal surplus will be found when the comparison base is some contracts, rather than all. Different minimum premiums result.

These considerations matter and the users of Kreps' formula need to consider how the limitations in his assumptions may distort their analyses. Fortunately, the distortions are not fatal. We can avoid the first five considerations listed above. Starting with similar, but broader assumptions, a more realistic model can support results much like Kreps' contributions. However, the last concern, circularity, is not directly avoided (at least, not yet).

Let's explore a model that allows an insurer the flexibility to pick a portfolio of risks so as to adjust its level of risk compared to its capital base. The alternative model, like Kreps', will omit tractable real world considerations including taxes and reserves and their associated investment income. A more complete model would reflect multiple risk factors, but is also deferred here for simplicity.

---

[5]See Gogol [5] and Mango [9] for illustrations of these differences and a suggested cure to the problem.

3. AN ALTERNATIVE MODEL

Nothing in the world is as soft and yielding as water.
As for dissolving the hard and inflexible nothing can surpass it.[6]

An insurer's job, like any firm's, is to maximize its worth, the expected present value of its future free cash flows. This present value reflects the riskiness of its business, among other things. We simplistically assume that the insurer's management's only decision variables are the portion of each proposed contract that the insurer will assume. That is, it can choose to assume between 0% and 100% of each contract that has been offered to it. Further, with Kreps, we assume that the insurer is a price-taker. Its individual decision does not change the price at which a contract is offered.

Like Kreps, let's also assume that our insurer is looking forward one period and examining how a proposed contract changes its probability of ruin. However, we do not assume an inflexible maximum probability of ruin. Instead we consider a fluid distribution of the insurer's future value to its owners (stockholders, or policyholders if a non-stock insurer). We use a nearly linear relationship to (GAAP) surplus that seems close to current market valuations:

$$V_1 = \text{Value of Insurer (at } t = 1)$$
$$= \begin{cases} 0, & \text{if} \quad \text{Surplus is less than some value, } G_1; \\ M \times \text{Surplus}, & \text{if} \quad \text{Surplus} \geq G_1. \end{cases} \tag{3.1}$$

Like a shark, when an insurer stops moving, it drowns. G_1, which is significantly greater than zero, represents the $t = 1$ surplus level below which the insurer ceases to be a going concern. The discontinuity point is likely much less than S_0, the current sur-

[6]The *Tao* [11, Chapter 78].

plus level. M, the insurer's book-value multiple,[7] is a number that is rarely below 1.0 nor often as high as 3.0.

If there were an absolutely efficient market in insurer capital, then $G_1 = 0$ and $M = 1$, because V_1 would always equal S_1 (or zero, when S_1 is negative). However, regulation and clients' security concerns limit the flexibility to move capital through insurers[8] and this allows market valuations higher than book values ($M > 1$). Let us assume that the *franchise value*, this capitalized value above the break-up value, $(M - 1)S_0$, is positive and much larger than the expected value of the losses that might be avoided by bankruptcy. Our insurer is solid now and underwrites believing it will stay that way.

We can analyze our insurer's microeconomic underwriting decision in light of this more general model. Our insurer has already selected a portfolio of contracts with premium P and random losses L, with known expectation, $\mathrm{E}(L)$. Our insurer is considering a new contract with premium p and random losses ℓ. The insurer will choose to insure some part of the risk, Q, between 0 and 1.

Thus its total premium will be $P + Qp$ and its total losses will be $L + Q\ell$. Ignoring investments, taxes and operating expenses for simplicity here, and assuming that the proposed contract expires in time for the measurement of the insurer's value that we assume occurs at $t = 1$, we find that the insurer's final surplus is:

$$S_1 = \max(0, S_0 + P - L + Q(p - \ell)), \tag{3.2}$$

where L and ℓ are random variables and S_0, P, and p are known to the insurer.

Consider how our insurer looks at the distribution and expectation of the net present value (NPV) of its total future value,

[7] Investment bankers often use book-value multiples for valuations of P/C insurers because these multiples are more stable than Price/Earnings ratios and also control for leverage differences.

[8] See Kneuer [7].

including the franchise value:

$$\text{NPV} = \begin{cases} 0, & \text{if} \quad S_1 < G_1; \\ \text{NPV}(M \times S_1), & \text{if} \quad S_1 \geq G_1. \end{cases} \tag{3.3}$$

$$\text{E(NPV)} = M \times \text{Prob}(S_1 \geq G_1) \times \text{E}(S_1 \mid S_1 \geq G_1)/(1+y). \tag{3.4}$$

Abbreviating, $\text{E(NPV)} = M \times \text{Z} \times \Psi/(1+y)$, where Z is the probability that $S_1 \geq G_1$, Ψ is the conditional expectation of S_1, given that $S_1 \geq G_1$, and y is Kreps' assumed management target yield rate, which approximates in concept the appropriately risk-adjusted market discount rate in effect between $t = 0$ and $t = 1$.

For any price on the proposed contract, our insurer seeks to maximize its current worth, its expected NPV, by choosing a value of Q. It will maximize this market value by differentiating E(NPV) over Q and examining the derivative at $Q = 0$. If the derivative is positive, the insurer will decide to assume at least some of the proposed contract. The insurer will decide to assume more as long as this derivative remains positive at higher values of Q:

$$d/dQ\,\text{E(NPV)} = d/dQ(\text{Z} \times M \times \Psi/(1+y)) \tag{3.5}$$

$$= M\left(\frac{\text{Z}'\Psi}{1+y} + \frac{\Psi'\text{Z}}{1+y} - \frac{\Psi \text{Z} y'}{(1+y)^2}\right) \tag{3.6}$$

where Z', Ψ', and y' are derivatives with respect to Q.

The interpretation of this formula is direct. The present value that our insurer expects to add (or subtract) by writing some of the proposed contract (in other words, the derivative with respect to Q) is:

- the increase (decrease) in the probability of remaining a going concern (Z') times the current present value of the firm as a going concern ($M\Psi/(1+y)$), plus

- the increase (decrease) in the current present value of the firm as a going concern ($M\Psi'/(1+y)$), times the current probability of remaining a going concern (Z), minus
- the current present value of the firm ($MZ\Psi/(1+y)$) times the relative increase (decrease) in the risk-adjustment in the market discount rate ($y'/(1+y)$).

When is d/dQ of E(NPV) positive? Since $1+y$ and M are both greater than zero for any conceivable insurer, this derivative has the same sign as:

$$Z'\Psi(1+y) + Z\Psi'(1+y) - Z\Psi y'. \tag{3.7}$$

$d/dQ\,$E(NPV) will be positive whenever

$$Z\Psi'(1+y) > \Psi Z y' - \Psi Z'(1+y). \tag{3.8}$$

Or since $Z > 0$,

$$\Psi' > \frac{\Psi Z y' - \Psi Z'(1+y)}{Z(1+y)} \tag{3.9}$$

$$= \Psi\left(\frac{y'}{1+y} - \frac{Z'}{Z}\right). \tag{3.10}$$

But Ψ' is just the increase in the expected value of the insurer (assuming it survives) caused by assuming some of the proposed contract.

Define $\hat{\mu} = \mathrm{E}(\ell \mid S_1 \geq G_1)$, the limited expectation of the losses on the proposed contract, given that our insurer is not impaired. Many insurers implicitly calculate something like $\hat{\mu}$ by modeling contract losses only under certain not-too-extreme scenarios.[9]

[9]For example, event-modeling software analyzes the probabilities of the 1906 San Francisco earthquake and the 1938 New England Hurricane, but not of the 1906 earthquake recurring in Boston! For some familiar loss distributions, such as the Pareto, μ may not be finite, so a limited mean is essential for any pricing analysis.

This definition allows us to find Ψ':

$$\Psi(Q) = \mathrm{E}(S_1 \mid S_1 \geq G_1) = \mathrm{E}(S_0 + P - L + Qp - Q\ell \mid S_1 \geq G_1) \tag{3.11}$$

$$= \mathrm{E}(S_0 + P - L \mid S_1 \geq G_1) + Q\mathrm{E}(p - \ell \mid S_1 \geq G_1), \qquad \text{or} \tag{3.12}$$

$$= \Psi(0) + Q(p - \hat{\mu}), \qquad \text{and thus} \tag{3.13}$$

$\Psi' = p - \hat{\mu}$. Substituting, we find that d/dQ of E(NPV) is positive when

$$\Psi' = p - \hat{\mu} > \Psi(y'/(1+y) - \mathrm{Z}'/\mathrm{Z}), \qquad \text{or,} \tag{3.14}$$

$$p > \hat{\mu} + \Psi(y'/(1+y) - \mathrm{Z}'/\mathrm{Z}). \tag{3.15}$$

If p is greater than the right-hand side then our insurer would accept more of the proposed contract, expecting to increase its own present value. This offers a minimum premium for the proposed contract without the concept of marginal surplus. The calculation also allows the probability of ruin to vary. The minimum premium is equal to the sum of:

- the losses that our insurer expects from the proposed contract, ignoring here any loss scenario that would impair it, plus,
- its expected amount of surplus (at $t = 1$) multiplied by
  - the relative increase in the discount on that future surplus caused by adding the risk of the additional contract, and
  - the relative decrease in the probability of surviving as a going concern, reflecting here those extreme loss scenarios not considered above in $\hat{\mu}$.

Like Kreps' result, this minimum premium consists of the sum of expected losses and a "reluctance" term that is positively related to the insurer's surplus and the variability of the proposed contract. $\Psi y'$ is analogous to Kreps' $z\sigma$; however, there is an additional component here ($\Psi Z'/Z$) reflecting the reduction in the probability of survival. While $\Psi Z'/Z$ is denominated in terms of the insurer's surplus, it is not Kreps' marginal surplus. Kreps allocates an amount of surplus to a contract, and while it is not expected to be lost, a marginal return is required because the surplus cannot be allocated to other uses. $\Psi Z'/Z$ is the expected surplus that will be lost by taking on the risk of the proposed contract. The minimum premium includes a charge for this expected loss of capital, not a return on it. The difference here is that the charge reflects principal lost to default ($\Psi Z'/Z$), versus only interest on principal outstanding (Kreps' $(y/(1+y))z\sigma$).

Under the alternative model, the reluctance term and the expected losses term are distinct and not in general dependent upon each other, as Kreps has also found. We will next examine how (re)insurers consider the risk of a proposed contract under this more general model of incentives. Then we will see separately how the marginal risk changes the probability of survival and the market discount rate. Combining these results produces a usable minimum premium under the more general model. The result has a strong symmetry with Kreps' simpler formula.

4. WHAT IS THE MARGINAL RISK OF A PROPOSED CONTRACT?

Think of the small as large and the few as many.
Confront the difficult while it is still easy;
accomplish the great task by a series of small acts.[10]

For simplicity, we have assumed that insurers are only concerned with one risk factor. (For national U.S. catastrophe reinsurance accounts that is a fair approximation.) Let's denote this one risk factor by R, and assume that it is a real-valued random variable that fully describes all of the common elements of risk in the insurer's portfolio. For simplicity here, let's also re-scale R to be positively correlated with L and have a standard devia-

[10] The *Tao*, [11, Chapter 63].

tion of one.[11] Further assume that the cumulative density function (c.d.f.) of R is continuous over its range, except perhaps at a finite number of points.

For the loss processes of the existing portfolio (L) and the proposed contract (ℓ), define L_0 and ℓ_0, the unsystematic parts of the loss processes, the parts that can't be explained by R.

$$L_0 = L - \frac{C}{\mathrm{Var}(R)}R, \qquad \text{and} \tag{4.1}$$

$$\ell_0 = \ell - \frac{c}{\mathrm{Var}(R)}R, \tag{4.2}$$

where $C = \mathrm{Cov}(L,R)$ and $c = \mathrm{Cov}(\ell,R)$.

Since we have assumed a single risk factor, $\mathrm{Cov}(L_0,\ell_0) = 0$. (If not, there is another external factor that affects at least two contracts and that must be quantified. We've restricted ourselves to a one-factor model for now.) It's also easy to show that $\mathrm{Cov}(L_0,R) = \mathrm{Cov}(\ell_0,R) = 0$. Finally, observe that L_0 is the sum of the many unsystematic risk elements of the contracts in the insurer's current portfolio: L_0 and $L_0 + Q\ell_0$ are normally distributed. While $L + Q\ell$ will not necessarily be normal, its c.d.f. will be continuous and differentiable for changes in either the mean or standard deviation of the loss process.

Our insurer has defined R to decompose the loss processes into a quantified external risk factor and the unsystematic part of its risks. It understands $\mathrm{Var}(L)$ in terms of the variances and covariances of L_0 and R.

$$\mathrm{Var}(L) = \mathrm{Var}(L_0) + [C/\mathrm{Var}(R)]^2\mathrm{Var}(R). \tag{4.3}$$

[11] For property catastrophe reinsurance coverage, this R might mean something like "aggregate insured property losses in the U.S. during the next year, in excess of \$2 billion per event, divided by \$5 billion." If the standard deviation of the annual excess losses is \$5 billion, as roughly true in the last decade, this variable has a standard deviation of one as required.

And similarly for $\mathrm{Var}(\ell)$:

$$\mathrm{Var}(\ell) = \mathrm{Var}(\ell_0) + [c/\mathrm{Var}(R)]^2\mathrm{Var}(R). \tag{4.4}$$

We can use these expressions to find the new portfolio variance, if our insurer assumes Q of the proposed contract.

$$\mathrm{Var}(L + Q\ell) = \mathrm{Var}(L) + Q^2\mathrm{Var}(\ell) + 2Q\,\mathrm{Cov}(L,\ell). \tag{4.5}$$

Since L_0 and ℓ_0 are uncorrelated with each other and R, we find

$$\mathrm{Cov}(L,\ell) = \mathrm{Cov}\left(L_0 + \frac{CR}{\mathrm{Var}(R)}, l_0 + \frac{cR}{\mathrm{Var}(R)}\right) \tag{4.6}$$

$$= \mathrm{Cov}\left(\frac{CR}{\mathrm{Var}(R)}, \frac{cR}{\mathrm{Var}(R)}\right) \tag{4.7}$$

$$= \frac{Cc}{\mathrm{Var}(R)^2}\mathrm{Cov}(R,R) = \frac{Cc}{\mathrm{Var}(R)^2}\mathrm{Var}(R) \tag{4.8}$$

and can substitute to show that

$$\begin{aligned}\mathrm{Var}(L + Q\ell) = {} & \mathrm{Var}(L_0) + [C/\mathrm{Var}(R)]^2\mathrm{Var}(R) + Q^2\mathrm{Var}(\ell_0) \\ & + Q^2[c^2/\mathrm{Var}(R)]^2\mathrm{Var}(R) \\ & + [2QCc/\mathrm{Var}(R)^2]\mathrm{Var}(R).\end{aligned} \tag{4.9}$$

Recall that we re-scaled R to make $\mathrm{Var}(R) = \mathrm{SD}(R)^2 = 1^2 = 1$, so

$$\mathrm{Var}(L + Q\ell) = \mathrm{Var}(L_0) + C^2 + Q^2\mathrm{Var}(\ell_0) + Q^2c^2 + 2QCc. \tag{4.10}$$

We can differentiate with respect to Q and find the marginal risk, which is the rate of increase in our insurer's portfolio variance with respect to changes in Q,

$$d/dQ\,\mathrm{Var}(L + Q\ell) = 2Q\,\mathrm{Var}(\ell_0) + 2Qc^2 + 2Cc. \tag{4.11}$$

5. MARGINAL RISK AND THE PROBABILITY OF SURVIVAL

If you realize that all things change,
There is nothing you will try to hold onto.
If you aren't afraid of dying,
There is nothing you can't achieve.[12]

One of the terms of the minimum premium is $\Psi Z'/Z$. To calculate this term, we need to see how Z, the probability of survival, changes with the marginal risk from assuming more of the proposed contract. Since G_1, P, and S_0 don't vary with Q, we can also view Z as Z^*, a function of only the mean (Ψ) and standard deviation (Λ) of the surplus amount. Within our assumptions both parameters depend on a single variable, Q,

$$Z(P,L,Q,p,\ell,G_1,S_0) = Z^*(\Psi(Q\ell,Qp),\Lambda(Q\ell)). \tag{5.1}$$

We find

$$\frac{dZ^*}{dQ} = \left(\frac{\partial Z^*}{\partial \Psi}(p-\hat{\mu}) + \frac{\partial Z^*}{\partial \Lambda}\frac{d\Lambda}{dQ}\right). \tag{5.2}$$

Clearly, greater resources always improve the probability of survival, so $\partial Z^*/\partial \Psi$ is positive; and increasing levels of variability can increase the chance of ruin, so $\partial Z^*/\partial \Lambda$ is negative, at least for the range of Z values that concern us, fairly solid companies.[13]

[12] The *Tao*, [11, Chapter 74].

[13] An unstable company may actually increase its survival probability by adding variance. The non-linear valuation caused by a floor of zero is equivalent to the shareholders owning an out-of-the-money put option. Any option increases in value as the variability increases. See Brealey and Myers [3, p. 498]. A more familiar illustration may be the "Hail Mary" passes by losing football teams during the last minutes of a game. While the marginal expected value of these plays (in yards) is very small, the increased variability significantly increases the teams' small probabilities of victory. It is equally important to note that winning teams don't throw "Hail Marys." They often "run out the clock," sacrificing all marginal gain to eliminate variance. As with our sound insurers, they believe that their $\partial Z/\partial \Lambda$ vastly outweighs $\partial Z/\partial \Psi$.

The standard deviation component depends on Q more strongly than the mean term does for two reasons: (1) $p - \hat{\mu}$ is small, while $d\Lambda/dQ$ is at least $2Cc$; and (2) for these solid companies, there's just more room for Z to go down than up. So by analyzing the Λ term the insurer sees whether Z'/Z (and thus the minimum premium) rises or falls as Q is increased above zero.

Evaluating the derivative of the variance (in Equation 4.11) at $Q = 0$,

$$d/dQ\,\mathrm{Var}(L + Q\ell) = 2Cc, \tag{5.3}$$

which is positive except in the not-often-found-in-nature case where the signs of C and c differ, a contract that could serve as a hedge against the existing book. Barring this curiosity, this contribution of Q to the portfolio variance is always positive. The sign of the derivative of the standard deviation is the same as the sign of the derivative of the variance.[14] So this result is also true for the marginal standard deviation of the combined loss process with respect to changes in Q: $d\Lambda/dQ$ is positive at $Q = 0$.

The non-systematic part of the risk of a proposed contract does not initially contribute to the marginal variance at all. But when Q is greater than zero, two additional positive terms ($Q\,\mathrm{Var}(\ell_0) + Qc^2$) add to the marginal variance. With a little more arithmetic, it is easy to show that the second derivative of the standard deviation with respect to Q is always positive. The marginal standard deviation is at its minimum at $Q = 0$ and increases monotonically and rapidly thereafter. As Q grows, the marginal standard deviation grows, and Z', the change in the probability of survival with respect to Q becomes more negative quickly. The ratio Z'/Z is monotonically decreasing,[15] at least for these solid companies.

[14] $\Lambda = \mathrm{Var}^{1/2}$. Differentiating with respect to Q, $\Lambda' = \frac{1}{2}(1/\Lambda)\mathrm{Var}'$. $\frac{1}{2}(1/\Lambda)$ is positive. The signs of Λ' and Var' are the same.

[15] If the loss processes are normal there is a direct proof of this conclusion. We have assumed that the probability of survival is quite high, so that the standard deviation has much more influence on changes in the survival probability than does the mean. We are

When we examine y', the change in the market discount rate caused by adding marginal risk, we will find that it cannot change the value of Q at which the minimum premium is lowest. The minimum premium at $Q = 0$ is its lowest value.

6. MARGINAL RISK AND THE MARKET DISCOUNT RATE

Money or happiness: which is more valuable?
Success or failure: which is more destructive?[16]

Under our frictionless, one-period, one-factor assumptions, the Capital Asset Pricing Model dictates[17] the market discount rate applied to the future earnings of our insurer,

$$y = r_f + \beta\Pi, \tag{6.1}$$

where r_f is the risk-free interest rate in effect between $t = 0$ and $t = 1$, β is the systematic risk of our insurer, and Π is the market risk premium.

---

interested in the derivative of the ratio Z'/Z with respect to changes in Q. $Z' = d/dQ\,Z$ is a function of the mean and Λ. However, the influence of the mean is approximately zero, so we can treat Z as a function only of Λ. Now we apply the chain rule

$$d/dQ(Z'(Q)/Z) = d/d\Lambda(Z'(\Lambda)/Z(\Lambda))\,d\Lambda/dQ,$$

with $Z'(\Lambda)$ evaluated at $\Lambda = \Lambda(Q)$. We have seen that Λ is an increasing function of Q, so $d\Lambda/dQ > 0$, and the sign of the derivative of Z'/Z with respect to Q is the same as the sign of the derivative with respect to Λ.

The derivative of a quotient has a positive denominator so $d/d\Lambda\,(Z'/Z)$ will be negative wherever $Z''Z - (Z')^2$ is. It suffices to show that $Z'' < (Z')^2$ because Z is no more than one. By differentiating the c.d.f. of the normal with respect to Λ twice, squaring the first derivative, and expanding both in powers of Λ we can compare Z'' and $(Z')^2$. When $S - G_1 + P$ is sufficiently large compared to $E(L)$ (greater than the mean = median more than suffices, i.e., a survival probability of at least 50%), we can compare and conclude that Z'' is always less than $(Z')^2$. So Z'/Z is a monotonically decreasing function of Λ, and thus also of Q.

Charles A. Thayer helped develop this proof and other derivations in this review.

[16] The *Tao*, [11, Chapter 44].

[17] These assumptions from Kreps, taken with familiar and reasonable assumptions about rationality, risk-free borrowing, and available information, meet the requirements of the CAPM. We conclude that it will apply here.

To understand the β of our insurer, we need to examine its market-based rate of return between $t = 0$ and $t = 1$

$$r(i) = \frac{V_1}{V_0} - 1 \tag{6.2}$$

$$= \begin{cases} \dfrac{S_0 + P - L}{S_0} - 1, & \text{if} \quad S_1 \geq G_1 \\ 0/S_0 - 1, & \text{if} \quad S_1 < G_1 \end{cases} \tag{6.3}$$

$$= \begin{cases} (P - L)/S_0, & \text{if} \quad S_1 \geq G_1 \\ -1 & \text{if} \quad S_1 < G_1. \end{cases} \tag{6.4}$$

The β of our insurer is:

$$\beta = \frac{\mathrm{Cov}(r(i), r(m))}{\mathrm{Var}(r(m))}, \qquad \text{where} \quad r(m) \text{ is the average return in the capital market.} \tag{6.5}$$

Our insurer ignores the small distortion caused by the possibility of its own impairment, so we can find

$$\beta = \frac{\mathrm{Cov}((P - L)/S_0, r(m))}{\mathrm{Var}(r(m))} \tag{6.6}$$

$$= -1/S_0 \frac{\mathrm{Cov}(L, r(m))}{\mathrm{Var}(r(m))}. \tag{6.7}$$

Now we can add the facts that

$$L = L_0 + CR \tag{6.8}$$

and that L_0 by assumption is independent of $r(m)$ (or else $r(m)$ is a risk factor, which we have assumed it isn't). So,

$$\beta = -\frac{C}{S_0} \times \frac{\mathrm{Cov}(R, r(m))}{\mathrm{Var}(r(m))} \tag{6.9}$$

$$= -\frac{C}{S_0} \beta_R. \tag{6.10}$$

β_R is the beta, the systematic risk measure, of R, our one external risk factor. Substituting in Equation 6.1,

$$y = r_f - \frac{C}{S_0}\beta_R\Pi, \tag{6.11}$$

$$\mathrm{Cov}(R, L + Q\ell) = \mathrm{Cov}(R, L) + Q\,\mathrm{Cov}(R, \ell) = C + Qc, \tag{6.12}$$

so, similarly, when we include Q,

$$y(Q) = r_f - \frac{C + Qc}{S_0}\beta_R\Pi, \tag{6.13}$$

and observe that

$$\frac{dy}{dQ} = -\frac{c}{S_0}\beta_R\Pi; \tag{6.14}$$

that is, the derivative is independent of the level of Q. This discount rate contribution will be insignificant if, as some investment bankers suggest, β_R is zero or small. Unfortunately, for very high-level catastrophe reinsurance, the available history suggests that catastrophe risk is not zero-beta.[18]

The change in the discount rate caused by assuming a marginal amount of risk does not depend on the amount already assumed. As promised, the discount rate term cannot affect the point at which the minimum premium is lowest.

---

[18] Kozik [8] notes practical and theoretical difficulties in computing and applying the betas applicable to the underwriting operations of insurers. This analysis is especially valid for diversifiable, low-level coverages. However, high-level reinsurance contracts address a small set of rare physical events, and the potential systemic correlations are both larger and clearer. If sizable, these systematic correlations are very relevent to the owners of insurance companies.

Interested readers may want to consider the notable falls both in the equity and bond markets and in affected currency values in the periods following the 1906 San Francisco earthquake in the United States and the 1995 Kobe earthquake in Japan. These two observations suggest that β_R can be significantly negative. Large physical catastrophes are correlated with losses in the capital markets.

7. THE MINIMUM PREMIUM IN THE ALTERNATIVE MODEL

All streams flow to the sea
Because it is lower than they are.
Humility gives it its power.[19]

Combining the contributions from the changes in the probability of survival and in the discount rate, and adding the assumption that the capital market valuation of the insurer is rational, although perhaps inefficient, we can solve for the insurer's minimum premium in an accessible way.

The minimum premium for a contract to be attractive to our insurer (any insurer) is (from Equation 3.15)

$$p = \hat{\mu} + \Psi(y'/(1+y) - (\mathrm{Z}'/\mathrm{Z})).$$

Substituting from Equation 5.2 and solving for p,

$$p = \hat{\mu} + \Psi\frac{y'/(1+y) - (\partial \mathrm{Z}^*/\partial\Lambda)(d\Lambda/dQ)/\mathrm{Z}}{1 + (\partial \mathrm{Z}^*/\partial\Psi)(\Psi/\mathrm{Z})}. \qquad (7.1)$$

This can be expressed differently when we further assume that the market valuation of our insurer is consistent with expectations, and that M is stable. Consistent expectations require that:

$$MS_0(1+y) = M\mathrm{E}(S_1) \qquad (7.2)$$

and further assuming that M is stable between $t = 0$ and 1 produces:

$$S_0(1+y) = \mathrm{E}(S_1 \mid S_1 \geq G_1)\mathrm{Prob}(S_1 \geq G_1), \qquad (7.3)$$

or abbreviating and regrouping,

$$\Psi = S_0(1+y)/\mathrm{Z}. \qquad (7.4)$$

This yields

$$p = \hat{\mu} + \frac{S_0 y' - \Psi(\partial \mathrm{Z}^*/\partial\Lambda)(d\Lambda/dQ)}{\mathrm{Z} + \Psi(\partial \mathrm{Z}^*/\partial\Psi)}; \qquad (7.5)$$

[19] The *Tao*, [11, Chapter 16].

and we know that (from Equation 6.14)

$$y' = (-c/S_0)\beta_R\Pi, \qquad \text{so}$$
$$p = \hat{\mu} + \frac{-c\beta_R\Pi - \Psi(\partial Z^*/\partial\Lambda)(d\Lambda/dQ)}{Z + \Psi(\partial Z^*/\partial\Psi)}. \tag{7.6}$$

This result applies at any level of Q, but we have seen that the lowest minimum premium occurs for a marginal participation. For our ideal price-taking insurer, with an offered premium near p, the marginal increase in its NPV quickly falls as the share of a proposed contract rises above zero. Our assumed insurer, like many real ones, maximizes its value by assuming and retaining very small parts of every possible risk.

The problem of insurers seeking geographic diversification can be restated from the insureds' perspective, as its dual problem of insurance risks seeking maximum spread among the world's insurers. If worldwide capacity meets the demand then our hypothetical contract would be fully placed with these ideal marginal participations. Q is approximately zero and the marginal variance of the contract (see Equation 5.3) for our insurer becomes

$$d/dQ\,\mathrm{Var}(L + Ql) = 2Cc, \tag{7.7}$$

and the marginal standard deviation is

$$d\Lambda/dQ = (\tfrac{1}{2})(1/\mathrm{SD}(L))(2Cc) \tag{7.8}$$
$$= Cc/\mathrm{SD}(L). \tag{7.9}$$

So the insurer's minimum premium becomes

$$p = \hat{\mu} + \frac{-c\beta_R\Pi - \Psi(\partial Z^*/\partial\Lambda)Cc/\mathrm{SD}(L)}{Z + \Psi(\partial Z^*/\partial\Psi)}. \tag{7.10}$$

We have seen that β_R and $\partial Z^*/\partial\Lambda$ are both less than zero, so the latter term is generally a positive number. The premium is the limited expected losses plus a risk load. The risk load depends on the covariance of the proposed contract with the risk factor of concern to the insurer, the capital market valuation for that risk, the insurer's capital structure, and the partial derivatives of its survival probability with respect to changes in its expected profits and variability. Using the normal distribution as a strong practical approximation,[20] there are closed-form expressions for these partial derivatives.

Since $\mathrm{SD}(R) = 1$, we see that

$$c = \mathrm{Corr}(\ell, R)\sigma \qquad \text{and} \qquad C = \mathrm{Corr}(L, R)\Lambda,$$

[20]The essential nature of insurance is the transfer and pooling of risks. In practice, catastrophe reinsurers track and control their risk accumulations in between six and more than thirty distinct zones. See the *1996 Annual Report* of CAT Limited for a clear example of the high end. Reinsurers' results are driven by the sum of these independent random processes. Their results will be close to normally distributed. (These underwriters can also rely on the exact derivation of the conclusion about decreasing values of Z'/Z in note 15.) If we define $T = S_0 + P - G_1$, then $Z^* = \mathrm{Prob}(L < T)$, where L is normally distributed with mean $W = \mathrm{E}(L)$ and standard deviation Λ

$$Z^* = \frac{1}{\sqrt{2\pi}\Lambda} \int_{-\infty}^{T} \exp\left(-\frac{1}{2}\left(\frac{x-W}{\Lambda}\right)^2\right) dx. \tag{20.1}$$

Since Ψ is independent of T and Λ, we can find the $\partial Z^*/\partial\Psi$ by bringing the differentiation within the integration

$$\frac{\partial Z^*}{\partial \Psi} = \frac{1}{\sqrt{2\pi}\Lambda} \int_{-\infty}^{T} \frac{d}{d\Psi} \exp\left(-\frac{1}{2}\left(\frac{x-W}{\Lambda}\right)^2\right) dx. \tag{20.2}$$

Over the range of integration, the conditional expectation of surplus, Ψ, is exactly and inversely related to the expectation of losses, W:

$$d/d\Psi = -d/dW \tag{20.3}$$

$$\frac{\partial Z^*}{\partial \Psi} = \frac{1}{\sqrt{2\pi}\Lambda} \int_{-\infty}^{T} -\frac{d}{dW} \exp\left(-\frac{1}{2}\left(\frac{x-W}{\Lambda}\right)^2\right) dx \tag{20.4}$$

$$= \frac{1}{\sqrt{2\pi}\Lambda} \int_{-\infty}^{T} \exp\left(-\frac{1}{2}\left(\frac{x-W}{\Lambda}\right)^2\right)\left(\frac{x-W}{\Lambda}\right) dx$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{T-W}{\Lambda}\right)^2\right). \tag{20.5}$$

$\partial Z^*/\partial\Psi$ also gives us $\partial Z^*/\partial\Lambda$. Since Z^* is a function of both W and Λ, we can express the two derivatives using the chain rule and find a simple relationship between them and

and the denominator is approximately one for our solid companies; so

$$p = \hat{\mu} + -(\beta_R \Pi + \Psi(\partial Z^*/\partial\Lambda)\mathrm{Corr}(L,R))\mathrm{Corr}(\ell,R)\sigma. \quad (7.11)$$

The alternative minimum premium formula roughly matches the dimensions in Kreps' analysis. $\mathscr{R}$, the reluctance, is directly related to the proposed contract's correlation with the relevant part of the existing portfolio, which is the risk factor of the insurer.[21]

$Z^{*\prime}((T-W)/\Lambda)$

$$\partial Z^*/\partial\Psi = -\partial Z^*/\partial \mathbf{W} \quad (20.6)$$

$$= -Z^{*\prime} d/d\mathbf{W}((T-W)/\Lambda) \quad (20.7)$$

$$= -Z^{*\prime}(-1/\Lambda) \quad (20.8)$$

$$= Z^{*\prime}/\Lambda, \qquad \text{and} \quad (20.9)$$

$$\partial Z^*/\partial\Lambda = Z^{*\prime} d/d\Lambda((T-W)/\Lambda) \quad (20.10)$$

$$= Z^{*\prime}(T-W) d/d\Lambda(1/\Lambda) \quad (20.11)$$

$$= -Z^{*\prime}(T-W)(1/\Lambda)^2 \quad (20.12)$$

$$= -(T-W)/\Lambda Z^{*\prime}/\Lambda, \qquad \text{or} \quad (20.13)$$

$$\partial Z^*/\partial\Lambda = -(T-W)/\Lambda\, \partial Z^*/\partial\Psi. \quad (20.14)$$

To illustrate, if an insurer believes that $(T-W)/\Lambda = 3.0$ (not unrealistically, $Z = .9975$, a one-in-four-hundred years probability of ruin) and that the standard deviation of the loss process on its existing contracts is, say, \$100,000,000, then

$$\frac{\partial Z^*}{\partial\Psi} = \frac{1}{\sqrt{2\pi}} \frac{1}{\$100{,}000{,}000 e^{4.5}} = +4.432 \times 10^{-11}/(\$ \text{ of mean}), \quad (20.15)$$

and

$$\frac{\partial Z^*}{\partial\Lambda} = \frac{1}{\sqrt{2\pi}} \frac{1}{\$100{,}000{,}000 e^{4.5}}(-3) = -1.130 \times 10^{-10}/(\$ \text{ of standard deviation}). \quad (20.16)$$

Results like these could be used in Equations 7.10 or 7.11 to solve for minimum premiums in a direct way.

[21] Bault [2] analyzes several common risk load approaches with different assumptions and concludes that each can be re-expressed as a covariance measure between the proposed

Of course, in the real world, regulation, occasional capacity shortages, frictional costs, information barriers, and some economies of scale will prevent perfect diversification. Higher prices result. However, this analysis provides a fair estimate of the market premium, establishes a lower bound, and shows a scale by which the costs of market inefficiencies can be seen.

8. NEXT STEPS

As it acts in the world, the Way is like the bending of a bow.
The top is bent downward; and the bottom is bent up.
It adjusts excess and deficiency so that there is perfect balance.
It takes from what is too much and gives to what isn't enough.[22]

This review has mimicked the development of Modern Portfolio Theory (MPT) and found a result like one of MPT's fundamental tenets. A diversified, rational, risk-averse insurer, like a similar investor, will accept a potential addition to its portfolio only after a comparison between the addition's systematic, non-diversifiable risk and its price in the market.

MPT goes on to show that since most investors price assets that way, the market pricing of assets must be based on only the value of their systematic risks. These investors get the best return for the least total risk (at market pricing) by distributing their portfolios in proportion to the asset distribution of the total market. Insurers find a similar optimal return for their risk: either writing a balanced worldwide spread or placing their riskier coverages with reinsurers who do.

Based on Kreps' and other recent results and the generalization added in this review, actuaries should be able to raise our knowledge of market risk pricing up to the level that MPT

---

contract and the insurer's surplus. This review attempts to show that this conclusion holds with realistic assumptions about insurers' incentives and the insurance and capital markets.

[22]The *Tao*, [11, Chapter 77]. For clarity, "Way" replaces "Tao" in Mitchell's translation.

has reached for asset pricing.[23] If reinsurers choose to diversify their exposures, as these results suggest they must and recent acquisitions[24] suggest they do, market pricing will be based only on each proposed contract's systematic risks. Any insurer whose capital structure or distribution of net risks varies significantly from the industry average will find that its minimum premiums will be higher than the market clearing prices in its areas of relative over-concentration. It will reduce its exposures (by reinsurance, securitization or direct volume reduction) until its minimum premiums fall to the market level. Since much of the industry follows this process, market prices will only be denominated by contracts' covariances with the one (actually more) risk factor(s) affecting the global insurance market.

In effect, every insurer trying to maximize its risk-adjusted NPV must act as if it desires a spread of net risks like the worldwide industry average. Again, regulation, returns to scale, and frictional and information costs prevent this in practice.

The minimum market premium for a proposed contract does not depend on its covariance with a particular group of an insurer's other contracts. It depends more on the covariances with the significant risks that influence the results of all possible contracts worldwide. No proposed contract is considered first. Or last. This final result eliminates the circularity in Kreps' analysis and mine.

---

[23]See Feldblum [4] for a very rigorous attempt at this analysis. However, this result only developed relative operating profit provisions and cannot develop specific targets without assumptions about allocated capitalization and cannot be reconciled to MPT results for equities because they consider different universes. See also Turner's article in Cummins and Harrington (Note 8, above), for an equilibrium analysis, but without product distinctions. A more general model, like Kreps' concept of marginal surplus requirements or the approach suggested in this review, can support a risk-specific price that is in equilibrium with the capital market valuations of the insurer.

[24]Numerous recent transactions, but note two common themes: property companies acquiring books on other continents (Cologne Re, Sphere Drake, SAFR, M&G, American Re, SOREMA-UK) to "balance all the buckets," and liability companies acquiring property operations (Tempest, GCR, IRI, Mid-Ocean, CAT Limited). Both can be explained by the search for a broader mix of the world's exposures.

9. CONCLUSIONS

1. Kreps' analysis is a significant addition to both the practice and theoretical understanding of reinsurance and insurance pricing in that
   - the risk load required for a contract to be attractive (or indifferent) to an insurer must be based on the risk of the proposed contract and not on its expected losses, and
   - the risk load an insurer must require depends on the covariance of the proposed contract with the insurer's existing risks (that is, with the product of the correlation between the proposed contract and others, and the standard deviation of the contract).
2. These results still hold with a more realistic model of insurers' incentives.
3. The required risk load also reflects the relative correlations between insurance risk factors and movements in the overall capital markets. This is true even though there are frictional barriers to moving capital through insurers.
4. The minimum amount of this required risk load occurs for marginal participations. Insurers thus have strong incentives to diversify. Since most do, market prices are based on the covariances of the proposed contract with the general risk factors exposing all other possible contracts, that is, the entire insured market. Risk loads should not reflect any diversifiable risks.

Knowing others is intelligence;
knowing yourself is true wisdom.
Mastering others is strength;
mastering yourself is true power.

If you realize that you have enough,
you are truly rich.[25]

[25]The *Tao*, [11, Chapter 33].

SUMMARY OF NOTATION

In Kreps' Original

μ, expected losses for proposed contract

$\mathscr{R}$, insurer's reluctance to assume contract

σ, standard deviation of losses for proposed contract

E, insurer's marginal expense

y, risk-adjusted discount rate

B, reinsured's "bank," amount reinsurer is willing to concede (proposing to extract) in renewal price.

Z, insurer's ratio of surplus to standard deviation of existing loss portfolio

Added in Discussion

V_i, market value of the insurer at time $t = i$

S_i, surplus, GAAP book value, at time $t = i$

G_1, minimum going-concern surplus level at $t = 1$

M, insurer's book-value multiple

P, premium for existing portfolio

L, random loss process for existing portfolio

p, premium for proposed contract

ℓ, random loss process for proposed contract

Q, decision variable, portion of the proposed contract assumed

μ, expected losses for the proposed contract, $E(l)$

$\hat{\mu}$, expected losses for the proposed contract limited to the scenarios where $S_1 \geq G_1$

R, a measure of the one external risk factor of concern to insurer, re-scaled here to be positively correlated with L and with standard deviation = 1

r_f, risk-free interest rate between $t = 0$ and $t = 1$

β, systematic risk measure of insurer's return

Π, market risk premium expected between $t = 0$ and $t = 1$

$r(i)$, insurer's market-based return between $t = 0$ and $t = 1$

$r(m)$, return on overall capital market between $t = 0$ and $t = 1$

β_R, systematic risk measure of R, the external risk factor

Abbreviations

NPV, risk-adjusted net present value at $t = 0$ of insurer's market value at $t = 1$

Z, probability that $S_1 \geq G_1$, that the insurer survives

$$\mathrm{Z} = f(P, L, Q, p, \ell, G_1, S_0) \qquad \text{and}$$

$$\mathrm{Z}^* = f(\Psi(Q\ell, Qp), \Lambda(Q\ell))$$

$\Psi = \mathrm{E}(S_1 \mid S_1 \geq G_1)$

$C = \mathrm{Cov}(L, R)$

$c = \mathrm{Cov}(\ell, R)$

$L_0 = L - (C/\mathrm{Var}(R))R = L - CR$

$\ell_0 = \ell - (c/\mathrm{Var}(R))R = \ell - cR$

$\Lambda = \mathrm{SD}(S_1) = \mathrm{SD}(L + Q\ell)$

REFERENCES

[1] Bass, Irene K. and C. K. Khury, "Surplus Allocation: An Oxymoron," *Insurer Financial Solvency*, Casualty Actuarial Society Discussion Paper Program, 1992, pp. 553–584.

[2] Bault, Todd R., Discussion of Feldblum: "Risk Loads for Insurers," *PCAS* LXXXII, 1995, pp. 78–96.

[3] Brealey, Richard A. and Stewart C. Myers, *Principles of Corporate Finance*, Fourth Ed., McGraw-Hill, 1991.

[4] Feldblum, Sholom, "Risk Loads for Insurers," *PCAS* LXXVII, 1990, pp. 160–195.

[5] Gogol, Daniel F., Discussion of Kreps: "Reinsurer Risk Loads from Marginal Surplus Requirements," *PCAS* LXXIX, 1992, pp. 362–366.

[6] Kneuer, Paul J., "Allocation of Surplus for a Multi-Line Insurer," *Financial Analysis of Insurance Companies*, Casualty Actuarial Society Discussion Paper Program, 1987, pp. 191–228.

[7] Kneuer, Paul J., Review of Cummins and Harrington: *Fair Rate of Return in Property–Liability Insurance*, (Kluwer-Nijhoff Publishing, 1987), *Actuarial Review*, May 1992.

[8] Kozik, Thomas J., "Underwriting Betas—The Shadow of Ghosts," *PCAS* LXXXI, 1994, pp. 303–329.

[9] Mango, Donald F., "An Application of Game Theory: Property Catastrophe Risk Load," *PCAS* LXXXV, 1998, pp. 157–186.

[10] Miller, Michael J. and Jerry W. Rapp, "An Evaluation of Surplus Allocation Methods Underlying Risk Based Capital Applications," *Insurer Financial Solvency*, Casualty Actuarial Society Discussion Paper Program, 1992, pp. 1–122.

[11] Mitchell, Stephen (translator), *Tau Te Ching, A New English Version with Forward and Notes*, Harper and Row, 1988.

[12] Roth, Richard J., Jr., "Analysis of Surplus and Rate of Return Without Using Leverage Ratios," *Insurer Financial Solvency*, Casualty Actuarial Society Discussion Paper Program, 1992, pp. 439–463.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXXIII

ESTIMATING THE PREMIUM ASSET ON RETROSPECTIVELY RATED POLICIES

MIRIAM PERKINS AND MICHAEL T. S. TENG

DISCUSSION BY SHOLOM FELDBLUM

1. INTRODUCTION

Perkins and Teng have provided us with a new and remarkably intuitive procedure for estimating the accrued retrospective premium asset: the PDLD (premium development to loss development) approach. This reserve is often significant—amounting to half a billion dollars or more for some of the major workers compensation carriers—and it has been difficult to accurately estimate with traditional procedures. The paper by Perkins and Teng should greatly enhance our actuarial repertoire.

Specifically, the PDLD method has several distinct advantages over other procedures:

1. It is modeled directly on the retrospective rating formula, so it is easily explained to underwriters and claims personnel who are familiar with retrospectively rated policies.
2. Its emphasis on the premium sensitivity in the retrospective rating formula parallels the risk-based capital loss-sensitive contract offset in the underwriting risk charges and the new loss-sensitive contract Part 7 of Schedule P. For regulators familiar with the risk-based capital formula and with the statutory accounting requirements, this loss reserving approach is a natural complement to the statutory procedures.

3. The procedure may prove particularly useful when changes in the retrospective rating plan parameters distort the indications of other methods.

There are few existing methods for estimating the accrued retrospective premium asset, and the indications are often highly uncertain. The PDLD approach will enable actuaries to estimate this asset more accurately.

This discussion has two parts.

1. The complexity of the reserve estimation procedures for the accrued retrospective premium asset often hides the rationale of these methods from the average reader. The first part of this discussion uses graphical representations of Fitzgibbon's method and of the PDLD method to show the rationale behind each method and to explain the advantages of the latter method.[1] We then show how to combine the better parts of the two methods to improve the PDLD procedure.
2. The second part of this discussion highlights the implications of the Perkins and Teng procedure for the calculation of the loss-sensitive contract offset to the underwriting risk charges in the risk-based capital formula and for the use of Schedule P, Part 7, to estimate premium sensitivity.[2]

2. THE PDLD PROCEDURE

This section addresses two issues:

1. How does the PDLD procedure differ intuitively from Fitzgibbon's procedure, and in what ways is it better?

[1]See Fitzgibbon [6], F. J. Hope [8], Unthoff [11], Berry [2], and Morell [10]. The term "Fitzgibbon's method" in the text includes the enhancements provided by Berry and Morell.

[2]The term "premium sensitivity" stems from the term "loss-sensitive contracts." This paper uses the term "premium responsiveness" to refer to the same phenomenon.

2. What aspects of Fitzgibbon's procedure can be added to the PDLD procedure to enhance its accuracy?

Let us begin our inquiry with a more fundamental question. Why not estimate the accrued retrospective premium asset the same way that we estimate loss reserves? That is, why not use a chain-ladder development procedure on historical triangles of either collected premium or billed premium? This would be the premium analogue to a chain-ladder development procedure using either paid losses or reported losses.

Indeed, Schedule P already does this. Part 6 of Schedule P shows historical triangles of exposure year earned premiums by line of business (for all types of contracts), and Part 7 of Schedule P shows historical triangles of policy year earned premium on loss-sensitive contracts (all lines of business combined). Why go through the complexities of Fitzgibbon's method or the PDLD method when a straightforward chain ladder development method would suffice?

The underlying rationale of Fitzgibbon's method and the PDLD method is that

a. estimates of ultimate incurred losses can be obtained sooner than estimates of retrospective premiums can be obtained, and

b. retrospective premiums depend on incurred losses.

In workers compensation, for instance, a good estimate of ultimate incurred losses is generally available soon after the expiration of the policy, since claims emerge rapidly and development on known claims is relatively stable. The first retrospective adjustment, however, occurs about six months after the expiration of the policy. The retrospective premium may not be billed and collected for an additional three months after the adjustment is done.

Using Fitzgibbon's method or the PDLD method, an initial estimate of the accrued retrospective premium asset can be produced soon after the policy expires, using the known loss information and the relationships between incurred losses and retrospective premium. Similarly, the accrued retrospective premium asset estimate can be updated each quarter, as new loss data becomes available. If a chain-ladder premium development procedure is used, however, the initial estimate cannot be produced until at least nine months after the policy expiration, and it can be updated only annually thereafter.

The reserve estimation procedures in both Fitzgibbon's method and the PDLD method are based upon the retrospective rating formula. They differ in the details, not the concept, although the details can be crucial for reserve estimation. Using graphs to clarify the methods, the two approaches will be compared and contrasted using the following steps:

- how premium is determined in the retrospective rating formula;

- how Fitzgibbon, followed by Berry, converts the premium determination procedure to a reserve estimation procedure;

- what problems arise in the reserve estimation procedure, and how Berry resorts to a second reserve estimation procedure to resolve them;

- how the PDLD procedure modifies the original Fitzgibbon procedure to solve the aforementioned problems, without having to resort to a second reserve estimation procedure; and

- how part of Fitzgibbon's procedure can be used to enhance the PDLD procedure, giving users the best of both worlds.

Retrospective Premium Determination

Fitzgibbon's method and the PDLD method both seek to replicate the premium determination procedure in the retrospective rating formula. Of course, a single reserving formula cannot perfectly replicate hundreds of slightly different rating plans. Nevertheless, the more successfully the reserving procedure can replicate the rating procedure, the more accurate will be the reserve estimates. So let us begin with the premium determination formula.

The retrospective premium is composed of two parts:

1. Part of the premium covers the incurred losses, as well as any expenses associated with these losses, such as loss adjustment expenses. However, not all losses enter the retrospective rating formula. There is a loss limit, which means that individual losses exceeding a certain amount—such as $250,000—do not affect the retrospective premium adjustments. In addition, state premium taxes, as well as other state assessments (such as involuntary market loads) are levied on the premiums, whether they are standard premiums or retrospective premium adjustments.

   The retrospective rating plan expresses this part of the premium as

   $$(\text{loss conversion factor}) \times (\text{incurred losses}) \times (\text{tax multiplier}),$$

   where the loss conversion factor (LCF) covers primarily loss adjustment expenses.

2. The other part of the premium covers company expenses and the insurance charge. Company expenses are all expenses that are not a direct function of losses, such as underwriting expenses and acquisition expenses.

The insurance charge results from the maximum and minimum limitations on the retrospective premium. Having a maximum premium, of course, is the whole purpose of insurance. The insured needs protection against the unanticipated large losses that it cannot prudently retain. But the insurer must collect premium to cover these large losses. So the insurance charge is the difference between

a. the expected loss (to the insurer) caused by the maximum premium and

b. the expected gain (to the insurer) caused by the minimum premium.

The expected loss is the average additional amount of premium that the insurer would have collected had there been no maximum premium limitation. The expected gain is the average amount of premium that it would not have collected had there been no minimum premium limitation.

This charge must also cover any premiums lost because of the loss limits, which cap the individual loss values entering the retrospective rating plan.[3]

As before, a provision must be added for state premium taxes and other state assessments. This part of the premium may be expressed as

$$[(\text{expense provision}) + (\text{insurance charge}) + (\text{excess loss charge})] \times (\text{tax multiplier}).$$

[3]The computation of the insurance charge is the standard Table M and Table L calculation. For the "formula" approach in the PDLD method, which can be used with Fitzgibbon's method as well, the reserving actuary may have to recompute certain Table M or Table L charges.

For simplicity, the first three components are combined into the basic premium, so the expression above can be restated as

$$\text{(basic premium)} \times \text{(tax multiplier)}.$$

Thus, the formula for the retrospective premium is

$$\text{Retrospective premium} = \text{(tax multiplier)} \times [\text{(basic premium)} + (\text{(loss conversion factor)} \times \text{(limited incurred losses)})].$$

The Reserving Formula

The formula above is the rationale for Fitzgibbon's reserve formula. Premium is assumed to be a linear function of the incurred losses, or

$$\text{Retrospective premium} = C + B \times \text{Losses}.$$

The pricing formula becomes the reserving formula. For application to an entire book of business, Fitzgibbon and Berry make two modifications to this basic equation:

1. They use ratios to standard premium. That is, they write

   $$\text{Retrospective premium} \div \text{Standard Premium} = K + B \times \text{Standard Loss Ratio},$$

   where $K = C \div \text{Standard Premium}$.

2. They examine the retrospective adjustment. In other words, they subtract unity from both sides of the equation above, to get

   $$\text{Retro Adjustment} = A + B \times \text{Standard Loss Ratio},$$

   where $A = K - 1$.

The Historical Regression

Fitzgibbon and Berry estimate the parameters *A* and *B* from a historical regression, using standard loss ratios and retrospective adjustments from mature policy years. But the attentive reader might observe that the two parameters in Fitzgibbon's formula depend on the parameters in the retrospective rating formula. So why do they use a regression analysis on past experience? Why don't they just walk over to the pricing actuary in the next office and ask what parameters are used in the retrospective rating plan?

Actuarial reserves are typically estimated on an aggregate basis, for all states, all insureds, all policy years. The parameters, however, vary from year to year, from state to state, and from plan to plan. For instance:

- A small insured may purchase a plan with a low maximum premium and therefore a large insurance charge, whereas a large insured may prefer a plan with a high maximum premium and a low insurance charge. Also, larger insureds may be offered plans with lower expense provisions, since their underwriting and acquisition expenses as a percentage of standard premium are lower than for smaller insureds.
- Premium taxes differ from state to state. In addition, some retrospective rating plans include involuntary market expense loads as a part of the tax multiplier, and the involuntary market loads vary widely among jurisdictions.
- The basic premium may vary from year to year. It may be low when interest rates are high and the insurer expects to earn its required profit margin from investment income. It may be higher when interest rates are low, or if the insurer uses a cash flow plan, such as a paid loss retro, so little investment income is retained by the insurer.

In theory, the reserving actuary could collect the hundreds of needed plan combinations and match these with the appropriate experience and calculate the reserve. Or the actuary, to save a few months of work, might determine the average parameters by means of a regression analysis on historical data.

This is what Fitzgibbon and Berry have done. The regression analysis calculates the average retrospective rating plan parameters from past experience. In fact, this method is probably more accurate than might be achieved by collecting all the parameters actually used in each state and each policy year for each insured. Most companies allow their underwriters and agents substantial flexibility in rating workers compensation contracts. The pricing actuary may recommend a basic premium charge of 30% of standard premium, but the underwriter or salesperson may reduce the basic premium charge to 25% of standard premium. The pricing actuary's recommended parameters may not match the plan parameters that are actually used in practice. The reserving actuary needs to know the premiums that are actually charged, not the pricing actuary's indicated premiums. So the reserving actuary turns to the regression analysis, not to the pricing actuary's rate book.[4]

---

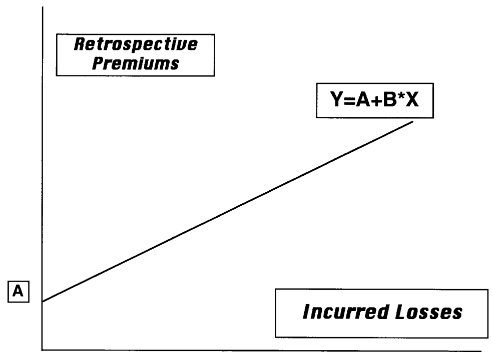
[4]How is it then, that Perkins and Teng manage to estimate PDLD ratios from the retrospective rating plan parameters in their formula approach? Moreover, they need to estimate more numbers than Fitzgibbon and Berry need to estimate, so how are they able to do this when Fitzgibbon and Berry found it unmanageable?

The answer is that the Perkins and Teng paper presents the method only. In practice, estimating the PDLD ratios from the retrospective rating plan parameters is exceedingly difficult, particularly if the company writes business in different states and for different types of insureds, if the company has changed its plan parameters over time, or if the company allows its underwriters and agents discretion in modifying the plan parameters to attract potentially good risks. Perhaps Ms. Perkins or Mr. Teng can elaborate on the relative ease or difficulty of estimating the PDLD ratios in various scenarios.

As pointed out by Robert Finger, the regression approach is not without its difficulties as well. Rating plan factors and aggregate loss ratios change over time, so a regression performed on historical data may not be equally applicable to current policies. Moreover, the observed values are actually the result of many changes at the individual plan level. The premium on individual plans is not a simple function of total incurred losses. For instance, premium may decrease on an adjustment when incurred losses increase, since there may be positive development on a claim that was already limited and negative development on claims that were below the per accident limit. See also Morell [10], which discusses this same issue.

FIGURE 1

FITZGIBBON'S METHOD



Graphical Representations

To see the difference between Fitzgibbon's method and the PDLD method, let us look at these procedures graphically. Fitzgibbon's method represents the relationship between the net earned premium[5] on the retrospectively rated book of business (as a percentage of standard premium) and the total incurred losses on this book of business (again, as a percentage of standard premium) as a straight line, as shown in Figure 1.[6] Algebraically, the straight line is $Y = A + B^*X$, where A is the constant factor and B is the slope factor.

One interpretation of this graph is as follows: if there are no incurred losses on this book of business, then the ratio of net premium to standard premium equals A. The constant factor A represents the basic premium percentage in the retrospective

[5]Net earned premium is earned premium after retrospective adjustments; see Feldblum [3].

[6]The figures on both axes of this graph are shown as ratios to standard earned premium. Alternatively, one could show both sets of figures as absolute dollar amounts. Berry uses ratios, though he shows the vertical axis as ratios of retrospective premium *returns* to standard premium. The three methods are equivalent.

rating formula.[7] As losses are incurred, and the loss ratio to standard premium increases, we move to the right and up along the straight line, and the net premium as a percentage of the standard premium increases. For each dollar of additional loss, the net retrospective premium increases by B dollars.

The slope factor B is the premium responsiveness for this book of business. The slope is not exactly unity, for several reasons. First, some losses exceed the loss limit, or they cause the retrospective premium to reach the maximum premium, even before the first adjustment, thereby reducing the slope of the line segment. Second, in some plans the minimum premium exceeds the basic premium. Third, a loss conversion factor and a tax multiplier are applied to the incurred losses in the retrospective rating formula, thereby changing the slope of the line segment. The combined effect depends on the "swing" of the plan. For plans with narrow swing, generally sold to small accounts, the slope would be less than unity. For plans with wide swing, generally sold to large accounts, the slope might be greater than unity.[8]

Projections versus Reality

The problem with this method, as Berry points out, is that it does not consider the emerging experience on the book of business itself. This emerging experience may differ from that expected from the graph for several reasons. First, the A and B factors are only estimated from the regression; they are not known with certainty. Moreover, they may vary from year to year. Second, the pattern of losses among the individual policies

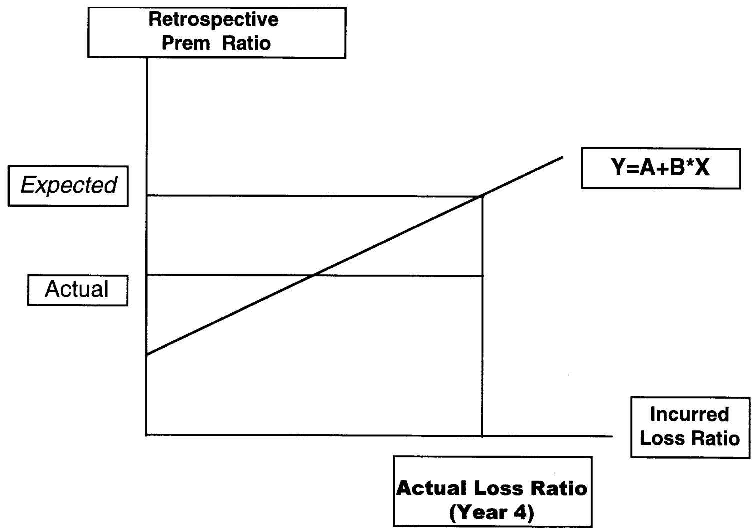
---

[7] Since the *A* factor is fitted by a regression on the aggregate book of business, it would not necessarily equal the basic factor on any particular plan.

[8] Fitzgibbon and Berry might say that this is not an exact interpretation of their regression line. Their regression line relates the *ultimate* loss ratio to the retrospective premium percentage. Their graph is not necessarily intended to represent the *movement* from no losses at policy inception to ultimate losses many years later. However, the purpose here is to highlight the contrast with the PDLD method, not to explain Fitzgibbon's method itself.

FIGURE 2

ACTUAL VERSUS EXPECTED RESULTS



affects the results. One large loss may have the same effect on the aggregate loss ratio as a dozen small losses. The effect on the net premium may differ because of loss limits and maximum premiums.

Suppose that after four years, the actual experience on this book of business shows less premium responsiveness than had been initially anticipated, as shown in Figure 2. The book of business is relatively mature after four years. The projection produced by this method does not change from year to year (as long as the incurred losses do not change), so it will continue to give an estimate of retrospective premium that is too high.

Berry's solution is to gradually discard this method, and to substitute a method that relies on the actual experience of the book of business (his "DR2" method). Initially, his reserve estimate relies entirely on this method. As time goes on, and more information becomes available from the actual book of business, he assigns progressively less weight to this method and more weight to his “DR2” method.

The Perkins and Teng Solution

Perkins and Teng transform Fitzgibbon’s graph to solve this problem. Think of Fitzgibbon’s graph in a slightly different fashion: as the movement over time of reported losses, net earned premium, and reported loss ratio. At policy inception, reported losses are \$0, so the reported loss ratio is 0% and the ratio of net premium to standard premium equals A, the constant factor in Fitzgibbon’s regression equation, or the Y-intercept in Fitzgibbon’s graph.

There are two ways to interpret the chart in Figure 1. Only the first of these reflects the intentions of Fitzgibbon and Berry. The second reflects the PDLD method. The alternative interpretations are:

1. the graph relates the ultimate loss ratio and the ultimate retrospective premium ratio among different books of business or different years of experience, or

2. the graph relates the reported loss ratio and the net earned premium at different points in time for a single book of business.

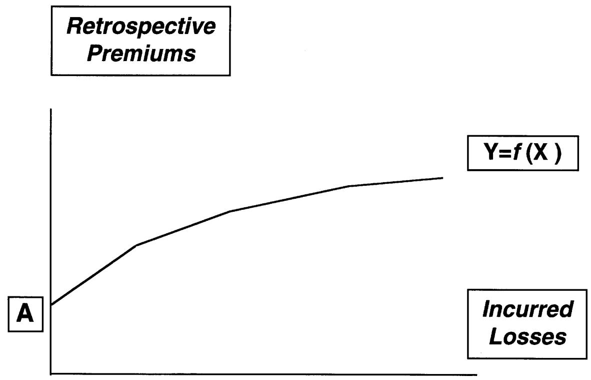
Decreasing Slopes

These two types of graphs seem similar. In truth, they look quite different. The first relationship is drawn by Fitzgibbon and Berry as a straight line. Actually, the curve is concave, as explained below, but a straight line is a close enough approximation for the majority of the curve.[9] The second relationship, however,

[9]It is a poor approximation at high loss ratios and at low loss ratios, though, where the maximum and minimum premium limitations flatten the curve. Fitzgibbon and Berry were aware of the approximation problems at the end points, and adjustments could always be made where necessary.

FIGURE 3

THE PERKINS AND TENG "PDLD" GRAPH



is not a straight line at all. Rather, it is a set of line segments, of steadily decreasing slope as we move to the right, as shown in Figure 3.[10]

The differing slopes of these line segments result from the loss limits and the maximum premiums in the retrospective rating plans. Most reported losses from policy inception until the first retrospective adjustment are rateable losses, which means that they are generally not truncated by the loss limit, and the retrospective premium is generally not capped by the maximum premium. The slope of the line segment is therefore close to unity. That is, for each dollar of reported loss, the insurer receives about a dollar of premium.

During subsequent periods, new reported losses stem from the emergence of IBNR claims and from development on known

[10]We use a series of line segments because retrospective adjustments are done annually, and the PDLD method reflects this by using line segments with different slopes for each adjustment period. In truth, a continuous concave curve better reflects reality, though it would not lead to a feasible reserving method.

claims. In workers compensation, for instance, new reported losses after the first adjustment may arise from the re-evaluation of a lower back sprain from a temporary total injury to a permanent total injury, with a corresponding re-estimation of the incurred loss from $25,000 to $500,000. This loss may be truncated by the loss limit in the retrospective rating formula, and the resulting retrospective premium may also be capped by the maximum premium.

This example is not contrived. On the contrary, it is quite common in workers compensation. Persons unfamiliar with industrial accidents often think of lifetime pension cases as quadriplegics or workers who have lost arms or legs in workplace accidents. Such injuries would be recognized immediately as high-cost, permanent total disabilities. These claims, which are recognized well before the first retrospective adjustment, are the ones that are most likely to be curtailed by the loss limits and maximum premium. This might lead some actuaries to think that the slope of the line segment in our graph should be flattest in the initial period.

In fact, accidents resulting in quadriplegia or the loss of arms or legs are rare. Most lifetime pension cases stem from sprains and strains and similar injuries that seem at first to be only temporary. After several years, when it becomes evident that the injured employee will not be returning to work, the claim is recorded as a permanent total injury and the benefit amount is re-estimated.[11]

We may state this as a general rule:[12]

1. *As a book of business matures, premium responsiveness on loss-sensitive contracts declines.*

---

[11] In the company at which the PDLD method was developed, fewer than 20% of claims that will ultimately be lifetime pension cases are recognized as such by the claims department at the first retrospective adjustment.

[12] As with any general rule, there are exceptions in particular instances.

In other words, as policies mature, a greater percentage of loss development is excluded from retrospective rating by the maximum premium and by the loss limit.

A second factor contributing to the declining slopes of the line segments is the overall increase in the reported loss ratio. It is not just that late-reported losses may be capped by the loss limit. Even a small claim will not increase the retrospective premium if the maximum premium has already been reached. Suppose the retrospective premium equals the maximum premium two years after policy inception. Then small claims reported during the first two years would have a premium responsiveness exceeding unity (because of the loss conversion factor and the tax multiplier), while small claims reported after the first two years would show a premium responsiveness of zero. We can state this second phenomenon as a general rule as well:

2. *At higher loss ratios, premium responsiveness on loss-sensitive contracts declines.*

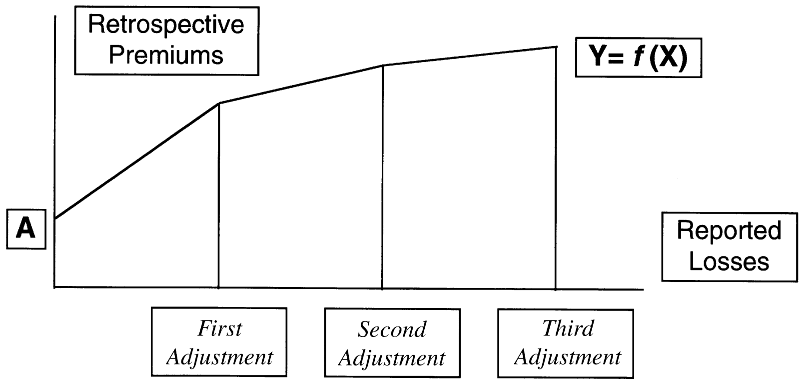
This last phenomenon relates to the overall loss ratio, not to the types of claims reported in any particular period. At higher overall loss ratios, more policyholders have reached their maximum premiums, so premium responsiveness is lower. Thus, it applies not only to the PDLD method, but to Fitzgibbon's method as well. That is, Fitzgibbon's graph is not really a straight line. In theory, it is a curve that is concave downwards, with steadily decreasing slope as the loss ratio increases.

Let us return to the PDLD method. At policy inception, the projected premium responsiveness graph is shown in Figure 4. Each line segment represents one period. The first line segment is from policy inception to the first retrospective adjustment, at about 21 months.[13] Subsequent periods are each one year long.

[13]The billing of retrospective premium generally lags the incurral of additional losses by about three months (on average) for an individual policy and by about nine months (on average) for a policy year. See below in the text for a full explanation of the lag times and effects that these may have on the observed premium responsiveness.

FIGURE 4

THE PDLD SEGMENTED GRAPH



The horizontal axis represents reported losses. For clarity, the graph is not drawn to scale. That is, the change in reported losses from policy inception to the first retrospective adjustment may be 50 percentage points or more in workers compensation, whereas the change in reported losses between adjustments at late maturities may be only a few percentage points. However, the graph shows all the line segments of equal length, so that the difference in their slopes can be seen clearly.

Actual versus Expected Experience

At the first adjustment, actual experience may differ in two ways from the experience that would be expected from the theoretical graph.

1. Actual reported losses may differ from the projected reported losses. For instance, at policy inception, the projected reported loss ratio to standard earned premium at 21 months may have been 55%. The actual reported loss ratio to standard earned premium at 21 months may be 50%.

2. The relationship between reported losses and retrospective premium may differ from that projected at policy inception. For instance, suppose that the Y-intercept in the graph is 20% and the slope of the first line segment is 1.100. Then for an actual reported loss ratio of 50% at the first retrospective adjustment, the ratio of net premium to standard premium is expected to be $20\% + 1.100 * 50\% = 75\%$. Suppose, however, that the actual ratio of net premium to standard premium at the first retrospective adjustment is only 72%.

These effects are shown in Figure 5 (not drawn precisely to scale).

- The projected experience at policy inception was for a reported loss ratio of 55% and a retrospective premium ratio of 80.5% [$= 20\% + 1.100 * 55\%$].
- For a reported loss ratio of 50% at the first retrospective adjustment, the graph projects a retrospective premium ratio of 75%.
- Actual experience at the first retrospective adjustment shows a reported loss ratio of 50% and a retrospective premium ratio of 72%.

The Perkins and Teng Assumptions

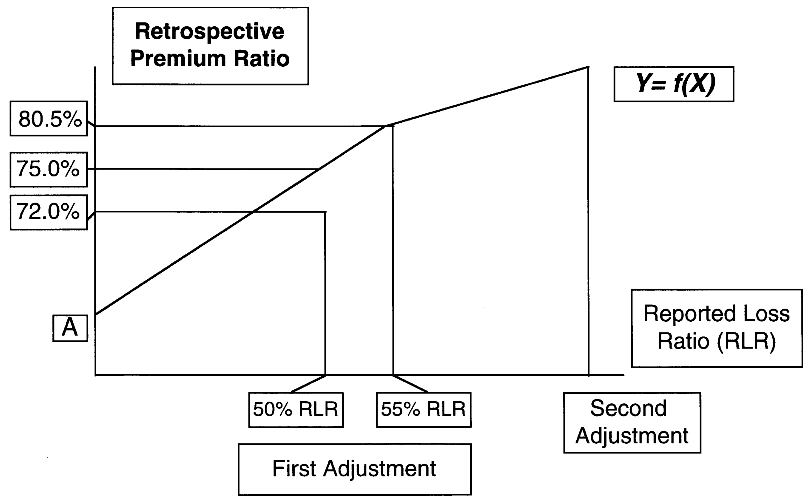
Two assumptions underlie the PDLD method. These are:

A. The premium responsiveness during subsequent adjustments is independent of the premium responsiveness during preceding adjustments.

B. The slope of the line segment depends on the time period, not on the beginning loss ratio or the beginning retrospective premium ratio.

FIGURE 5

PDLD METHOD: ACTUAL VERSUS EXPECTED RESULTS



We illustrate this for the first two line segments in Figure 5. Suppose the slope of the second line segment is 0.800. Think of the second line segment as an infinite number of parallel lines, all with slope of 0.800. At policy inception, we expected the second line segment to start at the point (55%, 80.5%) and to continue onwards with a slope of 0.800. As it turns out, the second line segment begins at the point (50%, 72%), but it still continues onwards with a slope of 0.800.

Compare the illustration with the two assumptions. We had expected a 75% retrospective premium ratio with a 50% reported loss ratio, but we actually get a 72% retro premium ratio. In other words, the slope of the first line segment is lower than we had originally expected. Nevertheless, we do not change our expectations for the slope of the second line segment. This is Assumption A.

The second assumption relates to when we change from the first line segment to the second line segment. From the appearance of the graph in Figure 5, one might think that we change when the reported loss ratio reaches 55%. That is not the meaning of the graph. Rather, we change at the first adjustment, regardless of the reported loss ratio at that time.

The manner in which the PDLD method solves Berry's problem should now be clear. Fitzgibbon's graph relates the *ultimate* loss ratio to the ultimate retrospective premium ratio. If actual experience differs from expected experience along the way, there is no way to get back on track. The PDLD method relates the *reported* loss ratio to the retrospective premium ratio. If actual experience differs from expected experience along the way, the next line segment begins at a starting point that corresponds to the actual experience.

The PDLD method quantifies the accrued retrospective premium asset in two steps.

1. Project the future loss development in each adjustment period.

2. Estimate the future premium revenue by the product of the future loss development in each period and the slope of the line segment in that period. The sum of these products is the accrued retrospective premium asset.

The PDLD method can be thought of as follows. The line segments represent a mountain being climbed, from the 0% reported loss ratio at policy inception to the ultimate loss ratio when all losses are settled. At each retrospective adjustment, the remaining part of the climb is shifted, both horizontally and vertically, but the shape of the climb is not changed (that is, the slopes of each line segment remain fixed).[14]

An Enhancement

In Figure 5, the first line segment begins at a point on the Y-axis representing the amount of retrospective premium when the reported loss ratio is 0%; that is, the Y-intercept is positive. This is the proper way to estimate the accrued retrospective premium asset. Perkins and Teng, however, have the first line segment passing through the origin; that is, the Y-intercept is 0. As a result, Perkins and Teng get a slope for the first line segment of 1.750. In fact, empirical data in their Exhibit 4, Sheet 1 for the most recent four quarters shows an average slope of 1.825.

Perkins and Teng's numbers combine two separate items: the basic premium ratio and the slope of the first line segment (when drawn properly). By failing to distinguish between these two elements, the method becomes less intuitive: how does one explain a slope of 1.825 or 1.750?

Similarly, the combination of these two elements leads to confusing interpretations. For instance, when discussing the cumulative premium development to loss development ratios (CPDLD), Perkins and Teng write:

> The CPDLD ratio tells how much premium an insurer can expect to collect for a dollar of loss that has yet to emerge. For instance, the first CPDLD ratio is 1.492, which means that each dollar of loss emerged provides the insurer one dollar and 49 cents of premium. The second CPDLD ratio is 0.556, which means that after the first retro adjustment, each additional dollar of loss provides the insurer 56 cents of premium.

[14]Actually, although the slopes of each line segment remain fixed, the length of the line segments may be changed. At each retrospective adjustment, Perkins and Teng re-estimate the losses expected to be reported in each subsequent period. These revisions, however, are generally minor.

The interpretation of the second CPDLD ratio is correct. The interpretation of the first CPDLD ratio, however, is mistaken. The first CPDLD ratio relates to all the expected losses from policy inception, at least according to the procedure in the Perkins and Teng paper.

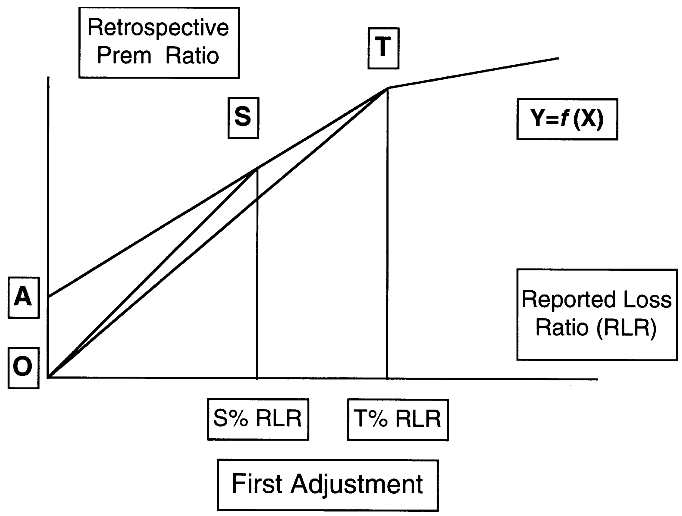
How should we interpret the 1.492 CPDLD ratio from policy inception to the first retrospective adjustment? Consider a relatively wide-swing retrospective rating plan: that is, a plan with high loss limits and maximum premiums. The amount of expected premium for each dollar of loss equals the loss conversion factor times the tax multiplier, minus a small amount for the non-rateable losses. This product may be about 1.200. The remainder of the first CPDLD ratio which Perkins and Teng calculate is the basic premium charge divided by the expected loss ratio (as a function of standard premium). For a basic premium charge of 25% and a standard loss ratio of 85%, this calculation gives $0.25 \div 0.85 = 0.294$. Adding 1.200 to 0.294 gives 1.494, which is about equal to the empirical figure which Perkins and Teng compute. In other words, when the basic premium charge is disentangled from the slopes of the line segments, the Perkins and Teng procedure corresponds intuitively with the actual retrospective rating formula.[15]

The failure to separate these two issues makes it harder for the actuary to analyze changes in the figures over time. For instance, what causes the steady rise in the slope of the first line segment from an average of 1.254 in policy year 1963 to an average of 1.825 in policy year 1992 (see Exhibit 4, Sheet 1 in the original paper)? Is it caused by a change in the average basic

[15]For a plan with significant loss limits or maximum premiums, the intuitive is analogous. The lower the loss limits, or the lower the maximum premium, the weaker will be the premium responsiveness, but the basic premium charge will be greater, because the insurance charge will be larger. These two effects will offset each other, since the insurance charge is calculated as the expected losses arising from the loss limits and maximum premiums.

FIGURE 6

LOSS REPORTING PATTERNS IN THE PDLD GRAPH



premium ratio, or is it caused by a change in premium responsiveness during the first period? These two factors are shown separately in the graphs drawn in this discussion, but they are not easily distinguished in the way that Perkins and Teng show their procedure.

This change could also be caused by a lengthening of the loss reporting pattern. This is an equally likely cause, and a graphical representation of it is illuminating.

In Figure 6, the basic premium ratio and the slope of the first line segment are not changed, but the percentage of losses expected to be reported before the first adjustment is decreased. That is, the expected ultimate loss ratio remains the same, but the expected reported loss ratio at the first adjustment decreases from T to S. The first line segment is therefore shorter, though it has the same slope. In the PDLD procedure, however, the slope of the first line segment appears to increase. That is, the slope from 0 to S is greater than the slope from 0 to T.[16]

Fortunately, it is simple to adjust the PDLD method to show the basic premium ratio separately from the true slope of the first line segment. One need only estimate the average basic premium charge as a ratio to the standard loss ratio, and then subtract this figure from the first CPDLD.

3. LOSS-SENSITIVE CONTRACTS AND UNDERWRITING RISK

Insurance serves several important economic functions, such as the transfer of the risk of financial loss from the consumer to the insurance company. Because of the unlimited nature of workers compensation benefits, a single severe workplace injury might financially impair a small employer. The transfer of this risk from the employer to the insurance company is a societal benefit of workers compensation insurance.

A societal downside to insurance is moral hazard. If there were no workers compensation insurance, then employers would take great pains to keep their workplaces as safe as possible, since they would shoulder any cost of workplace accidents. Insurance has two effects on employers' safety efforts. On the one hand, the loss engineering staffs of most workers compensation carriers can identify potential workplace hazards and improve employers' safety procedures. On the other hand, some employers become less concerned with employee safety, since they no longer bear all the costs.

An increase in moral hazard hurts both employees and employers. It hurts employees since workplace accidents may in-

[16]The effect is even more pronounced in the Perkins and Teng graph, which is drawn as a concave curve instead of a series of line segments.

crease. It hurts employers in numerous ways: there are training costs for new employees, work flows are interrupted, and workers compensation premiums increase to cover the higher loss costs.

Retrospectively rated contracts are an attempt to achieve the benefits of insurance while reducing the drawbacks. Employers are protected from the risk of large losses that might otherwise bankrupt the firm. But they still bear the cost of most other injuries, so moral hazard is kept low.

Insurance involves the transfer of risk from the consumer to the insurer. In retrospectively rated contracts, some of this risk is transferred back to the consumer. The NAIC has developed the loss-sensitive contract offset to the underwriting risk charges in the risk-based capital formula in order to reflect the fact that the risk on retrospectively rated contracts differs from the risk on prospectively rated contracts. Previous actuarial studies had not addressed this question, and the American Academy of Actuaries Task Force on Risk-Based Capital had little actuarial or statistical data to give to the NAIC.

The PDLD procedure, however, provides a direct answer. In fact, the Perkins and Teng paper sheds light on the potential limitations of both the risk-based capital loss-sensitive contract offset and the loss-sensitive contract exhibits in Part 7 of Schedule P.

Underwriting Risk

The insurance contract transfers the risk of random loss occurrences from the consumer to the insurance company. This risk is primarily process risk. For instance, suppose the consumer is an employer concerned with industrial accidents. The employer may estimate that there is a one in one hundred chance of a severe accident in his workplace this year. The primary risk that this employer faces is *not* that he has misestimated the probability—

that it is truly one in ninety, not one in one hundred. Nor is it the risk that the cost of such accidents may change, say from an average of $20,000 per accident to $25,000 per accident. Rather, the primary risk is that an accident will indeed occur this year in his workplace.

The risk to the insurance company is different. It is primarily parameter risk, not process risk. If the book of business is large enough, process risk effectively disappears. However, the risk that the probability of an accident is truly one in ninety, or the risk that the average cost of these accidents is truly $25,000, are serious concerns for the insurer. A relatively small error in the estimation of these parameters may wipe out the expected profits of the insurer.

Loss-sensitive contracts mitigate this risk for the insurance company. The insured is still protected against random large losses by the loss limit in the retrospective rating plan and by the maximum premium. Meanwhile, the insurance company is protected against the accumulation of more losses than expected, or a rise in the average cost per claim, by the responsiveness of retrospective premiums to incurred losses.[17]

Underwriting risk has two facets. Premium risk (or "written premium risk," in the NAIC risk-based capital terminology) is the risk that future premiums will prove inadequate to cover the future losses and expenses. This risk takes a variety of forms. For instance, there is a market risk that the competitive pressures of an underwriting cycle downturn will force premium rates below adequate levels. There is a regulatory/political risk that needed premium increases will not be approved or that new types of claims will be deemed compensable by the courts.

Reserving risk is the risk that the reserves held for accidents that have already occurred may prove inadequate. Once again,

[17]For a full discussion of the effects of loss-sensitive contracts on workers compensation reserving risk, see Hodes, Feldblum and Blumsohn [7].

this risk takes a variety of forms. For instance, there is the economic risk that a recession will cause injured employees to remain on disability for longer periods, since there may be no jobs to return to (workers compensation). Or there may be judicial risk, that courts or juries may grant higher awards to claimants (general liability).

Loss-Sensitive Contracts and Underwriting Risk

Loss-sensitive contracts reduce the risks to the insurer, since if losses are higher than expected, additional premiums are collected from the insureds. When the NAIC instituted its risk-based capital formula, which quantified the capital needed to guard against written premium risk and reserving risk, several large commercial lines insurers argued that a capital requirement that is appropriate for prospectively rated business is too high for retrospectively rated business, since the retrospective rating formula itself protects against unexpectedly high losses.

But how effective are these contracts in mitigating risk? In other words, how responsive are the premiums to unexpected losses?

If there were no loss limits or maximum premiums in the retrospective rating plans, the premium responsiveness would equal the product of the loss conversion factor and the tax multiplier. We term this 100% responsiveness, since the loss conversion factor generally covers loss-related expenses and the tax multiplier pays for premium taxes (and other state assessments) that depend upon the losses incurred or the premium collected. In other words, with 100% responsiveness, the insurer would get \$1.00 in extra premium for each \$1.00 in additional losses and loss-related expenses.

If there were no loss limits or maximum premiums in the retrospective rating plans, then the insurer would not be exposed to underwriting risk. If underwriting results are worse than expected, or if reserves develop adversely, the insurer would collect the full loss from the insured through retrospective premium adjustments. There remain some other risks, such as the credit risk that the insured will not be able to pay the retrospective premiums when they come due, but these risks are usually far smaller than the underwriting risk.

In practice, of course, there are loss limits and maximum premiums. Premium responsiveness is less than 100%. So the NAIC instituted a 30% loss-sensitive contract offset on primary insurance policies and a 15% loss-sensitive contract offset on reinsurance treaties. The loss-sensitive contract offset of 30% means that if the risk-based capital underwriting risk charge for a block of prospectively rated business is $\$X$, then the corresponding charge for the same book of business written on loss-sensitive contracts is $\$X * (1 - 30\%)$.[18]

In other words, the primary insurance loss-sensitive contract offset assumes (conservatively) that the premium responsiveness is only 30%. That is to say, for each \$1.00 in additional losses and loss-related expenses, \$0.30 of additional premium (on average) is collected.

The 30% figure was not based on definitive data because credible industry data on premium responsiveness was not available. The consulting firm Tillinghast/Towers Perrin conducted an industry-wide survey of 16 large writers of retrospectively rated contracts, and calculated an average premium responsiveness of 65%. The survey asked insurance companies how responsive they thought their loss-sensitive contracts were to unexpected loss emergence or unexpected loss development. The 65% was a rough average of the company estimates. Adjusting this figure downward for conservatism and for the potential

[18] For a complete description of the loss-sensitive contract offset in the risk-based capital formula, see Feldblum [5].

credit risk led to the 30% offset factor in the risk-based capital formula.[19]

In order to obtain industry data to more accurately estimate the loss-sensitive contract offset factor, the NAIC added Part 7 to Schedule P. The exhibits in this section of Schedule P are designed to allow the estimation of premium responsiveness on loss-sensitive contracts. These exhibits are a considerable advance over the information available previously, but they are far less useful than the information provided by reserving studies using the PDLD method.

In the future, insurance companies will seek to better quantify the effects of loss-sensitive contracts on underwriting risk, and state regulators will attempt more accurate estimations of the appropriate offset factor for these contracts. The study by Perkins and Teng highlights several areas that must be carefully considered.

Time Frames

The Schedule P Part 7 exhibits are the NAIC's attempt to quantify premium responsiveness, using the same method as Perkins and Teng, but with annual reporting of premiums and losses. The Perkins and Teng paper shows that the Schedule P results will be distorted in several ways, possibly to the extent that premium responsiveness will not be shown at all. Some of the problems can be corrected (in theory, at least) by means of the procedures in the Perkins and Teng paper; other distortions may be more difficult to remove.

---

[19]The rationale given by the Tillinghast study and adopted by the NAIC for the lower (15%) offset factor used for reinsurance treaties reflects the different types of loss-sensitive contracts generally used by primary companies and by reinsurers. The primary company retrospective rating plan adjusts the premiums billed for adverse loss experience. Some of these plans have extremely wide swings, in that the final premium may be as much as 100% more than the standard premium. Reinsurers generally use sliding scale commissions, in that the reinsurance commission remitted to the ceding company depends upon the loss experience on the book of business. Since the commission rate

The intended use of the Schedule P Part 7 exhibits is not explained in the Annual Statement Instructions, and few actuaries understand how these exhibits purport to quantify premium responsiveness. Let us first clarify the intention of this part of Schedule P with an illustration. We will then explain the problems with the statutory exhibits by a comparison with the Perkins and Teng paper.

The risk-based capital reserving risk charge is based on the loss reserves—both case and IBNR reserves—that are shown by the company's Schedule P, Part 2, minus Schedule P, Part 3. The reserving risk charge quantifies the capital needed to protect against the risk that these reserves may develop adversely in a worst-case scenario. The loss-sensitive contract offset factor reduces this capital requirement to reflect the additional premium that the insurer expects to receive in this worst-case scenario.

The dollar amount of adverse development of the loss reserve equals the dollar amount of adverse development of the incurred losses in Schedule P, Part 2. Part 7 of Schedule P displays incurred losses on loss-sensitive contracts and the corresponding adverse or favorable premium development relative to the adverse or favorable loss development.

An Illustration

An example should clarify this. Suppose we are given the extracts from Schedule P, Part 7A, Sections 2 through 5 shown in Table 1 (figures are in thousands of dollars). The actual exhibits contain more cells, but these extracts suffice to illustrate the quantification techniques. We wish to determine premium responsiveness from 24 to 36 months and from 36 to 48 months.

The sections of Schedule P, Part 7A, contain the following historical triangles, by policy year and valuation date, of experi-

is bounded below by 0%, and in many treaties it is bounded below by an even higher amount, the swing of the typical reinsurance treaty is much narrower than that of many primary retrospective rating plans.

TABLE 1

SCHEDULE P, PART 7A, SECTIONS 2, 3, 4, AND 5, SELECTED ENTRIES ($000)

Section 2	1994	1995	1996	1997
1994	1,000	2,200	2,400	2,500
1995		1,100	2,500	2,650
1996			1,200	3,000
1997				1,500

Section 3	1994	1995	1996	1997
1994	350	550	300	200
1995		400	600	450
1996			450	650
1997				500

Section 4	1994	1995	1996	1997
1994	1,500	3,150	3,300	3.350
1995		1,650	3,600	3,700
1996			1,800	4,200
1997				2,000

Section 5	1994	1995	1996	1997
1994	0	200	150	110
1995		0	210	155
1996			0	220
1997				0

ence on loss-sensitive contracts:[20]

- Section 2: Incurred losses and ALAE on loss-sensitive contracts
- Section 3: IBNR plus bulk loss and ALAE reserves on loss-sensitive contracts
- Section 4: Earned premium on loss-sensitive contracts

[20]For a full description of Schedule P, Part 7, see Feldblum [4].

- Section 5: Accrued retrospective premium reserves on loss-sensitive contracts.

This illustration is contrived. It is designed to show how Part 7 of Schedule P was intended to be used. We then examine how the Perkins and Teng paper explains the problems with this use of the Part 7 exhibits.

These exhibits are policy year exhibits, not accident year losses (as in Parts 2, 3, and 4 of Schedule P) or exposure year premiums (as in Part 6 of Schedule P). In Section 2 of Part 7, the incurred losses as of 24 months are about twice the incurred losses as of 12 months. This makes sense: the policy year 1994 incurred losses as of 12 months are those losses on policies written in 1994 that have occurred by December 31, 1994. These are about half of the policy year 1994 losses. By December 31, 1995, all of the policy year 1994 losses have occurred (though they have not necessarily all been reported by this time), so the 24 month figure is about twice as great as the 12 month figure.

The same is true for Section 4, showing the policy year earned premiums. By the end of the policy year, all the premiums have been written (though not necessarily collected), but only about half of these premiums have been earned.

This example assumes that the initial written premium for this block of business is the estimated ultimate net premium. Initially, there is no retrospective premium reserve. At the first retrospective adjustment, some premiums are returned to policyholders, since not all losses have yet been recorded, even though the insurer knows that there will probably be development on the reported losses. The accrued retrospective premium asset becomes positive after the first adjustment. For companies that charge initial premiums below the estimated ultimate net premium (for competitive reasons), the accrued retrospective premium asset will be positive from policy inception.

Quantifying Premium Responsiveness

Consider first the premium responsiveness from 24 to 36 months. Only policy years 1994 and 1995 in our illustration are mature enough to measure this.[21] For policy year 1994, losses develop from \$2.20 million to \$2.40 million from 24 months to 36 months, for a change of \$0.20 million. Premiums develop from \$3.15 million to \$3.30 million during the same time period, for a change of \$0.15 million. The premium responsiveness is \$0.15 million ÷ \$0.20 million, or 75%.

For policy year 1995, losses develop from \$2.50 million to \$2.65 million from 24 months to 36 months, for a change of \$0.15 million. Premiums develop from \$3.60 million to \$3.70 million during the same time period, for a change of \$0.10 million. The premium responsiveness is \$0.10 million ÷ \$0.15 million, or 67%.

As the estimated premium responsiveness from 24 months to 36 months, we might take the average of these two numbers. Alternatively, we might give more weight to the 1995 policy year, particularly if the rating plan parameters had changed in 1995.

For the premium responsiveness from 36 months to 48 months, only policy year 1994 is sufficiently mature to provide the needed figures. Losses develop from \$2.40 million to \$2.50 million from 36 months to 48 months, for a change of \$0.10 million. Premiums develop from \$3.30 million to \$3.35 million during the same time period, for a change of \$0.05 million. The premium responsiveness is \$0.05 million ÷ \$0.10 million, or 50%.

This is consistent with the Perkins and Teng paper. As reserves mature, premium responsiveness diminishes, since more losses are censored by the loss limit and more premiums are capped

[21] In an actual Schedule P, all earlier policy years would also show this relationship.

by the maximum premium. In addition, at later maturities, some retrospective rating plans are closed.

This example was designed to illustrate the intended use of the Schedule P exhibits; it would rarely be encountered in practice. The incurred losses here develop smoothly upward, and the premiums follow them equally smoothly. An adequately reserved company should show flat incurred losses along development periods, and similarly flat earned premiums. After all, these incurred losses include IBNR and bulk reserves, and the earned premiums include the accrued retrospective premium asset. The changes in incurred losses from period to period would be sometimes small and sometimes large, sometimes positive and sometimes negative, resulting primarily from random loss fluctuations. The changes in earned premiums from period to period would be equally variable, resulting again from random loss fluctuations as well as from censoring by the loss limits and capping by the premium maximums.[22]

We have two series of variable figures with means of zero, since favorable and adverse development are equally likely (in theory, at least). The ratios of these series will be even more variable, sometimes very high, sometimes very low, sometimes positive, and sometimes negative. These ratios may not tell us much about premium responsiveness.

Reported Losses and Billed Premium

As the Perkins and Teng paper shows, premium responsiveness does not deal with the relationship of changes in total earned premium to changes in total incurred losses. Rather, it deals with the relationship of changes in billed premium to changes in re-

[22]The date of recognition of additional losses or additional accrued retrospective premium reserves would add to the variability in the two series of changes, one of incurred losses and one of earned premiums. That is, the reserving actuary may recognize the potential increase in ultimate losses in one year, but he or she may not book the corresponding increase in the accrued retrospective premium reserves until some later time.

ported losses. Accordingly, Schedule P, Part 7 allows that analysis to be performed as well.

Section 2 of Part 7 shows incurred losses, and Section 3 shows IBNR and bulk reserves. The difference between Sections 2 and 3 represents reported losses.[23] Similarly, Section 4 shows total earned premiums, and Section 5 shows the net reserve for premium adjustments and accrued retrospective premiums. The difference between Sections 4 and 5 represents billed premium.

Let us repeat the premium responsiveness calculations using the simulated Schedule P, Part 7 exhibits provided above. For the premium responsiveness from 24 months to 36 months, we have data from policy years 1994 and 1995. For policy year 1994, reported losses develop from ($2.2 million–$0.55 million) at 24 months to ($2.4 million–$0.3 million) at 36 months, for a change of $0.45 million. Billed premium develops from ($3.15 million–$0.2 million) at 24 months to ($3.3 million–0.15 million) at 36 months, for a change of $0.20 million. Premium responsiveness from 24 months to 36 months is $0.20 million ÷ $0.45 million = 44.4%.

For policy year 1995, reported losses develop from ($2.50 million–$0.60 million) at 24 months to ($2.65 million–0.45 million) at 36 months, for a change of $0.30 million. Billed premium develops from ($3.6 million–$0.21 million) at 24 months to ($3.70 million–$0.155 million) at 36 months, for a change of $0.155 million. Premium responsiveness from 24 months to 36 months is $0.155 million ÷ $0.30 million = 51.7%.

Anticipated Emergence versus Unanticipated Development

These figures do indeed reflect reality, but is this reality related to the risk-based capital loss-sensitive contract offset factor?

---

[23]This is the same as the calculation of accident year reported losses as Part 2 of Schedule P minus Part 4 of Schedule P.

The risk-based capital reserving risk charge seeks to quantify the amount of capital needed to guard against unanticipated adverse development of loss reserves. For instance, if in a worst-case (but still reasonable) scenario, the company's reserves would develop adversely by $15 million, then the company should hold $15 million of capital to ensure its solvency.

The figures calculated in the preceding section measure the responsiveness of retrospective premiums to the emergence of anticipated losses. They do not tell us how responsive the retrospective premiums would be to the emergence of unanticipated losses.

An example should clarify this. Suppose we are examining the premium responsiveness from 24 months to 36 months on a workers compensation retrospectively rated plan with an average swing. Suppose that at 24 months the reported losses are $100 million, and the anticipated reported losses at 36 months are $120 million. The expected ultimate losses are $150 million.

From our hypothetical experience, we find a premium responsiveness for this period of 50%. That is to say, when reported losses increase by $20 million, the billed premium increases by $10 million. What are the implications for large and unanticipated adverse loss development, as envisioned in the risk-based capital worst-case scenario? For example, if the ultimate losses are re-estimated at $180 million at 36 months instead of $150 million, will the accrued retrospective premium asset increase by an additional $15 million, or 50% of the additional losses of $30 million?

Consider the real-world characteristics of the numerical example given above. The development of reported losses from $100 million to $120 million from 24 months to 36 months may be decomposed into several parts. One part is the lengthening of some temporary cases for another few months, or an increase in some medical benefits. This development is rateable, so premium responsiveness is high. Another part is the reclassification of some temporary total cases, such as lower back sprains, into lifetime pension cases, when it becomes clear that the injured employee will not be returning to work. Only some of this development is rateable, and the rest is truncated by the loss limits or the maximum premiums.

Large and unanticipated adverse loss development has a heavy proportion of this nonrateable element. The re-estimation of the ultimate losses from $150 million to $180 million may result from the re-classification of several back sprains as severe and permanent disabilities, or from a judicial or legislative decision that certain disease claims, or psychiatric claims, are compensable. These claims are generally large and they are paid over a long period of time. A large part of these claims may not be rateable.

The Perkins and Teng paper discusses these issues. As noted above in this discussion, the premium responsiveness depends on the maturity of the losses as well as on the average loss ratio in the block of business. The emergence of anticipated losses differs from the unanticipated adverse development of the expected losses in that:

- the anticipated losses are generally paid sooner than the unanticipated losses, and
- the anticipated losses generally represent a lower loss ratio than do the unanticipated losses.

Since the anticipated losses are generally paid sooner, they are accompanied by a stronger premium responsiveness. Since the anticipated losses are generally in a lower loss ratio environment, they are associated with a stronger premium responsiveness. In sum, the figures derived from the historical triangles in Schedule P, Part 7 may not be relevant to the scenarios with which risk-based capital is concerned.

Reserving Risk Offset versus Premium Risk Offset

The NAIC risk-based capital formula uses the same loss-sensitive contract offsets for reserving risk as for written premium risk: 30% for primary insurance contracts and 15% for reinsurance contracts. As the Perkins and Teng paper shows, the offset should be much higher for written premium risk than for reserving risk.[24]

For the written premium risk loss-sensitive contract offset, one must examine the first CPDLD factor in a Perkins and Teng study. However, one must separate the basic premium charge from the premium responsiveness to losses, or the offset factor will be overstated; see the discussion above for further explanation of this. Moreover, one must remove the effects of the loss conversion factor and the tax multiplier, which would also overstate the appropriate offset factor.

For the reserving risk loss-sensitive contract offset, one must examine the CPDLD factors at each maturity. One would then weight these CPDLD factors by the distribution of reserves at each maturity. As is true for the written premium risk loss-sensitive contract offset, one must remove the effects of the loss conversion factor and the tax multiplier.

The difference between premium responsiveness to the emergence of anticipated losses and premium responsiveness to unanticipated adverse loss development (or unanticipated adverse un-

---

[24]The appropriate figures depend on the types of plans sold by the insurance company. The indicated range of figures is wide, and the type of analysis used by Perkins and Teng must be applied to each company's book of business. For instance, for a workers compensation carrier selling wide-swing plans to large national accounts, the appropriate figures may be between 80% and 85% for the written premium risk loss-sensitive contract offset and between 60% and 65% for the reserving risk loss-sensitive contract offset. For a company selling narrow swing plans to small risks, the offsets are much smaller, extending down as far as the figures used in the NAIC risk-based capital formula. For a full analysis of premium sensitivity on plans sold to small accounts, see Bender [1] and Mahler [9].

derwriting results) can be significant. In the Perkins and Teng framework, the CPDLD's should be based on a book of business with an overall loss ratio equal to the worst-case year loss ratio in the NAIC risk-based capital scenario. Empirical data for such CPDLD's are not readily accessible. Approximations by curve-fitting techniques to the CPDLD's that are empirically available may have to be substituted.

Premium Billing Lags

Another section of the Perkins and Teng paper brings to light an equally significant problem with the Schedule P exhibits. When quantifying premium responsiveness, it is important to use corresponding premiums and losses. Premium billing occurs about 3 months after the retrospective adjustment. This implies that the premium billing lags the average loss occurrence by 3 to 15 months.

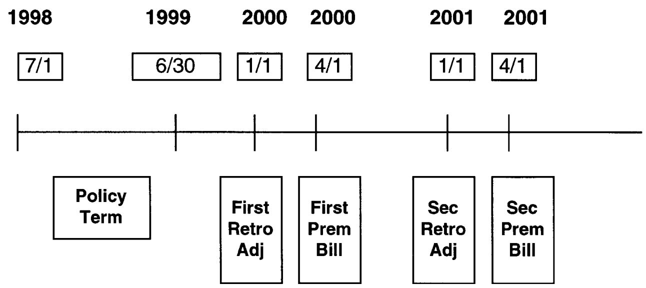
An example should clarify these figures. Suppose a policy is effective from July 1, 1998 through June 30, 1999. Retrospective adjustments are done six months after the policy's expiration and every 12 months subsequently. For this policy, the retrospective adjustments will be done on each January 1, starting with January 1, 2000. The resulting retrospective premium adjustment will be billed or returned to the policyholder on each April 1.

Each retrospective premium adjustment is driven by losses that are reported between 15 months and 3 months prior to the premium billing date. For this policy, losses reported between January 1 and December 31 affect the premium adjustment that will be billed on April 1. The schematic in Figure 7 shows this graphically.

The average lag between loss reporting and premium billing is 9 months. This is the lag used by Perkins and Teng. If one does not use any lag, as was the intention of the designers of Schedule P, Part 7, the results will be distorted. To see this most

FIGURE 7

PREMIUM AND LOSS DATES FOR RETROSPECTIVELY RATED POLICIES



clearly, suppose that:

- the retrospective premium billing is done on July 1,
- all losses occur on July 1,
- there is 100% premium responsiveness, and
- the annual incurred losses alternate between \$1,000 and \$0.

The Schedule P, Part 7, premium responsiveness test would show the following:

Year	1	2	3	4	5	6
Change in incurred losses	\$1000	\$0	\$1000	\$0	\$1000	\$0
Change in billed premium	—	\$1000	\$0	\$1000	\$0	\$1000

The premium billing shows up a year after the loss occurs. In this example, there is 100% premium responsiveness, but Schedule P, Part 7, shows a -100% premium responsiveness.[25]

[25] If X denotes the change in incurred losses, and Y is the change in billed premium, than 100% premium responsiveness is represented as $Y = 100\% * X$. This policy's experience shows a line of $Y = \$1000 - 100\% * X$. In the actual calculations of premium respon-

In practice, simplistic examinations of premium responsiveness may yield regression coefficients which are negative or seemingly random. The reserving actuary may conclude that the data are incorrect, when the true problem is an improper matching of premiums and losses.

The Perkins and Teng paper shows a possible solution to our problem. Ideally, one should use quarterly data, with a 9-month lag between premium billing dates and loss reporting dates. Few insurers have this data, and the costs of obtaining such data far outweigh any benefits from these exhibits. As a practical alternative, one should use a 12-month lag in the quantification of premium responsiveness. A 12-month lag is not ideal, but it is better than no lag at all. Moreover, this requires no change in the exhibit completion process: the same exhibits may be used, but the quantification procedure would be modified.

4. CONCLUSION

Miriam Perkins and Michael Teng have put together an excellent paper, based on eight years of carefully examining the accrued retrospective premium reserves in workers compensation, general liability, and commercial auto for one of the country's largest writers of retrospectively rated policies. They methodically analyzed how premium responsiveness changes by reserve maturity and by aggregate loss ratio, and they systematically tested the lags between loss reporting and premium billing in the company's book of business.

The Perkins and Teng procedure is important not just for reserve projections but also for risk analysis. Our profession has much to gain as other actuaries learn the techniques presented by Perkins and Teng and use them to quantify the risk and rewards of loss-sensitive contracts.

---

siveness, of course, one does not use successive adjustments for a single policy or block of policies, but successive calendar years for the same adjustment for successive blocks of policies. The underlying concepts are the same, though the schematic becomes more complex.

REFERENCES

[1] Bender, Robert K., "Aggregate Retrospective Premium Ratio as a Function of the Aggregate Incurred Loss Ratio," *PCAS* LXXXI, 1994, pp. 36–74.

[2] Berry, Charles H., "A Method for Setting Retro Reserves," *PCAS* LXVII, 1980, pp. 226–238.

[3] Feldblum, Sholom, "Workers' Compensation Ratemaking," CAS Examination 5 Study Note, 1993.

[4] Feldblum, Sholom, "Completing and Using Schedule P," *Regulation and the Casualty Actuary*, NAIC, 1997.

[5] Feldblum, Sholom, "NAIC Property-Casualty Risk-Based Capital Requirements," *PCAS* LXXXIII, 1996, pp. 297–435.

[6] Fitzgibbon, Walter J., Jr., "Reserving for Retrospective Returns," *PCAS* LII, 1965, pp. 203–214.

[7] Hodes, Douglas M., Sholom Feldblum, and Gary Blumsohn, "Workers Compensation Reserve Uncertainty," Casualty Actuarial Society *Forum*, Summer 1996, pp. 61–149.

[8] Hope, F. J., Discussion of "Reserving for Retrospective Returns," *PCAS* LIII, 1966, pp. 185–187.

[9] Mahler, Howard C., Discussion of "Aggregate Retrospective Premium Ratio as a Function of the Aggregate Incurred Loss Ratio," *PCAS* LXXXI, 1994, pp. 75–90.

[10] Morrell, Roy K., Discussion of "A Method for Setting Retro Reserves," *PCAS* LXVIII, 1981, pp. 107–110.

[11] Uhthoff, D. R., Discussion of "Reserving for Retrospective Returns," *PCAS* LIII, 1996, pp. 187–189.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXXIV

RETROSPECTIVE RATING: 1997 EXCESS LOSS FACTORS

WILLIAM R. GILLAM AND JOSE R. COURET

DISCUSSION BY HOWARD C. MAHLER

INTRODUCTION

This discussion will present some of the mathematical aspects of the effect of dispersion of loss development on excess ratios. It will be shown how the formulas developed in "Retrospective Rating: 1997 Excess Loss Factors" fit into this more general mathematical framework.

THE PROBLEM

Even if one included average loss development beyond fifth report in the estimation of excess ratios, there are at least two phenomena that would affect excess ratios that are not being considered. First, the different sizes of claims may have varying expected amounts of development. If larger claims had higher average development, this would raise the excess ratios for higher limits.

Secondly, there is a "dispersion" effect. Assume we have two claims of $1 million each that are expected on average to develop by 10%. It makes a difference whether we assume we'll have two claims each at $1.1 million or one claim at $1 million and one claim at $1.2 million. The ratio excess of $1.1 million will differ in the two cases.[1]

It is assumed for simplicity that there is no development on average; alternatively, any average development has already been

[1] In the former case it is zero, since there are no dollars excess of $1.1 million. In the latter case it is 0.1/2.2, since there are $1.2–$1.1 million = $.1 million dollars excess of $1.1 million, and total losses of $2.2 million.

incorporated into the size of loss distribution. However, some individual claims will develop more than average while others will develop less than average. In total, the average development factor is assumed to be unity.

SIMPLE EXAMPLE

Assume we have a piece-wise linear size of accident distribution such that:[2]

$$\begin{aligned} F(0) &= 0 \\ F(100) &= .90 \\ F(1{,}000) &= .99 \\ F(5{,}000) &= 1.00. \end{aligned}$$

Any size of loss distribution can be approximated sufficiently well by such an "ogive."[3] For actual applications one would have many more intervals, but this example will illustrate the principles involved.

The probability density function is:

$$f(x) = \begin{cases} .009 & 0 < x \le 100 \\ .0001 & 100 < x \le 1{,}000 \\ .0000025 & 1{,}000 < x \le 5{,}000 \\ 0 & x > 5{,}000. \end{cases}$$

One can compute the average size of claim as the sum of three integrals of $xf(x)$:

$$\begin{aligned} E[X] &= \int_0^{100} (.009)x\,dx + \int_{100}^{1000} (.0001)x\,dx \\ &\quad + \int_{1000}^{5000} (.0000025)x\,dx \\ &= 45 + 49.5 + 30 = 124.5. \end{aligned}$$

[2]Assume everything is in units of thousands of dollars. Thus, 5,000 actually corresponds to $5 million.

[3]See Hogg and Klugman [3].

The excess ratio at limit L can be computed as:

$$R(L) = \int_L^\infty (x - L) f(x)\,dx / E[X].$$

In this case we can compute the numerator as a sum of three terms:

$$\begin{aligned}
&\int_L^\infty (x - L) f(x)\,dx \\
&\qquad = (\text{If } L < 100) \int_L^{100} (x - L)(.009)\,dx \\
&\qquad\quad + (\text{If } L < 1000) \int_{\max[100,L]}^{1000} (x - L)(.0001)\,dx \\
&\qquad\quad + (\text{If } L < 5000) \int_{\max[1000,L]}^{5000} (x - L)(.0000025)\,dx.
\end{aligned}$$

If $L < 100$, let

$$\begin{aligned}
R_1(L) &= \int_L^{100} (x - L)\,dx \Big/ \int_0^{100} x\,dx \\
&= \text{excess ratio at } L \text{ if losses are uniformly distributed on the interval } [0, 100].
\end{aligned}$$

Note that $R_1(L) = 0$ if $L \geq 100$. Then the first term above is

$$R_1(L) \int_0^{100} .009x\,dx = R_1(L) E_1[X],$$

where $E_1[X] = \int_0^{100} (.009)x\,dx$ is the contribution to the overall mean from claims in the first interval. Then, working similarly with the other two intervals:

$$\int_0^\infty (x - L) f(x)\,dx = R_1(L)E_1[X] + R_2(L)E_2[X] + R_3(L)E_3[X],$$

$$R(L) = \frac{R_1(L)E_1[X] + R_2(L)E_2[X] + R_3(L)E_3[X]}{E_1[X] + E_2[X] + E_3[X]}.$$

Thus, the overall excess ratio can be expressed as a weighted average of excess ratios each computed as if the losses were uniformly distributed on an interval. The weights are the contributions to the overall mean of the claims in each interval. In this case, the weights are 45, 49.5, and 30 or 45/124.5, 49.5/124.5, and 30/124.5.

For example, for a limit of 70, the individual excess ratios are:[4] .09, .8727, and .9767. The weighted average is

$$R(70) = \frac{(45)(.09) + (49.5)(.8727) + (30)(.9767)}{124.5} = .6149.$$

Further, if the losses were uniform from 100 to 1000 then the excess ratio would be:

$$\frac{1}{900}\int_{100}^{1000}(x-70)\,dx \Big/ \frac{1}{900}\int_{100}^{1000} x\,dx = (550-70)/550$$
$$= 480/550 = .8727.$$

Table 1 shows the excess ratios for this simple example for several limits. As can be seen, in the absence of any loss development, the ratio excess of 5,000 is zero; there are no losses above 5,000.

SIMPLE DISPERSION

Assume for simplicity that each accident has an equal likelihood of developing in a manner such that it is divided[5] by either: .75, .833, 1, 1.25, or 1.5. Then the average development is

$$\frac{1}{5}\left(\frac{1}{.75} + \frac{1}{.833} + \frac{1}{1} + \frac{1}{1.25} + \frac{1}{1.5}\right) = 1.$$

---

[4]For losses distributed uniformly on $[a,b]$, for $b > L > a$, $R(L) = (b-L)^2/(b^2 - a^2)$; for $L < a$, $R(L) = 1 - 2L/(b+a)$; for $L > b$, $R(L) = 0$. For the interval [0,100] we have the first case. For the intervals [100,1,000] and [1,000,5,000] we have the second case.

[5]Loss development *divisors* are used in order to match the presentation in "Retrospective Rating: 1997 Excess Loss Factors." Loss development multipliers or factors could have been used equally well for the presentation.

TABLE 1

EXCESS RATIOS*

LIMIT	No Development	Simple Dispersion**	Gamma Dispersion***
50	.6888	.6813	.6945
100	.5582	.5597	.5684
500	.3012	.2932	.3076
1,000	.1606	.1632	.1721
2,000	.0904	.0856	.0930
3,000	.0402	.0394	.0459
4,000	.0100	.0156	.0190
5,000	.0000	.0045	.0069
6,000	.0000	.0005	.0024
7,000	.0000	.0000	.0008
8,000	.0000	.0000	.0003
9,000	.0000	.0000	.0001
10,000	.0000	.0000	.0000

*For simple piece-wise linear distribution with $F(0) = 0$, $F(100) = .9$, $F(1000) = .99$, $F(5000) = 1$.
**For five possibilities, see text. Mean development = 1; Variance of development = .060.
***For a gamma loss divisor with $\alpha = 16.67$, $\lambda = 15.67$, see text. Mean development = 1; Variance of development = .060.

Thus, the total expected losses are unaffected. The variance of the development is .060.

We can compute excess ratios for each of the five possibilities and average the results together. If all the accidents were divided by 1.25; i.e., multiplied by .8, then a limit of 100 is equivalent to a limit of 125 without any development. So the excess ratio for 100 for the developed losses can be computed as $R(125)$ for the original distribution.[6]

Thus, to compute the excess ratio for the developed losses for a limit of 100:

$$\hat{R}(100) = \tfrac{1}{5}(R(75) + R(83.3) + R(100) + R(125) + R(150))$$
$$= \tfrac{1}{5}(.6009 + .5817 + .5582 + .5384 + .5191) = .5597.$$

[6]If each of the accidents are divided by 1.25, then the ratio excess of a limit of 100 declines from .5582 to .5384. Reducing the size of accidents reduces the excess ratio over any fixed limit.

Similarly, we can compute the excess ratio for the developed losses for a limit of 5,000:

$$\begin{aligned}\hat{R}(5000) &= \tfrac{1}{5}(R(3750) + R(4165) + R(5000)\\ &\qquad + R(6250) + R(7500))\\ &= \tfrac{1}{5}(.0157 + .0070 + 0 + 0 + 0) = .0045.\end{aligned}$$

So the dispersion effect has now produced some losses excess of 5,000, without affecting the total expected losses.

As can be seen in Table 1, the dispersion effect raises the excess ratios for higher limits and alters those for lower limits. While this example could be changed to include more than 5 possibilities, the essence of the dispersion effect has been captured. However, if the possibilities were more dispersed around the mean; i.e., if the variance of the development were greater, then the impact of the dispersion would be greater.

CONTINUOUS LOSS DIVISORS APPLIED TO A UNIFORM DISTRIBUTION ON AN INTERVAL

What if, rather than five possible loss divisors, one had a continuous probability distribution?

Assume:

1. Losses are distributed uniformly on the interval $[a,b]$.
2. Losses will develop with loss divisors r given by a distribution $H(r)$, with density $h(r)$.[7]

Then, as shown in Appendix A, the distribution function for the developed losses y, is given by:

$$F(y) = [y/(b-a)]\{E(R;b/y) - E(R;a/y)\},$$

[7] It is assumed that $\int_0^\infty (h(r)/r)\,dr$ is finite, so that the overall loss development is finite. In the case where H is a gamma distribution, this requirement means that the shape parameter s must be greater than one.

where $E[R;L]$ is the limited expected value of the distribution of loss divisors, at a limit L.

Appendix A also shows that the density function can be written in a number of forms, as summarized below:

$$\begin{aligned} f(y) &= \frac{1}{b-a}\int_{a/y}^{b/y} rh(r)\,dr \\ &= \frac{1}{b-a}\{E[R;b/y]-(b/y)(1-H(b/y))-E[R;a/y] \\ &\qquad + (a/y)(1-H(a/y))\} \\ &= \frac{1}{b-a}\{E[R;b/y]-E[R;a/y]\} \\ &\quad + \frac{1}{y(b-a)}\{bH(b/y)-aH(a/y)\}-\frac{1}{y}. \end{aligned}$$

Further, Appendix A describes how one can use the density function and distribution function to calculate the excess ratio of the developed losses, as follows:

$$\begin{aligned} R(L) = \frac{1}{b^2-a^2}\Bigg\{ & b^2\int_0^{b/L} h(r)/r\,dr - a^2\int_0^{a/L} h(r)/r\,dr \\ & + 2aLH(a/L) - 2bLH(b/L) \\ & + L^2\int_{a/L}^{b/L} rh(r)\,dr\Bigg\} \Bigg/ \int_0^{\infty} h(r)/r\,dr. \end{aligned}$$

GAMMA DISPERSION APPLIED TO THE UNIFORM DISTRIBUTION

Assume that the loss divisor r is distributed according to a gamma distribution[8] with parameter s and l:

$$h(r) = \frac{l^s r^{s-1} e^{-lr}}{\Gamma(s)},$$

where $\Gamma(n) = (n-1)!$.

[8]Then the loss multipliers are distributed according to an inverse gamma. We assume $s > 1$, so that the overall loss development is finite.

Then, as shown in Appendix B, based on the general formula in Appendix A, if the losses are uniformly distributed on the interval $[a,b]$, after development the excess ratio for the limit L is given by:[9]

$$\begin{aligned} R(L) = {} & \frac{b^2}{b^2-a^2}\Gamma(s-1;lb/L) - \frac{a^2}{b^2-a^2}\Gamma(s-1;la/L) \\ & + \frac{2L(s-1)}{(b^2-a^2)l}\{a\Gamma(s;la/L) - b\Gamma(s;lb/L)\} \\ & + \frac{L^2(s-1)s}{(b^2-a^2)l^2}\{\Gamma(s+1;lb/L) - \Gamma(s+1;la/L)\}, \end{aligned}$$

where $\Gamma(s;y) = 1/\Gamma(s)\int_0^y t^{s-1}e^{-t}\,dt$ is the incomplete gamma function.

One can apply this "gamma dispersion" effect to a piece-wise linear size of accident distribution, such as in the prior example.

The mean development is the mean of an inverse gamma, $l/(s-1)$. For this discussion, the average development is unity, so we take $l = s-1$. The variance of the development is the variance of an inverse gamma, $l^2/\{(s-1)^2(s-2)\}$. For $l = s-1$, the variance is $1/s-2$. Thus, if one takes $s = 18.67$, (and $l = 17.67$) then the variance of the development is $1/16.67 = .060$, which matches that in the simple dispersion example. However, the gamma allows extreme possibilities (with a small probability), so one gets a somewhat different behavior than in the simple dispersion example.

As seen in Table 1, using the gamma dispersion for very high limits (7,000 and above) yields excess ratios that are now positive rather than zero. Gamma dispersion can have a particularly significant impact on very high limits, particularly if the variance is large.

[9]These are the formulas developed and shown in "Retrospective Rating: 1997 Excess Loss Factors."

Each excess ratio is computed as a weighted average of the excess ratios computed for losses uniformly distributed on each of the three assumed intervals. For example, for a limit of 2,000, the excess ratio for losses distributed uniformly from 1,000 to 5,000, with gamma dispersion with $s = 18.67$ and $l = 17.67$ is given by the formula from Appendix B:

$$
\begin{aligned}
R_3&(2000)\\
&= (1.04167)\Gamma(17.67;44.175) - (.04167)\Gamma(17.67;8.835)\\
&\quad + (.1667)\Gamma(18.67;8.835) - (.8333)\Gamma(18.67;44.175)\\
&\quad + (.1761)\Gamma(19.67;44.175) - (.1761)\Gamma(19.67;8.835)\\
&= (1.04167)(.9999980) - (.04167)(.0057148)\\
&\quad + (.1667)(.0011302) - (.8333)(.999987)\\
&\quad + (.1761)(.999949) - (.1761)(.0026)\\
&= .384.
\end{aligned}
$$

Similarly, for losses uniform from 100 to 1,000, $R_2(2000) = .00008$. For losses uniform from 0 to 100, $R_1(2000) = 10^{-19}$. Taking a weighted average, using weights of 45, 49.5, and 30, one obtains $R(2000) = .093$, as displayed in Table 1.

Note that the gamma distribution used in this example has a large value of s, the shape parameter. Therefore, the distribution of loss divisors is close to normal.[10] The distribution of loss divisors has a skewness of $2/\sqrt{s}$, which is only .49. The skewness of the distribution of loss multipliers is that of an inverse gamma: $4\sqrt{(s-2)}/(s-3) = 1.12$. If one were to take a different form of distribution with a larger skewness one would have a larger chance of extreme results. Therefore, in the case of very high limits, the excess ratios would be even larger.

[10]The distribution of loss multipliers is close to an inverse normal distribution.

DISTRIBUTION OF DEVELOPED LOSSES

The particular situations examined so far are a special case of a more general framework. As shown in Appendix C, if losses at latest report are distributed via $G(x)$ and the loss divisors r are distributed independently of x via density function $h(r)$,[11] then the distribution for the developed losses y is given by:

$$F(y) = \int_0^\infty G(yr)h(r)\,dr.$$

GAMMA LOSS DIVISORS APPLIED TO AN EXPONENTIAL DISTRIBUTION

For example, if $G(x)$ is an exponential distribution $G(x) = 1 - e^{-\lambda x}$ and the loss divisors are gamma distributed $h(r) = l^s r^{s-1} e^{-lr}/\Gamma(s)$, then

$$\begin{aligned} F(y) &= 1 - \frac{l^s}{\Gamma(s)} \int_0^\infty r^{s-1} e^{-lr} e^{-\lambda yr}\,dr \\ &= 1 - \frac{l^s}{\Gamma(s)} \frac{\Gamma(s)}{(l+\lambda y)^s} = 1 - \left(\frac{(l/\lambda)}{(l/\lambda) + y} \right)^s. \end{aligned}$$

Thus $F(y)$ has a Pareto distribution as per Hogg and Klugman [3], with shape parameter s and scale parameter l/λ. Thus, the excess ratio of the developed losses is that of a Pareto distribution:

$$R(L) = \left(\frac{(l/\lambda)}{(l/\lambda) + L} \right)^{s-1}.$$

MATHEMATICAL RELATION TO MIXED DISTRIBUTIONS

The calculation of the distribution of the developed losses is the same as that used to calculate the mixed distribution in the inverse gamma-exponential conjugate prior.[12] (An inverse gamma

---

[11] With $\int_0^\infty (h(r)/r)\,dr$ finite.

[12] See Herzog [2], or Venter [4]. The mixed distribution in the case of an inverse gamma—Exponential conjugate prior is a Pareto distribution, as obtained above.

distributed multiplier is the same as a gamma distributed divisor.) In general, if the loss multipliers and the loss distribution form any of the well known pairs[13] of prior distributions of the scale parameters of the conditional distributions and conditional distributions, then the developed losses will be given by the mixed distribution. For example, as shown in Venter [4], a Weibull loss distribution and a transformed gamma loss divisor[14] would produce a Burr distribution of developed losses. Thus, there are a number of mathematically convenient examples that might approximate a particular real world application.

GAMMA LOSS DIVISORS APPLIED TO PARETO LOSSES

Since the Pareto distribution is often used to model losses (or at least the larger losses), it would be valuable to be able to apply the concept of loss divisors to the Pareto distribution.

As shown in Appendix C, one can develop the mathematics of applying gamma loss divisors to losses distributed by a Pareto distribution with parameters α and λ: $F(x) = 1 - (\lambda/(\lambda + x))^{\alpha}$. As derived in Appendix C, the excess ratio for the developed losses is given by:

$$R(L) = \left(\frac{Xl}{L}\right)^{s-1} U(s-1, s+1-\alpha, \lambda l/L),$$

where U is a confluent hypergeometric function.[15]

It is also shown in Appendix C that when the average development is unity[16] then the excess ratios of the developed losses can be approximated by replacing λ in the Pareto by $\lambda' = \lambda(s-1)/(s-(\alpha/2)-1)$.

[13] Such as shown in Venter [4]. Venter displays a list of conjugate priors, but for the current application there is no requirement that it be a conjugate prior situation.

[14] An inverse transformed gamma loss multiplier.

[15] See Appendix D and *Handbook of Mathematical Functions* [1].

[16] Also, we need the shape parameter of the gamma, s, to be greater than $\alpha + 1$.

Table 2 and Figure 1 compare the excess ratios for a Pareto with $\alpha = 3.5$ and $\lambda = 1{,}000$, for the developed losses[17] with a gamma divisor with $s = 6$ and $l = 5$, and for an approximating Pareto with $\alpha = 3.5$ and $\lambda = 1{,}000(s-1)/(s-(\alpha/2)-1) = 1{,}538$. The excess ratios for the developed losses are larger than those for the undeveloped losses. The approximation using the rescaled Pareto yields excess ratios that are too high for the lower limits, but it does an excellent job of approximating the excess ratios for higher limits.

As shown in Appendix C, in the tail, the loss development[18] multiplies the excess ratios by a factor of approximately:

$$(s-1)^{\alpha-1}\Gamma(s-\alpha)/\Gamma(s-1) \approx \left((s-1)\Big/\left(s-\frac{\alpha}{2}-1\right)\right)^{\alpha-1}.$$

In this example, this factor is: $5^{2.5}\Gamma(2.5)/\Gamma(5) = 3.1$. Figure 2 shows how this adjustment factor varies as the shape parameters of the Pareto and gamma vary. As the shape parameter of the Pareto, α, gets smaller, the losses have a heavier tail and the impact of the dispersion increases. As the coefficient of variation of the gamma[19] increases, the impact of the dispersion increases.

In general, as the coefficient of variation of the loss divisors increases, the impact of the dispersion increases. As the coefficient of variation approaches zero, we approach the situation where each claim develops by the average amount and there is no effect of dispersion. Thus, in order to apply this technique, one of the key inputs would be the coefficient of variation of the loss divisors.

CONCLUSIONS

The effect of the dispersion of loss development beyond the latest available report can be incorporated into the calculation of

[17] Then $R(L) = (\lambda l/L)^{s-1}U(s-1, s+1-\alpha, \lambda l/L) = (L/5000)^{-5}U(5, 3.5, 5000/L)$.

[18] For gamma dispersion with $l = s-1$ so the average development is unity.

[19] The coefficient of variation is the standard deviation divided by the mean. For the gamma distribution with shape parameter s, the coefficient of variation is $1/\sqrt{s}$.

TABLE 2

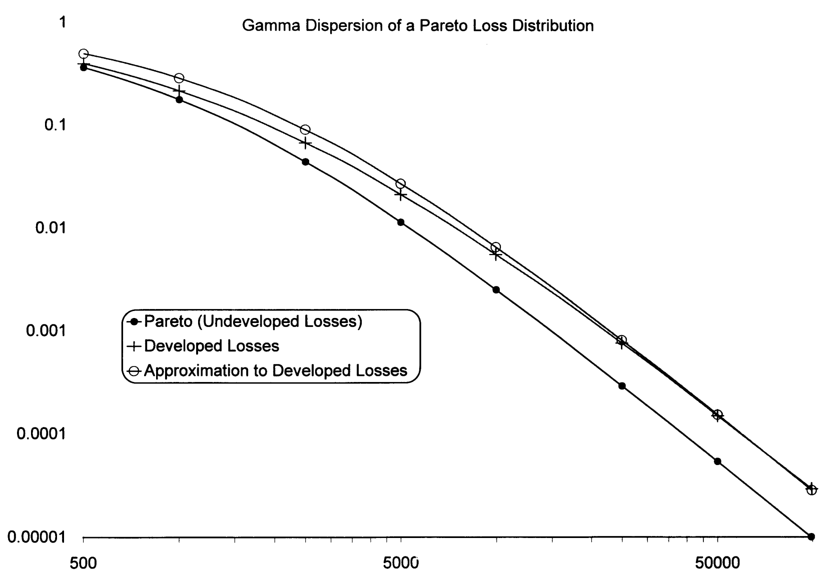
EXCESS RATIOS

LIMIT	Undeveloped Losses Pareto ($\alpha = 3.5, \lambda = 1{,}000$)	Developed Losses*	Approximating Developed Pareto ($\alpha = 3.5, \lambda = 1{,}538$)
500	.3629	.3960	.4947
1,000	.1768	.2152	.2859
2,500	.0436	.0668	.0895
5,000	.0113	.0211	.0268
10,000	.0025	.0055	.0065
25,000	.00029	.00076	.00081
50,000	.00005	.00015	.00015
100,000	.000010	.000029	.000028

*Assuming gamma loss divisor, with $s = 6$ and $l = 5$. $R(L) = (5000/L)^{2.5}U(5, 1.5; 5000/L)$.

FIGURE 1

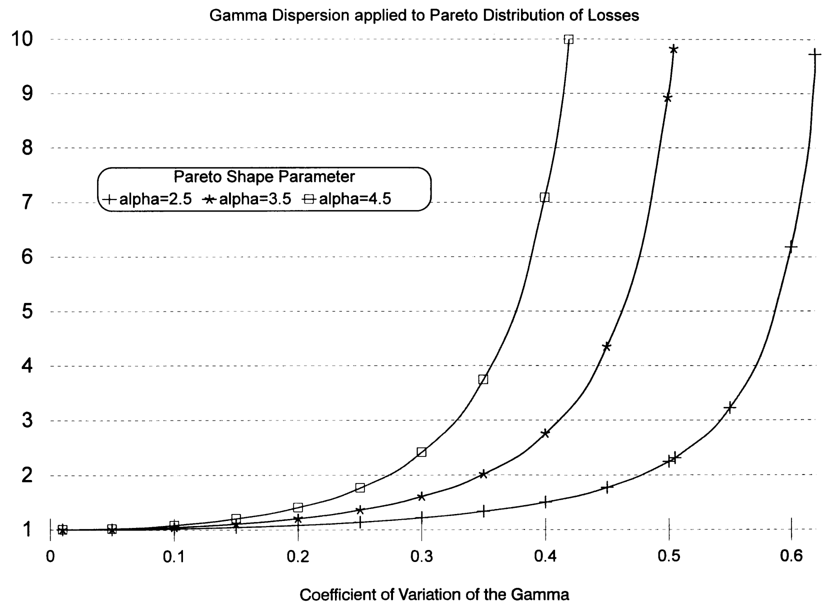
EXCESS RATIOS



Pareto Distirbution has parameters 3.5, 1000. Gamma Distribtuion has parameters 6, 5.
Approximation to Excess Ratios for Developed Losses is via a Pareto with parameters 3.5, 1538.

FIGURE 2

ADJUSTMENT FACTOR TO APPLY TO EXCESS RATIOS



excess ratios. In the case of loss dispersion which is (approximately) independent of size of loss, for many special cases one can calculate the distribution of the developed losses in closed form. In these cases, the excess ratios follow directly.

In other situations, one can approximate the loss distribution via a piece-wise linear distribution and then apply the effects of dispersion to each interval. Since on each interval the piece-wise linear approximation is a uniform distribution, one can apply the formulas developed in Appendix A. Then one can weight together the excess ratios for the developed losses from the individual intervals in order to get the excess ratio for all developed losses.

REFERENCES

[1] Abramowitz, Milton and Irene A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 1964, U.S. Department of Commerce.

[2] Herzog, Thomas N., *Introduction to Credibility Theory*, 1994, Mad River Books.

[3] Hogg, Robert V. and Stuart A. Klugman, *Loss Distributions*, John Wiley & Sons, 1984, New York.

[4] Venter, Gary G., "Credibility," Chapter 7 in *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990, pp. 375–483.

APPENDIX A

LOSS DIVISORS APPLIED TO A UNIFORM DISTRIBUTION ON AN INTERVAL

Assume:

Losses are distributed uniformly on the interval $[a,b]$. Losses will develop with loss divisors r given by a distribution $H(r)$ and density $h(r)$.

Then:

The distribution function for the developed losses y, is given by:

$$F(y) = (y/b-a)\{E[R;b/y] - E[R;a/y]\},$$

where $E[R;L]$ is the limited expected value of the distribution of loss divisors, at a limit L.

Proof:

The developed losses y are the ratio of the undeveloped losses x and the loss divisor r; $y = x/r$ or $x = yr$. Thus since x is uniform on $[a,b]$,[20] the conditional distribution of y given r is:

$$F(y \mid r) = \begin{cases} 0 & yr \le a \\ \dfrac{ry-a}{b-a} & a \le yr \le b \\ 1 & yr \ge b. \end{cases}$$

The unconditional distribution of y can be computed by integrating the conditional distribution of y given r times the assumed density function of r:

[20]For the uniform distribution on $[a,b]$, $F(x) = 0$ if $x \le a$, $F(x) = (x-a)/(b-a)$ if $a \le x \le b$, and $F(x) = 1$ if $x \ge b$.

$$
\begin{aligned}
F(y) &= \int_{r=0}^{\infty} F(y \mid r) h(r)\,dr \\
&= \int_{a/y}^{b/y} \left(\frac{ry-a}{b-a}\right) h(r)\,dr + \int_{b/y}^{\infty} h(r)\,dr \\
&= \frac{y}{b-a} \int_{a/y}^{b/y} r h(r)\,dr + \frac{a}{b-a}\left(H\left(\frac{a}{y}\right) - H\left(\frac{b}{y}\right)\right) \\
&\quad + 1 - H(b/y) \\
&= \frac{y}{b-a}\left\{ \int_0^{b/y} r h(r)\,dr - \int_0^{a/y} r h(r)\,dr + \frac{a}{y} H\left(\frac{a}{y}\right) - \frac{a}{y} \right. \\
&\qquad\qquad \left. + \frac{b}{y} - \frac{b}{y} H\left(\frac{b}{y}\right) \right\} \\
&= \frac{y}{b-a}\left\{ \left[\int_0^{b/y} r h(r) + \left(\frac{b}{y}\right)(1 - H(b/y))\right] \right. \\
&\qquad\qquad \left. - \left[\int_0^{a/y} r h(r) + \left(\frac{a}{y}\right)(1 - H(a/y))\right] \right\} \\
&= \frac{y}{b-a}\left\{ E\left[R; \frac{b}{y}\right] - E\left[R; \frac{a}{y}\right] \right\}.
\end{aligned}
$$

Similarly, we can get the density function $f(y)$. For conditional density at y given r is:

$$
f(y \mid r) = \begin{cases} 0 & yr \le a \\ \dfrac{r}{b-a} & a \le yr \le b \\ 0 & yr \ge b. \end{cases}
$$

The unconditional density at y can be computed by integrating the conditional density at y given r times the assumed density function of r:

$$\begin{aligned} f(y) &= \int_0^{\infty} f(y \mid r)h(r)\,dr \\ &= \int_{a/y}^{b/y} \frac{r}{b-a}h(r)\,dr \\ &= \frac{1}{b-a}\int_{a/y}^{b/y} rh(r)\,dr. \end{aligned}$$

We can put this type of integral in terms of limited expected values, since

$$\begin{aligned} E[R;r] &= \int_0^r rh(r)\,dr + r(1-H(r)) \\ f(y) &= \frac{1}{b-a}\{E[R;b/y] - (b/y)(1-H(b/y)) \\ &\qquad - E[R;a/y] + (a/y)(1-H(a/y))\} \\ &= \frac{1}{b-a}\{E[R;b/y] - E[R;a/y]\} \\ &\quad + \frac{1}{y(b-a)}\{bH(b/y) - aH(a/y)\} - \frac{1}{y}. \end{aligned}$$

One can use the density function and distribution function to calculate the excess ratio of the developed losses. The numerator of this excess ratio is the (developed) losses excess of L:

$$\int_L^{\infty}(y-L)f(y)\,dy = \int_L^{\infty} yf(y)\,dy - L(1-F(L)).$$

Since $f(y) = 1/(b-a)\int_{a/y}^{b/y} rh(r)\,dr$ we have

$$\int_L^{\infty} yf(y)\,dy = \frac{1}{b-a}\int_{y=L}^{\infty} y\int_{r=a/y}^{b/y} rh(r)\,dr\,dy.$$

Switching the order of integration:

$$\int_L^\infty yf(y)\,dy = \frac{1}{b-a}\int_{r=0}^{a/L}\int_{y=a/r}^{b/r} yrh(r)\,dy\,dr$$

$$+\frac{1}{b-a}\int_{r=a/L}^{b/L}\int_{y=L}^{b/r} yrh(r)\,dy\,dr$$

$$=\frac{1}{2(b-a)}\int_{r=0}^{a/L}\left(\frac{b^2}{r^2}-\frac{a^2}{r^2}\right)rh(r)\,dr$$

$$+\frac{1}{2(b-a)}\int_{r=a/L}^{b/L}\left(\frac{b^2}{r^2}-L^2\right)rh(r)\,dr$$

$$=\frac{b^2}{2(b-a)}\int_{r=0}^{b/L} h(r)/r\,dr-\frac{a^2}{2(b-a)}\int_{r=0}^{a/L} h(r)/r\,dr$$

$$-\frac{L^2}{2(b-a)}\int_{r=a/L}^{b/L} rh(r)\,dr.$$

In the course of deriving the form of the distribution function we had

$$F(y)=\frac{y}{b-a}\int_{a/y}^{b/y} rh(r)\,dr+1+\frac{a}{b-a}H\left(\frac{a}{y}\right)-\frac{b}{b-a}H\left(\frac{b}{y}\right).$$

Thus

$$1-F(L)=$$

$$\frac{b}{b-a}H(b/L)-\frac{a}{b-a}H(a/L)-\frac{L}{b-a}\int_{a/L}^{b/L} rh(r)\,dr.$$

Thus combining the terms, the numerator of the excess ratio is:

$$\int_L^\infty yf(y)\,dy - L(1-F(L))$$

$$= \frac{b^2}{2(b-a)}\int_0^{b/L} h(r)/r\,dr - \frac{a^2}{2(b-a)}\int_0^{a/L} h(r)/r\,dr$$

$$+ \frac{aL}{(b-a)}H(a/L) - \frac{bL}{b-a}H(b/L)$$

$$+ \frac{L^2}{2(b-a)}\int_{a/L}^{b/L} rh(r)\,dr.$$

The denominator of the excess ratio is:[21]

$$\int_0^\infty yf(y)\,dy = \lim_{L\to 0}\int_L^\infty yf(y)\,dy$$

$$= \frac{b^2-a^2}{2(b-a)}\int_0^\infty h(r)/r\,dr$$

$$= \frac{b+a}{2}\int_0^\infty h(r)/r\,dr.$$

Combining the numerator and denominator, the excess ratio (of the developed losses) at L is:

$$R(L) = \frac{1}{b^2-a^2}\left\{b^2\int_0^{b/L} h(r)/r\,dr - a^2\int_0^{a/L} h(r)/r\,dr\right.$$

$$+ 2aLH(a/L) - 2bLH(b/L)$$

$$\left. + L^2\int_{a/L}^{b/L} rh(r)\,dr\right\} \Big/ \int_0^\infty h(r)/r\,dr.$$

[21] The denominator of the excess ratio is the mean of the developed losses. It is equal to the product of the mean undeveloped losses $(b+a)/2$, and the average loss development $\int_0^\infty h(r)/r\,dr$.

APPENDIX B

Gamma Loss Divisors Applied to Losses Uniform on an Interval

For the situation discussed in Appendix A but for the specific case where the distribution of the loss divisors, $h(r)$, is a gamma distribution with parameters s and l:

$$\int_0^x h(r)r\,dr = \int_0^x l^s e^{-lr} r^s/\Gamma(s)\,dr = (l^s/\Gamma(s))\int_0^x e^{-lr} r^s\,dr$$

$$= (l^s/\Gamma(s))(\Gamma(s+1)/l^{s+1})\Gamma(s+1;lx)$$

$$= (s/l)\Gamma(s+1;lx)$$

$$H(x) = \int_0^x h(r)\,dr = \int_0^x l^s e^{-lr} r^{s-1}/\Gamma(s)\,dr = \Gamma(s;lx)$$

$$\int_0^x h(r)/r\,dr = \int_0^x l^s e^{-lr} r^{s-2}/\Gamma(s)\,dr$$

$$= (l^s/\Gamma(s))(\Gamma(s-1)/l^{s-1})\Gamma(s-1;lx)$$

$$= \frac{l}{s-1}\Gamma(s-1;lx)$$

$$\int_0^\infty h(r)/r\,dr = (l/s-1)\Gamma(s-1;\infty) = l(s-1).$$

Thus, using the formula from Appendix A, the excess ratio of the developed losses for limit L is in this case:

$$R(L) = \frac{b^2}{b^2-a^2}\Gamma(s-1;lb/L) - \frac{a^2}{b^2-a^2}\Gamma(s-1;la/L)$$
$$+ \frac{2L(s-1)}{(b^2-a^2)l}\{a\Gamma(s;la/L) - b\Gamma(s;lb/L)\}$$
$$+ \frac{L^2(s-1)s}{(b^2-a^2)l^2}\{\Gamma(s+1;lb/L) - \Gamma(s+1;la/L)\}.$$

For the loss divisors given by a gamma distribution with parameters s and l, we can plug in the limited expected value for the gamma distribution in terms of the incomplete gamma function:[22]

$$E[R;r] = \frac{s}{l}\Gamma(s+1;lr) + r[1-\Gamma(s;lr)].$$

Thus using the formula derived in Appendix A:

$$\begin{aligned} F(y) &= \frac{y}{b-a}\left\{E\left[R;\frac{b}{y}\right] - E\left[R;\frac{a}{y}\right]\right\} \\ &= \frac{ys}{l(b-a)}\left\{\Gamma\left(s+1;\frac{lb}{y}\right) - \Gamma\left(s+1;\frac{la}{y}\right)\right\} \\ &\quad + 1 + \frac{a}{b-a}\Gamma\left(s;\frac{la}{y}\right) - \frac{b}{b-a}\Gamma\left(s;\frac{lb}{y}\right). \end{aligned}$$

Also using the formula derived in Appendix A, the probability density function is given by:

$$\begin{aligned} f(y) &= \frac{1}{b-a}\int_{a/y}^{b/y} rh(r)\,dr \\ &= \frac{s}{(b-a)l}\{\Gamma(s+1;lb/y) - \Gamma(s+1;la/y)\}. \end{aligned}$$

[22] See Hogg and Klugman [3, page 226].

APPENDIX C

GAMMA LOSS DIVISORS APPLIED TO A PARETO DISTRIBUTION

Assume:

Losses are distributed (at latest report) on $(0,\infty)$ via a distribution function $G(x)$. Losses will develop with loss divisors r given by a density function $h(r)$.[23] (The distribution of r is independent of x.)

Then:

The distribution function for the developed losses y, is given by:

$$F(y) = \int_0^\infty G(yr)h(r)\,dr.$$

Proof:

The developed losses y are the ratio of the undeveloped losses x and the loss divisor r; $y = x/r$ or $x = yr$.

Given a value for r, the developed losses are less than y if the undeveloped losses are less than yr. Thus:

$$F(y \mid r) = G(yr).$$

Integrating over all possible values of r we have

$$F(y) = \int_0^\infty F(y \mid r)h(r)\,dr = \int_0^\infty G(yr)h(r)\,dr.$$

In the specific case where r follows a gamma distribution with parameters s and l and the undeveloped losses follow a Pareto distribution with parameters α and λ:

$$G(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha},$$

$$h(r) = l^s r^{s-1} e^{-lr}/\Gamma(s).$$

[23]It is assumed that $\int_0^\infty (h(r)/r)\,dr$ is finite, so that the average loss development is finite.

Then the distribution function for the developed losses is

$$\begin{aligned} F(y) &= \int_0^\infty h(r)G(yr)\,dr \\ &= \int_0^\infty \left(1-\left(\frac{\lambda}{\lambda+yr}\right)^\alpha\right) l^s r^{s-1} e^{-lr}/(\Gamma(s))\,dr \\ &= \int_0^\infty l^s r^{s-1} e^{-lr}/\Gamma(s)\,dr - \frac{\lambda^\alpha l^s}{\Gamma(s)}\int_0^\infty r^{s-1}e^{-lr}(\lambda+yr)^{-\alpha}\,dr. \end{aligned}$$

The first integral is unity,[24] while the second integral can be put in terms of confluent hypergeometric functions.[25]

Let $q = (y/\lambda)r$, then the second integral becomes

$$\begin{aligned} &\frac{\lambda^{s-\alpha}}{y^s}\int_{q=0}^\infty q^{s-1}e^{-\lambda l q/y}(1+q)^{-\alpha}\,dq \\ &\quad = \frac{\lambda^{s-\alpha}}{y^s}\Gamma(s)U(s,s+1-\alpha;\lambda l/y) \end{aligned}$$

where U is a confluent hypergeometric function such that[26]

$$U(a,b;z) = (1/\Gamma(a))\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}\,dt.$$

Thus the distribution function of the developed losses is:

$$\begin{aligned} F(y) &= 1 - \frac{\lambda^\alpha l^s}{\Gamma(s)}\frac{\lambda^{s-\alpha}}{y^s}\Gamma(s)U(s,s+1-\alpha;\lambda l/y) \\ &= 1 - \left(\frac{\lambda l}{y}\right)^s U(s,s+1-\alpha;\lambda l/y). \end{aligned}$$

Similarly one can compute the density function of the developed losses. Differentiating the distribution function one gets:

$$f(y) = \int_0^\infty r g(yr)h(r)\,dr.$$

[24] It is the cumulative distribution function of the gamma distribution at infinity.

[25] See Appendix D and *Handbook of Mathematical Functions* [1].

[26] See Equation 13.2.5 in *Handbook of Mathematical Functions* [1].

In the specific case where h is gamma and g is Pareto it turns out that the density of the developed losses is:

$$f(y) = \frac{s\alpha l^s \lambda^s}{y^{s+1}} U(s+1, s+1-\alpha; l\lambda/y).$$

This can be obtained either by substituting the specific form of y and h into the above integral or by differentiating $F(y)$, and making use of the facts that[27]

$$\frac{d}{dz}U(a,b;z) = -aU(a+1,b+1;z),$$
$$U(a-1,b;z) - zU(a,b+1;z) = (a-b)U(a,b;z),$$
$$\begin{aligned}
f(y) = \frac{d}{dy}F(y) &= \frac{d}{dy}\left(1 - \frac{\lambda^s l^s}{y^s} U\left(s, s+1-\alpha; \frac{\lambda l}{y}\right)\right) \\
&= \frac{\lambda^s l^s s}{y^{s+1}} U\left(s, s+1-\alpha; \frac{\lambda l}{y}\right) \\
&\quad - \frac{\lambda^s l^s}{y^s}\left(\frac{\lambda l}{y^2}\right) U'\left(s, s+1-\alpha; \frac{\lambda l}{y}\right) \\
&= \frac{s\lambda^s l^s}{y^{s+1}}\{U(s, s+1-\alpha; \lambda l/y) \\
&\qquad - (\lambda l/y) U(s+1, s+2-\alpha; \lambda l/y)\} \\
&= \frac{s\lambda^s l^s}{y^{s+1}}\{(s+1) - (s+1-\alpha)\} U(s+1, s+1-\alpha; \lambda l/y) \\
&= \frac{s\alpha\lambda^s l^s}{y^{s+1}} U(s+1, s+1-\alpha; \lambda l/y).
\end{aligned}$$

One can use the density function and distribution function to calculate the excess ratio of the developed losses. The numerator of this excess ratio is the total (developed) losses excess of L:

$$\int_L^\infty (y-L) f(y)\,dy = \int_L^\infty y f(y)\,dy - L(1 - F(L)),$$

[27] See Equations 13.4.21 and 13.4.18 in *Handbook of Mathematical Functions* [1].

$$\int_L^\infty yf(y)dy = s\alpha\lambda^s l^s \int_L^\infty y^{-s}U(s+1,s+1-\alpha;\lambda l/y)dy.$$

Letting $z = \lambda l/y$ this integral becomes

$$-s\alpha\int_0^{\lambda l/L} z^s U(s+1,s+1-\alpha;z)\frac{\lambda l}{-z^2}dz$$

$$= \lambda l s\alpha\int_0^{\lambda l/L} z^{s-2}U(s+1,s+1-\alpha;z)dz.$$

Using the theorem from Appendix D:

$$\int_L^\infty yf(y)dy = \frac{\lambda l s\alpha z^{s-1}}{s\alpha}\left[U(s,s+1-\alpha;z)+\frac{U(s-a,s+1-\alpha;z)}{(s-1)(\alpha-1)}\right]_{Z=0}^{\lambda l/L}$$

$$= \lambda l\left(\frac{\lambda l}{L}\right)^{s-1}\left(U\left(s,s+1-\alpha;\frac{\lambda l}{L}\right)+\frac{U(s-1,s+1-\alpha;\lambda l/L)}{(s-1)(\alpha-1)}\right).$$

Now

$$L(1-F(L)) = \frac{\lambda^s l^s}{L^{s-1}}U(s,s+1-\alpha;\lambda l/L).$$

Thus the numerator of the excess ratio is

$$\int_L^\infty yf(y)dy + L(1-F(L))$$

$$= \frac{\lambda^s l^s}{(s-1)(\alpha-1)L^{s-1}}U(s-1,s+1-\alpha;\lambda l/L).$$

The denominator of the excess ratio is the total (developed) losses or the mean of the undeveloped losses times the mean loss development. The former is the mean of the Pareto or $\lambda/(\alpha-1)$. The latter is the mean of the inverse gamma or $l/(s-1)$.

Thus the excess ratio is:

$$R(L) = \left(\frac{\lambda l}{L}\right)^{s-1} U(s-1,s+1-\alpha,\lambda l/L).$$

Note that this compares to the excess ratio for the undeveloped losses (given by a Pareto) of $(\lambda/\lambda + L)^{\alpha-1}$. For z small:[28]

$$U(a,b,z) \approx z^{1-b}\Gamma(b-1)/\Gamma(a) \quad b > 2.$$

Thus for large limits L, and $s > \alpha + 1$

$$\begin{aligned} R(L) &= \left(\frac{\lambda l}{L}\right)^{s-1} U(s-1, s+1-\alpha, \lambda l/L) \\ &\approx \left(\frac{\lambda l}{L}\right)^{s-1} \frac{\Gamma(s-\alpha)}{\Gamma(s-1)} \left(\frac{\lambda l}{L}\right)^{\alpha-s} \\ &= \left(\frac{\lambda l}{L}\right)^{\alpha-1} \Gamma(s-\alpha)/\Gamma(s-1). \end{aligned}$$

For the Pareto for large limits

$$R(L) = (\lambda/(\lambda + L))^{\alpha-1} \approx (\lambda/L)^{\alpha-1}.$$

Thus the ratio of the excess ratios for the developed and the undeveloped losses is approximately: $l^{\alpha-1}\Gamma(s-\alpha)/\Gamma(s-1)$. If the mean development is unity, then $l = s - 1$. Then this ratio is:

$$\begin{aligned} &(s-1)^{\alpha-1}/\{(s-\alpha-1)\cdots(s-2)\} \\ &\approx \left((s-1)\Big/\left(s-\frac{\alpha}{2}-1\right)\right)^{\alpha-1}. \end{aligned}$$

Since for a Pareto for large limits $R(L) \approx \lambda^{\alpha-1}/L^{\alpha-1}$ if one adjusts λ by multiplying by a factor of $(s-1)/(s-(\alpha/2)-1)$, then one will approximately multiply the excess ratios by the desired adjustment factor.

[28] See Equation 13.5.6, *Handbook of Mathematical Functions* [1].

APPENDIX D

CONFLUENT HYPERGEOMETRIC FUNCTIONS[29]

There are a number of related functions referred to as confluent hypergeometric functions. They can be usefully thought of as generalizations of the beta and gamma functions. They can be thought of as two parameter distributions. Let:

$$M(a,b,z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_0^1 e^{zt}t^{a-1}(1-t)^{b-a-1}\,dt,$$

$$U(a,b,z) = \frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}\,dt.$$

Then M can be computed using the following power series in z:

$$M(a,b;z) = 1 + \frac{az}{b} + \frac{a(a+1)z^2}{b(b+1)(2!} + \frac{a(a+1)(a+2)z^3}{b(b+1)(b+2)(3!)} + \cdots.$$

U can be computed as a combination of two values of M:

$$U(a,b;z) = \frac{\pi}{\sin\pi b}\left(\frac{M(a,b,z)}{\Gamma(1+a-b)\Gamma(b)} - \frac{z^{1-b}M(1+a-b,2-b,z)}{\Gamma(a)\Gamma(2-b)}\right).$$

U is related to the incomplete gamma function:

$$U(1-a,1-a;x) = e^x\Gamma(a;x).$$

Among the facts used in Appendix C are:

$$\frac{d}{dz}U(a,b;z) = -aU(a+1,b+1;z),$$

$$U(a-1,b;z) - zU(a,b+1;z) = (a-b)U(a,b;z).$$

For z small and $b > 2$, $U(a,b;z) \approx z^{1-b}\Gamma(b-1)/\Gamma(a)$.

[29] See *Handbook of Mathematical Functions* [1].

THEOREM

$$\int z^{a-3}U(a,b,z)\,dz$$

$$= -\frac{z^{a-2}}{(a-1)(b-a)}\frac{U(a-1,b,z)-U(a-2,b,z)}{\{(a-2)(b+1-a)\}}.$$

Given:

$$\frac{dU(a,b,z)}{dz} = -aU(a+1,b+1,z), \qquad \text{and}$$

$$zU(a,b+1,z) - U(a-1,b,z) = (b-a)U(a,b,z).$$

Proof:

Let

$$\nu = z^{a-2}(U(a-1,b,z) - U(a-2,b,z)/\{(a-2)(b+1-a)\})$$

$$\frac{d\nu}{dz} = (a-2)\nu/z + z^{a-2}\frac{-(a-1)U(a,b+1,z)+U(a-1,b+1,z)}{(b+1-a)}$$

$$= z^{a-3}\{(a-2)U(a-1,b,z) - U(a-2,b,z)/(b+1-a)$$
$$- (a-1)zU(a,b+1,z)$$
$$+ zU(\alpha-1,b+1,z)/(b+1-\alpha)\}$$

$$= z^{a-3}\{(zU(\alpha-1,b+1,z) - U(\alpha-2,b,z))/(b+1-\alpha)$$
$$- (a-1)(zU(a,b+1,z) - U(\alpha-1,b,z)) - U(\alpha-1,b,z)\}$$

$$= z^{a-3}(U(a-1,b,z) - (a-1)(b-a)U(a,b,z) - U(a-1,b,z))$$

$$= -z^{a-3}(a-1)(b-a)U(a,b,z). \qquad \text{Q.E.D.}$$

Errata for Discussion by Howard Mahler of "Retrospective Rating: 1997 Excess Loss Factors"

At the bottom of page 320, the equation for $\hat{R}(100)$ is incorrect.

This example of simple dispersion is an example of a mixture with five pieces.

The excess ratio of the mixture is a weighted average of individual excess ratios, with the weights the product of the means and the probabilities for each piece of the mixture.[1]

If the probability of each piece of a mixture is p_i, $\sum p_i = 1$, the mean of each piece of the mixture is m_i, and R_i is the excess ratio for each piece of the mixture, then $\hat{R}(L) = \sum p_i\, m_i\, R_i(L)$.

If each loss is divided by for example .75, then after development, the excess ratio at L is the same as the original excess ratio at .75 L.[2]
$R_i(L)$ is the excess ratio when the losses have all been divided by r_i.
Thus $R_i(L) = R(r_i\, L)$.

In the example on page 320, each mean is proportional to 1/divisor = $1/r_i$, and each probability is the same at 1/5. Thus the weights are: $(1/5)(1/r_i)$.

The sum of the weights is: $\sum (1/5)(1/r_i) = (1/5)(1/.75 + 1/.833 + 1/1 + 1/1.25 + 1/1.5) = 1$.[3]

Thus $\hat{R}(L) = \sum (1/5)(1/r_i)\, R(r_i\, L) = (1/5) \sum R(r_i\, L)\,/r_i$.

Therefore, the corrected equation at the bottom of page 320 is:

$$\hat{R}(100) = (1/5)\{R(75)/.75 + R(83.3)/.833 + R(100)/1 + R(125)/1.25 + R(150)/1.50\}$$
$$= (1/5)\{.6009/.75 + .5817/.833 + .5582/1 + .5384/1.25 + .5191/1.50\} = .5669.$$

Similarly, the corrected equation at the top of page 321 is:

$$\hat{R}(5000) = (1/5)\{R(3750)/.75 + R(4165)/.833 + R(5000)/1 + R(6250)/1.25 + R(7500)/1.50\}$$
$$= (1/5)\{.0157/.75 + .0070/.833 + 0/1 + 0/1.25 + 0/1.50\} = .0059.$$

---

[1] See page 154 of "Workers Compensation Excess Ratios: An Alternate Method of Estimation" by Mahler.
[2] If each loss is multiplied by 1/.75 = 1.333, this is mathematically the same as uniform inflation of 33.3%. Thus we can get the excess ratio after development, by taking the original excess ratio at the deflated value of L/1.333 = .75 L. Increasing the sizes of loss, increases the excess ratio over a fixed limit.
[3] Mahler chose these loss divisors so that the total expected losses are unaffected.

At page 324, some of the numerical values shown in the computation of $R_3(2000)$ are mixed up, although the final value is correct at 0.384 as shown.
It should have read:

$$R_3(2000) = (1.04167)(0.9999980) - (0.04167)(0.0057148) + (0.1667)(0.0026029) - (0.8333)(0.999995) + (0.1761)(0.999987) - (0.1761)(0.0011302) = 0.384.$$

Also, in Table 1 the excess ratios were computed for Gamma loss divisors with shape parameter 16.67 and inverse scale parameter 15.67. However, the text at page 323 refers to Gamma loss divisors with shape parameter s = 18.67 and inverse scale parameter l = 17.67; this distribution of loss divisors corresponds to a mean loss development of 1 and a variance of loss development of 0.060, matching the simple dispersion example.

Using the intended Gamma parameters of s = 18.67 and l = 17.67 changes the excess ratios in Table 1 slightly, although the pattern remains the same.

The values in the simple dispersion column of Table 1 at page 320 are revised in a similar manner to that for 5000.
The values in Gamma dispersion column of Table 1 at page 320 are revised based on a shape parameter of s = 18.67 and inverse scale parameter of l = 17.67.

Corrected Table 1
Excess Ratios

LIMIT	No Development	Simple Dispersion	Gamma Dispersion
50	.6888	.6949	.6939
100	.5582	.5669	.5673
500	.3012	.3080	.3069
1,000	.1606	.1705	.1709
2,000	.0904	.0931	.0927
3,000	.0402	.0462	.0453
4,000	.0100	.0194	.0182
5,000	.0000	.0059	.0062
6,000	.0000	.0007	.0020
7,000	.0000	.0000	.0006
8,000	.0000	.0000	.0002
9,000	.0000	.0000	.0001
10,000	.0000	.0000	.0000

As can be seen in corrected Table 1, the simple dispersion effect raises the excess ratios, especially at the higher limits.[4]

At page 326, the formula near the bottom of page should have λ in place of X:

$R(L) = (\lambda\, l/L)^{s-1}\, U(s-1, s+1-\alpha, \lambda\, l/L)$.

[4] It can be demonstrated that when dispersion has no overall effect, loss dispersion either increases an excess ratio or keeps it the same. In most practical applications, the excess ratio will be increased by loss dispersion.

ADDRESS TO NEW MEMBERS—MAY 18, 1998

P. ADGER WILLIAMS

Madame President, thank you for that kind introduction.

Good morning and congratulations to all the new Fellows and new members. I'd also like to thank Mavis for inviting me to give this address to new Fellows and Associates. When Mavis called to ask me to speak, she said that one of the reasons for asking me was that I was President of the CAS when she became a Fellow. I don't know whether that says more about how old I am or how young she is. Probably both! But it does say one thing loud and clear to you new Fellows. When one of you becomes President of the CAS, Mavis will be expecting a call.

Speaking to new members makes me think back to what appealed to me about becoming an actuary.

First of all, you could make a living while studying to become a professional. That was quite an incentive for me. And you didn't have to go to the "right" school or know somebody to get ahead. All you had to do was pass the exams.

Now that you have done that, there is a wide range of opportunities waiting for you. Your actuarial training gives you the most versatile basic training that can be found in insurance.

We can point with pride to CAS members who have become insurance commissioners, bureau heads, presidents, CEOs, and managers in a wide range of disciplines: data processing, finance, underwriting, marketing, and many others. Their actuarial training was a stepping stone to management. But most actuaries don't want to be managers! They want to be practicing actuaries. And you can have a fulfilling and, I might add, lucrative career as an actuarial professional doing actuarial work.

Before going any farther, I'd like to say a few words just to the new Associates. Get your Fellowship! Become an F.C.A.S.

At this point in your career, it's very easy to look around and say to yourself, "Why do I need to pass more exams? My present work doesn't require more knowledge, my work is interesting, and I'm really too busy to study. Besides, I have a great boss who would just as soon see me work more and study less. Most important of all, I've been getting regular promotions and raises in a great organization where I work with a great bunch of people."

Those are pretty good reasons to stop studying, aren't they? But let me tell you, times change, companies change, bosses change, friends change, or, more likely, leave, and, sometimes, worst of all, your job doesn't change. You wake up some morning and realize you're at a dead end, and you want a ticket out. That ticket is the professionalism that comes with being a Fellow … I see some of my friends smiling who have been in that situation. Having their Fellowship allowed them career choices that would have been unavailable otherwise.

So put yourself in a position to have that career choice. Get your Fellowship. *You'll never forgive yourself if you don't.*

Becoming a Fellow of the CAS bestows a unique status upon you as a professional in the actuarial community. It gives you the opportunity to choose any kind of career you wish. As I said before, you can become an executive, a manager, a consultant, or a pure research actuary—your professionalism gives you countless opportunities.

But professionalism also brings with it responsibility. Now I'm not talking about what you often hear some long-time actuaries saying, "The profession has been good to me so I want to give something back to the profession." That is a worthwhile sentiment, and I think it would be wonderful if all of you would put in some time working for the CAS or elsewhere in the profession. But that's not what I mean.

I'm talking about the responsibility that comes with the mantle of professionalism that has now been draped over your shoulders. You studied actuarial science to pass the exams. Much of that science is contained in the CAS *Proceedings* and other actuarial literature. It would seem that all you have to do is apply what you've learned. After all, the CAS has been around for nearly 85 years; how much could be left to discover or develop?

Early in my career, back in 1960, I got sidetracked into a data processing project to develop the first computer-communications system in the insurance industry. As we approached the time to go on the air in 1964 with our gigantic computers ("gigantic" in those days meant a 64K memory), a young man who had just joined the project said to me, "Gee, I wish I could have gotten into data processing early while there was still something new to be contributed."

Don't make the same mistake that young man made. We owe much to the work that has been done in the past, but we have just scratched the surface of actuarial science—especially casualty actuarial science. You should view yourselves as pioneers, entering the profession, not at the end or the middle, but at the beginning.

So the work that has gone before, the body of actuarial knowledge that has been developed, is both a gift and a legacy. Now it becomes your responsibility. Where that body of knowledge is worthy and deserving of your support, apply it and defend it. Where it's lacking, it's up to you to improve or replace it.

Somehow it just doesn't seem fair, does it? You've had your new actuarial designation for only 10 or 15 minutes, and you've already been given the responsibility for all the work that has gone before.

That's not all! There are several other responsibilities that come with your actuarial designation:

—The responsibility for the advancement of actuarial science. It's up to you to see to it that our science has substance. There are those who contend that what we do is an art, not a science. And we do have to be careful not to tie ourselves in knots with rigid rules that stifle actuarial innovation. At the same time, we can't let the desire for actuarial art lead us to actuarial anarchy.

—Next, there's the responsibility for actuarial standards that must march hand-in-hand with the advancement of the science. Here we have no choice. If we don't set our own standards, someone will set them for us. But we must set them in a way that gives us actuarial freedom within a framework of sensible boundaries.

—We also have the responsibility for communicating our knowledge in a way that it can be understood. Some of you may have read Stephen Hawking's book, *A Brief History of Time*, in which he attempts to explain, in simple terms, some of the most complex theories relating to the universe and the quest for a unified theory. But he concludes that even if the theory is discovered, it will do no good if it is understood by only a few scientists; it must be communicated so that it is, in his words, "understandable in broad principle by everyone." That is our task and our responsibility, to communicate in such a way that the least knowledgeable of our audience understands what we mean.

—Finally, there is the responsibility for actuarial integrity, which to me is the heart of professionalism. There are many directions your careers will take, positions in regulation, industry, consumerism, and many others. Throughout your career, you must remain keenly aware of when you're speaking as an actuary and when you're not. And you can never completely shed the responsibilities of professionalism when you're speaking in those areas where your training gives you that unique capability that identifies you as an actuary.

With all of these responsibilities, you're probably beginning to wonder what you get out of being an actuary. In the years ahead, as you look back on your career, you'll find that being an actuary really did put you in a unique position; a position to have a positive influence on your place of employment, a position to have a positive influence on your industry, a position to make a difference in your profession, and if you were willing to participate, a chance to be part of the rule-setting process rather than the rule-following process.

Look around you at those who are in your group of new Fellows and Associates. As the years go by they will form what I like to call an accumulation of actuarial fellowship which will become the continuity in your life. Ultimately I think you will find, as I did, that being an actuary is not only a profession, it is a process of life enrichment.

Thank you, and, once again, my congratulations to all of you!

MINUTES OF THE 1998 SPRING MEETING

May 17–20, 1998

MARRIOTT'S MARCO ISLAND RESORT AND GOLF CLUB

MARCO ISLAND, FLORIDA

Sunday, May 17, 1998

The Board of Directors held their regular quarterly meeting from noon to 5:00 p.m.

Registration was held from 4:00 p.m. to 6:00 p.m.

New Associates and their guests were honored with a special presentation from 5:30 p.m. to 6:30 p.m. Members of the 1998 Executive Council discussed their roles in the Society with the new members.

During the special reception, CAS President Mavis A. Walters introduced the CAS Board of Directors and the CAS Executive Council attending the Spring Meeting. Walters also introduced American Academy of Actuaries President Allan M. Kaufman, and Chairman of the Casualty Practice Council Michael L. Toothman.

A reception for all meeting attendees followed the new Associates reception and was held from 6:30 p.m. to 7:30 p.m.

Monday, May 18, 1998

Registration continued from 7:00 a.m. to 8:00 a.m.

The 1998 Business Session, which was held from 8:00 a.m. to 9:00 a.m., ushered in the first full day of activities for the 1998 Spring Meeting. Ms. Walters introduced the CAS Executive Council, the Board of Directors, and CAS past presidents who were in attendance, including Robert A. Anker (1996), Phillip N. Ben-Zvi (1985), Ronald L. Bornhuetter (1975), David P. Flynn

(1992), Michael Fusco (1989), Charles C. Hewitt Jr. (1972), Allan M. Kaufman (1994), W. James MacGinnitie (1979), LeRoy J. Simon (1971), Michael L. Toothman (1991), Michael A. Walters (1986), and P. Adger Williams (1977).

Ms. Walters also recognized special guests in the audience: Stephen P. D'Arcy, President-Elect of the American Risk and Insurance Association; and Allan M. Kaufman, President of the American Academy of Actuaries.

Curtis Gary Dean, Kevin B. Thompson, and Robert S. Miccolis announced the 118 new Associates and Steven G. Lehmann announced 18 new Fellows. The names of these individuals follow.

NEW FELLOWS

Michael K. Curry	Man-Gyu Hur	David Molyneux
Elizabeth B. DePaolo	Steven W. Larson	Vinay Nadkarni
Steven T. Harr	Andre L'Esperance	William Peter
Daniel F. Henke	Christina Link	John S. Peters
Thomas G. Hess	Michael K. McCutchan	Michael D. Price
Marie-Josée Huard	Thomas S. McIntyre	Michael J. Steward II

NEW ASSOCIATES

Mustafa Bin Ahmad	Hayden Burrus	François R. Dumontet
Nancy S. Allen	Thomas J. Chisholm	Mark Kelly Edmunds
Wendy L. Artecona	Wanchin W. Chou	Brian A. Evans
Carl X. Ashenbrenner	Christopher W. Cooney	Stephen C. Fiete
David S. Atkinson	Jonathan S. Curlee	Sarah J. Fore
Craig V. Avitabile	Loren R. Danielson	Mauricio Freyre
Phillip W. Banet	Timothy A. Davis	Timothy J. Friers
Emmanuil Bardis	Nancy K. DeGelleke	Bernard H. Gilden
Michael W. Barlow	Brian H. Deephouse	Sanjay Godhwani
Gina S. Binder	Michael B. Delvaux	Daniel C. Greer
Kevin M. Bingham	Karen D. Derstine	Daniel E. Greer
James D. Buntine	Sara P. Drexler	David J. Gronski
Alan Burns	Tammi B. Dulberger	Eric C. Hassel

Christopher R. Heim
Chad A. Henemyer
Melissa K. Higgins
Tina T. Huynh
Susan E. Innes
Claudine H. Kazanecki
Kelly Martin Kingston
James D. Kunce
Carl Lambert
Hugues Laquerre
Dennis H. Lawton
Manuel Alberto T. Leal
David Leblanc-Simard
Bradley R. LeBlond
Glen A. Leibowitz
Craig A. Levitz
John N. Levy
Shiu-Shiung Lin
Victoria S. Lusk
Allen S. Lynch Jr.
Stephen J. McAnena
Jennifer A. McCurry
Mark Z. McGill III
David P. Moore
Jennifer A. Moseley
Ethan Mowry
Jarow G. Myers
Seth W. Myers
Kari A. Nicholson
John E. Noble
Jason M. Nonis
Corine Nutting
Jean-François Ouellet
Kathryn A. Owsiany
Pierre Parenteau
M. Charles Parsons
Jeremy P. Pecora
Richard M. Pilotte
Glen-Roberts Pitruzzello
Christopher D. Randall
Hany Rifai
Brad E. Rigotty
Karen L. Rivara
Rebecca L. Roever
Nathan W. Root
Kimberly R. Rosen
Richard A. Rosengarten
Seth A. Ruff
Brian C. Ryder
James C. Sandor
Gary F. Scherer
Nathan A. Schwartz
Steven G. Searle
Meyer Shields
Aviva Shneider
Alastair Shore
Matthew R. Sondag
Benoit St-Aubin
Joy M. Suh
Karrie L. Swanson
Rachel R. Tallarini
Varsha A. Tantri
Glenda O. Tennis
Laura L. Thorne
Beth S. Tropp
Kris D. Troyer
Turgay F. Turnacioglu
Leslie A. Vernon
Kyle J. Vrieze
Matthew J. Wasta
Lynne K. Wehmueller
Christopher B. Wei
Scott Werfel
Dean A. Westpfahl
Thomas J. White
Vanessa C. Whitlam-Jones
Kendall P. Williams
Yoke Wai Wong
Linda Yang

Ms. Walters then introduced P. Adger Williams, a past president of the Society, who presented the Address to New Members.

Patrick J. Grannan, CAS Vice President–Programs and Communications, spoke to the meeting participants about the highlights of this meeting and what was planned in the program.

Susan T. Szkoda, CAS Vice President–Continuing Education announced that ten *Proceedings* papers and two discussions of *Proceedings* papers would be presented at this meeting. In all, ten papers were accepted for publication in the 1998 *Proceedings of the Casualty Actuarial Society*. One paper by Cross and Doucette was presented at this meeting but was published in the 1997 *Proceedings*. Another discussion of a *Proceedings* paper by Feldblum was accepted in spring 1998 but not presented.

John J. Kollar, chairperson of the Michelbacher Award Committee, gave a brief description of this year's Call Paper Program on Dynamic Analysis of Pricing Decisions. He announced that all of the call papers would be presented at this meeting. In addition, the papers were published in the 1998 CAS Discussion Paper Program and could be found on the CAS Web Site. Mr. Kollar presented the Michelbacher Prize to Richard L. Stein for his paper, "The Actuary As Product Manager in A Dynamic Product Analysis Environment." This award commemorates the work of Gustav F. Michelbacher and honors the authors of the best paper submitted in response to a call for discussion papers. The papers are judged by a specifically appointed committee on the basis of originality, research, readability, and completeness.

Ms. Walters then began the presentation of other awards. She explained that the CAS Harold W. Schloss Memorial Scholarship Fund benefits deserving and academically outstanding students in the actuarial program of the Department of Statistics and Actuarial Science at the University of Iowa. The student recipient is selected by the Trustees of the CAS Trust, based on the recommendation of the department chair at the University of Iowa. Ms. Walters announced that Changki Kim is the recipient of the 1997 CAS Harold W. Schloss Memorial Scholarship Fund. She will be presented with a $500 scholarship.

Ms. Walters then concluded the business session of the Spring Meeting by calling for a review of *Proceedings* papers.

Ms. Walters next introduced the featured speaker, Dr. Barry Asmus, who is senior economist with the National Center for Policy Analysis.

The first General Session was held from 10:30 a.m. to noon.

"The Future of the Actuary"

Moderator: W. James MacGinnitie
CFO & Senior Vice President
CNA

Panelists: Joan Lamm-Tennant, Ph.D.
Vice President
General Re-New England Asset Management, Inc.

Paul W. McCrossan
Partner
Eckler Partners, Ltd.

Peter R. Porrino
President and CEO
Consolidated International Group, Inc.

After a luncheon, the afternoon was devoted to presentations of concurrent sessions and discussion papers. The call papers presented from 1:15 p.m. to 2:45 p.m. were:

1. "Direct Marketing of Insurance Integration of Marketing, Pricing, and Underwriting"

   Author: Bruce D. Moore
   Partner
   Ernst and Young LLP

2. "The Actuary as Product Manager in A Dynamic Product Analysis Environment"

   Author: Richard L. Stein
   Senior Consultant
   CNA

The concurrent sessions presented from 1:15 p.m.–2:45 p.m. were:

1. Florida Commission on Hurricane Modeling

Moderator: Alice H. Gannon
Vice President
United Services Automobile Association

Panelists: Jack Nicholson, Ph.D.
Chief Operating Officer
Florida Hurricane Catastrophe Fund

Mark E. Johnson
Professor of Statistics and Director
Institute of Statistics
University of Central Florida

2. Managed Care and Its Impact on Workers Compensation

Moderator: Richard I. Fein
Principal
Coopers & Lybrand, L.L.P.

Panelists: David Appel, Ph.D.
Chief Economic Consultant
Milliman & Robertson, Inc.

Gloria Gebhard
Medical Policy Analyst
Minnesota Department of Labor & Industry

Layne M. Onufer
Principal
Ernst & Young LLP

3. International Relations and the CAS

Moderator: David G. Hartman
Senior Vice President and Managing Director
Chubb Group of Insurance Companies

Panelists: Members of the International Relations Committee of the CAS

4. An Extremely Important Application of Extreme Value Theory to Reinsurance Pricing

   Panelists: Gary S. Patrik
   Chief Actuary
   Swiss Reinsurance America Corporation

   Farrokh Guiahi
   Assistant Actuary
   Swiss Reinsurance America Corporation

Proceedings papers presented during this time were:

1. "Investment-Equivalent Reinsurance Pricing"

   Author: Rodney E. Kreps
   Executive Vice President and Chief Actuary
   Sedgwick Re Insurance Strategy, Inc.

2. Discussion of "Reinsurer Risk Loads for Marginal Surplus Requirements"

   (by Rodney E. Kreps, *PCAS* LXXVII, 1990, p. 196)

   Discussion by: Paul J. Kneuer
   Vice President and Actuary
   Holborn Corporation

3. "A Buyer's Guide for Options on a Catastrophe Index"

   Author: Glenn G. Meyers
   Assistant Vice President
   Insurance Services Office, Inc.

4. "A Comparison of Property/Casualty Insurance Financial Pricing Models"

   Authors: Stephen P. D'Arcy
   Professor
   Department of Finance–University of Illinois

Richard W. Gorvett
Assistant Professor
College of Insurance

After a refreshment break, presentations of call papers, concurrent sessions, and *Proceedings* papers continued from 3:15 p.m. to 4:45 p.m. Certain call papers and concurrent sessions presented earlier were repeated. Additional call papers presented during this time were:

1. "The Impact of Price Changes on Costs"

   Author: Russell L. Sutter
   Consulting Actuary
   Tillinghast-Towers Perrin

2. "Estimating the Actuarial Value of the Connecticut Second Injury Fund Loss Portfolio"

   Author: Abbe Sohne Bensimon
   Vice President
   General Reinsurance Corporation

3. "Actuarial Considerations in the Development of Agent Contingent Compensation Programs"

   Author: William J. VonSeggern
   Assistant Vice President and Actuary
   Milliman & Robertson, Inc.

   Lori E. Stoeberl
   Associate Actuary
   Milliman & Robertson, Inc.

An additional concurrent session presented from 2:45 p.m. to 3:15 p.m. was:

1. Strategic Issues and the AAA

   Moderator: Allan M. Kaufman
   Principal
   Milliman & Robertson, Inc.

Panelists: Stephen P. Lowe
Consulting Actuary
Tillinghast-Towers Perrin

John M. Purple
Consulting Actuary
Arthur Andersen LLP

Richard S. Robertson
Executive Vice President
Lincoln National Corporation

Proceedings papers presented during this time were:

1. "Measurement of Asbestos Bodily Injury Liabilities"

   Authors: Susan L. Cross
   Executive Vice President and Chief Actuary
   Tillinghast-Towers Perrin

   John P. Doucette
   Vice President
   European International Reinsurance Company Ltd.

2. "Workers Compensation Excess Ratios"

   Author: Howard C. Mahler
   Vice President and Actuary
   Workers Compensation Rating & Inspection Bureau of Massachusetts

3. Discussion of "Retrospective Rating: 1997 Excess Loss Factors"

   (by William R. Gillam and Jose Couret, *PCAS* LXXXIV, 1997, p. 450)

   Discussion by: Howard C. Mahler
   Vice President and Actuary
   Workers Compensation Rating & Inspection Bureau of Massachusetts

A reception for new Fellows and their guests was held from 5:30 p.m. to 6:30 p.m., and the general reception for all members and their guests was held from 6:30 p.m. to 7:30 p.m.

Tuesday, May 19, 1998

Registration continued from 7:00 a.m. to 8:00 a.m.

Two General Sessions were held from 8:00 a.m. to 9:30 a.m. The General Sessions presented were:

"Junk Science and Insurance"

Moderator:	Amy S. Bouska Consulting Actuary Tillinghast-Towers Perrin
Panelists:	John J. Delany III, Esquire Delany & O'Brien
	Michael Green Professor of Law University of Iowa College of Law
	Sorell L. Schwartz, Ph.D. Professor Emeritus Georgetown University

"Merging Business and Technological Strategies"

Moderator:	Michael L. Toothman Managing Partner Arthur Andersen LLP
Panelists:	Edward E. Bambauer Director, Financial Market-New England Arthur Andersen LLP
	Joel S. Weiner Senior Manager Coopers & Lybrand, L.L.P.

Clark M. Sykes
Vice President
Information Technology
Merchants Group, Inc.

Certain discussion papers and concurrent sessions that had been presented earlier during the meeting were repeated this morning from 10:00 a.m. to 11:30 a.m. Concurrent sessions presented during this time were:

1. Securitization 101

   Panelists: Richard W. Gorvett
   Assistant Professor
   College of Insurance

   Bryon G. Ehrhart
   President
   Aon Re Services

2. Latin-American Market Issues

   Moderator: Jay B. Morrow
   Vice President and Actuary
   American International Underwriters

   Panelists: Richard G. Cadugan
   Senior Manager
   Deloitte & Touche Consulting Group

   Susan J. Patschak
   Consulting Actuary
   Tillinghast-Towers Perrin

3. Codification of Statutory Accounting Principles

   Moderator: Andrew E. Kudera
   Senior Vice President & CFO
   CNA Risk Management

   Panelists: Richard J. Roth Jr.
   Chief Property/Casualty Actuary
   California Department of Insurance

Mark A. Parkin
Partner
Deloitte & Touche LLP

Jan A. Lommele
Principal
Deloitte & Touche LLP

4. Sport Utility Vehicles and Auto Insurance Costs

Moderator: Jerry W. Rapp
Consulting Actuary
Miller, Rapp, Herbers, & Terry, Inc.

Panelists: Kim Hazelbaker
Senior Vice President
Highway Loss Data Institute

Michael C. Dubin
Consulting Actuary
Milliman & Robertson, Inc.

Proceedings papers presented during this time were:

1. "The Impact of Investment Strategy on the Market Value and Pricing Decisions of a Property/Casualty Insurer"

Author: Trent R. Vaughn
Vice President, Actuarial Pricing
Empire Fire & Marine Insurance Company

2. "A Markov Chain Model of Shifting Risk Parameters"

Author: Howard C. Mahler
Vice President and Actuary
Workers Compensation Rating & Inspection Bureau of Massachusetts

3. "Smoothing Weather Losses: A Two-Sided Percentile Model"

Author: Curtis Gary Dean
Assistant Vice President and Actuary
American States Insurance Companies

David N. Hafling
Senior Vice President and Actuary
American States Insurance Companies

William F. Wilson
Associate Actuary
American States Insurance Companies

4. "An Application of Game Theory: Property Catastrophe Risk Load"

   Author: Donald F. Mango
   Assistant Vice President
   Zurich Centre ReSource, Ltd.

Various CAS committees met from 1:00 p.m. to 5:00 p.m. In addition, a limited attendance workshop was held from 1:00 p.m. to 4:00 p.m.:

"Getting Your Ideas Across"

Leader: Ira M. Blatt
President
Society of Insurance Trainers and Educators
Director of Training and Development
Insurance Services Office, Inc.

All members and guests enjoyed a buffet dinner at the resort from 5:30 p.m. to 9:00 p.m.

Wednesday, May 20, 1998

Certain concurrent sessions that had been presented earlier during the meeting were repeated this morning from 8:00 a.m. to 9:30 a.m. An additional concurrent session presented was:

1. Questions and Answers With the CAS Board of Directors

   Moderator: Steven G. Lehmann
   Principal and Consulting Actuary
   Miller, Rapp, Herbers & Terry, Inc.

Panelists:	Jerome A. Degerness Actuarial Officer St. Paul Fire & Marine Insurance Company
	Alice H. Gannon Vice President United Services Automobile Association
	Richard J. Roth Jr. Chief Property/Casualty Actuary California Department of Insurance

After a refreshment break, the final General Session was held from 10:00 a.m. to 9:30 a.m.:

"The Effect of Science and Technology on Risk Classification and Underwriting"

Moderator:	Michael A. Walters Principal Tillinghast-Towers Perrin
Panelists:	David J. Christianson Vice President Insurance Services Lutheran Brotherhood
	Chris Garson I/S Executive of Agent Marketing Progressive
	Ward Jungers Group Vice President & Chief Underwriting Officer CNA
	Philip O. Presley Chief Actuary Texas Department of Insurance

Mavis A. Walters officially adjourned the 1998 CAS Spring Meeting at 11:45 a.m. after closing remarks and an announcement of future CAS meetings.

Attendees of the 1998 CAS Spring Meeting

The 1998 CAS Spring Meeting was attended by 308 Fellows, 216 Associates, and 151 Guests. The names of the Fellows and Associates in attendance follow:

FELLOWS

Shawna Ackerman
Mark A. Addiego
Martin. Adler
Terry J. Alfuth
Timothy P. Aman
Robert A. Anker
Steven D. Armstrong
Lawrence J. Artes
Nolan E. Asch
Richard V. Atkinson
Anthony J. Balchunas
Timothy J. Banick
W. Brian Barnes
Allan R. Becker
Linda L. Bell
Douglas S. Benedict
Robert S. Bennett
Abbe S. Bensimon
Phillip N. Ben-Zvi
Michele P. Bernal
Lisa M. Besman
Wayne E. Blackburn
Gavin C. Blair
Jean-François Blais
Robert G. Blanco
Cara M. Blank
Daniel D. Blau
LeRoy A. Boison
Steven W. Book
Joseph A. Boor
Ronald L. Bornhuetter
François Boulanger
Pierre Bourassa
Amy S. Bouska
Christopher K. Bozman
John G. Bradshaw
Paul Braithwaite
Dale L. Brooks
Ward M. Brooks
Brian Z. Brown
Lisa J. Brubaker
Kirsten R. Brumley
James E. Buck
Mark E. Burgess
John E. Captain
Michael J. Caulfield
Francis D. Cerasoli
Scott K. Charbonneau
David R. Chernick
Stephan L. Christiansen
Francis X. Corr
Susan L. Cross
Diana M. Currie
Ross A. Currie
Michael K. Curry
Stephen P. D'Arcy
Ronald A. Dahlquist
Lawrence S. Davis
Curtis Gary Dean
Jerome A. Degerness
Jeffrey F. Deigl
Elizabeth B. DePaolo
Janet B. Dezube
Anthony M. DiDonato
Michael C. Dolan
Michael C. Dubin
Richard D. Easton
Grover M. Edie
Dale R. Edlefson
David Engles
Paul E. Ericksen
Philip A. Evensen
John S. Ewert

Doreen S. Faga
Michael A. Falcone
Dennis D. Fasking
Richard I. Fein
John R. Ferrara
Carole M. Ferrero
Mark E. Fiebrink
Daniel J. Flick
David P. Flynn
Richard L. Fox
Michael Fusco
Alice H. Gannon
Robert W. Gardner
David B. Gelinne
John F. Gibson
Bruce R. Gifford
Julie T. Gilbert
Judy A. Gillam
William R. Gillam
Bryan C. Gillespie
Gregory S. Girard
Mary K. Gise
Nicholas P. Giuntini
Olivia W. Giuntini
Donna L. Glenn
Daniel C. Goddard
Charles T. Goldie
Richard W. Gorvett
Odile Goyer
Gregory S. Grace
Patrick J. Grannan
Gregory T. Graves
Mari L. Gray
Eric L. Greenhill
Anne G. Greenwalt
Russell H. Greig
Linda M. Groh
Farrokh Guiahi
Terry D. Gusler
David N. Hafling
James A. Hall
Robert C. Hallstrom
Jeffrey L. Hanson
Steven T. Harr
Christopher L. Harris
David G. Hartman
Matthew T. Hayden
E. LeRoy Heer
Daniel F. Henke
Dennis R. Henry
Teresa J. Herderick
Richard J. Hertling
Thomas G. Hess
Charles C. Hewitt Jr.
James S. Higgins
Kathleen A. Hinds
Wayne Hommes
Beth M. Hostager
George A. Hroziencik
Marie-Josée Huard
M. Stanley Hughey
Man-Gyu Hur
John M. Hurley
Richard M. Jaeger
Ronald W. Jean
Eric J. Johnson
Kurt J. Johnson
Allan M. Kaufman
Tony J. Kellner
Allan A. Kerin
Kevin A. Kesby
Joe C. Kim
Michael F. Klein
Charles D. Kline
Paul J. Kneuer
Leon W. Koch
John J. Kollar
Mikhael I. Koski
Thomas J. Kozik
Gary R. Kratzer
Gustave A. Krause
Rodney E. Kreps
David J. Kretsch
John R. Kryczka
David R. Kunze
Blair W. Laddusaw
Dean K. Lamb
John A. Lamb
Michael A. LaMonica
Nicholas J. Lannutti
Michael D. Larson
Steven W. Larson
Paul W. Lavrey
Merlin R. Lehman
Steven G. Lehmann
Andre L'Esperance
Jean-Marc Leveille
George M. Levine
John J. Lewandowski
Christina Link
Richard A. Lino
Richard W. Lo
Andre Loisel
Jan A. Lommele
Stephen P. Lowe
W. James MacGinnitie
Howard C. Mahler
Maria Mahon

Donald F. Mango
Donald E. Manis
Leslie R. Marlo
Paul C. Martin
Robert D. McCarthy
James B. McCreesh
Gary P. McDonald
Richard T. McDonald
Liam M. McFarlane
Dennis T. McNeese
Stephen V. Merkey
Robert J. Meyer
Glenn G. Meyers
Robert S. Miccolis
Jon W. Michelson
David L. Miller
Robert L. Miller
David Molyneux
Richard B. Moncher
W. Dale Montgomery
Brian C. Moore
Bruce D. Moore
Jay B. Morrow
Nancy D. Mueller
Todd B. Munson
Giovanni A. Muzzarelli
Vinay Nadkarni
Kenneth J. Nemlick
Richard T. Newell
Aaron W. Newhoff
John Nissenbaum
Kathleen C. Nomicos
Peter M. Nonken
Keith R. Nystrom
Marc F. Oberholtzer
Robert G. Oien
Kathy A. Olcese
Layne M. Onufer
Melinda H. Oosten
Timothy A. Paddock
Richard D. Pagnozzi
Rudy A. Palenik
Donald D. Palmer
Nicholas H. Pastor
Chandrakant C. Patel
Gary S. Patrik
Charles C. Pearl
Julia L. Perrine
William Peter
John S. Peters
Michael Petrocik
Stephen W. Philbrick
Richard C. Plunkett
Brian D. Poole
Philip O. Presley
Arlie J. Proctor
Mark R. Proska
John M. Purple
Jerry W. Rapp
Ronald C. Retterath
John J. Reynolds
Donald A. Riggins
Brad M. Ritter
Tracey S. Ritter
Douglas S. Rivenburgh
A. Scott Romito
Allen D. Rosenbach
Deborah M. Rosenberg
Richard J. Roth
Bradley H. Rowe
James B. Rowland
David A. Russell
Sean W. Russell
Stuart G. Sadwin
Neal J. Schmidt
David C. Scholl
Joseph R. Schumi
Peter R. Schwanke
Jeffery J. Scott
Kim A. Scott
Peter Senak
Robert D. Share
Jeffrey P. Shirazi
Jerome J. Siewert
Melvin S. Silver
LeRoy J. Simon
David Skurnick
Michael B. Smith
Richard H. Snader
Linda D. Snook
Lori A. Snyder
David B. Sommer
Bruce R. Spidell
David Spiegler
Douglas W. Stang
Michael J. Steward
Edward C. Stone
Collin J. Suttie
Kathleen W. Terrill
Kevin B. Thompson
Mark L. Thompson
Michael L. Toothman
Janet A. Trafecanty
Theresa A. Turnacioglu
James F. Tygh
John V. Van de Water
Trent R. Vaughn

Gary G. Venter
Dale G. Vincent
William J. VonSeggern
Robert H. Wainscott
Lisa M. Walsh
Mavis A. Walters
Michael A. Walters
Kelly A. Wargo
L. Nicholas Weltmann
Geoffrey T. Werner
Mark Whitman
P. Adger Williams
John J. Winkleman
Michael L. Wiseman
Tad E. Womack
Paul E. Wulterkens
Barbara L. Yewell
Richard P. Yocius
Edward J. Yorty
Ralph T. Zimmer

ASSOCIATES

Nancy S. Allen
James A. Andler
Wendy L. Artecona
Carl X. Ashenbrenner
David S. Atkinson
Craig V. Avitabile
Phillip W. Banet
Emmanuil T. Bardis
Michael W. Barlow
Brian K. Bell
Eric D. Besman
Gina S. Binder
Kevin M. Bingham
Ann M. Bok
Sherri L. Border
Erik R. Bouvin
Steven A. Briggs
Alan Burns
Linda J. Burrill
Hayden H. Burrus
Kenrick A. Campbell
Tania J. Cassell
Jill C. Cecchini
Paul A. Chabarek
Thomas J. Chisholm
Philip S. Chou
Wanchin W. Chou
J. Paul Cochran
Vincent P. Connor
Christopher W. Cooney
Brian K. Cox
Jonathan S. Curlee
Richard J. Currie
Thomas V. Daley
Michael K. Daly
Loren R. Danielson
Timothy A. Davis
Nancy K. DeGelleke
Brian H. Deephouse
Michael B. Delvaux
Karen D. Derstine
Gordon F. Diss
Jeffrey E. Doffing
William A. Dowell
Sara P. Drexler
Peter F. Drogan
Stephen C. Dugan
Tammi B. Dulberger
François R. Dumontet
Rachel Dutil
Mark K. Edmunds
Robert P. Eramo
David L. Esposito
Brian A. Evans
Benedick Fidlow
Loy W. Fitz
Terrence A. Flanagan
Ross C. Fonticella
Jeffrey M. Forden
Mauricio Freyre
Timothy J. Friers
Mary B. Gaillard
Lynn A. Gehant
Bernard H. Gilden
Sanjay Godhwani
Terry L. Goldberg
Laurence B. Goldstein
John W. Gradwell
Gary Granoff
Daniel C. Greer
Daniel E. Greer
Roger E. Griffith
David J. Gronski
Jacqueline L. Gronski
William A. Guffey
Nasser Hadidi
Alexander A. Hammett
Eric C. Hassel

Ia F. Hauck
Susan E. Hayes
Thomas F. Head
Christopher R. Heim
Chad A. Henemyer
David E. Heppen
Joseph A. Herbers
Melissa K. Higgins
Thomas E. Hinds
Jane W. Hughes
Tina T. Huynh
Susan E. Innes
Ali Ishaq
James W. Jonske
Robert B. Katzman
Claudine H. Kazanecki
Martin T. King
Karen L. Krainz
Alexander Krutov
Sarah Krutov
James D. Kunce
Dennis H. Lawton
Manuel Alberto T. Leal
Kevin A. Lee
Stephen E. Lehecka
Glen A. Leibowitz
Daniel E. Lents
Charles Letourneau
Craig A. Levitz
John N. Levy
Barry I. Llewellyn
Lee C. Lloyd
Ronald P. Lowe
Allen S. Lynch
David J. Macesic
Sudershan Malik
Gabriel O. Maravankin
Joseph Marracello
Jason N. Masch
Bonnie C. Maxie
Stephen J. McAnena
Jennifer A. McCurry
David P. Moore
Stephen T. Morgan
Jennifer A. Moseley
Ethan C. Mowry
Robert A. Mueller
Seth W. Myers
Charles P. Neeson
Henry E. Newman
Tieyan T. Ni
John E. Noble
Jason M. Nonis
Corine Nutting
Christopher E. Olson
Rebecca R. Orsi
Jean-François Ouellet
Kathryn A. Owsiany
Teresa K. Paffenback
Pierre Parenteau
M. C. Parsons
Jeremy P. Pecora
Anthony G. Phillips
Amy A. Pitruzzello
Glen-Roberts Pitruzzello
Richard A. Plano
Gregory J. Poirier
Richard B. Puchalski
Patricia A. Pyle
Christopher D. Randall
Peter S. Rauner
James E. Rech
Hany Rifai
Brad E. Rigotty
Rebecca L. Roever
Nathan W. Root
Kim R. Rosen
Richard A. Rosengarten
Scott J. Roth
Seth A. Ruff
Stephen P. Russell
John P. Ryan
Shama S. Sabade
James C. Sandor
Michael Sansevero
Gary F. Scherer
Susan C. Schoenberger
Nathan A. Schwartz
Michael L. Scruggs
Steven G. Searle
Michael Shane
Aviva Shneider
Alastair C. Shore
Janet K. Silverman
Byron W. Smith
David C. Snow
Matthew R. Sondag
Benoit St-Aubin
Lori E. Stoeberl
Lisa M. Sukow
Brian K. Sullivan
C. S. Swalley
Karrie L. Swanson
Rachel R. Tallarini
Varsha A. Tantri
Glenda O. Tennis

Joseph P. Theisen
Robert W. Thompson
Dom M. Tobey
Jennifer M. Tornquist
Beth S. Tropp
Kris D. Troyer
Turgay F. Turnacioglu
Eric Vaith
Jeffrey A. Van Kley
Leslie A. Vernon
Jerome F. Vogel
David M. Vogt
Kyle J. Vrieze
Linda F. Ward
Matthew J. Wasta
Denise R. Webb
Lynne K. Wehmueller
Joel S. Weiner
Mark S. Wenger
Russell B. Wenitsky
Scott Werfel
Dean A. Westpfahl
Michael W. Whatley
Thomas J. White
Vanessa C. Whitlam-Jones
Jerelyn S. Williams
Kendall P. Williams
William F. Wilson
Bonnie S. Wittman
Robert F. Wolf
Jeffrey F. Woodcock
Linda Yang

Volume LXXXV, Part 2 **No. 163**

PROCEEDINGS
November 8, 9, 10, 11, 1998

PERSONAL AUTOMOBILE: COST DRIVERS, PRICING, AND PUBLIC POLICY

JOHN B. CONNERS AND SHOLOM FELDBLUM

Abstract

Traditional actuarial pricing procedures have focused on pre-accident driver attributes, vehicle characteristics, and garaging location in an effort to explain personal automobile loss cost "drivers." Although these traditional factors are important for statewide ratemaking in a static environment, they account for only part of the influences on auto insurance loss costs.

This paper draws on the industry research of the past fifteen years to present a more comprehensive four-dimensional framework for understanding auto insurance loss costs, comprising factors grouped into the following categories:

- *pre-accident driver attributes and vehicle characteristics;*

- *the external environment, such as road conditions and traffic density;*
- *compensation systems, such as tort liability versus no-fault; and*
- *post-accident factors, such as claimant characteristics, medical providers, and attorney representation.*

The paper shows the explanatory value of this framework as compared with the traditional decomposition of loss costs into frequency and severity components.

As an illustration, the paper shows how territory, which is sometimes considered a reflection of external conditions (such as road safety and traffic density), is more properly analyzed as a proxy for post-accident factors—specifically, the "treatment triangle" among claimants, medical providers, and attorneys in certain locations. The paper concludes with two proposed public policy reforms, demonstrating how the expanded four-dimensional framework for personal auto loss cost drivers facilitates the development of more efficacious methods for holding down auto insurance loss costs.

1. INTRODUCTION

Actuarial ratemaking sets policy premiums to cover anticipated losses and expenses. To estimate the needed premiums, the pricing actuary examines the "cost drivers"—that is, the factors that influence the expected future losses and expenses.

In the past, actuaries have concentrated on variables related to driver, vehicle, and geographic characteristics. Indeed, these are the factors most susceptible to policy rating, the traditional role of the casualty actuary.

Although this traditional approach produces accurate rates, it does not provide a full understanding of the underlying factors that influence automobile insurance loss costs. The recent studies of the Insurance Research Council (IRC, formerly AIRAC), the RAND Institute, and the Automobile Insurance Bureau of Massachusetts (AIB) illuminate a host of other factors that play significant roles in determining these costs.

This paper integrates the results of these studies into a comprehensive framework for analyzing personal automobile insurance loss costs. The framework looks at four dimensions that affect loss costs: (a) pre-accident driver attributes and vehicle characteristics, (b) the external environment, (c) compensation systems, and (d) post-accident factors. Section 6 shows how these four dimensions combine to influence territorial rates.

The implications for policy pricing are highlighted by comparison with the traditional "claim severity/claim frequency" paradigm, using national statistics compiled by the IRC and Massachusetts experience analyzed by the AIB. The importance of the expanded framework is further revealed by three other uses besides policy pricing:

- Several traditional classification dimensions are reinterpreted, underscoring their true effects on insurance loss costs. The IRC studies, for instance, show how territory is shifted from a factor related to the physical environment to a factor related to claimant characteristics.
- Changes in compensation systems can be more accurately priced. The AIB studies show how a simplistic prognosis of the 1989 Massachusetts no-fault reform vastly mis-estimated the true effects on loss frequency and loss severity.
- Public policy recommendations for lowering the cost and improving the efficiency of personal auto insurance are made more realistic and more effective.

These uses of the expanded framework for personal automobile insurance cost drivers reflect the widening role of the casualty actuary in today's insurance environment.

2. FRAMEWORK

Let us begin with the fundamental question faced by the pricing actuary:

> *An insurer issues a personal automobile insurance policy. What factors influence the loss cost expected from this policy?*

The traditional actuarial focus on ratemaking and classification systems, as well as a predilection for quantifiable data, has led to an emphasis on pre-accident factors—particularly driver, vehicle, and geographic characteristics—to the virtual exclusion of other factors that affect the insurer's payments. The likelihood and severity of an accident are considered to depend on driver attributes, vehicle characteristics, and garaging location. The amount of the claim and its monetary resolution stem directly from the physical aspects of the auto accident.

This perspective suffices for an insurance environment with an existing classification plan. It is insufficient for an actuary working with changing external conditions and compensation systems, or for an actuary refining classification plans, revising pricing procedures, or formulating public policy recommendations.

The expanded perspective in this paper groups loss cost drivers into four dimensions, as shown in Figure 1.

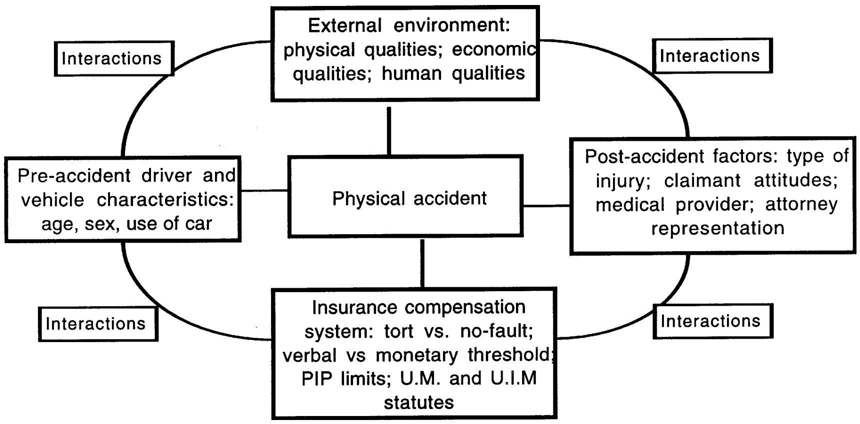
1. *Pre-Accident Driver Attributes and Vehicle Characteristics*

   Pre-accident characteristics include the traditional rating variables that are shown on the policy application:

   - *Driver attributes*, such as age, sex, marital status, driving record, driving experience, and driver education.
   - *Vehicle and vehicle use characteristics*, such as make and model of the car, horsepower, mileage driven, multi-car discounts, and vehicle use (e.g., drive to work vs. pleasure).

FIGURE 1

DIMENSIONS OF LOSS COST DRIVERS



- *Policy age*, such as new versus renewal policy.

These factors are used for setting rate relativities in existing classification schemes, since they are known to the insurer at policy inception and they can therefore be used to rate the policy. These factors are most important for predicting the occurrence of a physical event (e.g., an accident). Once that event occurs, the insurance payments (if any) depend on a number of post-accident factors and on the compensation system.

2. *The External Environment*

The external environment relates to non-insurance characteristics that affect claim frequency or claim severity. We group these factors into three categories:

- *Physical qualities*, such as traffic density, road hazards and maintenance, and safety regulations (such as speed limits and seat-belt statutes). The garaging location, or the rating territory, is often thought of as reflecting physical road qualities. In truth, territory affects auto claim costs primarily by its relationship to several post-accident factors, such as attorney representation, the nature of the medical providers, and claimant characteristics. As the discussion below indicates, territory is not simply a reflection of road characteristics and traffic density.[1]

- *Economic trends*, such as the argument that in prosperous years people drive more, purchase new vehicles, and take more vacations, leading to higher bodily injury accident frequencies.

- *Individual circumstances*, e.g., a higher proportion of poor residents in certain geographic areas may lead to more uninsured motorists and higher UM costs.

3. *Compensation Systems*

Auto injury compensation systems may be grouped into tort liability, no-fault, and add-on systems. Tort liability systems may be subdivided by the financial responsibility limits and by the type of comparative negligence rule. No-fault compensation systems may be subdivided by the type of tort threshold: pure, verbal, and monetary. Verbal thresholds may be further classified by their definitions. Monetary thresholds may be further classified by their magnitude. No-fault systems may also be classified by the personal injury protection (PIP) limits, by the type of benefits provided, and by the compensation rate (e.g., "75% of wage loss").[2]

---

[1]Physical factors may be important in particular instances, such as to explain a high accident frequency at a four way intersection with stop signs but no traffic light. They are less important in the aggregate. Two cities may have similar physical characteristics and similar accident rates but different claim frequencies.

[2]The types of auto compensation systems, and their resultant incentive effects, may also be categorized in relationship to other health care plans. For instance, traditional "fee for

The compensation system has a direct effect on claim frequency and claim severity, since a claim may be compensable under one system but not under another system. The compensation system has an incentive effect both on claim filing and on claim severity. For instance, claims may be built-up either to pass a monetary tort-threshold in a no-fault compensation system or to legitimize claims for pain and suffering awards in a tort liability system.

These incentive effects are sometimes subsumed under a broader "insurance lottery" perspective, which says that claim-filing behavior depends in part on the ease of pressing an insurance claim. States with strong anti-fraud statutes may greatly reduce claim frequency. The build-up of claims is useful only if it provides a greater net gain to the claimant and his or her associates. Incurring additional medical expenses in a no-fault state with a strong verbal tort threshold is sometimes pointless, if the type of injury does not allow a tort claim to be pursued.

Auto injury compensation systems are most important in explaining state-by-state differences in insurance costs. Not only the insurance compensation but also the occurrence of claims and the amount of economic damages depend on the state compensation system.

4. *Post-Accident Factors*

Studies of "classification efficiency" often fault traditional risk classification plans for failing to adequately explain the variance in insurance loss costs (see Spetzler, Casey, and Pezier [13], Giffin, Travis, and Owen [4], and Woll [16]). Indeed, the factors discussed above relate primarily to the occurrence of the physical event—i.e., cars colliding with one another. Other factors, such as the type of injury,

---

service" medical plans require the claimant to pay both a deductible and a coinsurance payment, and they restrict over-payment by the collateral source rule. Most auto insurance compensation systems, in contrast, have no such offsets.

the honesty of the claimant, attorney representation, and the type of medical treatment sought, are strong predictors of insurance claim costs.[3]

Post-accident factors relate to (i) whether an injury claim will be brought for the physical accident and (ii) the amount of the claim. These factors may be grouped into the following categories:

- *Type of injury*, such as soft-tissue injuries (back and neck sprains and strains) vs. fractures vs. more serious injuries. The hierarchy of injury types should distinguish between injuries that are more or less susceptible to "build-up" and potential fraud. For instance, a fracture is readily discernable, and the length of needed treatment is objectively determinable. Soft-tissue injuries are harder to validate, and there is less consensus on their appropriate treatment. If claim frequency depends (in part) on claim-filing behavior, and if claim severity depends (in part) on "build-up," then a hierarchy of injury types that differentiates claims by the criteria mentioned above is most useful for forecasting loss costs.
- *Type of medical practitioner*, such as physician vs. chiropractor vs. physical therapist, as well as type of treatment, such as hospital admission vs. outpatient treat-

[3]See, for instance, Weisberg and Derrig [15], particularly Tables 2 and 3 on page 133, Table 4 on page 135, and Table 6 on page 138. Weisberg and Derrig note (page 132) that

> *For claims that involved strains or sprains, variables that reflected the seriousness of the injury explained little of the variation in medical expenses. For pure strains/sprains our model R^2 was only .04 and for mixed claims with strains/sprains and "hard" injuries, the R^2 was .21. … However, when variables related to treatment utilization and claimant behavior were added in, the value of R^2 for strain/sprain claims jumped to .78 and that for mixed claims to .79.*

In general, claimants are more likely to engage attorneys in more serious cases. However, even when the degree of injury is comparable, attorney-represented cases are more likely to settle for higher amounts, though the net proceeds to the accident victim may not be higher (AIRAC [2], IRC [7, pp. 56–62]).

ment.[4] The type of injury and type of medical practitioner variables have two or more values for most claims. In other words, many auto liability claims allege both a sprain/strain and another type of injury. Similarly, many claimants see two or more types of medical practitioner, such as a physician in an emergency room setting and then a chiropractor for extended visits.[5]

- Whether the claimant is being *represented by an attorney*. In tort liability claims, plaintiffs' attorneys are generally compensated on a contingent fee basis. That is, the attorney receives a percentage of the court award or of the insurance compensation, such as 33%.

For bodily injury (BI) claims, the insurance company's settlement offer is often a multiple of the economic damages (generally medical bills and wage loss) suffered by the accident victim. The plaintiff's attorney

---

[4]The distributions of auto insurance claims by type of injury and type of medical practitioner differ from the distributions for standard health insurance. The distributions noted by Marter, Weisberg, and Derrig for claims reported in Lawrence, Massachusetts (an area suspected of widespread insurance fraud) are particularly revealing. Among the 1985–1986 Lawrence claims studied by Marter and Weisberg [12], 44 out of 48 were for sprains or strains (page 404). For these claims, moreover, 89% of the medical charges went to chiropractors, and only 10% went to physicians (page 407); see also Weisberg and Derrig [14].

The predisposition of some actuaries is to view the neck and back sprains treated by a chiropractor as a minor influence on auto insurance loss costs. The contrary is true. In certain areas, such claims are the principal loss cost drivers. Even in the rest of the country, strains and sprains are the predominant type of auto injury in bodily injury claims, and treatment by chiropractors and physical therapists is becoming increasingly common.

[5]The Insurance Research Council [7] has documented both the multiplicity of injuries and of medical practitioners as well as the trends in these statistics in recent years. In 1992, the average BI claimant reported about two different types of injury and was treated by about two different types of medical practitioners.

> *The growing share of claimants reporting multiple types of injuries also is reflected in the growth of the average number of different types of injuries reported by BI claimants. BI claimants reported an average of 1.92 types of injuries per person in 1992, up from 1.79 types of injuries per person in 1987 [7, p. 2].*
>
> *On average, BI claimants were treated by 1.95 different types of medical practitioners per person in 1992, up from 1.59 in 1987 [7, p. 3].*

has a financial incentive to encourage the "build-up" of the claim.[6] The IRC studies have consistently shown higher average costs for attorney-represented claims, even when the type of injury is held constant (see IRC [7, page 61]).

Perspectives regarding post-accident factors vary widely, and can be illustrated by looking at two extremes. The difference in viewpoint is essential for understanding the costs of the auto insurance system and for developing reforms to reduce this cost.

Suppose an accident victim in a no-fault state with a monetary tort threshold suffers a lower back sprain, sees a chiropractor 30 times, recovers the out-of-pocket expenses from PIP coverage, and files a BI claim which is handled by an attorney.

- From an innocent (perhaps "idealistic") perspective, the physical injury itself is the loss cost driver. The back sprain incurred in the auto accident motivates the victim to seek out a medical

[6]An illustration should clarify this. Suppose that an insurance company settles most BI cases for three times the economic damages: that is, the compensation for "pain and suffering" is about twice the sum of wage loss and health care bills. Suppose also that attorneys require 33% of the award for most BI claims.

If an accident victim without an attorney incurs $1,000 in medical bills, the total BI compensation would be $3,000, for a "net monetary gain" to the claimant of $2,000. If the claimant is represented by an attorney, who takes 33% of the award, or $1,000, the claimant receives only $1,000. However, if the attorney "encourages" the claimant to stay home from work or to incur greater medical bills (perhaps by recommending a medical practitioner who sets a longer course of treatment), so that the economic damages rise to $2,000 and the insurance compensation rises to $6,000, the attorney's fee becomes $2,000 and the claimant's "share" is back to $2,000, which is the amount of general damages when no attorney is involved. Many insurance company personnel and industry researchers believe that this accurately depicts the role played by many (though not all) attorneys. In other words, attorneys often drive up the cost of the system, with little benefit to claimants (assuming there are no other collateral sources of compensation, such as sick pay plans and private medical insurance). See also the discussion later in this paper regarding the overtreatment of many automobile accident claims.

In no-fault states, there is a second incentive to build up claims. Many states have monetary tort thresholds, which allow accident victims to press bodily injury claims only if medical bills exceed a stated amount. (Most of these states also have verbal thresholds, which allow BI claims for "serious" injuries even if medical bills are low.) Attorneys can provide little aid in PIP recoveries. However, by encouraging their clients to "build up" the medical bills to exceed the tort threshold, they can file BI claims for "pain and suffering."

practitioner competent to handle such injuries. The length of the needed treatment and the lack of reimbursement for non-economic damages under PIP coverage (such as "pain and suffering") motivate the victim to file a BI claim. The complexity of the insurance claim process and the uncertainties of BI compensation motivate the victim to seek an attorney's aid. The "innocent perspective" sees the claim as the direct result of the physical accident and the insurance compensation as independent of the honesty of the claimant, attorney, or medical practitioner.

- The cynical perspective sees the "entitlement philosophy," or "claims-consciousness," or the "insurance lottery" as the loss cost driver.[7] Whether the accident victim files an insurance claim, seeks treatment from a particular medical practitioner, or even "suffers" a back sprain is not dependent solely upon the physical events in the auto accident. Rather, the accident victim, seeking to benefit financially from the accident, sees an attorney, who encourages him or her to be examined by a medical practitioner who has a history of recommending extended treatment. The medical practitioner diagnoses the back sprain and recommends an extended course of treatment. Either the medical practitioner or the attorney notes that the medical expenses will be covered by PIP (as well as by other health insurance) and that the BI claim will pay for additional "pain and suffering" costs. The accident victim, the attorney, and the medical practitioner all benefit from the extended course of treatment.

In this "cynical perspective," the treatment provided was not solely the result of the physical accident. Rather, it is also affected by the desire of all three parties involved (the claimant,

---

[7]Casualty actuaries speak of "claims consciousness," which the IRC studies refer to as "claim-filing behavior." "Claim consciousness" has been measured by ratios of bodily injury to property damage claims. See the discussion of territory in the text. The "entitlement philosophy" is broader. Many accident victims, having paid thousands of dollars over the years for their auto insurance, now feel that they are entitled to recover their money from the "insurance industry."

the attorney, and the medical practitioner) to maximize the insurance compensation.

The difference in perspectives leads to differing public policy recommendations. The "innocent perspective" sees injury prevention as the key to reducing insurance costs. Injury prevention efforts include safety standards for new cars, safety inspections for older cars, mandatory seat belt laws, air bags, lower speed limits, and better policing of driving-while-intoxicated statutes. The "cynical perspective" sees the removal of the "lottery" incentives as the key to reducing insurance costs. Policy actions include anti-fraud units, peer review of medical practitioners, and verbal tort thresholds in no-fault states.

3. THE FREQUENCY-SEVERITY PARADIGM

The explanatory power of the expanded framework can be seen most clearly in contrast with the old frequency/severity paradigm. Previously, personal automobile loss cost drivers were viewed as inflation-induced changes in loss severity and as slow, long-term trends in loss frequency. The frequency trends have sometimes been modeled by econometric equations based on changes in gasoline prices, car density, and similar factors.[8]

Although this paradigm is an important component of actuarial ratemaking, it does not fully explain why claim frequency or claim severity may be changing, nor does it necessarily tell us what may be expected in the future. The expanded framework presented in this paper provides a broader perspective for viewing personal auto loss frequency and loss severity. It is particularly useful for understanding the causes of frequency and severity trends and for formulating public policy proposals to improve the auto insurance compensation system.

[8]The Insurance Services Office, for example, has studied the effects of various economic factors on automobile insurance claim frequency and it has suggested potential econometric models incorporating these factors.

Frequency

The Insurance Research Council studies of the mid-1990s, using data compiled by the Insurance Services Office (ISO) and the National Association of Independent Insurers (NAII), note that the countrywide property damage (PD) claim frequency decreased by 12% from 1987 to 1992. This is a measure of accident frequency; and it is consistent with fewer youthful drivers, greater public awareness of drunk drivers, and better quality cars.

Over the same time period, the frequency of bodily injury claims increased by 16%. Given the 12% decline in accident frequency, this is a 32% increase in bodily injury claims per physical accident.[9]

For bodily injury, the changes in claim-filing behavior among the public overwhelm the changes in physical accident frequency. The frequency drivers are not economic and environmental attributes like gasoline prices and car density. Rather, the primary causes lie in the claim and claimant characteristics dimension of the expanded framework:

- *Type of injury*: The greatest increase over this period was in "soft-tissue" injuries (sprains and strains). Moreover, sprains and strains are particularly dominant in urban areas, which also have the highest ratio of BI to PD claims. In fact, the May 1994 IRC study, *Paying for Auto Injuries* [9], concludes that "almost all of these additional injury claims are for difficult-to-verify injuries such as sprains and strains."
- *Type of medical practitioner*: The greatest increase over this period was in chiropractic treatment, especially for sprains and

[9]Formally, $32\% = [(1 + 16\%) \div (1 - 12\%)] - 1$. For the full IRC studies, see Insurance Research Council [7; 10]. See also Insurance Research Council [9]: "More people involved in auto accidents are making claims for injuries, even though accident rates have been declining. … Many states enacted seat belt laws during these years, resulting in substantial increases in seat belt use. Seat belts reduce the number and severity of injuries in auto crashes. Around the same time, states passed tougher drunk driving laws in response to growing public awareness of this problem. In addition, the federal government now requires additional safety standards for vehicles that make cars safer for passengers."

strains. Conversely, injuries requiring hospital stays have declined.

- *Attorney involvement*: Between 1977 and 1992, the percentage of claimants represented by lawyers rose from 31% to 46% for all injury coverages combined and from 47% to 57% for bodily injury claims (IRC [7, pp. 43–44]).[10]

- *Law changes*: In 1989, the threshold in Massachusetts for pursuing a BI liability claim was increased from $500 to $2,000. The traditional actuarial analysis would predict that the frequency of BI claims would decrease substantially, because injury claims with medical expenses between $500 and $2,000 would no longer be eligible for BI liability payments. In fact, the frequency reductions were minimal, because of incentive effects. The higher tort threshold encouraged accident victims (and their attorneys) to "build up" the medical expenses so that a bodily injury claim could be filed (see Marter and Weisberg [12]; Weisberg and Derrig [14]).

In sum, changes in claim and claimant characteristics are the key drivers for bodily injury claim frequency trends. Moreover, the claim frequency trends for BI coverage have been different from the corresponding claim frequency trends for property damage liability and for collision coverage, even though these trends ostensibly relate to the occurrence of the same auto accidents.

---

[10] Of additional concern to pricing actuaries are the relative differences by state, which are relevant for severity and frequency trends. Credibility weighting statewide severity and frequency trends with the corresponding countrywide trends is inappropriate if the statewide trends are affected by changes in claim and claimant characteristics and in the compensation system in ways that the countrywide figures are not affected.

The same phenomenon may be seen in workers compensation insurance. In the past, statewide medical benefit trends were credibility weighted with countrywide trends. However, trends were lower in states with medical fee schedules than in states without such schedules. (The existence of a state medical fee schedule might be considered a workers compensation counterpart to the medical practitioner dimension of the personal automobile framework here.) Now, the figures assigned the complement of credibility in workers compensation medical benefit trends depends on whether the state has a medical fee schedule.

Loss Severity

Actuaries have traditionally used two methods to project trends in loss severity.

A. Trend projections based on internal data fit observed average costs per claim to an exponential curve and assume that the same trend will continue in the future.

B. Trend projections based on external data correlate the historical average costs per claim with an economic index, such as the medical cost component of the CPI, and then estimate future claim severity based on the expected future values of the economic index.

Both methods work well in certain environments. The first method works well when the underlying trends are stable, so that past changes in loss severity are deemed to be unbiased predictors of future changes. The second method works well when loss cost trends are considered to be closely linked to recognized inflation indices.

In personal automobile bodily injury insurance, loss severity trends are composed of several influences, such as:

- *Trends in cost of treatment*. This includes both (a) medical cost inflation and (b) trends in utilization rates that are independent of the personal auto compensation system.[11]
- *Trends in loss frequency*. Severe automobile accidents lead to insurance claims regardless of the claim-filing proclivity of the accident victim. The growing influence of attorneys and the changing claim-filing behavior of the public lead to greater claim frequency for minor injuries, such as sprains and strains with no visible signs of impairment. These are often low cost

[11] For instance, even when the personal auto compensation system remains unchanged, the development of new medical procedures may engender greater utilization of services, medical malpractice suits may stimulate more "defensive medicine," and the increased use of chiropractic treatment and physical therapy may change the mix of claims.

claims. In other words, the factors that increase loss frequency often lead to decreases in average loss severity.[12] A change in expected frequency stemming from changes in claim or claimant characteristics should be partially offset by changes in expected severity.

- *Changes in compensation systems and in claim handling procedures.* Compare the discussion above on the tort threshold change in Massachusetts in 1989. The new low severity projections changed dramatically because a whole cohort of cases, which formerly had medical costs between $500 and $2,000, moved up to over $2,000, with higher pain and suffering awards (see Marter and Weisberg [12]; Weisberg and Derrig [15]).

4. PROXIES

Many of the traditional classification variables used today are proxies for the true ("causative") factors affecting insurance loss costs. To clarify the difference between a causative factor and a proxy, let us contrast life insurance with automobile insurance.

- Age is generally considered a physiological attribute that directly affects expected mortality rates, so it is used as a rating variable for life insurance underwriting and life annuity underwriting.
- Sex and age also have strong correlations with auto accident frequencies, so they are used to set auto insurance rate relativities. Indeed, a 17 year old unmarried male may have about the same mortality rate as a 30 year old married female, but he may have several times the auto bodily injury claim frequency rate that she has. Yet sex and age (except at advanced ages when bodily capabilities deteriorate) have little intrinsic relationship with accident propensity. Rather, they serve as prox-

[12] The IRC studies demonstrate this phenomenon. Among the BI, PD, and PIP coverages over the 1980 to 1993 period, BI had the greatest increase in claim frequency and the smallest increase in claim severity; see especially Insurance Research Council [10, chapters 1 and 2].

ies for other driver characteristics that are not easily defined or measured, such as "risk-taking" predilections or psychological maturity.

The use of territory as a proxy for external conditions, driver attributes, and claimant characteristics is discussed below.

5. INTERACTIONS

The factors in one dimension of the expanded framework presented here may interact with the factors in another dimension to determine expected loss costs. We illustrate with two examples.

- *Pre-accident underwriting attributes and compensation systems:* Age, sex, and marital status may be more important as rating variables in tort liability systems, which focus on the tortfeasor's "fault," than in no-fault compensation systems, in which all accident victims are compensated. Conversely, the applicant's income and employment status may be important in no-fault compensation systems with high PIP wage-loss limits.[13]
- *Claim characteristics and compensation system:* The "padding" of claims, or "build-up," can be stimulated by a no-fault compensation system with a low or moderate monetary tort threshold. The AIB studies by Marter, Weisberg, and Derrig referenced above show how the 1989 increase in the Massachusetts tort threshold increased the average number of outpatient visits to chiropractors, thereby resulting in more claimants exceeding the tort threshold.

The interactions of the four components of the expanded framework are essential for proper pricing and public policy recommendations, as discussed in the final section of this paper.

[13]The comments in the text relate to relative importance only. Thus, age, sex, and marital status are important for no-fault compensation systems as well, since young, unmarried, male drivers are not only more likely to cause accidents, they are also more likely to be injured in accidents. Similarly, income and employment status are important for tort liability systems as well, since unemployed persons with few assets are often "judgment proof" and therefore carry low liability limits of coverage.

6. TERRITORY AS A RATING VARIABLE

Territory is one of the chief variables used by U.S. insurers for automobile rate setting. Territory provides an excellent example of how pre-accident driver characteristics, the pre-accident physical environment, post-accident characteristics, and the compensation system all affect automobile insurance loss costs.

Pre-Accident Driver Characteristics

Pre-accident driver characteristics, such as age, sex, and marital status, do not generally have a direct effect on territorial relativities. Since the distributions by age and sex are relatively constant by territory, these variables do not affect territorial relativities.[14]

External Environment

The physical environment in an area can raise or lower the expected number of accidents. For instance, population density and vehicle density are often cited as explanatory variables for accident frequency on the assumption that, with more cars per square mile, there will be more accidents per car.

In a 1988 study, ISO and the NAII compared the variation in traffic density with the variation in PD claim frequencies.[15] Although the major cities in each state had traffic densities over ten times the statewide average, these cities had PD claim frequencies that were often only 10% higher than the statewide average.[16]

---

[14] An exception would be communities, such as retirement communities, where a disproportionate number of senior citizens reside. This lowers the average pure premium of the territory, but the class rating system should produce the correct overall territorial rate.

[15] Traffic density, or "vehicle density," is defined in the study as car registrations per square mile.

[16] For example, the 1988 study shows a traffic density for Chicago of 5,423 cars per square mile, versus the statewide average of 152 car registrations per square mile. Nevertheless, the PD claim frequency in Chicago was only 11.7% higher than the statewide average claim frequency. More recent data (Insurance Research Council [10]) shows a

In sum, traffic density does not explain much of the elevation of automobile claim frequencies in urban areas. In theory, accident frequencies might be expected to increase proportionately with traffic densities. In practice, traffic safety devices in urban areas, such as traffic lights, stop signs, and well-designed roads, by causing traffic to move at a somewhat lower speed, keep the increase in the accident frequency over the statewide average frequency to a relatively small percentage.

Table 1 shows 1993 property damage claim frequencies by state.[17] With only two exceptions, the states lie in a narrow range from 20% above to 25% below the countrywide average of four claims per 100 insured vehicles.

Several other attributes of the physical environment also affect automobile insurance rates. Automobile theft rates vary by geographic location. Higher theft rates in urban areas cause higher comprehensive losses and therefore higher premiums for comprehensive coverage. Similarly, the 1988 ISO/NAII study shows substantially higher uninsured motorist costs in many urban areas, presumably resulting, at least in part, from higher numbers of uninsured motorists. Finally, the cost of services provided by insurers, such as auto body shop repair costs and medical costs, varies by region; and they therefore affect territorial relativities.

Post-Accident Characteristics

The occurrence of an automobile accident is a physical event. The decision to press a BI claim once an accident has occurred, however, varies dramatically by state and even within a state.

The two dimensions of the expanded framework discussed directly above—pre-accident driver characteristics and pre-accident physical characteristics—relate to the occurrence of the

---

similar relativity, with the Chicago PD claim frequency being about 13% higher than the statewide average claim frequency.

[17] The data are from IRC [10, Figures 2–6].

TABLE 1

NUMBER OF PD CLAIMS PER 100 INSURED VEHICLES (1993)

Massachusetts	7.13	Indiana	3.98	California	3.65	South Carolina	3.38
District of Columbia	5.38	Nebraska	3.98	Oklahoma	3.64	Hawaii	3.38
Texas	4.76	Georgia	3.89	Kentucky	3.63	Vermont	3.36
Missouri	4.72	Alaska	3.89	Wisconsin	3.62	South Dakota	3.32
New York	4.67	Iowa	3.89	Arkansas	3.60	North Carolina	3.31
Illinois	4.35	Michigan	3.81	West Virginia	3.59	New Mexico	3.29
Rhode Island	4.23	Ohio	3.77	Virginia	3.54	Mississippi	3.26
Maryland	4.18	Nevada	3.76	Tennessee	3.54	Alabama	3.26
Connecticut	4.11	Minnesota	3.73	Colorado	3.52	North Dakota	3.26
Utah	4.09	Pennsylvania	3.70	New Jersey	3.50	Maine	3.23
Louisiana	4.05	Florida	3.69	Washington	3.45	Montana	3.19
Kansas	4.03	Arizona	3.68	Oregon	3.45	Wyoming	3.02
New Hampshire	4.02	Delaware	3.67	Idaho	3.39	Countrywide	4.00

accident itself. Post-accident characteristics relate to the probability of a claim being filed given that an accident has occurred.

We want to measure this probability for BI claims. Note carefully: we are not concerned with BI claim frequency or with automobile accident frequency. Rather, we are concerned with the probability of a BI claim being filed given that an accident has occurred where another driver could potentially be liable for damages.

We presume that the filing of a PD liability claim is influenced primarily by the nature of the physical accident, so relative PD claim frequency is a proxy for relative accident frequency where another driver could potentially be liable for damages. The ratio of BI claims per 100 PD claims serves as a measure of the propensity to press personal injury claims.[18] Table 2 shows the countrywide trend in this ratio over the past 15 years, from 18 BI claims per 100 PD claims in 1980 to over 29 BI claims in 1993.[19]

Our concern here is the relationship of this ratio to geographic location; that is, the variation in this ratio by state and by territory within state. Indeed, the BI/PD ratios vary greatly by state, as Table 3 shows. California, for instance, produces 61 BI claims for every 100 PD claims, whereas Wyoming, which is also a tort state, produces only 18 BI claims. (The effects of the compensation system are also evident from Table 3: the eight states with the lowest BI/PD ratios are all no-fault states.)

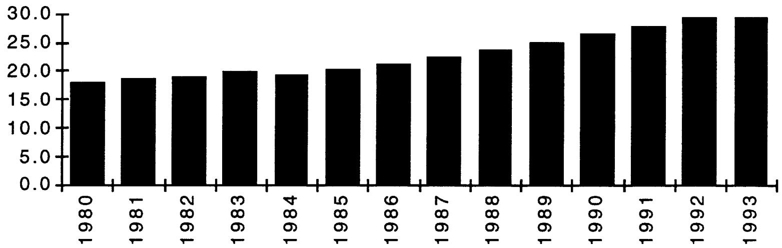
---

[18]The Institute for Civil Justice (RAND) uses a similar measure, the ratio of soft injury claims to hard injury claims; see Carroll, Abrahamse, and Vaiana [3, page 13]. The reasoning is similar to that underlying the BI/PD ratios. Hard injury claims, such as broken bones, will be pressed in almost all circumstances, whereas the number of soft injury claims, such as sprains and strains, depends in part on the propensity to file insurance claims. The Institute for Civil Justice estimates the cost to consumers from over-treatment and similar types of claim buildup and fraud to be between $13 billion and $18 billion a year [3, p. 3]. IRC estimates this cost to be between $5.2 billion and $6.3 billion a year [8, p. 23].

[19]The data for the exhibits in this section are derived from IRC studies. They are from both full tort states and no-fault states. These are BI liability claims; they do not include no-fault claims.

TABLE 2

CLAIMS PER 100 PD CLAIM



The trends in BI/PD ratios over time and the variations by territory highlight the strong effects of post-accident characteristics on auto insurance loss costs. In California, for instance, the 61% BI/PD ratio for 1993 marks a steady climb from a 31% BI/PD ratio in 1980.

A common perception is that the accident frequencies themselves vary greatly by territory, being far higher in urban areas than in rural areas. Although these differences in accident frequencies do exist, the preceding statement confuses two issues, and it misinterprets the reasons for the territorial differences. Often, the frequency of physical accidents and of PD liability claims is only marginally greater in metropolitan areas than in the surrounding region. Once the accident occurs, however, the BI claiming pattern is substantially different in metropolitan areas than in other parts of the state.

IRC data from 1989 through 1991 (IRC [10, App. B]) illustrate this phenomenon. For instance, the PD claim frequency during these years was about 10% higher in Los Angeles than in the rest of the state, but the BI/PD ratio was 98.8% in Los Angeles, versus 45.2% in the rest of the state. In other words, it was not accident frequency differences that were driving up BI liability costs in Los Angeles, but BI claim filing patterns that were causing the difference.

TABLE 3

NUMBER OF BI CLAIMS PER 100 PD CLAIMS (1993)

California	60.7	Massachusetts	34.8	West Virginia	26.9	Nebraska	19.5
Louisiana	49.4	Oregon	34.3	Indiana	26.0	Florida	19.1
South Carolina	46.8	North Carolina	34.1	Maine	26.0	South Dakota	18.5
Nevada	45.4	Arkansas	33.9	Idaho	25.6	Wyoming	17.6
Arizona	45.3	Georgia	33.6	Alabama	25.1	New York	16.3
Rhode Island	39.7	Virginia	31.3	Connecticut	24.9	Kentucky	15.9
Oklahoma	38.9	Illinois	30.4	Montana	24.3	Hawaii	13.9
District of Columbia	38.8	New Hampshire	29.8	Utah	22.2	Colorado	12.8
New Mexico	37.6	Delaware	29.1	Alaska	21.3	Minnesota	11.7
Washington	37.4	Ohio	28.1	New Jersey	21.2	Kansas	9.2
Texas	36.7	Tennessee	28.1	Vermont	20.9	Michigan	8.2
Maryland	35.5	Missouri	27.8	Pennsylvania	20.4	North Dakota	5.6
Mississippi	35.3	Wisconsin	27.4	Iowa	19.9	Countrywide	29.3

Although BI/PD ratios are generally higher in large metropolitan areas, a simple urban/rural dichotomy is not always a good proxy for the actual claim-filing patterns. For instance, during the 1989 through 1991 period, the state of Pennsylvania as a whole had a BI/PD ratio of 23%, the city of Pittsburgh had a ratio of 18%, and the city of Philadelphia had a ratio of 78%.

The attributes of territorial differences implicit in the discussion above have major implications for understanding auto bodily injury liability loss cost drivers:

- Loss cost differences by region are great, with some areas, whether urban centers or entire states, having high insurance costs and affordability concerns.
- Traffic congestion is *not* the primary determinant of these differences. In fact, the variations in PD claim frequencies are often minor between urban areas and the statewide average.
- Differences in the BI/PD ratios account for much of the variation in BI loss costs by region, with higher cost areas having higher BI/PD ratios.

Thus, once an accident occurs, the decision of whether to over-treat the injury, or even to seek medical treatment when no injury exists, is one of the major factors driving the cost differences between states for bodily injury coverage.

The Treatment Triangle

The over-treatment of automobile injuries in certain locations, as well as the treatment of non-existent injuries, results from the interaction between claimants, medical providers, and attorneys; and it depends upon the type of injury and the structure of the compensation system. Our emphasis in this paper is on the loss cost drivers affecting territorial relativities. In particular, the major factors affecting territorial relativities are *not* pre-accident driver characteristics or pre-accident physical characteristics. Rather, the post-accident characteristics and the compensation system attributes determine how automobile accidents affect insurance payments.

Television reports on the human toll of highway accidents leave us with grisly pictures of torn metal and mangled bodies, as if most automobile accidents resulted in severe injuries. In fact, the opposite is true. About 60% of BI claimants report their only injury to be a strain or a sprain, and another 23% claim to have suffered a strain or a sprain plus another injury (IRC [7, p. 19]). Most strain and sprain injuries are difficult to verify, their severity is hard to measure, and radically different treatment patterns may be recommended by medical providers.

For over-treatment of injuries to occur, it is necessary that all parties deciding on the course of treatment gain from the over-treatment. For injuries and illnesses *not* covered by automobile liability insurance or workers compensation insurance, the patient generally derives no financial gain from the medical treatment. Even if the patient has health insurance coverage (whether individual health insurance or employer provided group health insurance), the coverage simply reimburses the hospital costs or physicians' charges, and it often requires a co-payment from the patient.

Automobile bodily injury claims are different. BI liability awards consist of two parts: economic damages, such as medical costs or wage loss, and general damages, or "pain and suffering." Medical expenses comprise about three-fourths of economic damages. "Pain and suffering" damages are not objectively determinable on their own. Rather, the general damages are usually pegged as a multiple of the economic damages.

In sum, the medical expenses incurred by the claimant drive not only the insurance reimbursement for economic damages but also the insurance award for general damages. Each dollar of medical expenses incurred may translate into three dollars of insurance compensation.[20] In fact, many potential BI claims in the United States are not even pursued unless there is a sufficient amount of medical expense to support a "pain and suffering" claim.

Three parties are needed for excessive treatment to exist on a large scale, and the interactions of these parties are a major influence on territorial relativities. The three parties are:

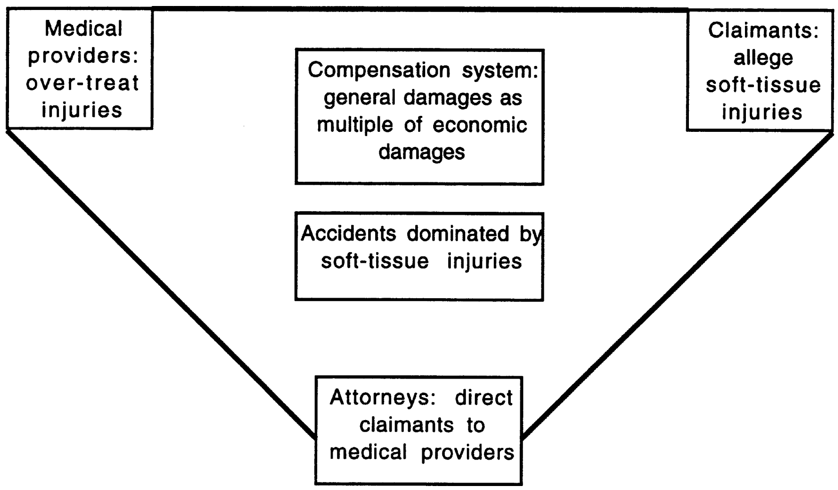
- Medical providers who aggressively treat even routine strain and sprain injuries in order to increase the medical expenses paid. The vast majority of medical providers, of course, do not engage in such over-treatment of minor injuries. Rather, a small coterie of medical providers who specialize in injuries covered by automobile liability and workers compensation insurance serve this function well.
- Accident victims willing to complain of soft-tissue injuries, even when objective medical impairment is non-existent or slight.
- A third party who can direct a willing accident victim to the proper medical provider. Most auto accident victims are not sufficiently aware of the auto liability compensation system to take full financial advantage of the system. In the United States, a relatively small number of attorneys who specialize in strain and sprain injuries in automobile liability and workers compensation insurance claims fulfill this function by directing potential BI claimants to medical providers willing to over-treat soft-tissue claims.

In automobile accident cases, excessive treatment of soft-tissue injuries inures to the financial benefit of the claimant, the medical provider, and the attorney, and to the detriment of the driving public who pay the premiums that fund these loss pay-

[20]The actual ratio, of course, varies by state and by year, since it is greatly influenced by the type of compensation system.

FIGURE 2

THE TREATMENT TRIANGLE



ments. This phenomenon raises the BI/PD ratios and is a major driver of auto insurance loss costs.

This treatment triangle is shown schematically in Figure 2.

This phenomenon is exceedingly difficult to police, even when insurers are aware of its existence in a given location. As long as the accident victim claims to be injured, the medical provider can continue the aggressive treatment pattern. To justify the recommendation of a particular medical provider, the attorney need only state that the medical provider is licensed by the state and has produced "good results." Sting operations are difficult to run, since a claimant who claims not to be injured will simply not be treated.

Evidence for over-treatment of automobile injuries is necessarily indirect, though in some locations it is compelling. The data from Massachusetts, where a detailed claim database has been in existence for four years, illustrate this point.

Were there no incentive to over-treat injuries, one would expect a wide dispersion of treatment costs for each provider, with some patients requiring substantial treatment while others require minimal treatment, depending on the severity of the injury. Moreover, one would expect that the number of BI claimants treated by a medical provider would be about half the number of PIP claimants, since all injuries need treatment, whereas a BI claim may be filed only if another driver was at fault.[21]

The automobile compensation system in Massachusetts has a $2,000 tort threshold. That is, a BI claim may be filed only[22] if the PIP medical expenses exceed $2,000. A small number of medical providers in Massachusetts have a large percentage of their patients suffering from automobile accident injuries who routinely require greater than $2,000 in treatment. The implication is that the course of treatment is being determined not by the type of injury but by the desire to reach the tort threshold in order to file a BI claim.

Similarly, among automobile accident victims being treated by these same medical providers, the number of BI plus uninsured motorist claimants is almost equal to the total number of PIP claimants. The implication is that patients are being referred to these medical providers for the primary purpose of building up the PIP expenses so that a liability suit can be pursued.

Compensation Systems and Benefit Levels

The type of compensation system and the level of benefits are reflected in the statewide rates and the territorial relativities. Changes in state laws require an analysis of the effectiveness of the current law and of the proposed law. For example, in an

---

[21] In fact, we would expect the number of BI claimants treated by a medical provider to be less than half the number of PIP claimants, since only those cases exceeding the tort threshold can lead to a BI claim (see below in the text).

[22] For certain types of injuries, such as significant scarring, fractures, and serious injuries, a BI claim may be filed even if medical expenses do not exceed $2,000. However, these types of severe injuries are relatively rare in auto accidents. When they do occur, the $2,000 tort threshold is quickly reached.

urban area, the current tort system or monetary tort threshold in a given state may lead to substantial medical overtreatment, with resultant high rates, in comparison to a suburban or rural area with little overtreatment. A law change that curtails this overtreatment would cause a larger percentage decrease in costs in the urban territory than in the suburban or rural territories.

Summary: Territory and the Four-Dimension Framework

Geographic location, or rating territory, has often been a difficult classification variable for the actuary to explain. Why should auto insurance policies cost more in California than in other states? Why does auto coverage cost so much more in certain urban areas?

Driver characteristics do not differ significantly from place to place. Physical conditions, such as road hazards and traffic density, have a minor effect on accident frequencies. They contribute only marginally to the observed loss cost differences by territory. Rather, geographic location and rating territory serve as proxies for powerful but often overlooked factors that drive auto insurance loss costs, particularly the treatment triangle phenomenon discussed here.

7. PRICING AND PUBLIC POLICY

The framework for analyzing personal automobile loss cost drivers presented in this paper has numerous ratemaking and public policy implications, ranging from territorial relativity analysis to pricing statutory amendments. In workers compensation, for instance, the pricing of statutory amendments is a finely honed actuarial tradition, well described in Fratello [4]. It is also half wrong, as shown by the consistent actuarial misestimates throughout the 1980s, since it covers only the direct effects of law changes, not the incentive effects.[23]

[23] See Gardner [5], as well as the numerous state specific studies from the Workers' Compensation Research Institute.

Compensation system reforms in personal auto insurance are often accompanied by mandatory rate rollbacks. If no changes are assumed in claim-filing behavior, then the cost effects of the reform may be grossly over- or under-estimated, as shown by the 1989 Massachusetts changes. It is vital for casualty actuaries to understand the complete system of personal auto loss cost drivers in order to accurately price system changes.

The availability and affordability of auto insurance are of public concern in many jurisdictions, and casualty actuaries are often called to testify on these issues. The actuary who knows what the existing rating plan indicates, but who does not understand why rates are higher in some territories than in others, or how the compensation system affects loss costs, makes a poor prognosticator. Rather, the actuary must measure and explain how claimant behavior and the compensation system interact with the traditional driver attributes, vehicle characteristics, and the external environment to determine the expected loss costs.

We provide two possibilities for public policy reforms to reduce automobile insurance loss costs that stem from the expanded framework in this paper. These are not the only possible reforms, but they are efficacious and practical proposals.[24]

Peer Review of Medical Treatment

The previous discussion of claim characteristics and of medical treatment indicates that one of the major factors contributing to the increases in bodily injury loss costs over the past decade has been the "build-up" of hard-to-verify soft-tissue injuries, generally with extended courses of treatment by a small number of chiropractors, physical therapists, and physicians, often

[24]Other reforms would be equally effective. For instance, most actuaries agree that movement from a tort liability compensation system to a no-fault system with a strong verbal tort threshold, as in Michigan, would reduce overall costs. However, there are strong interest groups opposing such a move, and who support instead such changes as epitomized by California's Proposition 103: rate rollbacks, classification restrictions, and prior approval, but no attack on the real problem of overtreatment.

orchestrated by attorneys experienced in such claims. Insurance claims adjusters are aware of the "padding" in these claims. Yet it is nearly impossible for claims adjusters to find objective evidence of unnecessary or inappropriate treatment, especially on any specific case.

Peer review of medical treatment in auto insurance claims, by state panels of physicians and other medical practitioners, could succeed in eliminating the worst abuse and stemming or reversing the upward trend in bodily injury loss costs. The state insurance department or the Board of Registration would appoint a panel of medical experts to review treatment patterns by individual medical providers. A substantial database of auto injury losses would be needed to properly identify such patterns. It is generally impossible to determine over-treatment by reviewing any one specific case since the severity of any soft-tissue strain or sprain is a subjective estimate. However, by reviewing all treatment by particular medical providers, patterns of overtreatment can be recognized. Medical practitioners would be more hesitant to provide excessive treatment on a consistent basis if they knew that their actions would be subject to professional review.

Consumer Representation

A second factor contributing to the increase in bodily injury loss costs over the past decade has been the rapid increase in attorney representation of insurance claims. If the attorney helps build up the economic damages, there is generally no net loss to the claimant despite the hefty contingency fee, and sometimes there is even a net gain. In addition, the attorney handles all the claim filing paperwork and negotiates with insurance loss adjusters. Both of these activities can be confusing to the average citizen, particularly in third party cases.

State insurance departments could provide claims representatives to handle claim filing and negotiation on behalf of auto accident victims who need aid in insurance matters. The claims representatives would be compensated by salary, so they would have no interest in building up claims. The insurance industry would defray the costs of these claims representatives.

All parties could gain. Claimants would have representation by state insurance officials, who could guide them through the claims process—*at minimal cost to the claimant*. Insurance companies would gain because the cost of such claims representatives is far less than the costs of claim "build-up." The general public would gain by lower insurance premiums and increased satisfaction with the insurance claim process. State insurance departments would gain because they would be offering additional and highly valued services.

8. CONCLUSION

Although claim severity and claim frequency trends are important tools for automobile insurance ratemaking, their explanatory power is limited. The ultimate cost of automobile insurance is a complex and changing mosaic of many diverse factors. Actuaries who understand these factors will be of great value to their companies, and they may eventually help design systems to control the cost of automobile insurance.

REFERENCES

[1] *Attorney Involvement in Auto Injury Claims*, Oak Brook, IL: All-Industry Research Advisory Council, 1988.

[2] *Compensation for Automobile Injuries in the United States*, Oak Brook, IL: All-Industry Research Advisory Council, March 1989.

[3] Carroll, Stephen, Allan Abrahamse and Mary Vaiana, *The Costs of Excess Medical Claims for Automobile Personal Injuries*, Santa Monica, CA: The Institute for Civil Justice, RAND, March 1995.

[4] Fratello, B., "The Workmen's Compensation Injury Table and Standard Wage Distribution Table—Their Development and Use in Workmen's Compensation Ratemaking," *PCAS* XLII, 1955, pp. 171–202.

[5] Gardner, John, *Return to Work Incentives: Lessons for Policymakers from Economic Studies*, Cambridge, MA: Workers' Compensation Research Institute, 1989.

[6] Giffin, Andrew F., Vincent Travis, and William Owen (Eds.), *Automobile Insurance Risk Classification: Equity and Accuracy*, Boston, MA: Massachusetts Division of Insurance, 1978.

[7] *Auto Injuries: Claiming Behavior and Its Impact on Insurance Costs*, Oak Brook, IL: Insurance Research Council, September 1994.

[8] *Fraud and Buildup in Auto Injury Claims: Pushing the Limits of the Auto Insurance System*, Wheaton, IL: Insurance Research Council, September 1996.

[9] *Paying for Auto Injuries: A Consumer Panel Survey of Auto Accident Victims*, Oak Brook, IL: Insurance Research Council, May 1994.

[10] *Trends in Auto Injury Claims*, Second Edition, Part One: Analysis of Claim Frequency, Wheaton, IL: Insurance Research Council, February 1995.

[11] Insurance Services Office and National Association of Independent Insurers, *Factors Affecting Urban Auto Insurance Costs*, New York, NY: Insurance Services Office, 1988.

[12] Marter, Sarah S., and Herbert I. Weisberg, "Medical Costs and Automobile Insurance: A Report on Bodily Injury Liability Claims in Massachusetts," *Journal of Insurance Regulation*, 9, 3, March 1991, pp. 381–422.

[13] Spetzler, Carl, Barbara Casey, and Jacques Pezier, *The Role of Risk Classification in Property and Casualty Insurance: A Study of the Risk Assessment Process*, Menlo Park, CA: Stanford Research Institute, 1976.

[14] Weisberg, Herbert I., and Richard A. Derrig, "Fraud and Automobile Insurance: A First Report on Bodily Injury Liability Claims in Massachusetts," *Journal of Insurance Regulation*, 9, 4, June 1991, pp. 497–541.

[15] Weisberg, Herbert I., and Richard A. Derrig, "Pricing Auto No-Fault and Bodily Injury Liability Coverages Using Micro-Data and Statistical Models," Casualty Actuarial Society *Forum Special Edition: 1993 Ratemaking Call Papers*, 1993, pp. 103–153.

[16] Woll, Richard G., "A Study of Risk Assessment," *PCAS* LXVI, 1979, pp. 84–138.

THE MECHANICS OF A STOCHASTIC CORPORATE FINANCIAL MODEL

GERALD S. KIRSCHNER AND WILLIAM C. SCHEEL

Abstract

Much has been written in recent years about the types of factors that should be considered in a dynamic financial analysis model. Much less has been written that actually provides a reader with an understanding of how the various pieces of a dynamic financial analysis model need to fit together. This paper is intended to provide a reader with a look "under the covers" at the structure of a model being used for dynamic financial analysis.

A second and equally important aspect of dynamic financial analysis is the determination of appropriate, or at least reasonable, parameters for different elements within the model's mechanical framework. Unfortunately, those organizations that have been the most active in the development of model parameters are in the uncomfortable position of having to choose between divulging the specifics of their parameterization studies, at the risk of losing a competitive advantage, or keeping the knowledge to themselves, to the long-term detriment of the actuarial profession's ability to effectively use models of this nature. The authors of this paper are no less constrained by our respective organizations. As such, we have largely excluded model parameterization from the subject matter of this paper.

1. INTRODUCTION

There are many facets to the problem of corporate financial model development. It is useful to begin with an analogy to the common actuarial problem of distinguishing between specification, parameter, and process risks. Specification risk relates to the question "Are the model structure and the selected probability distributions correct?" Parameter risk narrows the question to "Assuming the specification is correct, are the distributional parameters correct?" Lastly, process risk is concerned with randomness; i.e., answering the question "Assuming everything else is correct, what can happen in my universe of possible outcomes?"

One might quibble between modeling loss severity with a Weibull distribution instead of a lognormal distribution. Selecting from the universe of possible probability distributions in model design is coping with *specification* risk. In some situations, the specification risk may degenerate into subjective probability assessment—the knowledge set about a dynamic process may be so sparse that a rigorous description of the underlying probability distribution is not possible.[1] Even after this exercise is completed successfully, the analyst still must deal with describing the parameters of the chosen process model. This second stage investigation is the source of *parameter* risk. This risk involves the selection of incorrect parameters, even if the probability distribution is correctly chosen. This leaves only *process* risk to address. Ideally, process risk becomes insignificant under the weight of many, many recalculations of the model.

In financial modeling, there are many of these "risks," and the model designer should not be oblivious to them. The model designer must leap many hurdles while formulating a corporate financial model, particularly one for dynamic financial analysis (DFA). Examples of hurdles to be overcome or pitfalls to be avoided include:

1. The model can use the wrong equations when attempting to define causality or linkages among model constants and variables.

[1]The mathematics describing the fitting of distributions with only sparse knowledge of the underlying risk characteristics is described by Filshtein [4].

2. Important components of the operational or economic environment might be omitted.
3. Elements that should be rendered in a dynamic manner are kept static.
4. Model designers can be consumed by uncertainty regarding the dynamic behavior of those components deemed to be dynamic.
5. The model's accounting framework may be inaccurate.
6. The model could contain programming problems or other embedded divergent behavior.
7. It might not be possible to achieve a consensus among decision makers about the metrics (i.e., output results) of comparison.
8. Model results may not exhibit one clearly preferable alternative among different strategies under investigation.
9. Model results cannot be implemented or only can be implemented with constraints (e.g., the decision path that leads to the "best" long-term outcome is not feasible, either because it violates internal management operating constraints or regulatory boundaries).
10. The model can expand to consider such a wide array of possible situations, interrelationships, and outcomes that it becomes too time-consuming to use in a realistic and useful manner.

In summary, the risks include functional mis-specification of the model, errors in risk and process identification, and failure of the accounting framework to adequately divulge the metrics needed for decision making.

Let us begin with a disclaimer to all readers who hope to find an easy recipe for modeling. There is no panacea for model, functional, or dynamic variable mis-specification. Very often, there is not even a good place to start looking for a definition. With that in mind, we believe that (a) a definition that describes the event in question is better than no definition at all and (b) it is not worth quibbling over the finer points of parameterization—in the overall perspective of what we are trying to model, the error introduced by using a Weibull instead of a lognormal distribution to fit empirical claims severity data is not going to make or break our results.

We now turn to the three key concepts that form the basis for this paper:

- The model to be discussed is a *corporate financial model*; one that already has been deployed in the marketplace.
- The model is *stochastic*, with the capability of being made *dynamic.*
- There is often no clearly preferable solution among alternative decision paths.

2. KEY CONCEPTS

Corporate Financial Model

Day-to-day operations of a property/casualty insurance company include buying and selling assets, underwriting insurance policies, collecting premiums, administering claims, and running the insurance enterprise. A financial model of a property/casualty insurance enterprise needs to be able to model each of these operations separately and in conjunction with each other in order to produce realistic financial projections for the entity.

In order to perform a comprehensive dynamic financial analysis, a corporate financial model should have linkages and interrelationships between activity on the asset and liability sides of the business. For example, the model should:

- apply the same macroeconomic environmental conditions (e.g., interest rates, inflation rates, catastrophic events) across *all* operations of the company;

- allow investment decisions to be made after consideration of both operating needs and investment opportunities in the financial markets;
- look at the risk/return tradeoffs generated by both investment and operating decisions in the context of the entire company's risk/return spectrum rather than in isolation; and
- provide a universal set of metrics or decision criteria by which diverse company operations can be measured and managed.

These critical model components are couched in terms of one or more accounting frameworks (i.e., statutory, GAAP, or economic). The accounting mechanisms serve to organize the model's projected results into a readily understood and consistent set of outputs.

Stochastic vs. Static Corporate Financial Modeling

One of the purposes of a corporate financial model is to help company management understand how decisions made today can be expected to affect the company's financial well-being tomorrow. Traditionally, corporate financial modeling has relied on static evaluations of current and future events and predetermined cause and effect relationships. Static methods of analysis limit the ability to analyze the sensitivity of outputs to changes in input variables, especially if the number of input variables is large and the interrelationships among them are complex. Yet it is critical that strategic decisions be made with the understanding of how each decision impacts the following ones or how changes in the internal or external environment can alter the anticipated outcomes arising from each decision.

The essence of stochastic modeling is the ability to describe critical assumptions in terms of ranges of possible outcomes, rather than in terms of fixed values. Once each critical assumption is defined by a range of possible outcomes and the interrelationships among critical assumptions are mapped out, a series of model recalculations can be performed to obtain ranges of results that we can reasonably expect to see. The parameters used to model stochastic variables and the accounting interrelationships ultimately define the key criteria or metric variables that are of interest to management, regulators, and stockholders. Differences in financial results arising from alternative strategic decisions can be evaluated by replacing one set of strategic decisions with another, re-running the modeling exercise, and comparing the ranges of possible outcomes under each decision rule set.

A stochastic model should also be able to address *dynamic* modeling considerations. A dynamic modeling consideration is one that responds in a time-dependent manner to other events that are unfolding or have unfolded at an earlier point in the modeling environment. Dynamic modeling considerations might be as simple as adjusting the price adequacy of the premium in a line of business if previous years' loss ratios are higher than expected, or as complex as adjusting the mix of taxable and tax-exempt bonds in an investment portfolio in order to minimize tax payments. While dynamic modeling considerations of this nature are not discussed at great length in the remainder of the paper due to the individuality of their construction and application, the ability to implement such decision logic later is an important consideration in the construction of a dynamic financial analysis model.

Choosing Between Competing Strategic Decision Paths

Very often, a company is faced with deciding between two or more strategic options. Under some situations, one option may be clearly superior, while under other situations a different option is preferable. An evaluation of multiple alternative strategies understand their relative risk/reward tradeoffs provides the information needed to answer questions such as "What additional risks must I assume to achieve a higher long-run return on my investment?"

FIGURE 1

COMPARISON OF AFTER-TAX PORTFOLIO YIELD WITH CAPITAL GAINS

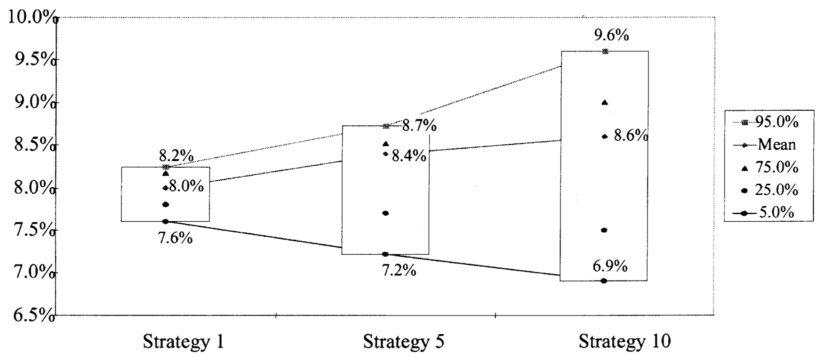


TABLE 1

ASSET COMPOSITION OF STRATEGIES IN FIGURE 1

Strategy 1	Strategy 5	Strategy 10
54% taxable bonds	80% taxable bonds	65% taxable bonds
31% tax-exempt bonds	0% tax-exempt bonds	0% tax-exempt bonds
0% stocks	15% stocks	30% stocks
4% cash	2% cash	3% cash
11% other	3% other	2% other

or "What is the probability of an important financial goal exceeding its expected value?" The answer to these questions should be the core DFA support for management decisions.

As an example, Figure 1 compares the three asset allocation strategies whose asset compositions are displayed in Table 1. The measurement criterion is the internal rate of return on the change in book value of all invested assets over the five-year projection horizon plus investment income, plus realized and unrealized capital gains, less the difference between the market value of assets maturing and sold and those purchased during the five years. Table 2 provides a numerical example of the metric displayed in Figure 1.

While there is no one alternative that is clearly superior, the picture illustrates that, in this case study, higher return is only achieved at the price of higher risk. The ultimate choice is a business decision; there is no alternative in this decision set that is superior to the others in all cases. This finding may seem to be a bane of dynamic financial analysis—there is no mechanically driven choice within a loosely defined utility framework. However, it points out the reality underlying strategic business decisions—it is not very often that one strategic direction is clearly superior to all others.

3. MODEL STRUCTURE OVERVIEW

The corporate financial model has been developed to include a minimum of one year of actual results and to produce pro-forma financial projections for the subsequent five years. For the purposes of simplification throughout the remainder of the article, it is assumed that the actual results are valued as of December 31, 1996 and that the projection period encompasses the years 1997 to 2001.

The model includes five separate and distinct components that must interact with each other in a structured and sequential manner. The components include

- an economic scenario generator,
- a projector of underwriting cash flows and accounting accruals,
- a projector of investment returns and asset valuations,
- a tax calculator, and
- a financial statement structure.

TABLE 2

NUMERICAL EXAMPLE OF A "PORTFOLIO YIELD WITH CAPITAL GAINS" CALCULATION

		1996	1997	1998	1999	2000	2001
(1)	**Investment Income**		40,000	38,000	44,000	43,000	45,000
(2)	**Realized capital gains**		2,000	−1,000	1,000	8,000	−3,000
(3)	**Unrealized capital gains**		10,000	−4,000	−10,000	26,000	−1,000
(4)	**Assets maturing or sold**		100,000	110,000	120,000	130,000	140,000
(5)	**Assets purchased**		120,000	135,000	150,000	165,000	180,000
(6) = (4) − (5)	**Net Sales**		−20,000	−25,000	−30,000	−35,000	−40,000
(7)	**Book Value**	−500,000					650,000
(8) = (1) + (2)+ (3) + (6) + (7)	**Cash Flows**	−500,000	32,000	8,000	5,000	42,000	651,000
	Internal rate of return (portfolio yield with capital gains)	8.66%					

FIGURE 2

CORPORATE MODEL STRUCTURE

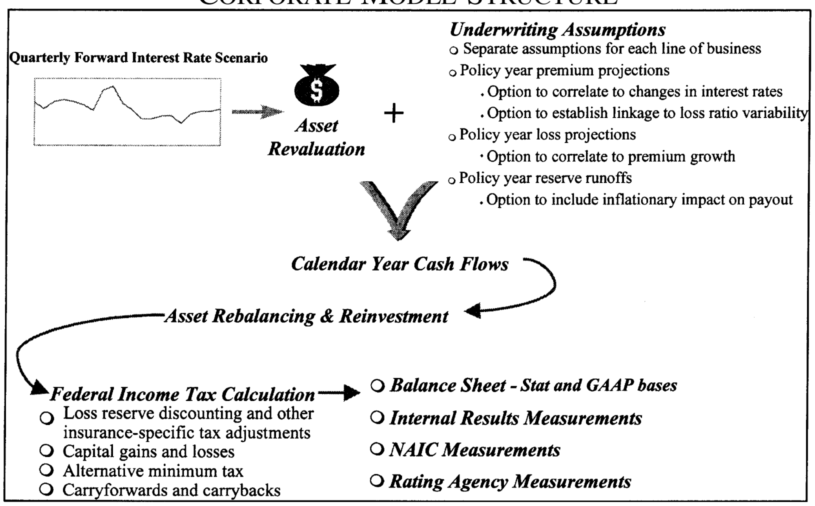


Figure 2 displays a flowchart of the period-by-period interactions of the five model sections.

The modeling starts with initial conditions—the beginning balance sheet, including accident year modeling of liabilities, knowledge of accruals, tax carry-backs and carry-forwards, costs and valuations of assets, and so forth. The following sequence of steps is replicated for each time period over which the model projects financial results:

1. Stochastically generate an economic scenario (interest rates, inflation, competitive conditions, etc.) for the next period.

2. Develop underwriting projections without consideration of the economic scenario (e.g., correlated, random effects on loss volume or severity that are independent of economic effects).

3. Overlay the economic scenario on top of the underwriting projections; quantify the effects of the economic scenario on the underwriting projections.
4. Apply the economic scenario to value existing assets.
5. Apply an asset rebalancing strategy based on current liability and asset conditions or on functions of previously observed or future expected ones.
6. Rebalance the portfolio of assets (and/or liabilities); i.e. buy and sell assets as needed.
7. Develop taxation effects and other fiscal period closing entries.
8. Tally assets and liabilities under the appropriate accounting scheme(s).
9. Create end-of-period financials, operating statistics, and metrics.

4. ECONOMIC SCENARIO

There is much literature describing models for the projection of economic scenarios. In fact, this may be the most well-documented of all DFA model parameters. The economic scenario model used in conjunction with the corporate financial model being discussed in this paper is of the family of "one-factor" interest rate projection models. It is closely based on the first of two interest rate generation algorithms described in a paper by James Tilley [8]. It is a one-factor lognormal model that reverts interest rates to short-term expectations. In other words, projected interest rates have a tendency to move from an initial seeding (e.g., the actual December 31, 1996 interest rate level) to an equilibrium that represents historic interest rate expectations in the short-term spectrum of the yield curve. Long-term interest rates, inflation rates, and projected movements in equity markets are produced by algorithms that relate each of these economic variables to short-term interest rates. One example of such a set of algorithms can be found in a paper written by Gary Venter in 1996 [9].

The economic scenario generation model is independent of the corporate financial model. Actions taken in the corporate financial model do not affect the future interest rate environments projected by the economic scenario generation model, but the future interest rate environments do impact the corporate financial model.

5. UNDERWRITING SECTION

The underwriting section performs seven basic tasks:

1. It converts held loss and allocated loss adjustment expense (ALAE) reserves into calendar year payouts.
2. It converts indicated redundancies or deficiencies in held loss reserves into calendar year payouts and captures the accounting impacts of reserve redundancy or deficiency emergence. Reserve redundancies or deficiencies can arise either from variability in the held reserves (i.e., the held reserves represent the best estimate of ultimate losses, but actual loss emergence might vary in some range around the best estimate) or from deliberately holding reserves at a level other than the best estimate.
3. It calculates the inflationary impact on loss payments arising from differences between a simulated future level of inflation and a level of inflation that was implicitly (or explicitly) assumed when the held reserve level was established.
4. It allows the emergence of reserve redundancies or deficiencies into the model's accounting results to be scheduled at the same rate or faster than the redundancies or deficiencies emerge into the model's cash flows.

5. It calculates any additional premium inflows that might be derived from policies already written (i.e., audit premium, premium from retrospectively rated policies) and earns premiums on in-force and new business according to a user-defined premium earning pattern.

6. It calculates discounted loss reserve levels for federal income tax calculations.

7. It provides the vehicle for entering a five year underwriting plan, including future premium inflows and associated loss and variable expense outflows at a line of business level of detail. (Only variable expenses are included in the line of business section. Fixed expenses are addressed in a different section of the model.)

Interrelationship of Held Reserves, Indicated Reserves, Payout Patterns, and Inflation on Income Statement and Cash Flow Projections

When the model was developed, it was decided that four factors needed to be considered in the loss reserve runoff structure

- the adequacy of held reserves,
- the speed with which reserves are paid out,
- the effects of unanticipated inflation on loss payments, and
- the way company management chooses to recognize adverse or favorable loss emergence through reserve additions or reserve takedowns.

The modeling structure allows each of the four pieces to be examined and modified separately for each line of business. This enables the model user to address each element with separate assumptions.

Conversion of Held Loss and ALAE Reserves Into Calendar Year Payouts

Predetermined loss and ALAE payout patterns are applied to *held* loss reserves and *held* ALAE reserves, respectively, to yield the calendar year payments that would be made, assuming the held reserves are correct. Each accident year's reserves are paid out over successive calendar years in accordance with the incremental calendar year payout percentages. The model should be flexible enough to allow for the application of different payout rates to different accident years, but in most situations only one underlying payout pattern will be needed.

An example of the conversion of held loss reserves at time T_0 (December 31, 1996) into payouts over the successive four years might look like the pattern shown in Table 3.

The calculation of the incremental calendar year payment is equal to:

$$\text{Held Reserve at Time } T_0 * \frac{\text{Incremental Payout Percentage at Time } T_i}{\sum_{i=1}^{n} \text{Incremental Payout Percentage at Time } T_i}.$$

In this example, the remaining incremental payout percentages for accident year 1995 are 20% at time T_1, 15% at time T_2, and 10% at time T_3. The calendar year 1997 payment amount equals the product of the \$8,000 T_0 held reserve and the 20% T_1 incremental payment percentage divided by the sum of the T_1 to T_3 payment percentages, or $\$8{,}000 * 20\% \div (20\% + 15\% + 10\%) = \$3{,}556$.

Treatment of Reserve Adequacy in the Model

The model addresses the relative level of reserve adequacy at the accident year level of detail within a line of business. It captures held reserves and their payout in one payout triangle,

TABLE 3

PAYOUT TRIANGLE

Payout Pattern						
		0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %		30%	25%	20%	15%	10%
Cumulative %		30%	55%	75%	90%	100%

Calendar Year Reserve Payouts						
Accident Year	Held Reserve	Calendar Year 1997	Calendar Year 1998	Calendar Year 1999	Calendar Year 2000	Calendar Year 2001
1993	2,000	2,000	0	0	0	0
1994	5,000	3,000	2,000	0	0	0
1995	8,000	3,556	2,667	1,778	0	0
1996	10,000	3,571	2,857	2,143	1,429	0
Calendar Year Total	25,000	12,127	7,524	3,921	1,429	0

as described in Table 3, and it captures any reserve redundancies or deficiencies and their payout in a second payout triangle, identical in format to the held reserve payout triangle.

The model assumes that the payout pattern being applied to held reserves is a correct representation of the rate at which exactly adequate reserves will become paid. However, held reserves are often not exactly adequate. If no adjustments for reserve inadequacies/redundancies are made, the resulting paid loss projections will understate/overstate actual future paid loss amounts.

The model addresses this by assuming that the sum of the held reserves and the indicated redundancy or deficiency amounts are equal to exactly adequate reserves. The model applies the same payout pattern to the indicated reserve deficiencies or redundancies (by accident year) as is used on the held reserves.[2] This ap-

[2]The assumption here is that the inadequacy or redundancy of held reserves is evenly spread across all outstanding claims. In reality, there are more likely to be differences in the relative adequacy levels of claims according to the length of time remaining until

proach projects incremental future paid loss amounts attributable to the difference between the level at which reserves should be held and what is being held. If held reserves are understated, the projections are positive (i.e., calendar year payments will be greater than held reserves would otherwise indicate), and, conversely, if held reserves are overstated, the projections are negative. The sum of the paid loss projections derived from the held reserve triangle and paid loss projections derived from the reserve redundancy/deficiency triangle equal the *correct* future payout amounts. Intuitively, this makes sense if one considers that, all other things being equal, there should be no impact on actual losses paid whether or not the held reserves at time zero are exactly equal to the future payment amounts.

Table 4 uses the payout pattern from Table 3 with modifications to the held reserves at time T_0. We can see how the assumption of equivalent payout patterns for held reserves and reserve redundancies/deficiencies results in the same calendar year payouts as if held reserves were equal to needed reserves.

Impact of Changes in Reserve Payout Speed

The model is structured so that changes in the speed with which reserves are paid out do not change the total amount to be paid out, only the timing with which the reserves are paid out. If a situation arises in which both the amount and timing of reserve payouts are impacted, the amount component would be captured through the reserve redundancy/deficiency triangle and the timing component would be captured through a shift in the reserve payout pattern.

The impact of a change in the payout pattern is therefore described as an "accordion effect" on calendar year underwriting cash flows. Any payout pattern change has the effect of stretching or compressing the cash flow pattern without changing the total

---

settlement, with those that are furthest from settlement being less adequately reserved than those that are closer to settlement.

amount to be paid out.[3] Table 5 uses an example to illustrate this phenomenon.

TABLE 4

CALENDAR YEAR RESERVE PAYOUTS

Accident Year	Held Reserve	Calendar Year 1997	Calendar Year 1998	Calendar Year 1999	Calendar Year 2000	Calendar Year 2001
1993	2,000	2,000	0	0	0	0
1994	4,000	2,400	1,600	0	0	0
1995	6,000	2,667	2,000	1,333	0	0
1996	8,000	2,857	2,286	1,714	1,143	0
Calendar Year Total	20,000	9,924	5,886	3,047	1,143	0

Accident Year	Reserve Deficiency	Calendar Year 1997	Calendar Year 1998	Calendar Year 1999	Calendar Year 2000	Calendar Year 2001
1993	0	0	0	0	0	0
1994	1,000	600	400	0	0	0
1995	2,000	889	667	444	0	0
1996	2,000	714	571	429	286	0
Calendar Year Total	5,000	2,203	1,638	873	286	0

Accident Year	Overall Total	Calendar Year 1997	Calendar Year 1998	Calendar Year 1999	Calendar Year 2000	Calendar Year 2001
1993	2,000	2,000	0	0	0	0
1994	5,000	3,000	2,000	0	0	0
1995	8,000	3,556	2,667	1,778	0	0
1996	10,000	3,571	2,857	2,143	1,429	0
Calendar Year Total	25,000	12,127	7,524	3,921	1,429	0

[3]Because the model assumes payout pattern variability might be induced through the generation of incremental payout period adjustment amounts, the payout pattern after adjustment might not sum to 100%. This is accounted for in the model through formulas that rescale the adjusted payout pattern to total 100%.

TABLE 5

IMPACT OF CHANGES IN PAYOUT SPEED

Original Payout Pattern

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	30%	25%	20%	15%	10%
Cumulative %	30%	55%	75%	90%	100%

Payout Pattern Adjustment Amounts

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	+10%	+10%	no change	−5%	−5%

Revised Payout Pattern, Prior To Rescaling

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	40%	35%	20%	10%	5%
Cumulative %	40%	75%	95%	105%	110%

Revised Payout Pattern, After Rescaling

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	36.4%	31.8%	18.2%	9.1%	4.5%
Cumulative %	36.4%	68.2%	86.4%	95.5%	100%

Revised Calendar Year Reserve Payouts

Accident Year	Held Reserve	Calendar Year 1997	Calendar Year 1998	Calendar Year 1999	Calendar Year 2000	Calendar Year 2001
1993	2,000	2,000	0	0	0	0
1994	5,000	3,333	1,667	0	0	0
1995	8,000	4,571	2,286	1,143	0	0
1996	10,000	5,000	2,857	1,429	714	0
Calendar Year Total	25,000	14,904	6,810	2,572	714	0

Table 5, based on the payout pattern used in Table 3, demonstrates the effects of changing the speed of the payout pattern.

Effects of Unanticipated Inflation on Loss Payments

By including the inflationary impacts on loss payouts, we incorporate a linkage between the macroeconomic environment affecting assets and the macroeconomic environment affecting losses. We are also in a position to examine the financial statement implications of unanticipated inflationary pressures on losses.

In addition to the speed of payment and inaccuracy in original reserve estimates, a third variable affecting losses in the model is inflation that affects claim costs through the claim payment date. It is assumed that the reserves quantified in the first two reserve triangles (the held reserve payout triangle and the reserve redundancy/deficiency payout triangle) implicitly or explicitly contemplate an *anticipated* level of inflation. Not contemplated in these two payment triangles are the differences between actual and anticipated payments arising from changes in inflation.[4]

As an example, suppose that if a claim were paid today, it would cost \$1,000. But, with an annual inflation rate for claims of this nature at five percent per annum, if the claims are paid five years later, then it will cost \$1,276 [$= 1{,}000 * (1.05)^5$]. The \$1,276 reserve is included in the reserves quantified in the held reserve and the reserve redundancy/deficiency payout triangle. Suppose now that the actual inflation rate increases to ten percent per annum in the third, fourth and fifth years. Now, the amount paid will be \$1,467, not \$1,276. The additional \$191 paid in year five is captured in the paid loss amounts for that year and also

[4]Robert Butsic, in his 1981 paper [2, pp. 58–102], describes two ways in which inflation can impact losses. One is through a claim's accident date and the second is through a claim's payment date. For the model described in this paper, the authors elected to contemplate and quantify the second way only; i.e., inflation impacting claims through the claim payment date.

shows up as adverse loss experience in the company's income statement in the fifth year.

Accounting for Reserve Additions or Takedowns

The first three components of the model's loss reserve runoff structure have been concerned with converting reserves into cash flows. The fourth component is concerned with the quantification of the accounting implications of reserve adjustments. It is not enough to know when a reserve redundancy or deficiency is converted into a cash event; the model must also know when a reserve redundancy or deficiency is recognized in the financial statements, either as an increase to held reserves or a decrease of held reserves. It should be noted that the model makes no provision for the recognition of unanticipated changes in loss payments arising from a change in the inflationary environment.

Returning to the numbers from Table 4, we examine a few different ways in which a reserve deficiency might manifest itself in a company's income statement. The reserve deficiency amounts and their calendar year payouts as shown in Table 4 were as follows:

Accident Year	Reserve Deficiency	Calendar Year 1997	Calendar Year 1998	Calendar Year 1999	Calendar Year 2000	Calendar Year 2001
1993	0	0	0	0	0	0
1994	1,000	600	400	0	0	0
1995	2,000	889	667	444	0	0
1996	2,000	714	571	429	286	0
Calendar Year Total	5,000	2,203	1,638	873	286	0

One approach to the income statement recognition of the reserve deficiency might be to recognize it as the deficiency emerges in loss payments. This approach would have no effect on the level of held reserves that appear on the company's balance sheet, but there would be an impact on the company's income statement in calendar years 1997 through 2000. The impact on the company's income statement in calendar year 1997 would be $2,203, the impact in calendar year 1998 would be $1,638, and so on.

A second approach might be to recognize the entire reserve deficiency in calendar year 1997. The 1997 financial statements would show a reserve increase of $2,797 over what would otherwise have been held at year end 1997, and the 1997 income statement would show incurred losses to be $5,000 higher than they would otherwise have been.

Suppose a company had one year of loss reserves on its books with the following additional information:

- Held reserves: $100,000, based on a reserve range of $80,000 to $105,000.
- Actual reserve need: $90,000.
- Reserve payout pattern: 25% over each of the next four years.
- Inflation level implicit in held reserves: 5%.
- Actual annual inflation rates: 5% in years 1 and 2, 8% in years 3 and 4.
- Company takes reserve redundancy into financial statements equally in years one and two by lowering reserves $5,000 in each year.

An example of the inter-relationship between the four elements (payout of held reserves, indicated reserve redundancy/deficiency emergence through payments, inflationary impacts on payments, and indicated reserve redundancy/deficiency emergence through the financial statements) is shown in Table 6 using the assumptions above.

TABLE 6

IMPACT OF RECOGNITION OF RESERVE REDUNDANCY/DEFICIENCY

	Year 0	Year 1	Year 2	Year 3	Year 4
1. Payout percentage		25%	25%	25%	25%
2. Expected inflation rate		5%	5%	5%	5%
3. Actual inflation rate		5%	5%	8%	8%
4. Held reserve cash flow	n/a	25,000	25,000	25,000	25,000
5. Redundancy cashflow	n/a	−2,500	−2,500	−2,500	−2,500
6. Inflationary impact	n/a	0	0	643	1,304
7. Reserve lowering	n/a	−5,000	−5,000	0	0
8. Held reserves	100,000	72,500	45,000	22,500	0
9. Net cash flow		22,500	22,500	23,143	23,804
10. Income statement impact (−gain/ + loss)		−5,000	−5,000	+643	+1,304

Line 6 formula:

$$\left[\left(\frac{\prod_{i=1}^{N}(1+\text{Line }3_{\text{Year }i})}{\prod_{i=1}^{N}(1+\text{Line }2_{\text{Year }i})}\right)-1\right]*(\text{Line }4_{\text{Year }i}+\text{Line }5_{\text{Year }i})$$

Example:

Year 3 inflationary impact

$$=\left[\frac{(1.05)(1.05)(1.08)}{(1.05)(1.05)(1.05)}-1\right]*(25{,}000-2{,}500)$$

Line 8 formula: $\text{Line }8_{\text{Year }i-1}-\text{Line }4_{\text{Year }i}+\text{Line }7_{\text{Year }i}-\text{Line }5_{\text{Year }i}$
Line 9 formula: $\text{Line }4_{\text{Year }i}+\text{Line }5_{\text{Year }i}+\text{Line }6_{\text{Year }i}$
Line 10 formula: $\text{Line }6_{\text{Year }i}+\text{Line }7_{\text{Year }i}$

C. K. Khury develops this concept in a similar manner [6, p. 14]. Khury notes that, given a balance sheet loss reserve liability at time T, "the actual experience corresponding to this estimate can generate two effects on the financial results of an insurer: (a) the effect of the difference between expected and actual claim payments, i.e. actual development and (b) the effect of any restatement of the remaining unpaid claim liability arising from changes in the underlying assumptions, i.e. change in expected development." Khury's first effect is comparable to the combined impacts of the redundancy cashflow and inflationary impacts on Lines 5 and 6 in Table 6. His second effect is comparable to the reserve lowering on Line 7 in Table 6.

Premiums: the Premium Writing, Earning and Collection Processes

Premiums can be earned on two types of business: those already written and those that will be written during the model's time horizon. For both, the premium earning process is identical. Premiums are earned according to a predefined earnings pattern that can be as short as one year or as long as twenty years. This earnings pattern is applied to *initial* policy year written premium levels, not ultimate policy year written premium[5] levels. Similarly, premiums can be collected for business already written and business that will be written during the model's time horizon. A second premium pattern, this one for premium collections, is applied to initial policy year written premium levels to determine when premiums are collected.

The model makes three simplifying assumptions:

1. All policies are annual policies.
2. The amount of premium collected in the calendar year in which the premium is written equals the initial written premium.[6]

---

[5]Differences between calendar year written premium and ultimate policy year written premium might arise because of premium audits and/or retrospective premium adjustments.

[6]The simplifying assumption that the amount of premium collected equals written premium in the calendar year in which the premiums were written ignores the existence of "cash-flow" premium collection arrangements. In reality, a premium collection arrangement could exist in which premiums are booked in one calendar year but are not fully collected until one or more years in the future. This would be most likely to occur on long-tail policies of large commercial accounts, such as workers compensation. The rules described in this paper do not work for these situations and would need to be modified to fit the actual premium collection/premium booking structure of the entity being modeled. One option is to leave the modeling of premiums unchanged at the line of business level,

3. If a line of business does not calculate a reserve for anticipated rate credits and retrospective adjustments in advance of actually collecting such adjustments, there is no change in the collected premium in the calendar year after the premium is written—any adjustments to collected premium occur in the third and subsequent calendar years.

The interrelationships between the writing, earning, and collection of premiums are explained by the following set of rules. Each of these rules are applied on a policy year by policy year basis:

- Written premium: In the first calendar year, written premium is set equal to the user-input initial written premium amount. If the line of business being modeled includes a provision for anticipated rate credits and retrospective reserve adjustments, then written premium is assumed to change by the calendar year change in collected premiums in the second and subsequent calendar years. If the line of business being modeled does not include a provision for anticipated rate credits and retrospective reserve adjustments, then written premium is assumed to change by the calendar year change in collected premiums in the third and subsequent calendar years.
- Earned premium: Premiums are earned by applying the user-input premium earning pattern to the initial written premium amount.
- Unearned premium: Unearned premium is calculated as:

$$\text{Written Premium} - \text{Earned Premium} \\ + \text{Prior Calendar Period Unearned Premium Reserve.}$$

---

but adjust the accounting portion of the model to reflect the impact of the delayed premium collection on the company as a whole. The accounting entries that would require “overrides” to accomplish this include the balance sheet entry or entries for uncollected premiums and the cash flow statement entry for total collected premium.

This develops an unearned premium reserve that is consistent with the one on the Underwriting and Expense Exhibit, Part 2A of the statutory Annual Statement (page 8 in the 1997 statutory Annual Statement), not one that is consistent with the unearned premium reserve displayed on the Liabilities, Surplus and Other Funds page (page 3) of the statutory Annual Statement. In the statutory Annual Statement, the difference between these two numbers is the reserve for rate credits and retrospective adjustments.

- Reserve for rate credits and retrospective adjustments, if applicable: In the first calendar year, this reserve is equal to:

  $$(\text{Ultimate earned premium} - \text{initial written premium}) * \left(\frac{\text{\% of initial written premium that is earned in CY 1}}{\text{ultimate earned premium/initial written premium}}\right).$$

  In the second and subsequent calendar years, this reserve is equal to the difference between the ultimate earned premium and the premium collected to date.

- Collected premium: The timing and amount of premium collections are calculated by applying a user-input premium collection pattern to the initial written premium amount.

- Uncollected premium: If the line of business being modeled does not include a provision for anticipated rate credits and retrospective reserve adjustments, then this is calculated as the difference between initial written premium and the premium collected to date. If the line of business being modeled includes a provision for anticipated rate credits and retrospective reserve adjustments, then the uncollected premium is set equal to the calculated rate credit or retrospective reserve adjustment.

Assume:

- Ultimate earned premium = initial written premium.

- Premium is collected as it is written.
- The line of business being modeled does not include a provision for anticipated rate credits and retrospective reserve adjustments.

The results of these assumptions are shown in Table 7.

Assume:

- Ultimate earned premium = 110% of initial written premium.
- No reserve for rate credits is used; i.e., the additional premium is written and earned when it is collected.

In the extended premium earning pattern, the ultimate policy year premium is greater than the initial written policy year premium, possibly due to the receipt of audit premium in the third calendar year after the start of the policy period. The extended premium earning and collection patterns account for this by totaling to 110% instead of 100%. The results are shown in Table 8.

Assume:

- Ultimate earned premium = 110% of initial written premium.
- A reserve for rate credits is used; i.e., the additional premium is earned at the same time the initial written premium is earned.

The results are shown in Table 9.

New Business Production

The model's new business production logic requires assumptions about future premium writing levels and associated loss and variable expense ratios. There is no one structure for this section that is suitable for every need, but certain capabilities and functional considerations can be generalized. These are discussed in the subsequent paragraphs.

TABLE 7

Simple Earning Pattern

Premium Earning Pattern, Applied to Initial Policy Year Written Premium

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	50%	50%	0%	0%	0%
Cumulative %	50%	100%	100%	100%	100%

Calendar Year Earning of Policy Year Premium Writings

Policy Year	Initial Written Premium	CY 1996 Earned Premium	CY 1997 Earned Premium	CY 1998 Earned Premium	CY 1999 Earned Premium	CY 2000 Earned Premium
1996	20,000	10,000	10,000	0	0	0
1997	25,000	n/a	12,500	12,500	0	0
1998	30,000	n/a	n/a	15,000	15,000	0
Calendar Year Total	75,000	10,000	22,500	27,500	15,000	0

Calendar Year Premium Collection

Policy Year	Initial Written Premium	CY 1996 Collected Premium	CY 1997 Collected Premium	CY 1998 Collected Premium	CY 1999 Collected Premium	CY 2000 Collected Premium
1996	20,000	20,000	0	0	0	0
1997	25,000	n/a	25,000	0	0	0
1998	30,000	n/a	n/a	30,000	0	0
Calendar Year Total	75,000	20,000	25,000	30,000	0	0

Calendar Year Accounting Results

Cal. Year	Written Premium	Earned Premium	Unearned Premium (AS p. 8)	Reserve for Rate Credits	Unearned Premium (AS p. 3)	Collected Premium	Uncollected Premium
1996	20,000	10,000	10,000	0	10,000	20,000	0
1997	25,000	22,500	12,500	0	12,500	25,000	0
1998	30,000	27,500	15,000	0	15,000	30,000	0
1999	0	15,000	0	0	0	0	0
2000	0	0	0	0	0	0	0

TABLE 8

EXTENDED EARNING PATTERN

Premium Earning Pattern, Applied to Initial Policy Year Written Premium

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	50%	50%	10%	0%	0%
Cumulative %	50%	100%	110%	110%	110%

Premium Collection Pattern, Applied to Initial Policy Year Written Premium

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	100%	0%	10%	0%	0%
Cumulative %	100%	100%	110%	110%	110%

Calendar Year Earning of Policy Year Premium Writings

Policy Year	Initial Written Premium	CY 1996 Earned Premium	CY 1997 Earned Premium	CY 1998 Earned Premium	CY 1999 Earned Premium	CY 2000 Earned Premium
1996	20,000	10,000	10,000	2,000	0	0
1997	25,000	n/a	12,500	12,500	2,500	0
1998	30,000	n/a	n/a	15,000	15,000	3,000
Calendar Year Total	75,000	10,000	22,500	29,500	17,500	3,000

Calendar Year Premium Collection

Policy Year	Initial Written Premium	CY 1996 Collected Premium	CY 1997 Collected Premium	CY 1998 Collected Premium	CY 1999 Collected Premium	CY 2000 Collected Premium
1996	20,000	20,000	0	2,000	0	0
1997	25,000	n/a	25,000	0	2,500	0
1998	30,000	n/a	n/a	30,000	0	3,000
Calendar Year Total	75,000	20,000	25,000	32,000	2,500	3,000

Calendar Year Accounting Results

Cal. Year	Written Premium	Earned Premium	Unearned Premium (AS p. 8)	Reserve for Rate Credits	Unearned Premium (AS p. 3)	Collected Premium	Uncollected Premium
1996	20,000	10,000	10,000	0	10,000	20,000	0
1997	25,000	22,500	12,500	0	12,500	25,000	0
1998*	32,000	29,500	15,000	0	15,000	32,000	0
1999*	2,500	17,500	0	0	0	2,500	0
2000*	3,000	3,000	0	0	0	3,000	0

*Note the change in written and earned premiums in calendar years 1998 through 2000. Both the written and earned amounts are increased by the premium earned three years after the start of the 1996, 1997 and 1998 policy periods.

TABLE 9

EXTENDED EARNING PATTERN WITH RESERVES

Premium Earning Pattern, Applied to Initial Policy Year Written Premium

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	55%	55%	0%	0%	0%
Cumulative %	55%	110%	110%	110%	110%

Premium Collection Pattern, Applied to Initial Policy Year Written Premium

	0 to 12 months	12 to 24 months	24 to 36 months	36 to 48 months	48 to 60 months
Incremental %	100%	0%	10%	0%	0%
Cumulative %	100%	100%	110%	110%	110%

Calendar Year Earning of Policy Year Premium Writings

Policy Year	Initial Written Premium	CY 1996 Earned Premium	CY 1997 Earned Premium	CY 1998 Earned Premium	CY 1999 Earned Premium	CY 2000 Earned Premium
1996	20,000	11,000	11,000	0	0	0
1997	25,000	n/a	13,750	13,750	0	0
1998	30,000	n/a	n/a	16,500	16,500	0
Calendar Year Total	75,000	11,000	24,750	30,250	16,500	0

Calendar Year Premium Collection

Policy Year	Initial Written Premium	CY 1996 Collected Premium	CY 1997 Collected Premium	CY 1998 Collected Premium	CY 1999 Collected Premium	CY 2000 Collected Premium
1996	20,000	20,000	0	2,000	0	0
1997	25,000	n/a	25,000	0	2,500	0
1998	30,000	n/a	n/a	30,000	0	3,000
Calendar Year Total	75,000	20,000	25,000	32,000	2,500	3,000

Calendar Year Accounting Results

Cal. Year	Written Premium	Earned Premium	Unearned Premium (AS p. 8)	Reserve for Rate Credits	Unearned Premium (AS p. 3)	Collected Premium	Uncollected Premium
1996*	20,000	11,000	9,000	1,000	10,000	20,000	1,000
1997*	25,000	24,750	9,250	3,250	12,500	25,000	3,250
1998*	32,000	30,250	11,000	4,000	15,000	32,000	4,000
1999*	2,500	16,500	−3,000	3,000	0	2,500	3,000
2000*	3,000	0	0	0	0	3,000	0

*In this situation, the accounting results have been altered in all years as a result of earning the additional premium over the same calendar periods as the initial premium is earned. The reserve for rate credits captures the amount of additional premium that is anticipated as "earned but not received." The only accounting entry that is not impacted is the unearned premium reserve that would appear on page three of the statutory Annual Statement.

New Business—Premium Volumes

The projection of future premium volumes can be as simple as a fixed set of assumptions or can be as complex as a system of assumptions that interrelate the relative amount of business that is expected to be retained each year, a company's internal growth objectives, the overall insurance market conditions, and company reactions to prior-year underwriting results. In general, it would seem that the more linkages that are established between new business production and other events being played out in the model, the better the model will be. The model then should be more reactive; it should do what the company itself might do when faced with similar circumstances. However, in some cases, the inclusion of additional dynamic elements in these linkages could lead to greater confusion in what the model is doing than is warranted by the additional realism that is gained.

Future Loss Ratios

When the model was being developed, two alternative approaches to developing future loss ratios were contemplated. One was to assume that, all other things being held constant, the loss ratios at time T and $T+1$ could be described as independent values selected at random from one statistical distribution. We call this a "force of loss" approach to loss ratio generation. The second approach assumes that the loss ratio at time $T + 1$ should be equal to the loss ratio at time T plus or minus a volatility parameter. The second approach assumes the loss ratio at time $T + 1$ is more or less dependent upon the loss ratio at time T, depending upon the size and shape of the volatility parameter. We call this approach to loss generation the "incremental volatility" approach.

From a theoretical standpoint, it seems that the force of loss assumption would be more valid for lines whose loss experience can be characterized as more directly attributable to external factors than to internal management decisions, or whose exposure base is highly volatile, or whose loss profile is one of low-frequency, high-severity claims. The incremental volatility approach would seem to be most appropriate for those lines of business that display a stable exposure base with a high retention of insureds from year to year, a tendency towards high-frequency, low-severity losses, and a small exposure to catastrophic loss. Examples of "force of loss" lines might include homeowners (if catastrophes are not explicitly separated from non-catastrophic claims), commercial liability or umbrella. Examples of "incremental volatility" lines might include personal automobile, commercial automobile, or any non-catastrophic portion of property lines.

From a practical standpoint, however, the force of loss approach is much simpler to program. All that is required to successfully implement a force of loss approach is to have a random number generator return values from a distribution that reasonably replicates the desired shape, spread, and mean of the loss ratios being modeled. To successfully implement an incremental volatility approach, formulas must be established that (a) cap the overall upwards or downwards movement to reasonable floor and ceiling values and (b) have "mean-reverting" tendencies (i.e., the incremental volatility in time $T + 1$ will be more likely to move the overall loss ratio towards the long-term mean than away from it), while still returning mean values that are consistent with the expected value.

It should be noted that both the force of loss and the incremental volatility approaches to loss ratio selection describe the loss ratio that would arise if there were no other changes occurring that have an impact on the final loss ratio. Other changes might include premium rate changes, inflationary increases in the premium exposure base, or inflationary impacts on loss costs. When the model is run stochastically, the final projected loss ratio is developed by first randomly sampling from the probability distribution that describes this force of loss, then modifying the random sample to reflect the other changes.

Exhibit 1 provides an example of the interrelationships between premium development and loss ratio development, including inflationary and rate impact influences. A more realistic rate change formula would consider many more parameters than the relationship between one year's actual and expected loss ratios.[7]

6. ASSET MODELING

There are two basic components to modeling assets: valuing assets and rebalancing an asset portfolio through the sale of existing assets and the purchase of new assets. The first component is concerned with determining the book and market value at time $T+1$ of assets the company owned at time T. The second component is concerned with how the asset portfolio owned at time T should be adjusted at time $T+1$, including the manner in which net cash inflows between times T and $T+1$ should be invested.

In order to perform these tasks, a model must be able to:

- quantify at any valuation date the book and market values of assets held at that point in time and
- quantify the amount of cash generated by the insurance operation between the previous asset revaluation and rebalancing and the current asset revaluation and rebalancing.

A model must also contain one or more decision algorithms that tell it what assets to sell at time $T+1$, if asset sales are needed or desired and what assets to purchase at time $T+1$.

Asset Categorization

In developing the model, we elected to evaluate assets at an aggregate level of detail. We feel that this approach is in keeping with the strategic nature of the questions the model is expected

[7]The formula for determining whether or not a rate change occurs, and by how much, is there for example only and is not a realistic rate change formula.

to address. We recognize that there are many situations when the level of aggregation described herein is neither sufficient nor appropriate for the desired analysis. At these times, a more refined, maybe even seriatim, approach to asset analysis may be needed. We leave the task of describing a seriatim approach to asset modeling to other papers and instead turn to a discussion of the manner in which assets have been aggregated in this model.

Bonds

The model assumes that bonds mature in no more than thirty years and that bonds are either taxable or tax-exempt, resulting in sixty possible bond categories.

The starting bond portfolio is sorted by tax status and maturity date into the sixty available bond categorizations. Sixty "proxy" bonds are then created from the underlying bond portfolio. Each proxy bond's market value, statement value, and par value are calculated as the sum of the values of the underlying bonds. The maturity date of each proxy bond is assumed to be equal to the midpoint of the calendar year in which the underlying bonds were to mature. Each proxy bond's coupon rate is a weighted average of the underlying bonds' coupon rates, using the par values for weights. The model assumes that each proxy bond will pay coupons semi-annually.

Suppose at December 31, 1996, the XYZ Company has an asset portfolio with the five bonds shown in Table 10.

Three proxy bonds would be created to summarize this portfolio, as follows:

	Maturity Date	Years to Maturity	Statement Value	Market Value	Par Value	Coupon Rate
Proxy 1	7/15/2000	3–4 years	3,000,000	3,009,000	2,950,000	6.6%
Proxy 2	7/15/2003	6–7 years	5,000,000	5,331,000	5,000,000	7.5%
Proxy 3	7/15/2010	13–14 years	7,000,000	7,608,000	7,000,000	7.5%

TABLE 10

ASSET PORTFOLIO

	Maturity Date	Years to Maturity	Statement Value	Market Value*	Par Value	Coupon Rate
Bond 1	6/15/2000	3–4 years	1,000,000	965,000	950,000	6.5%
Bond 2	9/30/2000	3–4 years	1,500,000	1,540,000	1,500,000	6.8%
Bond 3	12/30/2000	3–4 years	500,000	504,000	500,000	6.2%
Bond 4	7/15/2003	6–7 years	5,000,000	5,331,000	5,000,000	7.5%
Bond 5	1/1/2010	13–14 years	7,000,000	7,608,000	7,000,000	7.5%

*The market values are approximations that assume the bonds have no embedded options, no default risk, and that the "current," or December 31, 1996 risk-free interest rate for new bonds maturing in the year 2000 is 6.0%, for new bonds maturing in the year 2003 is 6.25%, and for new bonds maturing in the year 2010 is 6.5%.

Suppose one year has elapsed. The XYZ company has decided not to purchase any new bonds. The proxy bond portfolio now might look as follows:

	Maturity Date	Years to Maturity	Statement Value*	Market Value†	Par Value	Coupon Rate
Proxy 1	7/15/2000	2–3 years	2,987,500	3,015,000	2,950,000	6.6%
Proxy 2	7/15/2003	5–6 years	5,000,000	5,410,000	5,000,000	7.5%
Proxy 3	7/15/2010	12–13 years	7,000,000	8,019,000	7,000,000	7.5%

*The change in proxy one's statement value reflects the amortization of one year's bond premium. As a simplification, this example assumes that the bond premium will be amortized evenly over remaining time to maturity; i.e., one-fourth of the difference between the December 31, 1996 statement and par values. A more accurate approach would be to calculate the change in the present value of the bond based on the initial interest rate.

†The market values are approximations that assume the bonds have no embedded options, no default risk, and that the now "current," or December 31, 1997 risk-free interest rate for new bonds maturing in the year 2000 is 5.65%, for new bonds maturing in the year 2003 is 5.75%, and for new bonds maturing in the year 2010 is 5.80%.

Now suppose that the same one year has elapsed, but the company decides to purchase a new risk-free bond with a $1,000,000 par value that will mature in 2003. The model assumes that this bond is purchased at cost, so the statement value and the market value are equal to the $1,000,000 par value. The coupon rate for this bond is 5.75%.

The model recalculates Proxy Bond 2 as a weighted average of the old and new bond characteristics, resulting in a revised bond with the following information:

	Maturity Date	Years to Maturity	Statement Value	Market Value	Par Value	Coupon Rate
Proxy 2	7/15/2000	2–3 years	3,987,500	4,015,000	3,950,000	6.38%

Preferred and Common Stocks

Preferred and common stocks are aggregated into two groups, with one proxy equity for each group. The proxy equity reflects the total market value, book value and actual cost of the underlying equities within that group.

Assumptions with regard to the projection of future market values can be varied, but not the basic framework. We believe such a simplification is acceptable in most situations and that only when a company has a large preferred stock portfolio would it be inappropriate.

Within each of the preferred and common stock groups, the model assumes there exists an average dividend rate that can be applied to the proxy equity for that group. While it is theoretically possible that the rates might be the same, it would be more likely that the rate applied to the preferred stock group would be higher than that applied to the common stock group. The model further assumes that any unrealized capital gains or losses within a stock grouping are spread evenly across all of the underlying securities within the grouping. These assumptions are maintained as equities are bought and sold during each asset rebalancing.

Suppose the XYZ company had common stock holdings at December 31, 1996 of:

Statement Value of Proxy Equity	Market Value of Proxy Equity	Dividend Rate for Proxy Equity
1,800,000	2,500,000	2.0%

During 1997, the stock portfolio's market value increased by fifteen percent, to $2,875,000. During the year-end 1997 asset rebalancing, the XYZ Company decides to sell ten percent of its equity portfolio. The result is:

Statement Value of Retained Stocks in Proxy Equity	Market Value of Retained Stocks in Proxy Equity	Dividend Rate on Retained Stocks in Proxy Equity	Realized Capital Gains on Sold Portion of Proxy Equity
1,620,000	2,578,500	2.0%	$107,500

Alternatively, the XYZ Company might have decided to purchase additional equities at year-end 1997. Suppose that instead of selling ten percent of the 12/31/96 proxy equity, the XYZ Company decides to purchase an additional $1,000,000 of stocks. The model assumes the stocks purchased will have the same average dividend rate of the previously existing stock portfolio. The proxy equity is restated to reflect the newly purchased stocks as follows:

	Statement Value of Proxy Equity	Market Value of Proxy Equity	Dividend Rate for Proxy Equity
Prior to new stock purchases	1,800,000	2,875,000	2.0%
New purchases	1,000,000	1,000,000	2.0%
After inclusion of new purchases	2,800,000	3,875,000	2.0%

Real Estate

The model tracks real estate in two categories that are consistent with the NAIC Annual Statement: "properties occupied by the company" and "other properties." Besides the desire to match the model's asset categories as closely as possible to the Annual Statement, real estate is maintained as its own asset category in order to better address the accounting impacts of depreciation and capital improvements. The model's accounting and asset valuation structures are designed to allow annual depreciation of real estate assets to flow through the modeled company's balance sheet and income statement without affecting either calendar year cash flows or projected market values for the real estate. Capital improvements flow through the financial statements as a direct outflow from "cash" into the real estate's statement and market values.

Suppose the XYZ company owned real estate with a statement and market value of $10,000,000 at December 31, 1996. The property has annual depreciation of $500,000. Assume that:

(a) no real estate is bought or sold during 1997,

(b) no capital improvements are made to the property, and

(c) the market value remains unchanged from year-end 1996 to year-end 1997.

The December 31, 1997 balance sheet would show a statement value of $9,500,000 and a market value of $10,000,000. The income statement would reflect $500,000 of real estate expenses incurred during the year, and the cash flow statement would not be impacted at all.

If, instead, $1,000,000 of capital improvements were performed during 1997, the balance sheet would show a statement value of $10,500,000 and a market value of $11,000,000. The income statement would again reflect $500,000 of real estate expenses incurred during the year. The cash flow statement would reflect the conversion of $1,000,000 from "cash" to "real estate" through the accounting entry "cost of real estate acquired."

Short-Term Investments

Short-term investments are aggregated into one group and treated in a manner similar to cash. They are assumed to generate an investment income yield that is commensurate with the three-month risk-free interest rate projected by the economic scenario. The model assumes that the statement and market values of this asset group are identical.

All Other Asset Groups (Mortgage Loans, Collateral Loans, Other Invested Assets)

The model makes no special provisions for any other invested asset group. All assets invested in mortgage loans, collateral loans, and other invested assets are consolidated into one proxy asset for each group. A statement and a market value are entered and the model user can specify the annual investment income return anticipated from each of these proxy groups.

Asset Rebalancing

The model rebalances assets at the end of each calendar year. The amount of money that can be rebalanced at year-end equals the sum of:

- cash flow from operations during the year (premium collected less losses and underwriting expenses paid);
- investment income collected during the year, net of investment expenses paid during the year, including investment income derived from the insurance operation's average cash balances, which are deemed to be invested at the yield for "Cash" until the end of the year;
- the cash value of any bonds maturing during the calendar year; and
- the market value of all other invested assets at the end of the calendar year.

The asset rebalancing strategy is ad hoc; it is a user-defined strategy that defines how much money should be invested in any asset class at year-end. Examples of different asset rebalancing strategies might be: "invest 50% in taxable bonds, 30% in equities, and 20% in cash," or "invest 40% in taxable bonds, 40% in tax-exempt bonds, 15% in equities and 5% in cash." These ad hoc strategies are consistent with active investment management portfolio rebalancing where allocations among asset categories are important.

Depending on the rebalancing strategy, some existing assets may be sold and the proceeds reinvested to produce approximately the asset distribution dictated by the chosen asset strategy. The determination of whether to sell or buy assets in an asset class is based on a comparison of the market value of assets held in that class prior to rebalancing and the desired market value of that class. If the amount being held prior to rebalancing is greater than the desired amount, then some portion of assets in that class are sold. If the amount being held prior to rebalancing is less than the desired amount, then additional assets in that class are purchased.

An example might be as follows:

- Suppose we have $1,000 available for reinvestment.
- We want to invest the $1,000 in 3 asset classes; $500 in Asset Class 1, $300 in Asset Class 2 and $200 in Asset Class 3.
- Prior to rebalancing, we have $500 in Asset Class 1, $500 in Asset Class 2 and $0 in Asset Class 3.

The rebalancing algorithm compares the amount held in each asset class prior to rebalancing to the desired amount and causes the following asset redistributions to occur:

- Asset Class 1: no change. Held prior to rebalancing equals desired amount.
- Asset Class 2: Sell $200. New held amount equals $300.
- Asset Class 3: Buy $200. New held amount equals $200.

The final allocations are subject to modifications attributable to year-end closing transactions, primarily tax payments and the payment of investment expenses. The rebalancing can result in capital gains or losses, which are combined with operating results to determine the federal income tax liability for the year.

Calculation of Future Asset Market Values

One of the most important aspects of the model's asset calculations is determining new market values at time $T + 1$ for the assets held at time T. The new market values must be developed in concert with the projected interest rate environment. This calculation is essential because, as noted in the previous section, the model uses market values in the asset rebalancing algorithm as the basis for determining whether to sell some or all of the existing assets in any asset group or to buy additional assets for any asset group.

Different techniques are employed for valuing different categories of assets.

Calculating the Market Value at Time T $+ 1$ *of Bonds Owned at Time T*

Traditional bond valuation methods are used to calculate the market value at time $T + 1$ of bonds owned at time T. As described earlier, the model retains information about the pertinent characteristics of each proxy bond, namely amount and timing of coupon payments and principal repayment.[8] From this infor-

[8]The future cash flows of bonds held at December 31, 1996 are known because the bonds themselves are known quantities. We know their coupon rate and timing, their maturity date, and their par, book, and market values. This is sufficient information to project future cash flows arising from the December 31, 1996 bond portfolio.

The future cash flows of bonds purchased during 1997–2001 are known because we know (a) the risk-free interest rate environment at the time the bonds are purchased, (b) the risk factor that is added to the risk-free interest rate for each bond category, (c) the time to maturity of the bonds that are purchased, and (d) the total dollar amount of new investments in each bond category. With this information, we can calculate an appropriate coupon rate for each dollar of investment in each bond category. We make a mation, the future cash flows from each bond category can be developed. This future cash flow stream, in conjunction with the model-generated interest rate environment at time $T + 1$, is sufficient information to allow the creation of a market value for the bond category at time $T + 1$. It should be noted that this procedure does not contemplate the calculation of option-dependent market values, which are influenced not only by the interest rate at time $T + 1$, but the likelihood of exercising the bond option(s) at time $T + 2$, $T + 3$, etc.

Calculating Future Market Values for Equities

Future market values for equities are derived by projecting values for calendar year equity rates of return, and multiplying the market value of equities at time T by the projected rate of return during the $T + 1$ calendar period. For example, let us assume the XYZ company has a stock portfolio with a market value of $1,000,000 at time T. By some manner, we project this portfolio will generate a fifteen percent return during the upcoming calendar year. The model will calculate the market value of the stock portfolio at time $T + 1$ to be $1,150,000.

The more interesting aspect of this calculation is the way in which the portfolio's rate of return is developed. In some economic scenario generators, this process is embedded within the generator itself, so that the projected economic scenario "automatically" contains projected equity index returns that are correlated with interest and inflation rates. In our experience, the mathematics underlying this type of equity projection methodology tends to be proprietary to the entities that have developed the economic models.

In the absence of such an economic scenario generator, a few alternatives exist for projecting future equity returns. One is to base the rate of return on equities on a normally distributed ran-

---

simplifying assumption that new bonds are purchased at par, so the new bonds' market values at the time of purchase equal their book, par and statement values. We now have sufficient information to project future cash flows arising from new bond purchases.

dom variable with a mean market return and standard deviation based on investor expectations. This alternative uncouples equity pricing from changes in interest rates and inflation and is a conventional random walk model. A second alternative is to postulate a relationship between equity returns and interest and inflation rates so that future equity returns can be related to future projections of interest rates and inflation rates. As noted earlier, an example of a postulated relationship between interest rates and equity projections that also attempts to incorporate a time-dependant element is described in Gary Venter's paper [9].

Another method of relating equity returns to the projected interest rate environment is through the use of the Capital Asset Pricing Model (CAPM). Recall that the CAPM formula is $R = R_f + \beta(R_m - R_f)$, where

R = the expected return on a given stock,

R_f = the risk-free interest rate, such as the rate on Treasury bills,

R_m = the overall expected market return, and

β quantifies the undiversifiable or systematic risk associated with the stock (or stock portfolio) in question.

$(R_m - R_f)$ can also be thought of as the market risk premium, or the amount by which the return on stocks is expected to exceed the risk-free rate.

Under this approach, changes in the risk-free rate of return lead directly to changes in the projected equity return. The magnitude of the change felt by the company is driven by the volatility of the company's stock portfolio (β) in relation to the movement in the index portfolio, R_m. Unlike the Venter algorithm, no attempt is being made here to include a time-dependency.

An example of the relative differences in projected returns is shown in Table 11. The example assumes the risk-free rate of return at time T is 6% and that the expected return of the stock market as a whole, R_m, is 15%.

TABLE 11

PROJECTED EQUITY RETURN

	Risk Free Rates		
Betas	6%	8%	4%
1.00	15.00%	15.00%	15.00%
1.50	19.50%	18.50%	20.50%
0.50	10.50%	11.50%	9.50%

When $\beta = 1$, the projected equity return is identical to the projected return that would be achieved by basing the rate of return on a normally distributed random variable with a mean market return and standard deviation based on investor expectations.

It is worth noting that when using the CAPM equation, and assuming a value greater than zero for β, an inverse relationship is developed between changes in interest rates and equity returns. Other authors have postulated the appropriateness of such a relationship.[9]

Revaluing All Other Assets

As was noted earlier, the model makes no special provisions for any other invested asset group other than bonds and stocks. Short-term investments owned at time T are assumed to mature before time $T+1$, and as such are valued as cash at time $T+1$. The market value of real estate at time $T+1$ is changed only if it is specified that capital improvements were made to the property during the $T+1$st time period. For all other invested assets, it is up to the model user to specify when and how the market value of each asset class will change from time T to time $T+1$.

Amortization of Bond Original Issue Discount

Because new bond purchases are made at par, it is assumed that only the starting bond portfolio can have a difference be-

[9] See Becker [1], Feldblum [3], or Hodes et al. [5].

tween par value and amortized cost. The reader might recall that the starting bond portfolio is summarized into a set of proxy bonds, one for each maturity grouping. The difference between par value and amortized cost is calculated separately for each proxy bond. This difference is assumed to be "original issue discount," deriving from bond purchases at either a premium or a discount. The original issue discount is amortized over the proxy bond's remaining time to maturity.

7. TAX ALGORITHMS

Some financial models, rather than addressing the complexities of tax algorithms, will stop short of developing after-tax financial statements. We believe this presents a misleading view of the world. Consequently, we believe it is an important and worthwhile endeavor to have a model include tax calculations that are in keeping with the financial statements being developed. If the model develops only statutory financial statements, then it is sufficient that the model address only current federal income tax considerations. If the model develops GAAP financial statements as well as statutory ones, the model should also address deferred federal income tax considerations.

Current Income Taxes

Current income taxes are calculated in accordance with insurance company tax procedures [7, Chapter 13]. Current taxes are calculated by adjusting current year statutory net income as follows:

1. Increase or (decrease) current year net income by 20% of the change in the unearned premium reserve.
2. Increase or (decrease) current year net income by the difference in the amount of tax discount in held reserves.[10]

[10]The model is seeded with historical tax discount factors, either industry, company-specific or a combination of the two, depending upon what tax discount factor elections were made in 1987 and 1992. Projected future discount rates are developed using either

3. Decrease current year net income by 85% of the amount of tax-exempt investment income earned during the year.
4. Reduce current year net income by 59.5% of the amount of dividends received from common and preferred stock (the dividends received deduction is 70%, but 15% of the deduction must be added back into net income for tax purposes).
5. Apply a 35% tax rate to the resulting taxable net income amount.

Alternative minimum taxes also are calculated for the current year by increasing taxable net income by 75% of the amount of tax-exempt investment income and dividends received deduction excluded from regular taxable net income and multiplying the resulting alternative minimum taxable net income by the 20% AMT tax rate.

These calculations develop the preliminary current year tax position. If a projection year develops an operating loss, that loss is compared against the three prior calendar years to see if it can be used to offset prior years' operating gains. If not, it is retained for possible use as an operating loss carryforward, to be applied against operating gains in a later projection year.

Deferred Income Taxes for GAAP Accounting

The major components of the deferred income tax calculation are the tax discount in held loss reserves, deferred taxes on deferred acquisition expenses, and deferred taxes on unrealized gains or losses on equities and bonds available for sale or trade. The GAAP income statement includes the calendar year change in the portion of the deferred tax asset arising from the tax discount in held loss reserves, the deferred taxes on deferred

---

pre-seeded industry payout patterns or company-specific payout patterns derived from the line of business underwriting structure and a rolling sixty-month average interest rate that is linked to the model's projected risk-free interest rate projections.

acquisition expenses, and the deferred tax asset or liability arising from unrealized capital gains or losses on that portion of the bond portfolio available for trade.

8. THE FINAL STEP: FINANCIAL STATEMENT PRESENTATION

It is our belief that a model must begin by quantifying cash flows—if cash flows cannot be projected in a reasonably accurate manner, it does not matter how accurately the accounting accruals are developed. In keeping with this belief, we have tried to present mechanics that allow a model to establish with some amount of realism the details of insurance company asset and liability cash flows. It now remains to build up the balance sheet and income statement structure around the cash flows. We proceed by creating a series of general-ledger type accrual accounting entries that extend the underlying cash-basis modeling. We populate these accounting entries by relating them to elements of the insurance company operations that have already been modeled by the underwriting or asset valuation components. Some examples of the types of ratios and the underwriting or asset valuation components to which they might reasonably be related are as follows:

Item	Relationship to:
Agents' balances in course of collection	Written premium, possibly by line of business
Reinsurance recoverable on paid loss	Calendar year paid loss
Interest income due and accrued*	Interest income earned during the year
Expenses due and unpaid	Calendar year expenses incurred
Taxes due (federal or state)	Ratio to calendar year taxes incurred
Provision for reinsurance (Schedule F penalty)-unpaid loss and LAE portion	Year-ending ceded reinsurance balances due or net loss reserves
Provision for reinsurance (Schedule F penalty)-paid loss and LAE portion	Calendar year paid loss

*Depending on the level of detail and sophistication with which assets are modeled, this accrual item may be calculated as part of the asset valuation process. However, if assets are analyzed at even a moderate level of aggregation, this accrual item will need to be estimated instead of calculated directly.

After establishing the asset and liability accrual accounts, we begin the process of creating a full-fledged income statement and cash flow statement. Each of these must contain formulas that respond correctly to changes in the asset and liability accrual values. Much of the information needed to perform this task is described in various chapters of the *Property-Casualty Insurance Accounting* textbook [7]. This is not a step in the model development process to be taken lightly or to be treated superficially. Ultimately, a financial model's results will be shared with many non-actuaries. For them, the test of whether or not the model is (a) believable and (b) worthy of relying on for decision-making will rest in the model's ability to communicate valuable information through the medium of standardized financial statements. A model with sound underlying fundamentals can be undone by such seemingly trivial issues as balance sheets and income statements with minor discrepancies in surplus reconciliation amounts or an incorrect treatment of accounting entries. Accounting rigor also provides model developers with a way of verifying and validating the correctness of the underlying logic so that model users can be comfortable that the model has sound fundamentals.

For those readers already engaged in model development, we hope that this paper provides some ideas on alternative ways of addressing specific modeling issues. For those readers not yet engaged in the model development process, we hope that this paper provides some useful concepts to keep in mind when thinking about the ways in which a financial model might be structured for different organizations. Just as catastrophe models have come to be viewed as a necessity for companies writing property insurance, we believe that DFA-type models will soon be viewed as a necessary tool for examining the overall strategic direction of insurance enterprises. As computer capabilities expand the toolkit available to the actuarial profession, it becomes reasonable to contemplate and actually develop ever more sophisticated and realistic models that will be useful in guiding insurance company decisions.

REFERENCES

[1] Becker, David, "Stylized Historical Facts Regarding Treasury Interest Rates from 1955 to 1994," Technical Report, Lincoln National Life, 1995.

[2] Butsic, Robert, "The Effect of Inflation on Losses and Premiums for Property-Liability Insurers," *Inflation Implications for Property-Casualty Insurance*, Casualty Actuarial Society Discussion Paper Program, 1981, pp. 58–102.

[3] Feldblum, Sholom, "Asset Liability Matching for Property Casualty Insurers," *Continuing Education Committee 1989 Call Papers on Valuation Issues*, pp. 142–143.

[4] Filshtein, Eugene L., "Converting Experts' Knowledge into Dynamic Variable Distributions for Monte Carlo Simulation," *Contingencies*, Volume 8, Number 1, American Academy of Actuaries, January/February 1996, pp. 56–60.

[5] Hodes, Douglas M., et al. "The Financial Modeling of Property/Casualty Insurance Companies," Casualty Actuarial Society *Forum*, Spring 1996, pp. 4–88.

[6] Khury, C. K., "Loss Reserves: Performance Standards," *PCAS* LXVII, 1980, pp. 1–21.

[7] *Property-Casualty Insurance Accounting*, Sixth Edition, Insurance Accounting and Systems Association, July 1994.

[8] Tilley, James A., "An Actuarial Layman's Guide to Building Stochastic Interest Rate Generators," *Transactions of the Society of Actuaries*, Volume XLIV, 1992, pp. 509–538.

[9] Venter, Gary, "Modeling the Evolution of Interest Rates: The Key to DFA Asset Models," Casualty Actuarial Society *Forum*, Summer 1997, Volume 2, pp. 135–164.

[10] A. M. Best Company, *Best's Aggregates and Averages, Property-Casualty*, 1996 Edition.

EXHIBIT 1

PART 1

INSURANCE INDUSTRY COMPOSITE WORKERS COMPENSATION LINE

	Actuals 1996	Projected 1997	Projected 1998	Projected 1999	Projected 2000	Projected 2001
INCOME STATEMENT RELATED ASSUMPTIONS						
Premiums						
1. Net written premiums prior to rate, competitive impacts:		18,656	18,656	17,441	17,842	19,648
2. Rate impact on net written premiums:		n/a	−10.0%	−2.5%	6.7%	−10.0%
3. Competitive impact on net written premiums:		n/a	n/a	n/a	n/a	n/a
4. Other impact on net written premiums:		n/a	3.9%	4.9%	3.2%	2.7%
5. Final expected net written premiums = (1) * [1 + (2)] * [1 + (4)]	18,656	18,656	17,441	17,842	19,648	18,161
Net Loss & ALAE Ratio, net of subrogation/salvage						
6. Net loss & ALAE ratio, prior to rate, inflation impacts		71.5%	71.6%	82.4%	75.5%	79.0%
7. Impact due to premium rate changes		0.0%	9.5%	3.6%	−5.2%	8.4%
8. Inflationary impact		n/a	2.3%	2.3%	1.2%	1.2%
9. Other impact		n/a	n/a	n/a	n/a	n/a
10. Final expected loss & ALAE ratio = (6) * [1 + (7)] * [1 + (8)]		71.5%	80.2%	87.1%	72.6%	86.7%
Inflation and interest rate information						
11. Inflation rate that is implicitly embedded in premium growth levels		3.2%	3.2%	3.2%	3.2%	3.2%
12. Inflation rate that is implicitly embedded in loss payout pattern		4.7%	4.7%	4.7%	4.7%	4.7%
13. Risk free interest rate underlying asset valuations in current scenario		7.3%	9.2%	6.0%	5.1%	3.9%
Premium earning information						
14. Percent of written premium that is earned during first twelve months		86.5%	86.5%	86.5%	86.5%	86.5%
15. Percent of written premium that is earned during second twelve months		13.5%	13.5%	13.5%	13.5%	13.5%

EXHIBIT 1

PART 2

FORMULA EXPLANATIONS

2. **Rate impact on net written premiums:** Model is assuming an expected loss ratio of 81.9%. If the prior year's loss ratio is less than 81.9%, a rate decrease is implemented. The rate decrease is the lesser of a 10% decrease and the complement of the ratio of the prior year's loss ratio and 81.9%. For example, the projected 1997 loss ratio equals 71.5%, so the rate change in 1998 is the smaller of −10% and (71.5%/81.9%), or −12.7%. A similar formula exists for rate increases if the prior year's loss ratio exceeds 81.9%. The 81.9% expected loss ratio was derived from the ten year average industry loss and ALAE ratio. The 10% cap was implemented based on judgment. *This formula is not meant to be representative of a realistic rate change formula. It is used for example only. A realistic rate change formula would consider many more factors than just the relationship of one year's actual loss and expected loss ratios.*

3. **Competitive impact on net written premiums:** This premium adjustment element is not being used in this example. It could be used to incorporate an underwriting cycle element in pricing.

4. **Other impact on net written premiums:** This factor is used to quantify the impact of wage inflation on premium levels. Based on Bureau of Labor Statistics data, wage inflation has averaged 3.2% over the past ten years. For example purposes, it was assumed that the risk-free interest rate over the past ten years has averaged 6.0%. The projected wage inflation impact is equal to [(3.2%/6.0%) ∗ prior year projected risk free interest rates in (13)].
Example: 1999 premium inflation impact = (3.2%/6.0%) ∗ 9.2% = 4.9%.

6. **Net loss and ALAE ratio, prior to rate, inflation impacts:** This is a stochastically generated loss ratio. The distributional parameters were developed from historical industry loss ratios. This is the underlying "force of loss" that is associated with the policies being earned during the year. This is the loss ratio that would develop, absent any other influences on the loss ratio, such as premium rate changes, premium inflation, and loss inflation.

7. **Impact on loss ratio from premium rate changes:**

$$\frac{1}{\{[(1+\text{current year rate change \% from }(2))*(14)]+[(1+\text{prior year rate change \% from }(2))*(15)]\}}-1$$

Example: 1999 impact = 1/{[(1 − 2.5%) ∗ 86.5%] + [(1 − 10.0%) ∗ 13.5%]} − 1 = 3.6%.

EXHIBIT 1

PART 2

(Continued)

8. **Inflationary impact on loss ratio:** This is the result of inflationary pressure on loss costs, partly offset by inflationary increases in premium volume. Based on Bureau of Labor Statistics data, loss inflation has averaged 4.7% over the past ten years. Again, for example purposes, it was assumed that the historical risk-free interest rate over the past ten years has averaged 6.0%. Formula:

$$\frac{\{[(4.7\%/6.0\%) * \text{prior year risk-free interest rate in (13)}] + 1\}}{\{[(1 + \text{current year prem inflation \% from (4)}) * (14)] + [(1 + \text{prior year prem inflation \% from (4)}) * (15)]\}} - 1$$

Example:

$$1999 \text{ impact} = \frac{\{[(4.7\%/6.0\%) * 9.2\%] + 1\}}{\{[(1 + 4.9\%) * (86.5\%)] + [(1 + 3.9\%) * (13.5\%)]\}} - 1 = 2.3\%$$

9. **Other impact on net written premiums:** This element is not being used in this example. It could be used as the counterpart to item (3) in the premium development calculation.

11. **Inflation rate that is implicitly embedded in the premium growth levels:** This is the ten-year average wage inflation statistic from the Bureau of Labor Statistics Consumer Price Index.

12. **Inflation rate that is implicitly embedded in the loss payout pattern:** This is an average of the ten-year average wage inflation statistic from the Bureau of Labor Statistics Producer Price Index and the medical care inflation index from the Bureau of Labor Statistics Consumer Price Index.

13. **Risk-free interest rate underlying asset valuations in current scenario:** This is a stochastically generated future interest rate path, based on a one factor mean-reverting interest rate model.

14. **Percent of written premium that is earned in first twelve months:** This was calculated from industry statistics [10].

15. **Percent of written premium that is earned in second twelve months:** The complement of (14).

CREDIBILITY WITH SHIFTING RISK PARAMETERS, RISK HETEROGENEITY, AND PARAMETER UNCERTAINTY

HOWARD C. MAHLER

Abstract

This paper explores the important effects on credibility of three phenomena: shifting risk parameters, risk heterogeneity, and parameter uncertainty. When any of these phenomena are significant, the Bühlmann credibility formula no longer applies.

Covariance structures corresponding to these phenomena both separately and in combination are shown. Linear equations for the corresponding credibilities are derived.

Possible applications to classification ratemaking, overall rate indication calculation, and experience rating are illustrated in detail. The procedure for estimating the parameters of the covariance structure is discussed for each situation. Illustrative credibilities are then calculated for each situation.

ACKNOWLEDGEMENT

The author would like to thank Curtis Gary Dean for his many extremely helpful suggestions and corrections. I also wish to thank Dotty Culleton and Silvia Fitcher for typing this paper.

1. INTRODUCTION

In Mahler [1] Markov chains were used to model shifting risk parameters. This model was applied to calculate credibilities in four situations. This paper will expand on that work in a number of important areas.

The phenomena of parameter uncertainty and risk heterogeneity will be incorporated. The behavior of credibility as the size of risk changes will be explored. Possible implications for ratemaking, classification pricing, and experience rating will be discussed.

The three phenomena examined in this paper can be defined as follows:

Shifting Risk Parameters: The parameters defining the risk process for an individual insured are not constant over time. There are (a series of perhaps small) permanent changes to the individual insured's risk process as one looks over several years.

Risk Heterogeneity: An insured is a sum of subunits, and not all of the subunits have the same risk process.

Parameter Uncertainty: There are random fluctuations from year to year in the risk processes of insureds. Parameter uncertainty involves fluctuations that affect most or all insureds somewhat similarly, regardless of size.

Each phenomena can be understood and distinguished in the context of the dice examples to be presented.[1] Insurance examples of each phenomena include:

Shifting Risk Parameters: An automobile insured's risk parameters might shift if a major new road were opened in his locality or if he changed the location to which he commutes to work. Similarly, the automobile experience of a town relative to the rest of the state could shift as that town becomes more densely populated.

Risk Heterogeneity: A workers compensation insured may own several factories that have somewhat different risk characteristics.

[1] See Table 3 for a summary of the dice examples.

Parameter Uncertainty: Automobile insureds' risk processes might vary depending on the severity of the winter weather in each year.

The so-called Bühlmann credibility formula is:

$$Z = E/(E + K), \tag{1.1}$$

where E is a measure of size of risk and K is the Bühlmann credibility parameter.

As will be shown, these three phenomena have different effects on the covariance structure between years of data and the resulting credibilities. In the presence of any or all of these three phenomena, the credibility formula in Equation 1.1 does not hold.

Section 2 reviews the results of Mahler [1] relating to shifting risk parameters over time. Section 3 extends the simple dice example from Mahler [1] in order to incorporate parameter uncertainty. Then parameter uncertainty and shifting risk parameters are combined in one model. Section 4 extends the dice example to include risk heterogeneity. Then the model is expanded to include both risk heterogeneity and parameter uncertainty or risk heterogeneity and shifting risk parameters.

In Section 5 the model is expanded to include all three phenomena. The general form of the covariances is given. Section 6 illustrates the calculations of credibilities for various situations. The credibilities for very small risks are discussed. The effect of varying volumes of data by year is discussed. Finally, the case in which no weight is given to the grand mean is discussed.

Section 7 shows how the techniques developed in the prior sections might be applied to the calculation of classification rate relativities. Section 8 extends the results in Section 7 to the use of data from outside the state. Section 9 shows how these techniques might be applied to the calculation of an overall rate indication. Section 10 shows how these techniques might be applied to experience rating. Section 11 covers miscellaneous subjects. Section 12 contains conclusions and a summary.

In order to calculate credibilities there are three steps necessary. First, we must specify the covariance structure between years of data. This structure will vary depending on the phenomena that are important as well as the particular situation.[2] The different covariance structures are listed in Table 1. The general form of the covariance structure is given by Equations 5.10 and 5.11. Second, we must estimate and/or select the parameters appearing in the covariance structure. Finally, we must solve the appropriate set of linear equations for the credibilities. Table 2 lists the different sets of linear equations for the credibilities.

1.1. Bühlmann Credibility[3]

The Bühlmann credibility formula, Equation 1.1, is the least squares credibility corresponding to the following covariance structure between years of data:

$$\mathrm{Cov}[X_i, X_j] = \tau^2 + (\eta^2/E)\delta_{ij}, \tag{1.2}$$

where η^2 is the Expected Value of the Process Variance (for a risk of size 1),

τ^2 is the Variance of the Hypothetical Means,

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j, \end{cases}$$

and E is some measure of size of risk. If the Bühlmann credibility parameter is defined as $K = \eta^2/\tau^2$, then Equation 1.2 can be rewritten as:

$$\mathrm{Cov}[X_i, X_j] = \tau^2\{1 + (K/E)\delta_{ij}\}. \tag{1.3}$$

---

[2]For example, are we dealing with a single split experience rating plan?

[3]Bühlmann credibility is discussed, for example, in Mayerson [2], Hewitt [3], Hewitt [4], Philbrick [5], Herzog [6], Venter [7], Klugman, Panjer and Willmot [8], and Mahler [9].

TABLE 1

EQUATIONS FOR DIFFERENT COVARIANCE STRUCTURES

Shifting Risk Parameters	Risk Heterogeneity	Parameter Uncertainty	Varying Sizes of Risk	$\mathrm{Cov}[X_i, X_j]$		Credibility Assigned to Y Years of Data, Each of Size E[1]
No	No	No	No	$\tau^2\{1 + K\delta_{ij}\}$	Bühlmann	$Y/(Y+K)$
No	No	No	Yes	$\tau^2\{1 + (K/E)\delta_{ij}\}$	(1.2)	$EY/(EY+K)$
Yes	No	No	No	$\tau^2\{\lambda^{\|i-j\|} + K\delta_{ij}\}$	(3.15)	Not Applicable[2]
Yes	No	No	Yes	$\tau^2\{\lambda^{\|i-j\|} + (K/E)\delta_{ij}\}$	(3.16)	Not Applicable[2]
No	No	Yes	Yes	$\tau^2\{1 + ((K/E)+J)\delta_{ij}\}$	(3.5)	$EY/(E(Y+J)+K)$
Yes	No	Yes	Yes	$\tau^2\{\lambda^{\|i-j\|} + ((K/E)+J)\delta_{ij}\}$	(3.20)	Not Applicable[2]
No	Yes	No	Yes	$r^2\{1 + (I/E) + (K/E)\delta_{ij}\}$	(4.3)	$(E+I)Y/(EY+YI+K)$
No	Yes	Yes	Yes	$r^2\{1 + (I/E) + ((K/E)+J)\delta_{ij}\}$	(4.13)	$(E+I)Y/(E(Y+J)+YI+K)$
Yes	Yes	No	Yes	$r^2\{\rho^{\|i-j\|} + (I/E)\gamma^{\|i-j\|} + (K/E)\delta_{ij}\}$	(4.34)	Not Applicable[2]
Yes	Yes	Yes	Yes	$r^2\{\rho^{\|i-j\|} + (I/E)\gamma^{\|i-j\|} + ((K/E)+J)\delta_{ij}\}$	(5.5)	Not Applicable[2]

[1] In those cases where there are not different sizes of risk, take $E = 1$.
[2] Credibilities are the result of solving a set of linear equations, as listed in Table 2. They do not have a simple algebraic form.

In the presence of shifting risk parameters, risk heterogeneity, parameter uncertainty, and varying sizes of risk, the general covariance structure is given by Equations 5.10 and 5.11:

$$\mathrm{Cov}[X_i, X_j] = r^2\left\{\rho^{|i-j|} + \gamma^{|i-j|} I/\sqrt{E_i E_j} + \delta_{ij}\left(K/\sqrt{E_i E_j} + J\right)\right\}, \qquad \sqrt{E_i E_j} \geq \Omega \tag{5.10}$$

$$\mathrm{Cov}[X_i, X_j] = r^2\left\{\rho^{|i-j|} + \gamma^{|i-j|} I/\Omega + \delta_{ij}\left(K/\sqrt{E_i E_j} + J\right)\right\}, \qquad \sqrt{E_i E_j} \leq \Omega. \tag{5.11}$$

TABLE 2

LINEAR EQUATIONS TO SOLVE FOR CREDIBILITIES

Situation	
Y years of data X_i being used to predict Year $Y+\Delta$. Weight to the overall mean.	$\sum_{i=1}^{Y} \mathrm{Cov}[X_i, X_k] Z_i = \mathrm{Cov}[X_k, X_{Y+\Delta}]$, $k = 1,2,\ldots Y$ (2.4)
Y years of data X_i being used to predict Year $Y+\Delta$. No weight to the overall mean.	$\sum_{i=1}^{Y} \mathrm{Cov}[X_i, X_k] Z_i = \mathrm{Cov}[X_k, X_{Y+\Delta}] + \lambda/2$, $k = 1,2,\ldots Y$ $\sum_{i=1}^{Y} Z_i = 1$ (6.7)
Y years of classification data, both from within and outside the state, being used to predict classification relativities for Year $Y+\Delta$. No weight to the overall mean. S_{ij} = covariances within the state. T_{ij} = covariances outside the state. U_{ij} = covariances between state and outside the state.	$\sum_j Z_j S_{ij} + \sum_j W_j U_{ij} = \frac{\lambda}{2} + S_{i,Y+\Delta}$, $i = 1,2,\ldots Y$ $\sum_i Z_i U_{ij} + \sum_i W_i T_{ij} = \frac{\lambda}{2} + U_{Y+\Delta,j}$, $j = 1,2,\ldots Y$ $\sum_i Z_i + \sum_j W_j = 1$ (8.1)
Y years of experience rating data, primary and excess, being used to predict Year $Y+\Delta$. S_{ij} = covariances of primary losses. T_{ij} = covariances of excess losses. U_{ij} = covariances between primary and excess losses.	$\sum_{i=1}^{Y} (Z_{Pi} S_{ik} + Z_{Xi} U_{ki}) = S_{k,Y+\Delta} + U_{k,Y+\Delta}$, $k = 1,2,\ldots Y$ (10.12) $\sum_{i=1}^{Y} (Z_{Pi} U_{ik} + Z_{Xi} T_{ki}) = U_{Y+\Delta,k} + T_{k,Y+\Delta}$, $k = 1,2,\ldots Y$ (10.13)

In those situations where size of risk is not important, Equation 1.3 could be rewritten by setting $E = 1$:

$$\mathrm{Cov}[X_i, X_j] = \tau^2\{1 + K\,\delta_{ij}\}. \tag{1.4}$$

For Y years of data each of size E, the covariance structure given by Equation 1.3 corresponds to a Bühlmann/least squares credibility assigned to these Y years of data of:[4]

$$Z = \frac{EY}{EY + K}. \tag{1.5}$$

As displayed in Table 1, in the presence of any or all of the three phenomena discussed above, the simple covariance structure of Equation 1.3 and the simple credibility formula of Equation 1.5 no longer apply.

2. SHIFTING RISK PARAMETERS

The parameters defining the risk process for an individual insured are not constant over time. For example, for automobile insurance the expected claims frequency of an insured compared to the average changes over time. Mahler [1] presents a Markov chain model of shifting risk parameters which quantifies the effects of shifts over time in the risk process of an insured via the covariances between years of data.

2.1. Covariances, Shifting Risk Parameters

For this Markov chain model, in most cases the covariances can be approximated by:[5]

$$\mathrm{Cov}[X_i, X_j] = \tau^2 \lambda^{|i-j|} + \delta_{ij}\eta^2, \tag{2.1}$$

where

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j, \end{cases}$$

η^2 is the Expected Value of the Process Variance,

τ^2 is the Variance of the Hypothetical Means,

[4]See Section 3.1 for an example of the calculation of Bühlmann credibility. The Bühlmann credibility is calculated as $\mathrm{Cov}[X_i, X_j]/\mathrm{Cov}[X_i, X_j]$. This is the ratio of the variance of the hypothetical means to the expected value of the process variance (each for Y years of data each of size E).

[5]This is Equation 7.1 in Mahler [1].

and λ is the dominant eigenvalue (other than unity) of the transpose of the transition matrix of the Markov chain.

X has different meanings depending on the application. X_i could be the claim frequency for an insured in year i, the loss ratio for a state in year i, the relativity for a class in year i, the die roll in trial i, etc.

From Equation 2.1,

$$\text{Var}[X] = \text{Cov}[X,X] = \tau^2 + \eta^2, \qquad \text{and}$$

$$\text{Total Variance} = \text{VHM} + \text{EPV},$$

the usual relationship that the total variance can be split into the Variance of the Hypothetical Means and the Expected Value of the Process Variance.

As the separation between years of data increases, the (expected) covariance and correlation between years declines.

For example, if τ^2 = VHM = 1,000, η^2 = EPV = 5,000, and $\lambda = .9$, then the variance-covariance matrix given by Equation 2.1 for four consecutive years of data would be:

6,000	900	810	729
900	6,000	900	810
810	900	6,000	900
729	810	900	6,000

This contrasts with the situation in the absence of shifting risk parameters; if $\lambda = 1$, then the variance-covariance matrix has entries of 6,000 along the diagonal and 1,000 off the diagonal. With no shifting risk parameters, Equation 2.1 reduces to the usual Bühlmann covariance structure $\text{Cov}[X_i, X_j] = \tau^2 + \delta_{ij}\eta^2$.

2.2. *Rate of Shifting Risk Parameters*

It is not vital to understand the precise derivation of λ; rather it is important to understand that λ quantifies the rate at which the parameters shift. The smaller λ is, the faster the parameters shift. The closer λ is to unity, the slower the parameters shift. In the limit for $\lambda = 1$, there is no shifting of parameters.

The "half-life" is a useful way to quantify the rate of shifting parameters. The half-life is defined as the length of time necessary for the correlations between years to have declined by a factor of one-half:

$$\begin{aligned} \lambda^{\text{half-life}} &= .5, \\ \text{half-life} &= \frac{\ln .5}{\ln \lambda} = \frac{-.693}{\ln \lambda}. \end{aligned} \tag{2.2}$$

The longer the half-life, the slower the rate of shifting parameters over time.

2.3. Correlations Between Years of Data, Shifting Risk Parameters

If the Markov chain model holds, the correlations between different years of data should decline approximately exponentially. For $i \neq j$, Equation 2.1 gives $\mathrm{Cov}[X_i, X_j] = \tau^2 \lambda^{|i-j|}$.

Thus, as the distance between years grows, the expected covariance between the data from those years declines. Another feature of the Markov chain model is that even though the risk parameters of individuals vary over time, the overall portfolio of insureds looks (relatively) stable from year to year. Specifically, Equation 2.1 gives the same variance for each year of data, $\mathrm{Var}[X_i] = \mathrm{Var}[X_j] = \tau^2 + \eta^2$.

Therefore, the correlations between different years of data are:

$$\begin{aligned} \mathrm{Corr}[X_i, X_j] &= \left(\frac{\tau^2}{\tau^2 + \eta^2} \right) \lambda^{|i-j|}, \qquad \text{and} \\ \ln \mathrm{Corr}[X_i, X_j] &= \ln \left(\frac{\tau^2}{\tau^2 + \eta^2} \right) + |i-j| \ln \lambda, \qquad i \neq j. \end{aligned} \tag{2.3}$$

Therefore, if the Markov chain model holds, the log-correlations for years separated by a given amount should decline approximately linearly. The slope of this line is (approximately) $\ln\lambda$. The intercept is approximately

$$\ln\left(\frac{\tau^2}{\tau^2+\eta^2}\right).$$

Note that $\tau^2/(\tau^2+\eta^2)$ = VHM/Total Variance = credibility in the absence of shifting risk parameters.

Thus given a data set, we can determine whether this (simple) Markov chain model might be appropriate. We determine whether the log-correlations as a function of the separation between years (not including zero separation) can be approximated by a straight line.[6] Then we can estimate the parameter λ and the ratio $\tau^2/(\tau^2+\eta^2)$ from the slope and intercept of the fitted straight line.

2.4. Credibilities, Shifting Risk Parameters

These estimates can be used in turn to estimate credibilities. If we have data X_i from years $1,2,\ldots Y$ and are estimating year $Y+\Delta$, then the least squares credibilities Z_i to be assigned to individual years of data are found by solving the Y linear equations in Y unknowns:[7]

$$\sum_{i=1}^{Y}\mathrm{Cov}[X_i,X_k]Z_i = \mathrm{Cov}[X_k,X_{Y+\Delta}], \qquad k=1,2,\ldots Y. \tag{2.4}$$

3. PARAMETER UNCERTAINTY

Parameter uncertainty and its effect on credibilities is discussed in Meyers [10], Mahler [11] and Mahler [12]. Random

---

[6]In many cases there is a large amount of random fluctuation so even if the expected log-correlations are precisely along a straight line, the log-correlations estimated from the data will vary widely around a straight line. See Figure 10 in Mahler [1].

[7]See Equations 2.8 in Mahler [1].

fluctuations occur from year to year in the risk processes of insureds. Parameter uncertainty involves fluctuations that affect most or all insureds somewhat similarly, regardless of size.

While the distinction between parameter uncertainty and shifting risk parameters is not always clear-cut, parameter uncertainty tends to involve fluctuations not related to the insured while shifting risk parameters tend to involve (a series of perhaps small) permanent changes to the individual insured's risk process. For example, shifting risk parameters would occur if a workers compensation insured implemented a new safety program.

An example of parameter uncertainty occurs in workers compensation insurance, where the level of losses is affected by economic events that affect even very large employers. This creates a potential random fluctuation in the loss potential above and beyond what we normally think of as the process variance. The important feature is that while the large size of an employer reduces the impact of the random fluctuations inherent in observed accidents per year, it either does not reduce or only partially reduces the impact of (seemingly) random changes in the overall economy.

There is a kernel of uncertainty in the frequency of workers compensation claims that will not be reduced by observing more workers during a single year. In these circumstances, the credibilities as a function of the size of risk E will not be of the Bühlmann form $E/(E+K)$.

The covariance structure in the presence of parameter uncertainty is somewhat more complicated, as shown in Equations 3.4 and 3.5. When both parameter uncertainty and shifting risk parameters are present, the covariance structure, as shown in Equations 3.19 and 3.20, contains a combination of the features of each phenomenon separately. These covariance structures will be developed in the context of the simple dice example from

TABLE 3

VARIOUS DICE EXAMPLES

Shifting Risk Parameters	Risk Heterogeneity	Parameter Uncertainty	Section(s)	People	Different Colors of Dice
No	No	No	3.1	Joe	No
No	No	Yes	3.2, 3.4	Joe, Mary	No
Yes	No	No	3.5	Joe, Beth	No
Yes	No	Yes	3.6	Joe, Mary, Beth	No
No	Yes	No	4.1	Joe	Yes
No	Yes	Yes	4.4	Joe, Mary	Yes
Yes	Yes	No	4.9	Joe, Rose, Gwen	Yes
Yes	Yes	Yes	5.1	Joe, Mary, Rose, Gwen	Yes

Joe initially selects either *N* identical dice in the cases without different colors of dice, or *N* identical red dice and *N* possibly different green dice.

Mary flips a coin prior to each trial (year).

Beth, prior to each trial, may alter all the dice from one type to another. (For example, 6-sided dice could be switched to 4-sided dice.)

Rose, prior to each trial, may alter the type of all the red dice.

Gwen, prior to each trial, may alter the type of one or more of the green dice; Gwen acts independently on each green die.

Mahler [1]. Table 3 summarizes the various examples that will be presented.

Section 3.1 will present this simple dice example. Section 3.2 will expand on the dice example in order to incorporate parameter uncertainty. Section 3.3 will discuss how this example relates to parameter uncertainty in general. Section 3.4 will expand the example to observing several years of data. Section 3.5 will introduce shifting risk parameters into the example, in the absence of parameter uncertainty. Section 3.6 extends the example to include both parameter uncertainty and shifting risk parameters. Section 3.7 compares the credibilities corresponding to the various covariance structures discussed. Many readers may find it helpful to go directly to this graphical comparison of results.

3.1. Simple Dice Example, No Shifting Risk Parameters, No Parameter Uncertainty

Assume Joe selects N dice of the same type and rolls them. Assume Joe selected either four-sided,[8] six-sided[9] or eight-sided dice[10] with a priori probabilities of 25%, 50%, and 25%, respectively. Joe tells you how many dice he rolled and the resulting sum, but you do not know which type of dice Joe selected. Joe will roll the same dice again.

The process variances for 4, 6, and 8-sided dice are respectively 1.25, 2.92, and 5.25. Therefore, the expected value of the process variance (for one die) is (25%)(1.25) + (50%)(2.92) + (25%)(5.25) = 3.08. The means for 4, 6, and 8-sided dice are respectively 2.5, 3.5, and 4.5. Therefore, the a priori overall mean is (25%)(2.5) + (50%)(3.5) + (25%)(4.5) = 3.5. The variance of the hypothetical means is .500.

In this case, the Bühlmann credibility for estimating the sum of the next roll of the dice can be written as:

$$Z = \frac{N}{N+K} \tag{3.1}$$

where

$$K = \frac{\text{Expected Value of the Process Variance (for } N = 1)}{\text{Variance of the Hypothetical Means (for } N = 1)} = \frac{\eta^2}{\tau^2} = \frac{3.08}{.5} = 6.16.$$

The credibility Z is to be applied to the data (the sum of Joe's dice), while the complement of credibility $1 - Z$ is to be applied to the a priori grand mean of 3.5.

[8] With numbers 1, 2, 3, and 4 on the faces.
[9] With numbers 1 through 6 on the faces.
[10] With numbers 1 through 8 on the faces.

3.2. Parameter Uncertainty, Dice Example

Assume that in a modification of the previous example, Mary flips a single coin[11] and adds the result to each of Joe's N die rolls.[12] Each head adds $\frac{1}{2}$ to the result of a die, while each tail subtracts $\frac{1}{2}$ from the result of a die. You are again told the result of the combination of Joe's and Mary's actions but see neither the coin nor the dice.

While the addition of a coin flip does not change any of the means, the overall risk process has changed. The amount of credibility we would assign to a single observation has also changed. As will be shown, there is a fundamental change in the behavior of the credibility as a function of N, the number of dice per roll.

The expected value of the process variance is the sum of the expected value of the process variances from Joe's and Mary's actions, since these processes are independent. The expected value of the process variance of Mary's actions is $.25N^2$ since we multiply the result of a single coin flip by N and since $\mathrm{Var}[NX] = N^2\,\mathrm{Var}[X]$. Thus, since the EPV for Joe's action is $3.08N$, the overall expected value of the process variance is $3.08N + .25N^2$.

The hypothetical means have not been changed by the introduction of the coin flips. Therefore, the variance of the hypothetical means remains $.5N^2$.

This covariance structure can be written as:

$$\mathrm{Cov}[X_i, X_j] = .5N^2 + (3.08N + .25N^2)\,\delta_{ij}. \qquad (3.2)$$

Equation 3.2 can be rewritten for more general situations than this specific dice example. It will be useful to substitute E, representing some measure of size of risk such as expected losses, for N, the number of dice that Joe rolls in this specific example.

[11] For simplicity, we assume the coins are fair, with equal probability of heads or tails.

[12] Equivalently, one could add N times the result of the single coin flip to the sum of the die rolls.

If η^2 = expected value of the process variance = 3.08, u^2 = variance due to parameter uncertainty = 0.25, τ^2 = variance of the hypothetical means = 0.5, and E is a measure of size of risk, then Equation 3.2 can be rewritten as

$$\mathrm{Cov}[X_i, X_j] = \tau^2 E^2 + (\eta^2 E + u^2 E^2)\delta_{ij}. \tag{3.3}$$

Suppose that, instead of the sum of the dice, one were estimating the average per die rolled, in a manner analogous to claim frequency, claim severity or pure premium. Then, since the quantity of interest is divided by E, all the variances and covariances in Equation 3.3 are divided by E^2:

$$\mathrm{Cov}[X_i, X_j] = \tau^2 + (\eta^2/E + u^2)\delta_{ij}. \tag{3.4}$$

Letting $J = u^2/\tau^2$ and $K = \eta^2/\tau^2$, Equation 3.4 can be rewritten as:

$$\mathrm{Cov}[X_i, X_j] = \tau^2\{1 + ((K/E) + J)\delta_{ij}\}. \tag{3.5}$$

Equations 3.4 and 3.5 are the covariances in the presence of parameter uncertainty. A new parameter J has been introduced in addition to Bühlmann's K.

The credibility is the variance of the hypothetical means for N dice divided by the sum of the variance of the hypothetical means for N dice and the expected value of the process variance for N dice:

$$Z = \frac{.5N^2}{3.08N + .25N^2 + .5N^2} = \frac{N}{1.5N + 6.16}. \tag{3.6}$$

With $J = u^2/\tau^2 = .25/.5 = .5$ and $K = \eta^2/\tau^2 = 3.08/.5 = 6.16$, this is of the form:[13]

$$Z = \frac{N}{(1+J)N + K}, \qquad J > 0 \quad \text{and} \quad K > 0. \tag{3.7}$$

The form of the credibility as a function of size is fundamentally different. As $N \to \infty$, $Z \to 1/(1+J) < 1$. Therefore, no

[13]The notation in Meyers [10], Mahler [11] and Mahler [12] has been changed so that J there is called $1+J$ here. As will be seen, this cosmetic difference makes it easier to write the formulas involving more than one year of data.

matter how many dice Joe rolls, the credibility assigned to the observation stays less than $1/(1+J)$ or $1/1.5 = 67\%$ in this case. The fact that Joe is rolling more and more dice cannot eliminate the noise added by Mary's single coin flip, which is added to each and every die, and thus cannot increase the credibility beyond 67%.[14]

This is an example of the phenomenon of parameter uncertainty. We can think of this risk process as Joe selects (at random) which type of dice to roll and then Mary's coin flip alters the parameters of the risk process. If for example, Joe selects 6-sided dice, then prior to Mary's coin flip we are uncertain whether this time the expected value of Joe's roll is $3N$ or $4N$. Once Mary flips her coin, if it is tails, the expected value of Joe's roll is $3N$ (after subtracting $.5N$) and if it is heads, the expected value of Joe's roll is $4N$ (after adding $.5N$). The variance of this parameter uncertainty is $.25N^2$.

The value of J which quantifies the impact of parameter uncertainty in the credibility formula was:

$$J = .25/.5 = .5 = \frac{\text{variance due to parameter uncertainty}}{\text{variance of the hypothetical means}} = \frac{u^2}{\tau^2}. \tag{3.8}$$

The larger the J, the greater the impact of parameter uncertainty.

3.3. *Parameter Uncertainty in General*[15]

When parameter uncertainty is important, the within class variance will have two pieces. The "good" piece increases as N and is the expected value of the process variance in the absence of parameter uncertainty. The "bad" piece increases as N^2 and is the variance introduced by parameter uncertainty. Unlike

---

[14] If instead Mary had flipped N coins, one for each die rolled by Joe, then the credibilities would not have behaved in this manner. Instead they would have followed the usual Bühlmann formula, in this case, $Z = N/(N + 6.66)$. The Bühlmann credibility parameter would have been $6.16 + .5 = 6.66$.

[15] See Meyers [10] and Mahler [11].

the good piece, the bad piece increases as quickly as the variance between classes, which also increases as N^2. Thus taking many observations (in a single year) will not get rid of the effect of parameter uncertainty.

This effect is assumed to be due to the different possible states of the universe. Taking more observations will not get rid of the variation inherent in the universe.

In the simple example, Mary's single coin flip represented this random variation in the universe from year to year. In the case of workers compensation insurance, changes in the economy affect the relative costs of claims. These changes can affect firms with 1,000 workers as much as those with 100 workers. Such changes are therefore expected to affect the risk process in a manner similar to Mary's single coin flip (although there is a continuous spectrum of possible states of the economy).

If parameter uncertainty has an important impact on workers compensation insurance, one would expect the credibility to be of the form of Equation 3.7:

$$\frac{E}{(1+J)E+K}, \qquad J > 0, \quad K > 0,$$

where E represents the size of risk. This is one of the refinements introduced in the NCCI's Revised Experience Rating Plan.[16]

3.4. *Dice Example, Several Years of Data*

The dice example with parameter uncertainty will be extended to the situation in which more than one year of data is observed.

Assume Joe selects N dice of a given type and rolls them in each of Y years, while Mary flips a separate coin each year. Then the expected value of the process variance is Y times what it was for a single year: $3.08NY + .25N^2Y$. The variance of the

[16] See Gillam [13], and Mahler [12].

hypothetical means is Y^2 times what it was for a single year: $.5N^2Y^2$. Therefore, the credibility is:

$$Z = \frac{.5N^2Y^2}{.5N^2Y^2 + 3.08NY + .25N^2Y}$$
$$Z = \frac{NY}{NY + .5N + 6.16} \qquad (3.9)$$
$$Z = \frac{NY}{N(Y + .5) + 6.16}.$$

In general, for size of risk E and number of years Y Equation 3.9 can be written as:

$$Z = \frac{EY}{E(Y + J) + K}, \qquad J > 0, \quad K > 0. \qquad (3.10)$$

The credibility has a different form. In the presence of parameter uncertainty, the accumulation of Y separate years does not enter into the formula in the same way as size of risk E. There is the "extra" term involving E, where E is multiplied by J, which is the ratio of the variance due to parameter uncertainty divided by the variance of the hypothetical mean. For one year of data Equation 3.10 reduces to Equation 3.7, the previous result for parameter uncertainty in a single year, which for this example is $Z = E/(1.5E + 6.16)$.

For any fixed number of years, Z has the form $E/$(Linear in E), although the values of the coefficients depend on Y. For fixed size of the insured E, the formula reduces to the usual Bühlmann formula in terms of Y, the number of years. For fixed E as $Y \to \infty$, $Z \to 1$. Increasing the number of years of observations overcomes the impact of parameter uncertainty. We can in fact average over the different assumed random states of the universe in each year by averaging over time.

Observing a fleet of 100 cars for 10 years is *not* the same as observing a similar fleet of 1,000 cars for a single year. In the latter case, we cannot average out those aspects peculiar to that one individual year. For example, a gasoline shortage due to an oil embargo or a severe winter might produce unusual results in an individual year regardless of the size of the fleet.

In summary, in the presence of parameter uncertainty, one must carefully distinguish between size of risk and number of years of data.

3.5. Shifting Parameters Over Time, Dice Example

Shifting risk parameters over time were discussed in Section 2. In Mahler [1], shifting risk parameters were introduced into the simple dice example in Section 3.1 by altering the risk process as follows:

Joe selects a die and rolls it. Then prior to the next trial, Beth may at random replace that die with another die. Assume Beth's replacement process works such that:

1. A four-sided die will be replaced 20% of the time by a six-sided die.[17]
2. A six-sided die will be replaced 10% of the time by a four-sided die and 15% of the time by an eight-sided die.
3. An eight-sided die will be replaced 30% of the time by a six-sided die.

Then the process repeats: Joe rolls a die and Beth (possibly) replaces the die.

Beth's risk process is just a simple example of a Markov chain. See Appendix A for a discussion of Markov chains. There are three "states": 4-sided die, 6-sided die, and 8-sided die. For each trial there is a new, possibly different, state. The probability of being in a state depends only on the state for the previous trial.

[17] The remaining 80% of the time the die is left alone.

Beth's Markov chain is completely described by the "transition probabilities" between the states.

Generally, the transition probabilities for a Markov chain are arranged in a matrix P. For Beth's "risk process," the matrix of transition probabilities is:

	Four	Six	Eight
Four	.80	.20	0
Six	.10	.75	.15
Eight	0	.30	.70

As shown in Mahler [1], in this case, the covariances between years of data are given by:

$$\mathrm{Cov}[X_i, X_j] = (.468)(.769)^{|i-j|} + (.032)(.481)^{|i-j|} + \delta_{ij}(3.08). \tag{3.11}$$

In general, for years of data X_i and X_j:

$$\mathrm{Cov}[X_i, X_j] = \sum_{k>1} \zeta_k \lambda_k^{|i-j|} + \delta_{ij}\eta^2, \tag{3.12}$$

where η^2 is the Expected Value of the Process Variance, $\delta_{ij} = 0$ for $i \neq j$ and 1 for $i = j$, λ_k are the eigenvalues of the transpose of the transition matrix and the ζ_k are a function of the transition matrix P and the means of the states.[18] In general,

$$\sum_{k>1} \zeta_k = \tau^2 = \text{variance of the hypothetical means.}$$

Equation 3.11 can be approximated by:

$$\mathrm{Cov}[X_i, X_j] \approx (.5)(.769)^{|i-j|} + \delta_{ij}(3.08). \tag{3.13}$$

Equation 3.13 can be written in general as:

$$\mathrm{Cov}[X_i, X_j] \approx \tau^2 \lambda^{|i-j|} + \delta_{ij}\eta^2. \tag{3.14}$$

[18] See Mahler [1].

where λ is the dominant eigenvalue of the transpose of the transition matrix (other than unity), τ^2 is the variance of the hypothetical means, and η^2 is the expected value of the process variance. Taking as before $K = \eta^2/\tau^2$, Equation 3.14 could be rewritten as:

$$\mathrm{Cov}[X_i, X_j] = \tau^2\{\lambda^{|i-j|} + K\delta_{ij}\}. \qquad (3.15)$$

For a size of risk E, Equation 3.15 becomes:

$$\mathrm{Cov}[X_i, X_j] = \tau^2\{\lambda^{|i-j|} + (K/E)\delta_{ij}\}. \qquad (3.16)$$

One could then use Equation 2.4 to solve linear equations for the credibilities.

In the absence of shifting risk parameters $\lambda = 1$ and Equation 3.16 becomes the usual Bühlmann covariance structure, Equation 1.3:

$$\mathrm{Cov}[X_i, X_j] = \tau^2\{1 + (K/E)\delta_{ij}\}.$$

3.6. Combining Parameter Uncertainty and Shifting Risk Parameters

Let us now combine the models of parameter uncertainty and shifting risk parameters. Assume that Joe selects N dice (of the same kind) and rolls them. Mary then flips a coin and adds the result ($+1/2$ if heads and $-1/2$ if tails) to the result of each die. The sum is the result of one trial or year. After each trial, Beth (possibly) changes the type of all N dice, with transition matrix P.

Beth does not affect the variance of a single year. As discussed previously in the example involving just Joe and Mary, the total variance of a year of data for this example is $(3.08N + .25N^2) + .5N^2 = 3.08N + .75N^2$.

The covariances between different years are what they were in the absence of Mary, because Mary's action in one year is independent of her action in another year.

Therefore, the covariances of the years of data are for this example:[19]

$$\begin{aligned}\mathrm{Cov}[X_i, X_j] = \{(.468)(.769^{|i-j|}) + (.032)(.481^{|i-j|})\}N^2 \\ + \{.25N^2 + 3.08N\}\,\delta_{ij}. \qquad (3.17)\end{aligned}$$

In general, where E is a measure of the size of the insured:

$$\mathrm{Cov}[X_i, X_j] = \left\{\sum_{k>1} \zeta_k \lambda_k^{|i-j|}\right\} E^2 + \delta_{ij}\{\eta^2 E + u^2 E^2\}. \qquad (3.18)$$

where η^2 = Expected Value of the Process Variance and u^2 = variance due to parameter uncertainty. Given Y years of data, we can solve Y linear equations in Y unknowns, Equations 2.4, for the credibilities to be assigned to each year of data. Note that the solution is the same if we divide all of the variances and covariances by E^2:

$$\begin{aligned}\mathrm{Cov}[X_i, X_j]/E^2 = \sum_{k>1} \zeta_k \lambda_k^{|i-j|} \\ + \delta_{ij}\{(\text{Variance Due to Parameter Uncertainty}) \\ + ((\text{Expected Value of the Process Variance})/E)\}.\end{aligned}$$

This isolates the effect of the size of risk E. As will be discussed subsequently, this is the form that will apply in insurance applications where one is estimating claim frequency rather than total number of claims, pure premiums rather than total losses, etc.

As was done previously, the covariances can be approximated in terms of λ, the dominant eigenvalue of the transpose of the transition matrix (other than unity). For claim frequency, pure premiums, etc., the covariances in the presence of parameter un-

[19]Note that for $i = j$, $\mathrm{Cov}[X_i, X_j] = \mathrm{Var}[X_i] = (.468 + .032)N^2 + .25N^2 + 3.08N = 3.08N + .75N^2$, as stated above.

certainty and shifting risk parameters are then approximately

$$\mathrm{Cov}[X_i, X_j] \approx \tau^2 \lambda^{|i-j|} + (\eta^2/E + u^2)\delta_{ij}. \tag{3.19}$$

Taking as before $J = u^2/\tau^2$ and $K = \eta^2/\tau^2$, Equation 3.19 can be rewritten as:

$$\mathrm{Cov}[X_i, X_j] = \tau^2\{\lambda^{|i-j|} + (K/E + J)\delta_{ij}\}. \tag{3.20}$$

In the absence of shifting risk parameters over time, $\lambda = 1$ and Equation 3.20 would reduce to Equation 3.5. In the absence of parameter uncertainty, $J = 0$ and Equation 3.20 would reduce to Equation 3.16. In the absence of both phenomena Equation 3.20 would reduce to the usual Bühlmann covariance structure. These covariance structures are compared in Table 1.

3.7. Graphical Comparison of Results

Assuming the covariances given by Equation 3.20, we can solve Equation 2.4 for the corresponding credibilities. This has been done for the dice example, which had parameters $J = .5$, $K = 6.16$, and $\lambda = .769$.

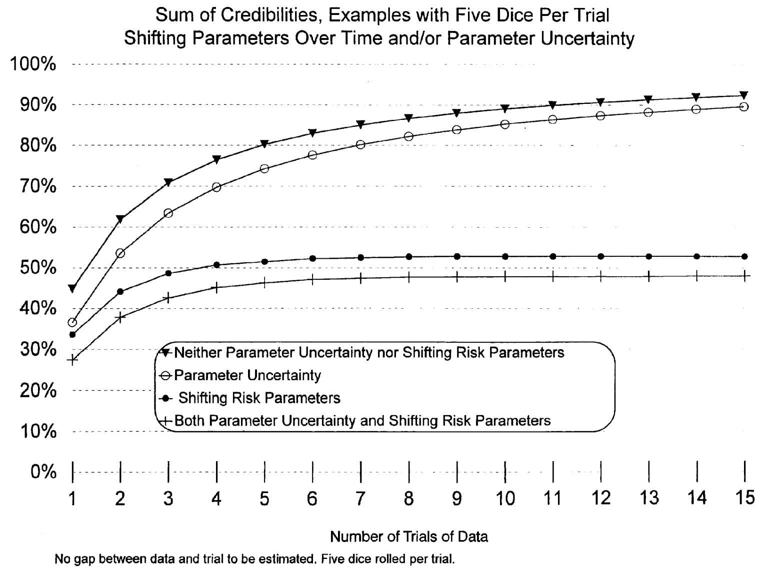
Figure 1 compares the behavior of the credibilities with and without parameter uncertainty as well as with and without shifting risk parameters over time, for five dice per year.[20] In general, both phenomena reduce the credibility assigned to the data by introducing additional noise to the results.

In this particular case with five dice, it so happens that each phenomenon individually results in roughly the same credibility being assigned to a single year of data.[21] Yet we see a radically different behavior as the number of years increases. With just parameter uncertainty, in the limit the effect of parameter uncertainty vanishes; the sum of the credibilities approaches unity.

[20] In Equation 3.20, $E = 5$.

[21] The relative importance of parameter uncertainty increases as the number of dice increases. In this case $Z = Y/(Y + .5 + 6.16/N)$ for Y years and N dice with parameter uncertainty but no shifting risk parameters.

FIGURE 1



With just shifting risk parameters over time, the sum of the credibilities approaches a limit strictly less than unity.[22]

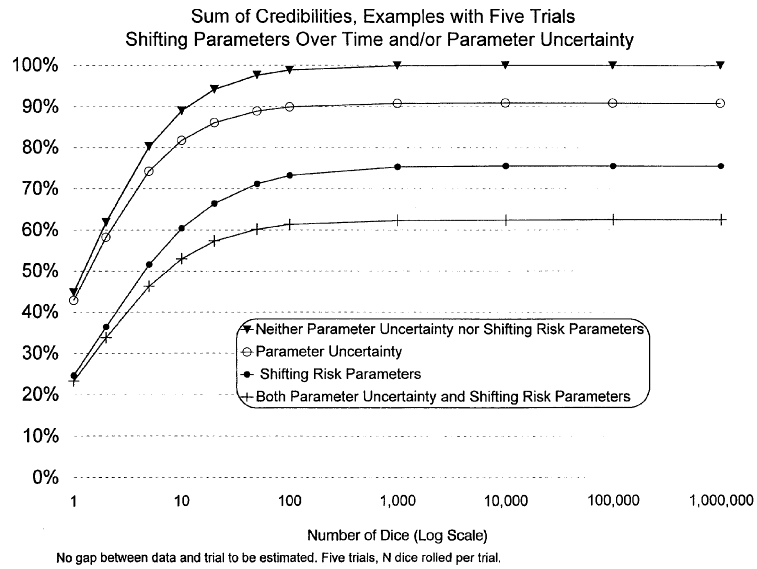
The credibilities in the presence of both phenomena are lower than those with only one of the phenomena. These credibilities approach an even lower limit as the number of years approaches infinity than when we had solely shifting risk parameters.[23] While similar behavior would be expected in general, the details will depend on the amount of parameter uncertainty and the speed at which the parameters shift.

Figure 2 compares for 5 years of data the dependence of the sum of the credibilities on the number of dice per year with the presence or absence of the two phenomena. As expected, with no shifting or parameter uncertainty, we get the usual Bühlmann

[22]In this example, the sum of the credibilities approaches .528.

[23]In this case, with both phenomena present, this limit is .480.

FIGURE 2



credibility, which goes to unity as the number of dice approaches infinity.[24] With parameter uncertainty, the credibilities are somewhat less. Also, as the number of dice approaches infinity, the credibility approaches a limit less than unity.[25]

With shifting risk parameters over time, the credibilities are less than in the absence of shifting risk parameters. As seen in Figure 2, as the number of dice approaches infinity, the credibilities approach a value less than unity.[26] With both phenomena present, the credibilities are lower.[27]

---

[24]In this case, $Z = 5/(5 + 6.16/N)$ for the sum of the credibilities for 5 years.

[25]In this case, $Z = 5/(5.5 + 6.16/N)$ which approaches $1/1.1 = 90.9\%$ as N approaches infinity. Using 5 years of data, one cannot get rid of the effects of parameter uncertainty (although it has less effect than if one relied on fewer than 5 years).

[26]In this case, the limit is .755.

[27]As the number of dice $N \rightarrow \infty$, the sum of the credibilities in this case approaches the limit .625.

4. RISK HETEROGENEITY

The phenomenon of risk heterogeneity and its effect on credibilities was discussed in Mahler [11] and Mahler [12]. As stated in Hewitt [14], "For loss ratio distribution purposes—two $50,000 risks don't make a $100,000 risk. Nor is a $100,000 risk for one year the same as a $50,000 risk for two years." Risk heterogeneity involves an insured which is a sum of individual subunits, where not all the subunits have the same risk process.

Assume we have a large workers' compensation insured. It might consist of several locations or several factories. It is reasonable to assume that the factories making up this insured will be affected by some of the same efforts of management. Therefore, if one factory has better than average expected losses for its mix of classifications, it is likely that another factory that is part of the same insured will have better than average expected losses.

Thus, the combined experience of the different factories has higher credibility for experience rating than the experience of a single factory. However, since the factories also differ in some ways, the larger risk is to some extent heterogeneous. The credibility will not increase as quickly as if the factories were identical; the credibilities are not of the form: $Z = E/(E + K)$.

In general, subunits are combined into one overall insured.[28] If the subunits of the overall insured have the same risk process,[29] then we have the familiar Bühlmann assumptions as in the simple dice example. If on the other hand the subunits of the overall insured are selected at random from the total available population, then there is no increase in the experience rating

[28] The term "subunit" is intended to be vague. It is intended to convey the general concept rather than a particular situation.

[29] "Risk process" refers to the random process that generates the observed quantity of interest. So in the dice example, it would be determined by the number of sides of the dice being rolled. In a Poisson frequency example it would be determined by the average frequency.

credibility of the overall risk compared to its subunits. If the subunits are more similar to each other on average than the total available population, then there is some increase in the experience rating credibility as risk size increases, but not as quickly as in the Bühlmann case where $Z = E/(E + K)$.

As with the prior phenomena, the behavior in the presence of risk heterogeneity will be demonstrated via the simple dice example from Mahler [1]; the example in Section 3.1 will be expanded upon in Section 4.1 in order to incorporate risk heterogeneity.

Section 4.2 discusses risk heterogeneity in general. Equation 4.3 is the corresponding covariance structure. Section 4.3 discusses a refinement to this covariance structure for very small risks.

In Section 4.4, the phenomena of parameter uncertainty and risk heterogeneity are combined in the dice example. Equation 4.13 is the corresponding covariance structure for insurance applications. Section 4.5 gives formulas for credibility in the absence of shifting risk parameters. Section 4.6 discusses a refinement for very small risks to the covariance structure with risk heterogeneity and parameter uncertainty. Sections 4.7 and 4.8 illustrate how this refinement might be applied to workers compensation experience rating.

In Section 4.9, the phenomena of risk heterogeneity and shifting risk parameters are combined in the dice example. Equation 4.34 is the corresponding covariance structure for insurance applications. Section 4.10 discusses the behavior with size of risk for this covariance structure for risk heterogeneity and shifting risk parameters.

4.1. Risk Heterogeneity, Dice Example

As before Joe selects dice, either four-sided, six-sided or eight-sided dice. However, he selects N red dice all of one type and N green dice of possibly different types. Then Joe rolls the dice and tells you the result:

$(1-h)$ (the sum of N red dice) + (h) (the sum of N green dice), where h is a known parameter $0 \leq h \leq 1$.

Assume Joe selected the type of red dice as either four-sided, six-sided, or eight-sided, with a priori probabilities of 25%, 50%, and 25%, respectively. All N of the red dice are of the same type.

Joe independently selected the type of each green die as either four-sided, six-sided, or eight-sided, with a priori probabilities[30] of 25%, 50%, and 25%. The N green dice will usually be a mixture of the three types.

The important feature that distinguishes this example from the prior examples is the different manner in which the green dice are selected compared to the red dice. The N red dice are identical, while the N green dice are a random mixture.

Thus, the green and red dice contribute differently to the variance of the hypothetical means. For a single die with means of 2.5, 3.5 or 4.5 selected with probabilities 25%, 50% and 25%, the variance of the hypothetical means is 0.5. For N identical dice each hypothetical mean is multiplied by N, so the variance of the hypothetical means for the sum of the N red dice is $.5N^2$. For N randomly selected dice the variances add. For the sum of the N green dice the variance of the hypothetical means is $.5N$.

Since the green and red dice are chosen independently of each other, the variance of the hypothetical means for $(1-h)$ (N red dice) + h (N green dice) is:

$$(1-h)^2(.5N^2) + h^2(.5N) = .5N^2(1-h)^2 + .5Nh^2.$$

This is the key effect of risk heterogeneity: the variance of the hypothetical means increases more slowly than the square of the risk size.

[30] The a priori probabilities for the green dice and red dice were selected to be equal solely for simplicity of illustration. This is not an essential feature.

Therefore, one feature of risk heterogeneity is that there is less variation between larger risks than between smaller risks. Specifically, for the dice example, the coefficient of variation[31] of the hypothetical means is $(\sqrt{.5}/3.5)\sqrt{(1-h)^2 + h^2/N}$, which decreases to a positive constant as N increases.

Here, as in Mahler [11] and Mahler [12], the variance of the hypothetical means increases in a combination of a linear and a quadratic term. The question of how quickly the variance of the hypothetical means increases goes back to the origins of workers compensation experience rating.[32] While a rate of increase between linear and quadratic was indicated, the assumption of a quadratic increase was used for practical reasons. This led to the now famous formula for credibility, $Z = E/(E+K)$, which was used for experience rating workers compensation, as discussed in Whitney [15] and Michelbacher [16].

The expected value of the process variance for a single die is 3.08. For the sum of N green dice or N red dice, the expected value of the process variance is $N(3.08)$, since the die rolls are independent. The expected value of the process variance for $(1-h)$ (N red dice) + h (N green dice) is: $(1-h)^2 N(3.08) + h^2 N(3.08)$.

This model might have some applicability to large commercial insureds. For example, assume a commercial automobile fleet involves N drivers. There are many features such as driver selection, driver training, vehicle maintenance, use of vehicle, etc., that are likely to cause the N drivers' risk processes to be more similar than those of the general population of drivers for similarly classified fleets. On the other hand, the N drivers are unlikely to each have the exact same risk process.

In the dice example, each driver's result could be taken as $(1-h)$ (roll of a red die) + h (roll of a green die). Then the red

---

[31] The coefficient of variation is the standard deviation divided by the mean. The overall mean in the dice example is 3.5.

[32] Whitney [15, p. 287] states that the variance of the hypothetical means seemed to increase as $P^{5/4}$, where P was the loss pure premium.

die captures that part of the risk process that is similar across the particular fleet[33] while the green die captures those aspects that mirror the variation across the total classification to which this fleet belongs. The smaller h, the more similar the drivers' risk processes across the fleet, and the smaller the impact of risk heterogeneity.

The credibility is:

$$\begin{aligned} Z &= \text{VHM}/(\text{VHM} + \text{EPV}) \\ &= \frac{.5N^2(1-h)^2 + .5Nh^2}{.5N^2(1-h)^2 + .5Nh^2 + 3.08N(1-h)^2 + 3.08Nh^2} \\ &= \frac{N + \left(\frac{h}{1-h}\right)^2}{N + \left(\frac{h}{1-h}\right)^2 + 6.16 + 6.16\left(\frac{h}{1-h}\right)^2}. \end{aligned} \tag{4.1}$$

If $h = 0$, then $Z = N/(N + 6.16)$, the familiar Bühlmann result in the absence of risk heterogeneity, as in Equation 3.1.

If $h = 1$, then $Z = 1/(1 + 6.16) = 14\%$, the Bühlmann credibility for a single die. If the subunits are chosen totally at random, ($h = 1$), then there is no increase in credibility with size of risk.

Let $I = h^2/(1-h)^2$ while $K = 6.16$, the usual Bühlmann credibility parameter in this case. Then we can rewrite Equation 4.1 as:

$$Z = \frac{N + I}{N + I + K + IK}. \tag{4.2}$$

Equation 4.2 is of the same general form as given in Mahler [11] and Mahler [12].[34] The additional parameter I is zero in the absence of risk heterogeneity. In the presence of risk heterogeneity $I > 0$, and the credibility is of the form: (size + constant)/(size + different constant).

---

[33] While the red dice are identical, the outcomes of the rolls are independent. They represent the same risk process, *not* the same outcome of that risk process.

[34] However, the definition of the parameters is not precisely the same.

While there are some specific assumptions that could be altered,[35] this is one reasonable model which captures the key effect of risk heterogeneity; the Variance of Hypothetical Means has a piece which increases more slowly than N^2 does.

4.2. Risk Heterogeneity in General

The key impact of risk heterogeneity in general is that the covariance between years of claim counts, losses, etc. increases more slowly than the square of the size of risk. Put another way, the covariance between years of claim frequency, pure premiums, etc. decreases with the size of risk. Here, as in Mahler [11] and Mahler [12], the assumption will be made of a covariance structure in the presence of risk heterogeneity of:

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + I/E + (K/E)\delta_{ij}\}, \qquad I, K \geq 0. \quad (4.3)$$

Between different years, Equation 4.3 gives a covariance of $r^2\{1 + I/E\}$, which has one term independent of size of risk and one term that declines as one over the size of risk. If $I = 0$, there is no risk heterogeneity, and Equation 4.3 reduces to the usual Bühlmann covariance structure.

In Equation 4.3, the Variance of the Hypothetical Mean frequencies, pure premiums, etc. is $r^2\{1 + I/E\}$. Assuming the mean claim frequency, pure premium, etc. is (largely) independent of the risk size E, then the coefficient of variation of the hypothetical means declines as E increases. As measured by the coefficient of variation of the hypothetical means, larger insureds are more similar to each other than smaller insureds are to each other. Larger insureds are likely to be a sum of somewhat dissimilar subunits; if we added up enough randomly selected subunits, then we would approach the overall average. Thus with risk heterogeneity, in some sense insureds get closer to average as they get very large.

[35] For example, h, the parameter that quantifies the heterogeneity, was assumed to not depend on N.

For one year of data, substituting the covariance structure given by Equation 4.3 into Equation 2.4 gives the following equation for the credibility:[36]

$$(1 + I/E + K/E)Z = 1 + I/E.$$

Thus, the credibility is of the form:

$$\text{(size + constant)/(size + different constant)},$$

$$Z = \frac{E + I}{E + I + K}. \tag{4.4}$$

With three years of data, all of size E, Equations 2.4 become the following 3 linear equations in three unknowns:

$$(1 + I/E + K/E)Z_1 + (1 + I/E)Z_2 + (1 + I/E)Z_3 = 1 + I/E,$$

$$(1 + I/E)Z_1 + (1 + I/E + K/E)Z_2 + (1 + I/E)Z_3 = 1 + I/E, \quad \text{and}$$

$$(1 + I/E)Z_1 + (1 + I/E)Z_2 + (1 + I/E + K/E)Z_3 = 1 + I/E.$$

This has solution:

$$Z_1 = Z_2 = Z_3 = \frac{E + I}{3E + 3I + K}.$$

If we let Z be the sum of these three credibilities,

$$Z = Z_1 + Z_2 + Z_3 = \frac{(E + I)3}{E3 + 3I + K}.$$

If instead of 3 years of data we have Y years of data, all of size E, then the sum of the credibilities obtained by solving Equations 2.4 is:

$$Z = \frac{(E + I)Y}{EY + YI + K}. \tag{4.5}$$

Equation 4.3 for the covariance structure and Equation 4.5 for the credibility have the same general behavior as in the dice

[36]The factors of r^2 on each side of the equation cancel out, and have no effect on the credibility.

example with risk heterogeneity, although the parameters are somewhat different. Equations 4.3 and 4.5 are in the form that will later be applied to insurance examples. Also, the covariance structure in Equation 4.3 will form the basis for the covariance structure when other phenomena besides risk heterogeneity are present.

4.3. Very Small Risks and Risk Heterogeneity

For the phenomena of risk heterogeneity we will now introduce a refinement for very small sizes of risk. In the dice example in Section 4.1, risk heterogeneity only applies for risks above a certain size, those with more than one die.

Similarly, in insurance examples we might expect that the effects of risk heterogeneity will apply only above a certain size. For commercial automobile insurance, this might be when there is more than one vehicle or more than five vehicles. For workers compensation insurance, this minimum size might be more than one worker, more than a dozen workers, or more than one location. In general, below a certain size, we might expect that there are no subunits which are being grouped and, therefore, no risk heterogeneity. In any case, we will assume there is some minimum size, Ω, which depends on the particular application, below which the phenomena of risk heterogeneity does not apply.

Then for sizes of risk less than Ω, Equation 4.5 will not give the appropriate credibility. It will give too much credibility to the very smallest risks; as $E \to 0$ in Equation 4.5, $Z \to (I/(I + K/Y)) > 0$.

In practical applications we can apply special caps to the effect of credibility for small risks.[37] In the NCCI Revised Experience Rating Plan for workers compensation insurance, there are caps

[37] In general one should cap the effects of credibility. See for example Mahler [17] and Mahler [18].

on the maximum debit for small risks.[38] In addition, below a certain size risks are not eligible for experience rating.[39]

It is worthwhile to explore the expected behavior of the credibilities for very small risks. For experience rating one may devise a simplified merit rating plan to apply to smaller risks. For classification rating one must assign the data of every class a credibility, no matter how small the volume of data.

We will assume a covariance structure and derive a formula for the credibilities that apply for risks of the smallest sizes. Equation 4.3 is assumed to be valid for risks of size $\geq \Omega$

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + I/E + (K/E)\delta_{ij}\}, \qquad E \geq \Omega. \quad (4.6)$$

For $E = \Omega$:

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + I/\Omega + (K/\Omega)\delta_{ij}\}.$$

We assume that for $E < \Omega$, the term related to risk heterogeneity, I/Ω, does not decline as the risk size declines below Ω, and thus acts as if the risk was homogeneous.[40] In other words:

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + I/\Omega + (K/E)\delta_{ij}\}, \qquad E \leq \Omega. \quad (4.7)$$

Thus, for risks of size less than Ω, the Variance of the Hypothetical Means is $r^2 + r^2 I/\Omega$, independent of size. This is the type of behavior we expect in the absence of risk heterogeneity.[41] While the dice example was useful for developing the ideas in

---

[38] See Mahler [12]. Recently the maximum debit has been revised. It is now given via a continuous formula for all sizes: $1 + (.00005)[E + 2E/g]$, where g is NCCI's state specific parameter.

[39] The minimum is based on premiums and varies by state. For example, for Massachusetts it is currently \$5,500 in annual premium.

[40] The term related to risk homogeneity, r^2, is independent of the size of risk, and thus below Ω remains the same.

[41] Although r^2 can be thought of as the piece of the VHM which is related to risk homogeneity, the VHM for small risks is assumed to be $r^2 + r^2 I/\Omega$. If one desired, one could reparametrize the covariances setting $\tau^2 = r^2 + r^2 I/\Omega$ and then use τ^2 rather than r^2. However, such a reparametrization would not in and of itself alter the credibilities.

this paper, it has its limitations. In the dice example N can never be less than one.

Using the covariance structure given by Equation 4.6, for $E \geq \Omega$ the credibilities are given by Equation 4.5

$$Z = \frac{Y(E+I)}{YE+YI+K}, \qquad E \geq \Omega. \tag{4.8}$$

However, for $E \leq \Omega$, the covariances are given by Equation 4.7, and the solution to Equations 2.4 is, in the absence of shifting risk parameters and parameter uncertainty:

$$\begin{aligned} Z &= \frac{Y(1+I/\Omega)}{Y\{1+(I/\Omega)\}+(K/E)} \\ Z &= \frac{YE}{YE+K'}, \qquad E \leq \Omega \end{aligned} \tag{4.9}$$

where $K' = K(\Omega/I+\Omega)$.

Equation 4.9 is of the same form as the Bühlmann credibility formula, but with the parameter K adjusted by a factor of $\Omega/(I+\Omega)$.

The credibilities given by Equation 4.9 approach zero as the risk size approaches zero. As expected, for $E = \Omega$, Equations 4.8 and 4.9 give the same credibility:

$$\begin{aligned} Z &= \frac{Y(\Omega+I)}{(Y)\Omega+YI+K} = \frac{Y(\Omega+I)}{Y(\Omega+I)+K} \\ &= \frac{Y}{Y+K\left(\dfrac{\Omega}{\Omega+I}\right)\left(\dfrac{1}{\Omega}\right)} \\ &= \frac{Y}{Y+K'/\Omega} = \frac{Y\Omega}{Y\Omega+K'}. \end{aligned}$$

Equations 4.8 and 4.9 together combine the usual Bühlmann credibility formula for small risks with that applicable in the presence of risk heterogeneity for large risks.

4.4. Risk Heterogeneity and Parameter Uncertainty, Dice Example

The dice models of risk heterogeneity and parameter uncertainty can be easily combined. Joe picks N identical red dice and N randomly selected green dice as in Section 4.1, and Mary flips a coin as in Section 3.2. Then the result is:

$$(1-h)(\text{Sum of } N \text{ Red Dice}) + h(\text{Sum of } N \text{ Green Dice}) + N(\text{Coin Flip}),$$

where the coin flip is counted as $-\frac{1}{2}$ if tails and $+\frac{1}{2}$ if heads.

Then, per Sections 3.2 and 4.1, the Expected Value of the Process Variance is the sum of Joe and Mary's individual process variances:

$$(3.08)(1-h)^2 N + (3.08)h^2 N + .25N^2.$$

The presence of the coin flips has not altered the hypothetical means. Therefore, according to Section 4.1, the variance of the hypothetical means is:

$$.5N^2(1-h)^2 + .5Nh^2.$$

The EPV and VHM can be combined into the covariance structure:

$$\text{Cov}[X_i, X_j] = .5N^2(1-h)^2 + .5Nh^2 + \{(3.08)(1-h)^2 N + (3.08)h^2 N + .25N^2\}\,\delta_{ij}. \tag{4.10}$$

The credibility is:

$$\begin{aligned} Z &= \text{VHM}/(\text{VHM}+\text{EPV}) \\ &= \frac{.5N^2(1-h)^2 + .5Nh^2}{.5N^2(1-h)^2 + .5Nh^2 + 3.08N(1-h)^2 + 3.08Nh^2 + .25N^2} \\ &= \frac{N + \left(\dfrac{h}{1-h}\right)^2}{N + \left(\dfrac{h}{1-h}\right)^2 + 6.16 + 6.16\left(\dfrac{h}{1-h}\right)^2 + \dfrac{.5N}{(1-h)^2}}. \end{aligned} \tag{4.11}$$

As before let $I = h^2/(1-h)^2$ while $K = 6.16$, the usual Bühlmann credibility parameter. Let $J = .5/(1-h)^2$, which for $h = 0$ reduces to the situation in Section 3.2 where J was .5. Then Equation 4.11 can be rewritten as:

$$Z = \frac{N + I}{N(1+J) + I + K + IK}. \tag{4.12}$$

4.5. *Credibilities, No Shifting Risk Parameters*

For insurance applications to frequency, pure premiums, etc., it will be useful to rewrite the covariance structure in Equation 4.10 with a somewhat different parametrization than in the dice example. Combining the features of Equations 3.5 and 4.3, the covariance structure with risk heterogeneity and parameter uncertainty is:[42]

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + (I/E) + ((K/E) + J)\delta_{ij}\}, \qquad I, J, K \geq 0. \tag{4.13}$$

When one uses Y years of data to predict a future year, Equations 2.4 become with the covariances from Equation 4.13:

$$(K/E + J)Z_i + \sum_{j=1}^{Y}(1 + I/E)Z_j = 1 + I/E, \qquad i = 1, 2, \ldots, Y.$$

By symmetry the credibilities for the individual years, Z_i, are all equal.

Let $Z_i = Z/Y$, where $Z = \sum Z_i$, the total credibility applied to the data.[43] Then the sum of the credibilities for Y years of data

[42] Where as before E is the size of risk, I quantifies risk heterogeneity, J quantifies parameter uncertainty, and K is the Bühlmann credibility parameter. The size of risk enters as $1/E$ since we are estimating quantities such as frequency or pure premiums rather than the sum of die rolls, the total number of claims, the total losses.

[43] The credibility applied to each year is in this case the total credibility divided by the number of years.

each of size E is:

$$Z = \frac{(E+I)Y}{E(Y+J)+YI+K}; \qquad I,J,K \geq 0. \tag{4.14}$$

For one year of data, $Y = 1$, Equation 4.14 becomes:[44]

$$Z = \frac{E+I}{E(1+J)+I+K}. \tag{4.15}$$

While Equation 4.14 with $Y = 1$ differs slightly from Equation 4.12, they have the same essential form as a function of size of risk.

Equation 4.14 is of the same general form as given in Mahler [11] and Mahler [12].[45] In the absence of parameter uncertainty, J=0 and the credibility is given by Equation 4.5. In the presence of parameter uncertainty, $J > 0$. In the absence of risk heterogeneity $I = 0$ and the credibility is given by Equation 3.10. In the presence of risk heterogeneity, $I > 0$.

The parameter I largely affects the credibilities for smaller risks. The parameter J largely affects the credibilities for larger risks. The maximum credibility as the size of risk approaches infinity is $Y/(Y+J) < 1$. The credibility is of the form: (linear function of size)/(linear function of size).

Equation 4.14 for the credibility in the presence of risk heterogeneity and parameter uncertainty is the form used in the NCCI Revised Experience Rating Plan for workers compensation. The primary and excess credibilities depend on a state specific parameter g as follows:[46]

$$\begin{aligned} Z_p &= (E' + 700g)/(1.1E' + 3270g), \qquad \text{and} \\ Z_x &= (E' + 5{,}100g)/(1.75E' + 208{,}925g), \end{aligned} \tag{4.16}$$

[44]This is the same general form of the credibilities in the presence of risk heterogeneity and parameter uncertainty, shown in Mahler [12]. This is the same basic form as Equation 4.4 of Mahler [12], with a slightly different treatment of the parameters I and K.

[45]However, the definition of the parameters is not precisely the same.

[46]See Mahler [12]. The parameter g is the average cost per case divided by 1,000; g is rounded to the nearest 0.05. Recently the NCCI has revised the excess parameters

where E' is the expected losses for the sum of 3 years of data. E' is the equivalent of $3E = YE$ in Equation 5.9. Equations 4.16 are the same as Equation 4.14 with $Y = 3$ and the parameters:[47]

	Primary	Excess
I	$700g/3$	$1{,}700g$
J	.3	2.25
K	$2{,}570g$	$203{,}825g$

(4.17)

Note that as $E \to 0$ in Equation 4.14, $Z \to YI/(YI + K)$. Thus the minimum credibility is $1/(1 + (K/IY))$. This is greater than zero for $I > 0$. For the NCCI Revised Experience Rating Plan the minimum primary credibility is $1/(1 + (2{,}570/700)) = 21.4\%$. The minimum excess credibility is $1/(1 + (203{,}825/5{,}100)) = 2.4\%$.

As $E \to \infty$ in Equation 4.14, $Z \to (YE)/(Y + J)E = Y/(Y + J) = 1/(1 + J/Y)$. This is less than 1 for $J > 0$. For the NCCI Revised Experience Plan, the maximum primary credibility is $1/(1 + .3/3) = 1/1.1 = 90.9\%$. The maximum excess credibility is $1/1.75 = 57.1\%$.

Without parameter uncertainty, $J = 0$ and Equation 4.14 becomes:

$$Z = \frac{(E + I)Y}{EY + YI + K} = \frac{E + I}{E + I + K/Y}. \qquad (4.18)$$

somewhat to take effect during 1998 and later. $J_x = 1.125$ rather than 2.25. $K_x = 150{,}000g$ rather than $203{,}825g$. In addition, only 30% of Medical Only losses will be included in experience rating.

[47]This differs from the values shown in Mahler [12] due to the somewhat different treatment of the parameters here. The important point is that the credibilities are of the form Linear/Linear. The Revised Experience Rating Plan was developed under the direction of Gary Venter while he was at the National Council on Compensation Insurance. As described in Gillam [13], this was the form of credibilities that worked well in the tests performed by the NCCI. Note that while in Section 10 of the current paper explicit recognition of the impact of the covariance of primary and excess losses is taken, this was not the case in the derivation of the credibilities in the NCCI Revised Experience Rating Plan.

For one year of data (and only risk heterogeneity) Equation 4.18 becomes:

$$Z = \frac{E+I}{E+I+K}. \tag{4.19}$$

Without risk heterogeneity $I = 0$, and Equation 4.14 becomes:

$$Z = \frac{YE}{E(Y+J)+K}. \tag{4.20}$$

For one year of data (and only parameter uncertainty) Equation 4.19 becomes:

$$Z = \frac{E}{E(1+J)+K}. \tag{4.21}$$

4.6. Very Small Risks, Risk Heterogeneity and Parameter Uncertainty

As in Section 4.3, we will introduce a refinement for very small sizes of risk. In the dice example, risk heterogeneity applies only for risks above a certain size, those with more than one die. Similarly, in insurance examples we might expect that the effects of risk heterogeneity will apply only above a certain size.

We will assume a covariance structure and derive a formula for the credibilities that apply for risks of the smallest sizes. Equation 4.13 is assumed to be valid for risks of size $\geq \Omega$:

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + (I/E) + ((K/E) + J)\delta_{ij}\}, \qquad E \geq \Omega. \tag{4.22}$$

For $E = \Omega$:

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + (I/\Omega) + ((K/\Omega) + J)\delta_{ij}\}.$$

We assume that for $E < \Omega$, the term related to risk heterogeneity, I/Ω, does not decline as the risk size declines below Ω, and thus acts as if the risk were homogeneous.[48] In other

[48] The term related to risk homogeneity, r^2, is independent of the size of risk, and thus below Ω remains the same.

words:

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + (I/\Omega) + ((K/E) + J)\delta_{ij}\}, \qquad E \le \Omega. \tag{4.23}$$

Using the covariance structure given by Equation 4.22, for $E \ge \Omega$ the credibilities are given by Equation 4.14:

$$Z = \frac{Y(E+I)}{(Y+J)E + YI + K}, \qquad E \ge \Omega. \tag{4.24}$$

However, for $E \le \Omega$, the covariances are given by Equation 4.23, and the solution to Equations 2.4 is, in the absence of shifting risk parameters:

$$\begin{aligned} Z &= \frac{Y(1 + (I/\Omega))}{Y(1 + (I/\Omega)) + (K/E) + J} \\ &= \frac{YE((I+\Omega)/\Omega)}{YE((I+\Omega)/\Omega) + JE + K} \\ &= \frac{YE}{(Y + J')E + K'}, \qquad E \le \Omega \end{aligned} \tag{4.25}$$

where

$$J' = J\left(\frac{\Omega}{I+\Omega}\right) \qquad \text{and} \qquad K' = K\left(\frac{\Omega}{I+\Omega}\right).$$

Equation 4.25 is of the same form as Equation 4.20, but with the parameters J and K each adjusted by a factor of $\Omega/(I+\Omega)$. This is the Bühlmann credibility formula with an additional parameter J' to account for parameter uncertainty. For very small risks, the parameter J' has very little effect; thus Equation 4.25 gives approximately the same result as the usual Bühlmann credibility formula.

The credibilities given by Equation 4.25 approach zero as the risk size approaches zero. As expected, for $E = \Omega$, Equations 4.24 and 4.25 give the same credibility:

$$
\begin{aligned}
Z &= \frac{Y(\Omega+I)}{(Y+J)\Omega+YI+K} = \frac{Y(\Omega+I)}{Y(\Omega+I)+J\Omega+K} \\
&= \frac{Y}{Y+J\left(\dfrac{\Omega}{\Omega+I}\right)+K\left(\dfrac{\Omega}{\Omega+I}\right)\left(\dfrac{1}{\Omega}\right)} \\
&= \frac{Y}{Y+J'+K'/\Omega} = \frac{Y\Omega}{(Y+J')\Omega+K'}.
\end{aligned}
$$

4.7. Very Small Risks, Workers Compensation Experience Rating

For example, consider the NCCI Revised Experience Rating Plan with parameters given in Equations 4.17 and $Y = 3$. Take solely for illustrative purposes $\Omega = \$1{,}000g$. If $g = 2$, corresponding to an average claim size of \$2,000, then $\Omega = \$2{,}000$. This would correspond to \$6,000 in expected losses[49] over 3 years. Assuming the expected loss rate is about 40% of the manual rate, then \$6,000 in expected losses corresponds to about \$15,000 in premium over 3 years.

This would be among the smallest risks eligible for experience rating. Nevertheless, let us ignore the eligibility criterion, and compare the primary credibilities given by Equations 4.14 and 4.25 for risks with expected annual losses less than $\Omega = 1{,}000g$. For $g = 2$, we get parameters in Equation 4.17 of:

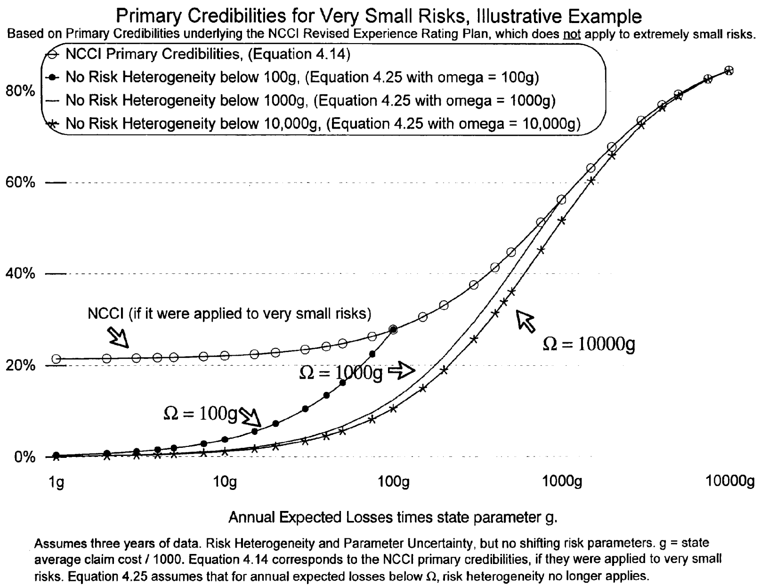
	Primary	Excess
I	466.67	3,400
J	0.3	2.25
K	5,140	407,650

Using Equation 4.14 with $Y = 3$, we get primary credibilities of:

$$
Z_p = \frac{3E + 1{,}400}{3.3E + 6{,}540}.
$$

[49] At first, second, and third reports as limited for experience rating.

FIGURE 3



For example, for $E = 100g = 200$, the primary credibility would be $2{,}000/7{,}200 = 27.8\%$. In contrast, using Equation 4.25 with $Y = 3$, $\Omega = 2{,}000$, $J' = J\ (\Omega/(I+\Omega)) = .243$, and $K' = K(\Omega/(I+\Omega)) = 4167.6$, we get primary credibilities of:

$$Z_p = \frac{3E}{3.243E + 4167.6}, \qquad E \leq 2{,}000.$$

For example, for $E = 100g = 200$, the primary credibility would be $600/4816.2 = 12.5\%$.

As shown in Figure 3, the credibilities given by Equation 4.25 decline quickly to zero, while those from Equation 4.14 have a minimum value of $YI/(YI+K) = 1{,}400/(1{,}400 + 5{,}140) = .214$.

For example, for expected annual losses of $100g$, the primary credibilities are 27.8% from Equation 4.14 and 12.5% from Equation 4.25. For $1{,}000g$ the credibilities from the two equations are equal. Similarly for $E = 100g = 200$, the excess credibilities are 2.6% from Equation 4.14 and .4% from Equation 4.25.

For $E = 200$, weighting together the primary and excess credibilities, assuming a D-ratio[50] of roughly 30%, produces credibilities of 10% from Equation 4.14 and 4% from Equation 4.25. The contrast is even greater for much smaller risks.

Expected Annual	NCCI Formulas[51]		Alternate Formula 4.25 with $\Omega = 1{,}000g$	
Losses	Z_p	Z_x	Z_p	Z_x
$10g$	22.1%	2.5%	1.4%	0.04%
$100g$	27.8%	2.6%	12.5%	0.4%
$1{,}000g$	56.6%	3.8%	56.6%	3.8%

The lower credibilities from Equation 4.25 make much more sense for very small risks. For $g = 2$, $10g = \$20$ in expected annual losses.[52] The alternative formula corresponding to Equation 4.25 gives a credibility of .4% (assuming a D-ratio of roughly 30%),[53] which at least has a possibility of being reasonable. The NCCI formulas corresponding to Equation 4.14 are not applied to such small risks, nor could they be. The resulting credibility of 8.4% (assuming a D-ratio of roughly 30%)[54] is way too high. Thus, the refinement to the covariance structure for very small risks, as in Equation 4.23, is at least a step in the right direction towards obtaining reasonable experience rating credibilities for very small risks.

[50]The D-ratio is the ratio of primary losses to primary plus excess losses.

[51]Equation 4.16 with $g = 2$ and E' equal to three times expected annual losses.

[52]A single full-time clerical employee might have \$20 or more in expected annual losses for workers compensation. This is very far below the size of risk that is experience rated.

[53]$(1.4\%)(30\%) + (.04\%)(1 - 30\%) = .448\%$, where from the table $Z_p = 1.4\%$ and $Z_x = .04\%$.

[54]$(22.1\%)(30\%) + (2.5\%)(1 - 30\%) = 8.38\%$, where from the table $Z_p = 22.1\%$ and $Z_x = 2.5\%$.

Figure 3 also displays the result of choosing $\Omega = 10{,}000g$ rather than $\Omega = 1{,}000g$. The credibilities are relatively insensitive to the choice between these two values of Ω. Either value of Ω used with Equation 4.25 allows a smooth transition down to zero from the NCCI credibilities for very small risks. The transition at $E = \Omega$ between Formulas 4.24 and 4.25 will be smoothest when the slopes at $E = \Omega$ are similar.

If the credibility Z is given by Formula 4.24, then

$$\frac{dZ}{dE} = \frac{Y(K - IJ)}{((Y+J)E + YI + K)^2} = \frac{Y(K - IJ)}{(YE(1 + I/E) + JE + K)^2}.$$

If instead the credibility is given by Formula 4.25, then

$$\frac{dZ}{dE} = \frac{YK'}{((Y+J')E + K')^2} = \frac{YK(1 + I/\Omega)}{(YE(1 + I/\Omega) + JE + K)^2}.$$

At $E = \Omega$, the denominators of the derivatives of the two formulas for Z are equal.

Thus, it follows that at $E = \Omega$, the ratio of the derivative with respect to E of Z given by Formula 4.25 to the derivative with respect to E of Z given by Formula 4.24 is: $(1 + I/\Omega)/(1 - IJ/K)$. The transition will be smoothest when the slopes of the curves are close, which occurs when this ratio of derivatives is close to unity.[55] In most applications IJ/K will be small, and thus $1/(1 - IJ/K)$ will be close to unity.[56] Thus, if Ω is at least $5I$, the ratio of derivatives at Ω will be close to unity, producing a smooth transition between the two credibility formulas.

Figure 3 also displays the result of choosing $\Omega = 100g$. This value would not allow a smooth transition between the two credibility formulas. The credibilities using $\Omega = 100g$ differ significantly from those obtained from using $\Omega = 1{,}000g$. Which value of Ω is most appropriate is an empirical question whose answer

[55] The ratio is greater than unity since $1 + I/\Omega > 1$ and $1 - IJ/K < 1$.

[56] For the NCCI Revised Experience Rating Plan, IJ/K is .027 for primary losses and .019 for excess losses.

depends on obtaining as much information as possible about the covariance structure in the particular situation.

4.8. W and B Values, Workers Compensation Experience Rating

In workers compensation experience rating it is common to display tables of W (weighting) and B (ballast) values rather than primary and excess credibilities.[57] The primary and excess credibilities are then given in terms of W, B and expected losses as:[58]

$$Z_p = \text{Expected Losses}/(\text{Expected Losses} + B), \qquad \text{and}$$

$$Z_x = WZ_p.$$

Thus, B acts like a Bühlmann credibility parameter, except that B varies by size of risk. W quantifies for a given size of risk how much smaller the excess credibility is than the primary credibility. For three years of data, each with expected annual losses of E, $Z_p = 3E/(3E + B)$.

We can calculate the ballast value B that corresponds to the primary credibilities calculated in the prior section. Prior to the imposition of a minimum value, $B = 3E(1/Z_p - 1)$, where E is the expected annual losses and Z_p is the primary credibility.[59]

Using Equation 4.25, which assumes risk homogeneity below risk size Ω, with parameters $I_p = 466.67$, $J_p = 0.3$, $K_p = 5{,}140$ and $\Omega_p = 2{,}000$ from the prior section, we can calculate the primary credibility and corresponding value of B. For example, for expected annual losses of $E = 200$, $Z_p = 12.5\%$ and thus $B = 600((1/.125) - 1) = 4{,}200$. Keeping the other parameters fixed, we can alter Ω_p, resulting in different graphs of B versus E, as shown in Figure 4.[60]

---

[57] See Gillam and Snader [19], Gillam [13] or Mahler [12].

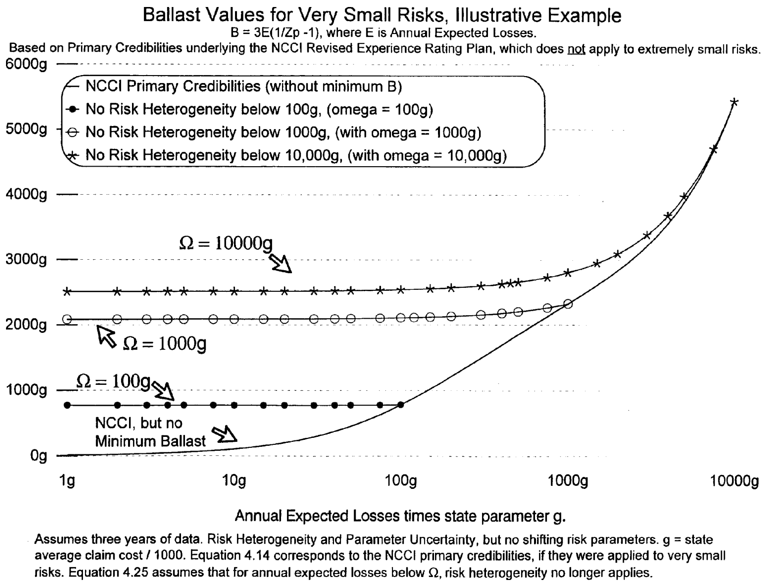
[58] Mahler [12] relates these equations to Equations 4.16.

[59] Thus, for 3 years of data, with expected losses $3E$,

$$3E/(3E + B) = 3E/(3E + 3E \cdot (1/Z_p - 1)) = Z_p.$$

[60] For example, for $\Omega_p = 1{,}000g = 2{,}000$ and $E = 100g = 200$, $B = 2{,}100g = 4{,}200$.

FIGURE 4



As seen in Figure 4, the assumption of no risk heterogeneity below Ω, with respect to primary credibilities, corresponds approximately to the imposition of a minimum ballast value. As the value of Ω varies from $100g$ to $10{,}000g$, the minimum B varies from around $800g$ to $2{,}500g$.

For $E \le \Omega_p$ from Equation 4.25 we have

$$Z_p = \frac{YE}{(Y+J_p')E + K_p'} \approx \frac{YE}{YE + K_p'} \qquad \text{where} \quad K_p' = K_p\left(\frac{\Omega_p}{I_p + \Omega_p}\right)$$

and J_p' is small.

For $E \le \Omega$, the credibilities approximately follow the usual Bühlmann formula, thus the minimum ballast value should be approximately

$$K_p\left(\frac{\Omega_p}{I_p + \Omega_p}\right).$$

For $I_p = 700g/3$, $K_p = 2{,}570g$, and illustrative values of Ω we get:

Ω_p	$\Omega_p/(I_p + \Omega_p)$	K_p	K'_p
$100g$	30.0%	$2{,}570g$	$771g$
$300g$	56.3%	$2{,}570g$	$1{,}447g$
$1{,}000g$	81.1%	$2{,}570g$	$2{,}084g$
$3{,}000g$	92.8%	$2{,}570g$	$2{,}385g$
$10{,}000g$	97.7%	$2{,}570g$	$2{,}511g$

Thus, for this range of values for Ω, the range of minimum ballast values K'_p is from about $800g$ to $2{,}500g$.[61] In any case, some minimum ballast value is appropriate regardless of the value of Ω. The minimum B should be a function of the state specific parameter g, and must be less than K_p.

Similarly the weighting value W is equal to $W = Z_x/Z_p$. For $E \leq \Omega_p$ and $E \leq \Omega_x$, using Equation 4.25, $W \approx (YE/(YE + K'_x))/(YE/(YE + K'_p)) = (YE + K'_p)/(YE + K'_x)$. As the size of risk goes to zero, $E \to 0$,

$$W \to K'_p/K'_x = (K_p/K_x)(\Omega_p/\Omega_x)\frac{(I_x + \Omega_x)}{(I_p + \Omega_p)}.$$

If, for example, we were to take $\Omega_x = \Omega_p = 10{,}000g$, then using the current NCCI values $I_p = 700g/3$, $K_p = 2{,}570g$, $I_x = 1{,}700g$, and $K_x = 203{,}825g$, the minimum W value would be .014; this compares to a current minimum W of .07.

4.9. Risk Heterogeneity and Shifting Risk Parameters, Dice Example

In this section, the phenomenon of shifting risk parameters will be added to the model in Section 4.1.

[61] The NCCI has introduced a minimum B of $2{,}500g$, which as seen here corresponds to $\Omega \approx 10{,}000g$.

Joe initially selects N identical red dice and N possibly different green dice.

Prior to each trial, Rose may alter the type of all the red dice. Prior to each trial, Gwen may alter the type of one or more of the green dice; Gwen acts independently on each green die. Then Joe rolls all the dice and the result is taken as: $(1-h)$ (the sum of the N red dice) $+ h$ (the sum of the N green dice.)

Assume that Rose's replacement of red dice follows the transition matrix $\boldsymbol{R}$:

$$\boldsymbol{R} = \begin{matrix} .96 & .04 & 0 \\ .02 & .95 & .03 \\ 0 & .06 & .94 \end{matrix}$$

Thus, if the red dice are 6-sided, there is a 2% chance Rose will change them to 4-sided, a 3% chance Rose will change them to 8-sided, and a 95% chance Rose will leave them alone.

Similarly, assume that Gwen's replacement of individual green dice follows the transition matrix $\boldsymbol{G}$:

$$\boldsymbol{G} = \begin{matrix} .60 & .40 & 0 \\ .20 & .50 & .30 \\ 0 & .60 & .40 \end{matrix}$$

Gwen is ten times as likely to switch dice as is Rose.[62] Thus, the parameters of the green dice shift more swiftly than those of the red dice.[63] The dominant eigenvalue[64] (other than unity) of the transpose of $\boldsymbol{R}$ is $\rho = .954$, with a half-life of 15 trials. The dominant eigenvalue[65] (other than unity) of the transpose of $\boldsymbol{G}$ is $\gamma = .537$ with a half-life of 1.1 trials. The transition matrices $\boldsymbol{G}$ and $\boldsymbol{R}$ have been chosen such that they each have the same stationary distribution:[66] .25, .50, and .25.

[62]We have chosen this simple relation for illustrative purposes. Gwen could switch dice at any rate relative to Rose.

[63]One could just as easily model the reverse situation.

[64]The three eigenvalues of $\boldsymbol{R}$ are 1, .954 and .896.

[65]The three eigenvalues of $\boldsymbol{G}$ are 1, .537 and $-.037$.

[66]One could model a somewhat more complicated situation where the green and red dice had different stationary distributions.

For now take the simplest case in which Joe rolls a single die of each color, $N = 1$. (The next section will deal with the more general case of $N \geq 1$.)

As shown in Mahler [1], the covariance of trials X_i and X_j for either a single red or green die is given by Equation 3.12:

$$\mathrm{Cov}[X_i, X_j] = \sum_{k>1} \zeta_k \lambda_k^{|i-j|} + \delta_{ij}\eta^2$$

where η^2 is the Expected Value of the Process Variance, $\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ij} = 1$ for $i = j$, λ_k are the eigenvalues of the transition matrix and the ζ_k are a function of the transition matrix and the means of the different dice.[67]

For transition matrix $\boldsymbol{R}$:

k	λ_k	ζ_k
1	1	12.25
2	.954	.4676
3	.896	.0324

For transition matrix $\boldsymbol{G}$:

k	λ_k	ζ_k
1	1	12.25
2	.537	.4676
3	−.037	.0324

Note that since we have chosen the same basic structure for the shifting of the green and red dice the ζ values are the same. Also note that $\sum_{k>1} \zeta_k = .5$ = Variance of the Hypothetical Means in the absence of shifting risk parameters. The eigenvalues are different, reflecting the different rates of shifting parameters.

In this case the expected value of the process variance = η^2 = 3.08. Thus, for the red dice the covariance between trials of data

[67] The dice in this case are the different states of the Markov chain. See Mahler [1].

is:

$$\mathrm{Cov}[Y_i, Y_j] = (.4676)(.954^{|i-j|}) + (.0324)(.896^{|i-j|}) + 3.08\,\delta_{ij}. \tag{4.26}$$

For the green dice the covariance between trials of data is:

$$\mathrm{Cov}[W_i, W_j] = (.4676)(.537^{|i-j|}) + (.0324)((-.037)^{|i-j|}) + 3.08\,\delta_{ij}. \tag{4.27}$$

Equation 4.27 can be approximated as:

$$\mathrm{Cov}[W_i, W_j] \approx (.5)(.537)^{|i-j|} + 3.08\,\delta_{ij}. \tag{4.28}$$

Similarly, Equation 4.26 can be approximated as:[68]

$$\mathrm{Cov}[Y_i, Y_j] \approx (.5)(.954)^{|i-j|} + 3.08\,\delta_{ij}. \tag{4.29}$$

Equations 4.28 and 4.29 are each of the form given by Equation 3.14:

$$\mathrm{Cov}[X_i, X_j] \approx \tau^2 \lambda^{|i-j|} + \eta^2\,\delta_{ij}. \tag{4.30}$$

In both cases the Variance of the Hypothetical Means[69] $= \tau^2 = .5$ while the Expected Value of the Process Variance $= \eta^2 = 3.08$.

Let Y_i = result of a red die, W_i = result of a green die, and $X_i = (1-h)Y_i + hW_i$ = result of a trial (for one die of each kind). Then

$$\begin{aligned}\mathrm{Cov}[X_i, X_j] &= \mathrm{Cov}[(1-h)Y_i + hW_i, (1-h)Y_j + hW_j] \\ &= (1-h)^2\,\mathrm{Cov}[Y_i, Y_j] + (1-h)h\,\mathrm{Cov}[Y_i, W_j] \\ &\quad + (1-h)h\,\mathrm{Cov}[W_i, Y_j] + h^2\mathrm{Cov}[W_i, W_j].\end{aligned}$$

[68]Depending on the particular example, putting the covariance in terms of the principal eigenvalue other than unity will represent more or less of an approximation. For example, for the green dice, the approximate covariances from Equation 4.28 for separations of 1, 2, and 3 trials are .2685, .1442, and .0774. These compare to the exact covariances from Equation 4.27 of .2499, .1349 and .0724. On the other hand, the approximation of Equation 4.26 by Equation 4.29 is an example where the approximate covariances are close to the exact covariances.

[69]In the absence of shifting risk parameters.

However, the green and red die are independent of each other, so that

$$\mathrm{Cov}[Y_i, W_j] = \mathrm{Cov}[W_i, Y_j] = 0.$$

Thus, $\mathrm{Cov}[X_i, X_j] = (1-h)^2 \mathrm{Cov}[Y_i, Y_j] + h^2 \mathrm{Cov}[W_i, W_j]$.

$$\mathrm{Cov}[X_i, X_j] \approx (1-h)^2(.5)(.954^{|i-j|}) + h^2(.5)(.537^{|i-j|}) + 3.08(1-h)^2\,\delta_{ij} + 3.08h^2\,\delta_{ij}. \qquad (4.31)$$

In general Equation 4.31 can be written as:

$$\mathrm{Cov}[X_i, X_j] \approx (1-h)^2\tau_1^2\rho^{|i-j|} + h^2\tau_2^2\gamma^{|i-j|} + (1-h)^2\eta_1^2\,\delta_{ij+h^2}\eta_2^2\,\delta_{ij} \qquad (4.32)$$

where we have allowed for possibly different values of the variance of the hypothetical means[70] τ_1^2 and τ_2^2, as well as possibly different values of the expected value of the process variance η_1^2 and η_2^2, for the "red" and "green" risk processes.

4.10. Behavior by Size of Risk with Risk Heterogeneity and Shifting Risk Parameters

According to Section 4.1, the green and red dice contribute differently to the Variance of the Hypothetical Means and to the covariances as the number of dice N increases. For the sum of N identical red dice, the VHM is $.5N^2 = N^2\tau_1^2$. For the sum of N possibly different green dice, the VHM is $.5N = N\tau_2^2$. In both case the EPV $= N\eta^2 = 3.08N$.

Thus, for N dice, Equation 4.32 becomes:

$$\mathrm{Cov}[X_i, X_j] = (1-h)^2N^2\tau_1^2\rho^{|i-j|} + h^2N\tau_2^2\gamma^{|i-j|} + (1-h)^2N\eta_1^2\,\delta_{ij} + h^2N\eta_2^2\,\delta_{ij}. \qquad (4.33)$$

---

[70] In the absence of shifting risk parameters.

For insurance applications to frequency, pure premiums, etc., it will be useful to rewrite Equation 4.33 as:[71]

$$\mathrm{Cov}[X_i, X_j] = r^2\{\rho^{|i-j|} + (I/E)\gamma^{|i-j|} + (K/E)\delta_{ij}\},$$
$$1 \geq \rho, \gamma \geq 0 \quad I, K \geq 0. \qquad (4.34)$$

Equation 4.34 for the covariances in the presence of shifting risk parameters and risk heterogeneity combines the features of Equation 3.16 with shifting risk parameters and Equation 4.3 with risk heterogeneity.

Equation 4.33 displays the typical behavior in the presence of risk heterogeneity ($h > 0$); there is a piece of the variance of hypothetical means that increases as N^2 and a piece that increases only as N, the size of risk. Therefore, the relative importance of the two dominant eigenvalues ρ and γ varies by size of risk N. For N large, ρ is relatively more important than for N small. Thus for large size risks the log-correlations decline at a rate of approximately ρ. For medium size risks, the decline rate will be between ρ and γ. For very small risks, the decline rate should be approximately γ. This same behavior also holds for Equation 4.34.

For the dice example, $\rho = .954$ and $\gamma = .537$, thus larger risks should have their log-correlations decline approximately with a slope of $\ln .954$,[72] while smaller risks would see their log-correlations decline more quickly. For $h = .8$, Figure 5 shows the behavior for various sizes of risk. The correlations are both smaller for fewer numbers of dice and decline more quickly as the separation of years increases.

For this example, plugging into Equation 4.33, the values $h = .8$, $\tau_1^2 = \tau_2^2 = .5$, $\eta_1^2 = \eta_2^2 = 3.08$, we obtain:

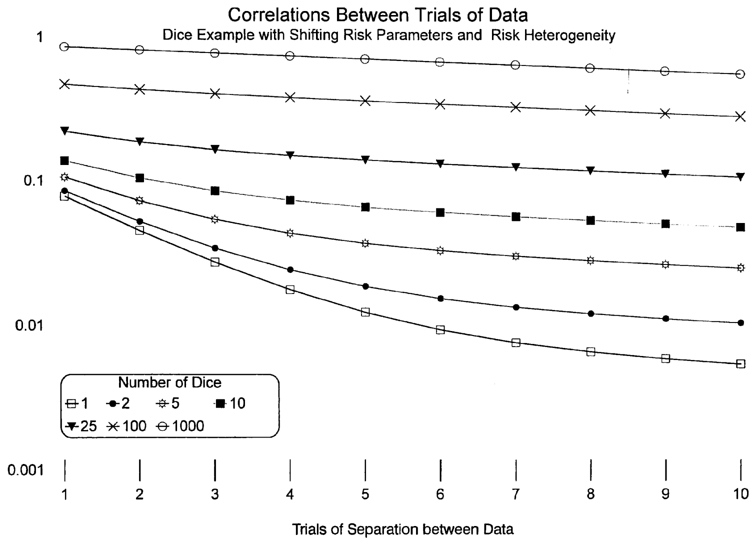
$$\mathrm{Cov}[X_i, X_j] = .02N^2 .954^{|i-j|} + .32N .537^{|i-j|} + 2.0944N\,\delta_{ij}.$$

---

[71] Where as before E is the size of risk, I quantifies risk heterogeneity and K is the Bühlmann credibility parameter. ρ and γ quantify the rate(s) of shifting of risk parameters.

[72] The correlation declines approximately as $.954^{|i-j|}$, thus, its log declines approximately as $|i-j|(\ln .954)$.

FIGURE 5



Note h=.8, indicating a large impact from risk heterogeneity. The two dominant eigenvalues are .954 and .537.

Thus, $\mathrm{Var}[X] = \mathrm{Cov}[X,X] = .02N^2 + 2.4144N$.

Thus, for this example the correlations between years are given by:

$$\mathrm{Corr}[X_i, X_j] = \frac{.02N^2 .954^{|i-j|} + .32N .537^{|i-j|}}{.02N^2 + 2.414N}, \qquad i \neq j. \tag{4.35}$$

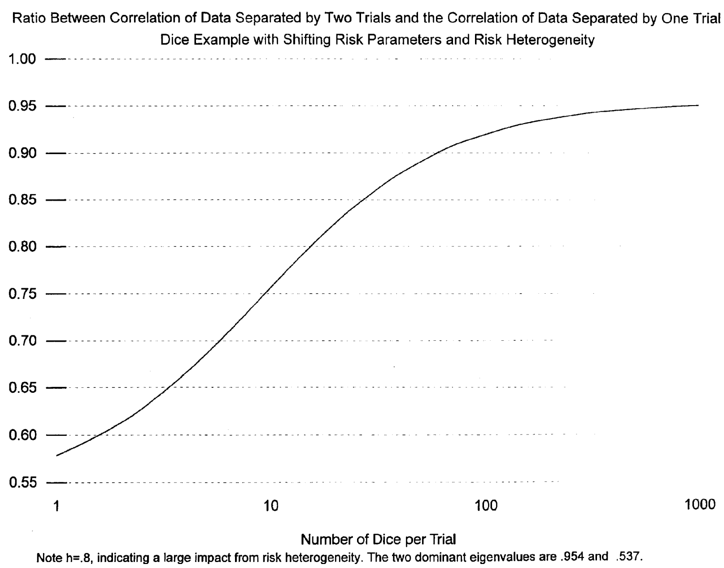
Figure 6 shows the ratio of $\mathrm{Corr}[X_i, X_i + 2]$ to $\mathrm{Corr}[X_i, X_i + 1]$.[73] As the number of dice increases this ratio gets closer to $\rho = .954$. In this example, larger risks have less quickly shifting risk parameters over time.[74]

---

[73]Figure 6 shows the approximation given by Equation 4.35. The more exact results that would be obtained starting with Equations 4.26 and 4.27 including terms for all the eigenvalues, would display the same behavior.

[74]If the transition matrices for Gwen and Rose had been reversed, then the larger risks would have had more quickly shifting risk parameters than smaller risks.

FIGURE 6



For general insurance applications, Equation 4.35 would become for the covariances written as in Equation 4.34:

$$\mathrm{Corr}[X_i, X_j] = \frac{E\rho^{|i-j|} + I\gamma^{|i-j|}}{E + I + K}, \qquad i \neq j. \tag{4.36}$$

In Equation 4.36, as $E \to \infty$, $\mathrm{Corr}[X_i, X_j] \to \rho^{|i-j|}$, while as $E \to 0$, $\mathrm{Corr}[X_i, X_j] \to \gamma^{|i-j|} I/(I + K)$. As will be discussed subsequently, examining the behavior of the correlations between years of data as the separation between years and the size of risk vary will allow one to estimate the parameters of the covariance structure which are needed to calculate credibilities.

5. SHIFTING RISK PARAMETERS, RISK HETEROGENEITY, AND PARAMETER UNCERTAINTY

In this section, the effects of shifting risk parameters, risk heterogeneity and parameter uncertainty will be combined. In

Section 5.1, the three phenomena will be combined for the dice example. The model will be put into a form useful for insurance applications in Section 5.2. Section 5.3 will incorporate the previously discussed refinement to the covariance structure for very small risks. Section 5.4 will discuss all three phenomena in the context of Philbrick's target shooting example.

5.1. All Three Phenomena, Dice Example

Combining the examples in Sections 3.6, 4.4, and 4.9 we can incorporate shifting risk parameters, risk heterogeneity, and parameter uncertainty.

Joe initially selects N identical red dice and N possibly different green dice. Prior to each trial, Rose may alter the type of all the red dice. Prior to each trial, Gwen may alter the type of one or more of the green dice; Gwen acts independently on each green die.

For each trial Joe rolls all the dice and Mary flips a coin. The result of a trial is:

$$(1-h)(\text{Sum of } N \text{ Red Dice}) + h(\text{Sum of } N \text{ Green Dice}) + N(\text{Result of Coin Flip})$$

where the coin flip is counted as $-\frac{1}{2}$ for tails and $\frac{1}{2}$ for heads.

The presence of the coin flip does not alter the hypothetical means. However, as in Section 4.4, the Expected Value of the Process Variance is $(3.08)(1-h)^2N + (3.08)h^2N + .25N^2$. Combining this with the Variance of the Hypothetical Means from Section 4.10, the covariance between the results of trials i and j is:

$$\mathrm{Cov}[X_i, X_j] = (1-h)^2N^2(.5).954^{|i-j|} + h^2N(.5).537^{|i-j|} + (1-h)^2N(3.08)\delta_{ij} + h^2N(3.08)\delta_{ij} + (.25)N^2\delta_{ij}. \quad (5.1)$$

Equation 5.1 can be written more generally as:

$$\mathrm{Cov}[X_i, X_j] = (1-h)^2 N^2 \tau_1^2 \rho^{|i-j|} + h^2 N \tau_2^2 \gamma^{|i-j|} + (1-h)^2 N \eta_1^2 \delta_{ij} + h^2 N \eta_2^2 \delta_{ij} + u^2 N^2 \delta_{ij}. \tag{5.2}$$

In insurance we normally are interested in quantities such as claim frequency[75] or pure premium,[76] which have the volume of data in the denominator. This introduces a factor of $1/\text{volume}^2$ into the variances and covariances.

In the dice example, this would be the equivalent of the result of a trial being the previously defined "result of a trial" divided by N:

$$\frac{1}{N}\left\{\begin{array}{l}(1-h)(\text{Sum of } N \text{ Red Dice}) + h(\text{Sum of } N \text{ Green Dice}) \\ +N(\text{Result of Coin Flip})\end{array}\right\}.$$

In that case, Equation (5.2) is modified to:

$$\mathrm{Cov}[X_i, X_j] = (1-h)^2 \tau_1^2 \rho^{|i-j|} + h^2 \tau_2^2 \gamma^{|i-j|}/N + (1-h)^2 \eta_1^2 \delta_{ij} + h^2 \eta_2^2 \delta_{ij} + u^2 \delta_{ij}. \tag{5.3}$$

There are those portions of the covariance that are independent of size of risk and those portions such as the process variance which decline with size of risk, when dealing with claim frequencies, pure premiums, etc.

5.2. General Form of Covariances, All Three Phenomena

Equation 5.3 contains four different types of terms. There are those that decrease as the inverse of the size of risk N and those that do not depend on N. There are those involving δ_{ij} that are related to the process variance and are not present in the covariance between different years. On the other hand, there

[75] Frequency = claims/exposures.

[76] Pure Premiums = losses/exposures.

are those involving $\lambda^{|i-j|}$ that are related to the variance of the hypothetical means.

In specific examples, the key elements will be the speed with which parameters shift and thus the half-lives of ρ and γ, and the relative weights of each of the four types of terms. With this in mind it will be worthwhile to rewrite Equation 5.3. Let $r^2 = (1-h)^2\tau_1^2$, $g^2 = h^2\tau_2^2$, $e^2 = (1-h)^2\eta_1^2 + h^2\eta_2^2$, and rather than N use E as some appropriate measure of size of risk.[77]

Then Equation 5.3 becomes:

$$\begin{aligned} \mathrm{Cov}[X_i, X_j] &= r^2\rho^{|i-j|} + g^2\gamma^{|i-j|}/E + \delta_{ij}(e^2/E + u^2) \\ \mathrm{Var}[X] = \mathrm{Cov}[X, X] &= r^2 + g^2/E + e^2/E + u^2. \end{aligned} \tag{5.4}$$

As before letting $I = g^2/r^2$, $J = u^2/r^2$ and $K = e^2/r^2$, then

$$\mathrm{Cov}[X_i, X_j] = r^2\{\rho^{|i-j|} + \gamma^{|i-j|}(I/E) + (J + K/E)\delta_{ij}\}. \tag{5.5}$$

Thus, the correlations are:

$$\mathrm{Corr}[X_i, X_j] = \frac{E\rho^{|i-j|} + I\gamma^{|i-j|}}{E(1+J) + K + I}. \tag{5.6}$$

For large risks the term with $\rho^{|i-j|}$ will dominate, while for small risks the term with $\gamma^{|i-j|}$ will dominate. For large risks the log-correlations will decline as ρ, while for small risks the log-correlations will decline as γ. For risks of medium size the decline will be between ρ and γ.

Thus, this model will be particularly useful when and if there are different decline rates in correlations by size of risk.[78] ρ can be estimated from the slopes for large risks of the log-correlations versus separations. γ can be estimated from the slopes for small risks of the log-correlations. The size of I can be

---

[77] For example, E could be expected losses in workers compensation experience rating.

[78] Where the rate of decline in the correlations is not dependent on size of risk, one can set $\rho = \gamma$.

estimated by what constitutes a "medium-size risk," where the decline rate of the log covariances are about halfway between ρ and γ. At that size $E \approx I$.

As we take larger and larger risks, Equation 5.6 for the correlations approaches

$$\lim_{E\to\infty} \mathrm{Corr}[X_i, X_j] = \frac{\rho^{|i-j|}}{1+J}.$$

Thus, we can estimate J, quantifying the impact of parameter uncertainty, by examining for large risks the correlations between years. For example, if we fit an exponential regression to such correlations versus the separations, then the intercept can be used to estimate J. For large risks:

$$\ln \mathrm{Corr}[X_i, X_j] \approx -\ln(1+J) + |i-j| \ln \rho.$$

For any size:

$$\ln \mathrm{Corr}[X_i, X_j] = \ln(E\rho^{|i-j|} + I\gamma^{|i-j|}) - \ln(E(1+J) + K + I).$$

Assuming a fixed set of parameters I, J, K, ρ and γ, then for a fixed size of risk E, the second term is constant, while the first term depends on the separation between years $|i-j|$. We expect the decline rate to be some rate between ρ and γ, depending on the relative sizes of E and I. Very approximately:[79]

$$\ln(E\rho^{|i-j|} + I\gamma^{|i-j|}) \approx |i-j| \ln\left(\frac{E\rho + I\gamma}{E+I}\right) + \ln(E+I)$$

Thus,

$$\ln \mathrm{Corr}[X_i, X_j] \approx |i-j| \ln\left(\frac{E\rho + I\gamma}{E+I}\right) + \ln\left(\frac{E+I}{E(1+J)+K+I}\right).$$

Thus, if we fit an exponential least squares regression to the correlations by separations ≥ 1, we would expect to have a slope between ρ and γ and an "intercept" of $(E+I)/(E(1+J)+K+I)$.

[79] For $E = 0$, $\ln(I\gamma^{|i-j|}) = |i-j|\ln\gamma + \ln I$. For $I = 0$, $\ln(E\rho^{|i-j|}) = |i-j|\ln\rho + \ln E$. For $|i-j| = 1$, the approximation is exact. The approximation is poor when $|i-j|$ is large, ρ and γ differ substantially, and E is approximately the same as I.

This intercept[80] is equal to the credibility for a single year of data in the absence of shifting risk parameters, as in Equation 4.15.

We can therefore approximate some of the necessary parameters from the behavior of the observed correlations as the size of risk and number of years of separation vary.

For each of various sizes of risk we can fit exponential least squares regressions to the correlations for years separated by one year or more. The intercept for each size category is an estimate of the credibility of one year of data in the absence of shifting risk parameters over time. These credibilities by size of risk can be used to estimate the parameters I, J and K. The slope (exponential rate of decline) of the correlations varies between γ and ρ as the size of risk increases. At an intermediate size of about I, the slope will be about halfway between γ and ρ.

In the situation where the years X_i and X_j have different expected volumes of data E_i and E_j, Equation 5.5 can be generalized to:

$$\mathrm{Cov}[X_i, X_j] = r^2 \left\{ \rho^{|i-j|} + \gamma^{|i-j|} I / \sqrt{E_i E_j} + \left(J + K / \sqrt{E_i E_j} \right) \delta_{ij} \right\} \tag{5.7}$$

In the covariance, those terms that were divided by E in Equation 5.5 are now in Equation 5.7 divided by the geometric average of the sizes of risk, $\sqrt{E_i E_j}$. If $E_i = E_j = E$, then $\sqrt{E_i E_j} = E$, so that Equation 5.7 would reduce to Equation 5.5. The use of the square root function in the generalization was motivated by the $\sqrt{\mathrm{VAR}[X_1]\mathrm{VAR}[X_2]}$ that appears in the denominator of the correlation of X_1 and X_2.

Equations 5.5 or 5.7 can be used to calculate all of the covariances necessary to solve Equations 2.4 for the credibilities.

An example of how to calculate the credibilities in general will be given in Section 6. However, prior to that it is worthwhile to

[80]For convenience in this paper, $(E + I)/(E(I + J) + K + J)$, rather than the natural log of that quantity, will be referred to as the intercept.

generalize Equation 5.7 for the covariance in order to take into account the different behavior of very small risks.

5.3. Very Small Risks, General Covariance Structure

In Sections 4.3 and 4.6 a refinement for very small sizes of risk was introduced. In this section, this refinement will be introduced into the general covariance structure.

The same logic concerning risk heterogeneity and very small risks applies as well when both parameter uncertainty and shifting risk parameters are considered. If we assume risk heterogeneity does not apply for $E \leq \Omega$, then Equation 5.4 for the covariances is split into two separate equations, per Equations 4.22 and 4.23.

For $E \geq \Omega$, Equation 5.5 holds:

$$\text{Cov}[X_i, X_j] = r^2\{\rho^{|i-j|} + \gamma^{|i-j|}(I/E) + \delta_{ij}((K/E) + J)\}, \quad E \geq \Omega. \tag{5.8}$$

For $E \leq \Omega$, the term involving I takes on its value at $E = \Omega$:

$$\text{Cov}[X_i, X_j] = r^2\{\rho^{|i-j|} + \gamma^{|i-j|}(I/\Omega) + \delta_{ij}((K/E) + J)\}, \quad E \leq \Omega. \tag{5.9}$$

In the situation where the years X_i and X_j have different expected volumes E_i and E_j, Equations 5.8 and 5.9 can be generalized to:[81]

$$\text{Cov}[X_i, X_j] = r^2\left\{\rho^{|i-j|} + \gamma^{|i-j|} I/\sqrt{E_i E_j} + \delta_{ij}\left(K/\sqrt{E_i E_j} + J\right)\right\}, \quad \sqrt{E_i E_j} \geq \Omega \tag{5.10}$$

[81] It should be noted that in Equations 5.10 to 5.11 the expression $\sqrt{E_i E_j}$ only enters due to the presence of risk heterogeneity. This results in terms such as $I/\sqrt{E_i E_j}$. In contrast, where $\sqrt{E_i E_j}$ divides K it is multiplied by δ_{ij}. These terms are zero unless $i = j$, so $\sqrt{E_i E_j}$ could be replaced in these terms by either E_i or E_j. This simplification in notation is conventional in the absence of risk heterogeneity.

$$\mathrm{Cov}[X_i, X_j] = r^2 \left\{ \rho^{|i-j|} + \gamma^{|i-j|} I/\Omega + \delta_{ij} \left(K/\sqrt{E_i E_j} + J \right) \right\}, \qquad \sqrt{E_i E_j} \le \Omega. \tag{5.11}$$

5.4. Philbrick's Target Shooting Example

Philbrick [5] explains credibility concepts by using a target shooting example. There are four marksmen each shooting at his own target. Each marksman's shots are assumed to be distributed around his target, with expected mean equal to his target. Once we observe a shot or shots from a single *unknown* marksman, we could use Bühlmann credibility to estimate the location of the next shot from the *same* marksman.

The key features of Bühlmann credibility are explained by Philbrick as follows by altering the initial conditions of the target shooting example:

Feature of Target Shooting Example	Mathematical Quantification	Bühlmann Credibility
Better Marksmen	Smaller EPV	Larger
Targets Further Apart	Larger VHM	Larger
More Shots	Larger N	Larger

These mathematical relationships also follow from Bühlmann's credibility formula, Equation 1.1.

We can modify the example in Philbrick to include each of the three phenomena discussed in this paper.

In Philbrick, it is assumed that each marksman continues to shoot at his target.[82] Within a single example in Philbrick, the risk parameters do not shift over time. If instead there were a small random chance that between each shot a marksman would

[82]It is also assumed within an example that the targets are stationary, the marksmen remain the same and do not get better or worse, nor do the marksmen move closer to or further from the targets.

switch targets, then one would have shifting risk parameters over time.[83] In this case, the credibility assigned to a given shot would be less than if each marksman always shot at the same target. The informational content of a shot for purposes of predicting the next shot from the same marksman has been reduced by the presence of shifting parameters over time.

The Philbrick example can also be altered in order to incorporate risk heterogeneity. Assume we have teams of marksmen. Assume each marksman on a team shoots at his own target. Assume that while members of a team each shoot at a possibly different target, the members of a team are more likely to shoot at the same target than are marksmen who are not members of the same team. For example, the six members of Team 1 might shoot at targets A, A, A, B, C, and D respectively. For the purpose of predicting the next shot, the informational content of a given number of shots from Team 1 is less than if all the members of the team always had the same target. Risk heterogeneity has reduced the credibility assigned to a given number of shots.

Assume, for example, as the teams got bigger each additional marksman in Team 1 was assigned target A half the time and targets B, C, and D one-sixth of the time. Then as the teams got bigger, the credibility assigned to a set of shots, one per team member, would not be the same as the Bühlmann case in which each team member shot at the same target. With risk heterogeneity the credibility would increase more slowly as the teams increase in size; the incremental informational content of another team member is less when they do not all shoot at the same target.

As discussed previously, in the presence of risk heterogeneity, the credibilities are given by Equation 4.4:

$$Z = \frac{E + I}{E + I + K}.$$

[83] This is analogous in the dice example to Beth possibly replacing dice between the rolls.

The derivative of Z with respect to the size of risk E is $K/(E+I+K)^2$. This derivative decreases as I increases; the greater the impact of risk heterogeneity, the more slowly the credibility increases with size of risk.

Finally, the Philbrick example can be altered to incorporate parameter uncertainty. Again assume that there are teams of marksmen, but each marksman shoots at the same target. Assume that for each round of shots, one per team member, every team member uses the same rifle. However, between rounds the rifle is replaced by another. Further assume the rifles look alike but some shoot high, some shoot low and to the left, etc. Also assume the marksmen on a team do not communicate with each other, nor adjust their aim based on their teammate's shots, so that all team members are equally affected by the peculiarities of the given rifle. The errors introduced by the switching rifles reduce the informational content of the shots; in the presence of parameter uncertainty less credibility is assigned to the data, holding all else equal.

In addition, adding more team members can never eliminate the effect of an individual, randomly chosen rifle. In the presence of parameter uncertainty the credibility of a single year of data does not approach unity as the risk size increases; rather in Equation 3.7 the credibility goes to $1/(1+J)$ as the risk size approaches infinity.

However, by observing many rounds of shots, assuming the errors of the rifles average to zero, one can eliminate their impact. In the presence of parameter uncertainty (and no shifting risk parameters over time), the credibility of a given size of risk goes to unity as the number of years goes to infinity; the credibilities in Equation 3.10 go to unity as the number of years increases.

Clearly, we could modify the Philbrick target shooting example to incorporate two or all three of the phenomena discussed in this paper.

6. ILLUSTRATIVE EXAMPLES OF CALCULATING CREDIBILITIES

This section will present illustrative examples of calculating credibilities based on the general covariance structure presented in Section 5. Section 6.1 deals with large risks, while Section 6.2 includes the refinement to the covariance structure for very small risks. Section 6.3 shows how differing volumes of data by year would affect the credibilities. Section 6.4 shows an example in which no weight is given to the overall mean.

6.1. An Example of Calculating Credibilities, Large Risks

As an example, take the following illustrative values in Equations 5.5 or 5.7 for the covariances in the presence of all three phenomena:[84]

$\rho = .9$ (rate of shifting parameters related to risk homogeneity),

$\gamma = .7$ (rate of shifting parameters related to risk heterogeneity),

$e^2 = 9{,}000$ (expected value of process variance without parameter uncertainty),

$u^2 = 2$ (variance related to parameter uncertainty),

$r^2 = 3$ (portion of variance of hypothetical means related to risk homogeneity),

$g^2 = 4{,}000$ (portion of variance of hypothetical means related to risk heterogeneity),

$I = g^2/r^2 = 1{,}333$,

$J = u^2/r^2 = .6667$, and

$K = e^2/r^2 = 3{,}000$.

[84]These values were chosen solely to present an example. Note that if one multiplies e^2, u^2, r^2 and g^2 all by the same constant, then all the covariances are multiplied by that same constant, but the credibilities are unchanged.

Assuming each year of data has equal volume E, Equation 5.5 becomes:

$$\text{Cov}[X_i, X_j] = (3).9^{|i-j|} + (4{,}000/E).7^{|i-j|} + \delta_{ij}(9{,}000/E + 2). \qquad (6.1)$$

Thus, the variance is:

$$\text{Var}[X_i] = \text{Cov}[X_i, X_i] = (13{,}000/E) + 5.$$

The covariance between years of data separated by two years is:

$$\text{Cov}[X_1, X_3] = (1{,}960/E) + 2.43.$$

For 4 years of data each of volume E, the variance-covariance matrix is:

$(13{,}000/E) + 5$	$(2{,}800/E) + 2.7$	$(1{,}960/E) + 2.43$	$(1{,}372/E) + 2.187$
$(2{,}800/E) + 2.7$	$(13{,}000/E) + 5$	$(2{,}800/E) + 2.7$	$(1{,}960/E) + 2.43$
$(1{,}960/E) + 2.43$	$(2{,}800/E) + 2.7$	$(13{,}000/E) + 5$	$(2{,}800/E) + 2.7$
$(1{,}372/E) + 2.187$	$(1{,}960/E) + 2.43$	$(2{,}800/E) + 2.7$	$(13{,}000/E) + 5.$

For example, if $E = 1000$ then the variance-covariance matrix is:

18	5.5	4.39	3.559
5.5	18	5.5	4.39
4.39	5.5	18	5.5
3.559	4.39	5.5	18.

Assume we are using three years of data to estimate the fourth year directly following them. Then Equations 2.4 for the credibilities to assign to each of the three years of data are:

$$18Z_1 + 5.5Z_2 + 4.39Z_3 = 3.559,$$
$$5.5Z_1 + 18Z_2 + 5.5Z_3 = 4.39, \qquad \text{and} \qquad (6.2)$$
$$4.39Z_1 + 5.5Z_2 + 18Z_3 = 5.5.$$

Equations 6.2 are three linear equations in three unknowns, with solution:

$$Z_1 = 9.62\%,$$
$$Z_2 = 14.15\%, \qquad \text{and}$$
$$Z_3 = 23.88\%,$$

where Z_1 is the credibility assigned to the oldest year of data and Z_3 is the credibility assigned to the most recent year of data. Note that $Z_1 + Z_2 + Z_3 = 47.65\% < 100\%$. The remaining weight of 52.35% is given to the grand mean.[85]

It should be noted that Equations 2.4 for the credibilities[86] were derived so as to minimize the expected squared error of the estimate. As derived in Mahler [1][87] the expected squared difference between the estimate and observation as a function of the variance-covariance matrix and the credibilities is:

$$V(Z) = \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j C_{ij} - 2\sum_{i=1}^{Y} C_{i,Y+\Delta} Z_i + C_{Y+\Delta,Y+\Delta}. \quad (6.3)$$

In this particular case for $E = 1{,}000$, we get for various selected values of the credibilities the following expected squared errors:

Z_1	Z_2	Z_3	$V(Z)$
0	0	0	18
1/3	1/3	1/3	18.454
1/2	0	0	18.941
0	1/2	0	18.110
0	0	1/2	17.000
9.62%	14.15%	23.88%	15.722

Thus, the use of (the optimal least squares) credibilities of 9.62%, 14.15%, 23.88% does indeed seem to have reduced the expected squared errors.[88]

Figure 7 shows how the sum of the credibilities for three years of data varies with size of risk. In addition to the case where all

---

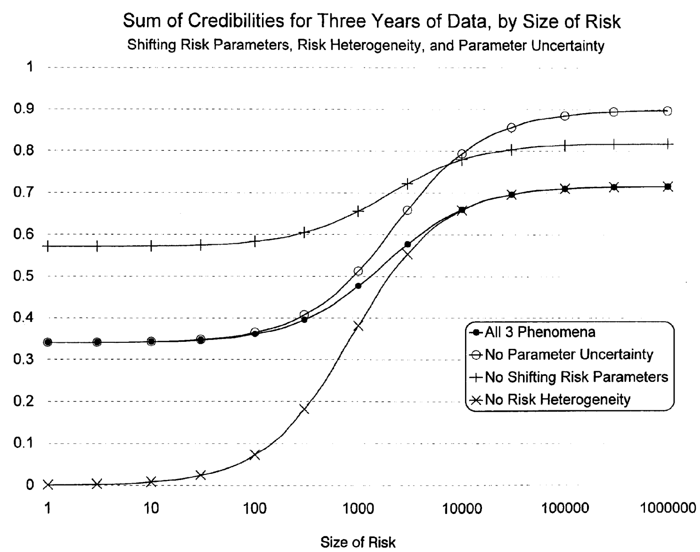
[85]The situation in which no weight is given to the grand mean is discussed below.

[86]Which are Equations 6.2 for this specific example with $E = 1{,}000$.

[87]See Appendix C in Mahler [1]. The derivation parallels that in Appendix B of the current paper.

[88]In this case, the expected squared error is about $15.722 \div 18 = 87\%$ of what one would obtain by ignoring the observations (assigning the observations zero credibility).

FIGURE 7



No gap between data and year to be estimated. See the example in Section 6.1.

three phenomena are present, cases are shown in which only two of the phenomena are present.

For no parameter uncertainty, J is set equal to zero rather than .6667. For large risks the credibility is higher than in the presence of parameter uncertainty. Nevertheless, the maximum credibility is less than 100%, due to the impact of shifting risk parameters over time.

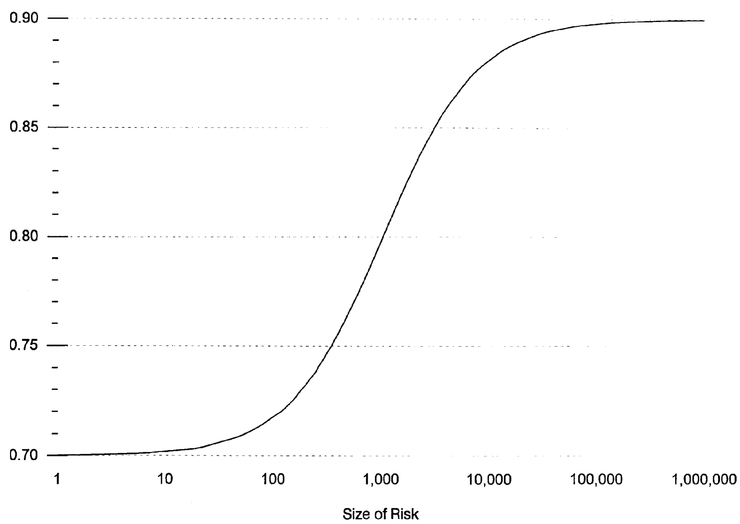
For no shifting risk parameters, $\rho = \gamma = 1$ rather than $\rho = .9$ and $\gamma = .7$. Credibilities are higher. The credibilities are given by Equation 4.14.

For no risk heterogeneity, I is set equal to zero rather than 1333. With risk homogeneity the credibilities go to zero as the risk size declines.[89]

[89] As discussed in Section 5.3, Equation 5.5 and the resulting credibilities are not appropriate for very small risks.

FIGURE 8

Ratio Between Correlation of Data Separated by Two Years and the Correlation of Data Separated by One Year



The two rates of decline of the log correlations are: $\rho = .9$ and $\gamma = .7$. See the example in Section 6.1.

The decline rate of the correlations is close to $\rho = .9$ for large risks and close to $\gamma = .7$ for small risks. Specifically, the ratio of the correlation between years separated by two years to the correlation between years separated by one year is:

$$\mathrm{Corr}[X_1, X_3]/\mathrm{Corr}[X_1, X_2] = (2.43E + 1{,}960)/(2.7E + 2{,}800). \quad (6.4)$$

Figure 8 shows how this decline rate given by Equation 6.4 varies by size of risk.

In general if the covariances are given by Equation 5.5, we expect this decline rate to be given by:

$$\mathrm{Corr}[X_1, X_3]/\mathrm{Corr}[X_1, X_2] = (\rho^2 E + I\gamma^2)/(\rho E + I\gamma). \quad (6.5)$$

If $\rho > \gamma$, then we expect to see something like Figure 8. If instead $\rho < \gamma$, we expect the curve to decrease from γ to ρ as the size increases.

The intermediate size at which the decline rate is about equally distant between ρ and γ is approximately I. This could be used to estimate I from data. In the example, $I = 1{,}333$. In Figure 8 for this size the decline rate is .81, roughly halfway between $\rho = .9$ and $\gamma = .7$.

6.2. Credibilities, Small Risks

In the example in Section 6.1, let us assume there is no risk heterogeneity for $E \leq \Omega = 100$. Then the covariances and credibilities are different for $E < 100$ than they were in Section 6.1.

For $E \leq 100$, the covariances are given by Equation 5.9:

$$\mathrm{Cov}[X_i, X_j] = (3)(.9^{|i-j|}) + (40)(.7^{|i-j|}) + \delta_{ij}(9{,}000/E + 2).$$

For $E \geq 100$, the covariances are given by Equation 5.8:

$$\mathrm{Cov}[X_i, X_j] = (3)(.9^{|i-j|}) + (4{,}000/E)(.7^{|i-j|}) + \delta_{ij}(9{,}000/E + 2).$$

For example, for $E = 10$, the variance-covariance matrix is:

945	30.7	22.03	15.907
30.7	945	30.7	22.03
22.03	30.7	945	30.7
15.907	22.03	30.7	945.

Assume we are using three years of data (each with $E = 10$), in order to estimate the fourth year directly following them. Then Equations 2.4 for the credibilities to assign to each of the three years of data are:

$$\begin{aligned} 945Z_1 + 30.7Z_2 + 22.03Z_3 &= 15.907, \\ 30.7Z_1 + 945Z_2 + 30.7Z_3 &= 22.03, \quad \text{and} \\ 22.03Z_1 + 30.7Z_2 + 945Z_3 &= 30.7. \end{aligned} \tag{6.6}$$

Equations 6.6 are three linear equations in three unknowns, with solutions:

$$Z_1 = 1.5\%, \quad Z_2 = 2.2\%, \quad \text{and} \quad Z_3 = 3.1\%$$

where Z_1 is the credibility assigned to the oldest year of data. The remaining weight not given to any of the years of data is given to the grand mean.

These credibilities assuming no risk heterogeneity below $E = 100$ are significantly smaller than those derived from Equation 5.5, which assumes risk heterogeneity for all sizes of risk. For $E = 10$, using Equation 5.5 to calculate the covariances rather than Equation 5.9 would result in credibilities of:

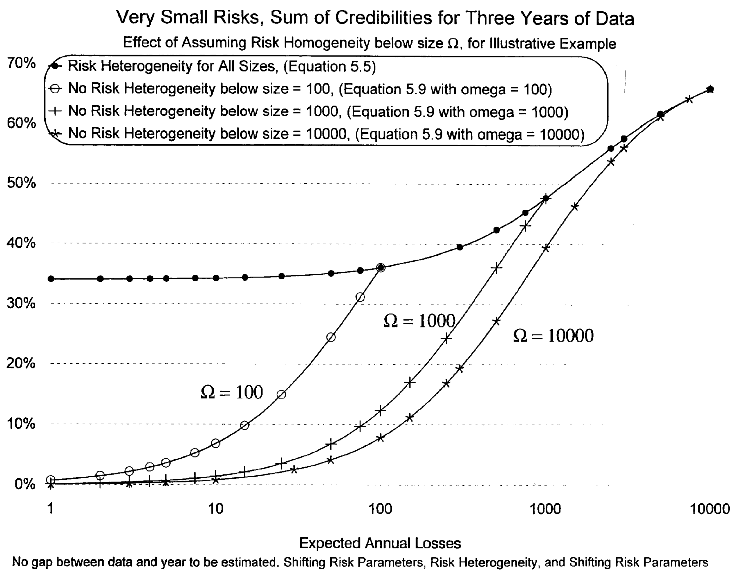
$$Z_1 = 5.7\%,$$
$$Z_2 = 9.9\%, \qquad \text{and}$$
$$Z_3 = 18.6\%.$$

Equation 5.9 produces credibilities that decline to zero as the risk size decreases in a manner similar to the usual Bühlmann formula, in contrast to Equation 4.14. Figure 9 contrasts this behavior for very small sizes, assuming $\Omega = 100$. Shown are the sum of the credibilities for three years of data as calculated above. For example, for $E = 10$, the credibilities for three years of data with risk heterogeneity sum to 34.2%, while those without risk heterogeneity (below $E = \Omega = 100$) sum to 6.8%. As E gets even smaller, in the presence of risk heterogeneity, the sum of the credibilities remains about 34%, while in the absence of risk heterogeneity it goes to zero.

Intuitively the credibility should approach zero as the size of risk approaches zero. Without the refinement discussed in Sections 4.3, 4.6 and 5.3, the covariance structure incorporating risk heterogeneity would produce credibilities that make no sense to actuaries. Credibility formulas such as Equation 4.14 or covariance structures such as Equation 5.5 should not be applied to very small risks.

Also shown in Figure 9 are the results of using Equation 5.9 with the alternate values $\Omega = 1{,}000$ or $\Omega = 10{,}000$ rather than $\Omega = 100$. In this case, the credibilities using the latter value

FIGURE 9



are significantly different than using either of the two former values.

In this example $I = 1{,}333$. This is the parameter related to risk heterogeneity, and it controls the behavior of the credibilities that result from Equation 5.5 for small risks. For $E < I$ these credibilities start leveling off significantly. Taking Ω significantly less than I, as for example 100 compared to 1,333, starts the steep descent to zero of the credibilities resulting from Equation 5.9 from an otherwise very small slope. In contrast, taking Ω either roughly equal to or greater than I, starts the descent in a much smoother manner, as is the case for $\Omega = 1{,}000$ or 10,000.

6.3. *Credibilities for Years with Differing Volumes of Data*

Returning to the example in Section 6.1, assume that the three years have differing volumes of data. Assume $E_1 = 600$, $E_2 = 1{,}600$, and $E_3 = 800$, where E_1 is the most distant of the three years. Using the inputs from before, Equation 5.7 becomes:

$$\mathrm{Cov}[X_i, X_j] = (3)(.9^{|i-j|}) + (4{,}000)(.7^{|i-j|})/\sqrt{E_i E_j} + \delta_{ij}\left(\left(9{,}000/\sqrt{E_i E_j}\right) + 2\right).$$

Assume that the year to be estimated will have a volume of data $E_4 = 1{,}000$, the average of the observed years.[90] Then the variance-covariance matrix is:

26.667	5.558	5.259	3.958
5.558	13.125	5.175	3.98
5.259	5.175	21.25	5.83
3.958	3.98	5.83	18

The credibilities are given by the solution to Equations 2.4:

$$Z_1 = 6.68\%,$$
$$Z_2 = 19.16\%, \quad \text{and}$$
$$Z_3 = 21.12\%.$$

Thus, as expected, years 1 and 3 with their smaller volumes are given less credibility than in Section 6.1, while year 2 with its larger volume of data is given more credibility than before.

It is interesting to note that in the presence of risk heterogeneity[91] the credibilities depend on the assumed volume of data for the year being estimated, year 4.

	$E_4 = 100$	$E_4 = 1{,}000$	$E_4 = 10{,}000$
Z_1	13.15%	6.68%	4.64%
Z_2	31.18%	19.16%	15.36%
Z_3	48.44%	21.12%	12.47%

[90] While $\mathrm{Var}[X_4]$ will not enter into the equations for the credibility, $\mathrm{Cov}[X_1, X_4]$ and similar terms will. $\mathrm{Cov}[X_1, X_4]$ depends on E_4, due to the presence of risk heterogeneity. In the absence of risk heterogeneity, one need not assume a value for E_4.

[91] Whether or not there are shifting risk parameters over time.

When E_4 is large, the covariances of the data years with the year to be estimated are smaller, and therefore we assign less credibility.[92]

As E_4 gets larger, we are assuming the insured will be larger in year 4, the year to be predicted. As discussed previously, one implication of risk heterogeneity is that larger insureds are in some sense more similar to average than are smaller insureds. The less distinct insureds are from average, the less credibility we give to the data from individual insureds and the more weight we give to the overall average.[93] Thus, if E_4 is larger, we give less credibility to this insured's data and more weight to the overall average.

For mechanical applications of the methodology,[94] we would probably just assume that the volume of data in the future would be some average of that observed in the recent past for that insured. In this example, we might assume as above that:

$$E_4 = (E_1 + E_2 + E_3)/3 = 1{,}000.$$

6.4. Credibilities, No Weight Given to the Grand Mean

So far we have assumed that the complement of credibility is given to the grand mean. In some cases the grand mean either does not exist or is not used. In those situations, we can have the credibilities be constrained to add to 100%.

Assume that we are using three years of data to estimate the fourth year directly after them, but that no weight is given to the grand mean. Then Equations 2.4 no longer apply.

---

[92] This differs from the Bühlmann case in which the covariances between the claim frequencies of different years are assumed to be independent of the size of risk.

[93] In the target shooting example in Philbrick [5], as the targets get closer together less credibility is given to each observed shot.

[94] For example, if one were performing many thousands of experience ratings by computer.

As shown in Appendix B, the general equations for credibility when no weight is applied to the grand mean are:[95]

$$\sum_{i=1}^{Y} \mathrm{Cov}[X_i, X_k] Z_i = \mathrm{Cov}[X_k, X_{Y+\Delta}] + \frac{\lambda}{2}, \qquad k = 1,\ldots,Y$$
$$\sum_{i=1}^{Y} Z_i = 1, \tag{6.7}$$

where λ is a Lagrange Multiplier.[96]

For the covariances used in the previous Section 6.1, with $E = 1{,}000$, Equations 6.7 become:

$$18Z_1 + 5.5Z_2 + 4.39Z_3 = 3.559 + \lambda/2,$$
$$5.5Z_1 + 18Z_2 + 5.5Z_3 = 4.39 + \lambda/2,$$
$$4.39Z_1 + 5.5Z_2 + 18Z_3 = 5.5 + \lambda/2, \qquad \text{and}$$
$$Z_1 + Z_2 + Z_3 = 1.$$

These are four linear equations in four unknowns.[97] The solution is:[98]

$$Z_1 = 27.60\%,$$
$$Z_2 = 30.53\%, \qquad \text{and}$$
$$Z_3 = 41.86\%.$$

We note that $Z_1 + Z_2 + Z_3 = 1$ as desired. The most recent year is given weight $41.86\% > 27.60\%$, the weight given to the most distant year.

[95] See Equation 11.7 in Mahler [20].

[96] The Lagrange Multiplier is introduced due to the constraint equation $\Sigma Z_i = 1$. Note that λ is used to denote the Lagrange Multiplier here and was used to denote the dominant eigenvalue in prior sections. λ is commonly used in both these roles, but the reader should not be confused. There is no connection between these two separate uses of the same Greek letter.

[97] Although we are really not particularly interested in the value of the Lagrange Multiplier.

[98] The Lagrange Multiplier $\lambda = 9.853$.

Usually, as the size of risk increases, the need for stability in the estimation procedure declines, so that we give more weight to recent years of data. However, in this case that is counteracted to some extent by the assumption that large risks have more stable risk parameters over time.[99] Thus the estimation procedure can afford to be less responsive.

In this example, this leads to the credibilities being relatively insensitive to the size of risk:

	Size of Risk		
	1	1,000	1 Million
Z_1	28.23%	27.60%	24.93%
Z_2	30.60%	30.53%	30.21%
Z_3	41.17%	41.86%	44.86%

If we switch the rates of shifting parameters and instead takes $\rho = .7$ and $\gamma = .9$, we get a significantly different behavior by size of risk:

	Size of Risk		
	1	1,000	1 Million
Z_1	30.32%	27.96%	21.96%
Z_2	32.34%	30.87%	25.81%
Z_3	37.34%	41.17%	52.23%

As risk size increases, the weight given to the recent year increases more substantially than before. In general, the dependence of credibility on size of risk will depend significantly on the relative magnitudes of ρ and γ.

[99]Larger risks correspond to a decline rate in the log-correlations of $\rho = .90$ rather than $\gamma = .70$.

7. CLASSIFICATION RATE RELATIVITIES

In this section, the ideas developed so far will be applied to a simplified version of the estimation of classification rate relativities.[100] While the example draws from workers compensation, it is intended to illustrate the general applicable concepts rather than the details of workers compensation insurance.

Section 7.1 defines rate relativities. Section 7.2 describes the classification data examined. Section 7.3 describes the covariance structure and explains how correlations were estimated. Section 7.4 describes how regressions were fit to the correlations in order to estimate the parameters γ, ρ, I and J. Section 7.5 describes how the parameter K was estimated. Section 7.6 describes how the parameter Ω was selected.

Section 7.7 calculates credibilities with no weight given to the overall mean. Section 7.8 calculates credibilities with weight given to the overall mean. Section 7.9 discusses using prior estimates of the class relativities.

Section 7.10 discusses the impact of maturity of data in general. Section 7.11 gives an example of the impact of maturity on correlations while Section 7.12 gives the corresponding credibilities.

7.1. Rate Relativities

Assume that we are trying to estimate for a number of individual classes the expected pure premiums relative to the average for that group of classes. Further, assume we will do so by weighting together the observed relativities for that class over several recent years.[101] If R_{ic} is the relativity for year i, for class

[100]For an introduction to classification ratemaking see, for example, the Risk Classification chapter of *Foundations of Casualty Actuarial Science* [21].

[101]This is a simplification of how we might get indicated pure premiums by classification for workers compensation insurance. In that case, the relative pure premiums by class would be compared to those for an industry group. Also, the "serious," "non-serious"

c, then the estimate of the relativity for that class for year $N + \Delta$ is: $\sum_{i=1}^{N} Z_{ic} R_{ic}$, where $\sum_{i=1}^{N} Z_{ic} = 1$. This is the situation covered by Equations 6.7.

If instead we gave the complement of credibility to the grand mean, which in this example is a relativity of unity, then Equations 2.4 would apply instead of Equations 6.7. In either case, in order to estimate credibilities the key step will be the estimation of the (expected) covariances between years of data.

7.2. Classification Data

The data to be examined is 13 (consecutive) years of classification experience in one state for workers compensation insurance.[102] For each class we will use its payroll and losses to compute its pure premium relative to its industry group for that year. If L_{ic} is the loss[103] and P_{ic} the payroll,[104] then the relative pure premium in year i for class c is:[105]

$$R_{ic} = (L_{ic}/P_{ic}) \Bigg/ \left(\sum_c L_{ic} \Bigg/ \sum_c P_{ic} \right). \tag{7.1}$$

In order to estimate the behavior of the covariances by size of class, the data for the Manufacturing and Goods and Services industry groups will be examined.[106] The Manufacturing industry

and "medical" pure premiums might be treated separately. See Kallop [22] and Feldblum [23]. In addition, we might rely on "National" as well as state pure premiums by class. See Harwayne [24].

[102] See Appendix C for details on the data set examined.

[103] In this illustration, the losses are paid losses plus case reserves, at latest report, for medical plus indemnity, without any limitation by claim size.

[104] Payroll is in units of \$100.

[105] Note that the relativity of an individual class within an industry group depends both on the experience of that class, the experience of the other classes, as well as the exposures by class within the industry group. Thus, a given class relativity may change over time for a number of different reasons, some of which may have little to do with the individual class.

[106] Currently five industry groups are most commonly used for workers compensation ratemaking: Manufacturing, Construction, Office and Clerical, Goods and Services, and Miscellaneous.

group will be particularly useful since it has about 270 separate classes of various sizes. The Goods and Services industry group, with only about 100 separate classes, will not allow as detailed a breakdown by size of class.[107]

7.3. Covariance Structure

The covariance structure will be assumed to be that given by Equations 5.10 and 5.11. However, for estimation purposes we will use the simpler Equations 5.8 and 5.9, which ignore the varying volume of data by year for a class.[108]

For an industry group we compute the relative pure premiums for each class for each year. Then we can compute the covariances and correlations between the different years. By examining the behavior of these covariances and correlations as the size of class and the number of years of separation vary, we can roughly estimate the parameters that appear in the covariance Formulas 5.8 and 5.9.

For this purpose, we will restrict our attention to one size category of class at a time.[109] There are a number of ways to categorize the volume of data. This example uses an estimate of the average annual expected losses for a class based on its reported payroll.[110] Other reasonable measures of volume should produce roughly similar results.

For each such size category, we estimate the covariance between any two years of observed relative pure premiums R_{ic} and R_{jc} for $c = 1,\ldots,k$ where there are k classes in the size

[107]The Office and Clerical industry group has only around 14 classes. The Construction industry group has about 71 classes. The Miscellaneous industry group has about 49 classes.

[108]As will be seen, the estimation process is sufficiently imprecise that this simplification is appropriate.

[109]Nevertheless, the pure premiums are relative to the *entire* industry group, regardless of size of class.

[110]The details of how the expected losses were estimated for each class for each year are described in Appendix C.

FIGURE 10

category:[111]

$$\mathrm{Cov}[R_{ic}, R_{jc}] \approx \frac{\sum_{c=1}^{k} \sqrt{P_{ic}P_{jc}} R_{ic} R_{jc}}{\sum_{c=1}^{k} \sqrt{P_{ic}P_{jc}}} - \frac{\sum_{c=1}^{k} P_{ic} R_{ic}}{\sum_{c=1}^{k} P_{ic}} \frac{\sum_{c=1}^{k} P_{jc} R_{jc}}{\sum_{c=1}^{k} P_{jc}}. \tag{7.2}$$

The payrolls P_{ic} have been used as weights, in order to take into account the fact that for some classes the volume of data may be radically different by year. The variances are estimated in the same manner. Then as usual the estimated correlations are:

$$\mathrm{Corr}[R_{ic}, R_{jc}] = \mathrm{Cov}[R_{ic}, R_{jc}] / \sqrt{\mathrm{Var}[R_{ic}]\mathrm{Var}[R_{jc}]}. \tag{7.3}$$

For example, Figure 10 shows the observed correlations for the Manufacturing classes with expected annual losses between

[111]Recall that $\mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y]$.

\$300,000 and \$1 million. There are a total of 61 such classes. With 13 separate years of data, we can estimate $(13)(12)/2 = 78$ correlations. These correlations correspond to a separation of between one year and twelve years. We note considerable random fluctuation. Nevertheless, as the separation grows the correlations tend to decline.

7.4. *Fitting Regressions to Correlations, Estimating γ, ρ, I, and J*

Figure 10 shows the results of fitting a linear regression to the logs of these correlations. The fitted curve is (approximately) $y = (.46)(.94^x)$. The y-intercept is .46, and the decline rate or slope is .94.

Thus, we might estimate for this size of class the decline rate is about .94.[112] In the assumed covariance model this corresponds to some sort of weighted average of γ and ρ, with the weights depending on the size of risk E and the variances g^2 and r^2.

On the other hand the intercept of .46 represents an estimate of the credibility (of a single year of data) in the absence of shifting risk parameters. That is, using Equation 4.15,

$$\frac{E + I}{E(1 + J) + I + K} \approx .46 \qquad \text{for} \quad E \approx \$650{,}000.$$

Similar regressions were fit to the correlations for various size categories. However, in order to improve stability, the correlations for the same separations were first averaged.[113] So for example, the 12 correlations for one year of separation in Figure 10 average to .498. Then a weighted regression was fit to

[112] The slope of the log-correlations is about $\ln .94$.

[113] The averaging of the correlations prior to the regression versus time lag is not necessarily the best procedure to employ in this particular application, let alone in general. Ideally one would identify the variables causing the wide dispersion in observed correlations between individual years of data, as seen, for example, in Figure 10. However, I was unable to do so, beyond convincing myself that some substantial portion of this dispersion was a result of the process variance inherent to a data set of this size. While for the illustrative example here the technique used seems sufficient, it would be preferable to find a technique that directly makes use of all the available data. This is a potential area for future research, which could lead to a sharper estimate of the time dependence.

the logs of these average correlations,[114] with weights equal to the number of observed correlations of that separation. For data shown in Figure 10, this would result in a very similar fitted curve:[115] $y = (.46)(.94^x)$.

The results for the Manufacturing industry group, for six size categories with substantial number of classes are:

Expected Annual Losses ($000)	Number of Classes	"Slope"	Intercept
10 to 30	22	1.109	.075
30 to 100	40	.758	.329
100 to 300	37	.979	.375
300 to 1,000	61	.944	.469
1,000 to 3,000	50	.977	.744
3,000 to 10,000	13	.887	.911

The intercepts reflect a general pattern of increasing credibility with size of class, as expected. The smallest and largest size categories have too few classes to reliably estimate correlations.[116] Thus one should not rely on the estimated slopes; the estimated intercepts for these categories are less reliable than those for the other size categories.

For the four size categories with a large number of classes, there is some indication that the "slope" is closer to unity for large classes than for small classes. This data provides a weak indication that the risk parameters of larger classes shift more slowly than those of smaller classes.

The results of fitting regressions to the correlations of two size categories for Goods and Services classes are:

[114] If the average correlation was negative as occasionally happened, that separation was not included in the regression.

[115] The curve is the same in this case to the number of decimal places displayed.

[116] Also for the smallest size category, there is a lot of random fluctuation in the pure premiums of the classes.

Expected Annual Losses ($000)	Number of Classes	"Slope"	Intercept
100 to 1,000	38	.938	.605
1,000 to 10,000	38	.994	.837

The same general pattern applies, but with only two size categories we cannot infer much.

As discussed in Section 5.2, we expect the decline rate of the correlations to be approximately $(E\rho + I\gamma)/(E + I)$. Note that this actually applies only when $E \geq \Omega$. For $E \leq \Omega$ the decline rate of the correlations should be approximately $(\Omega\rho + I\gamma)/(\Omega + I)$.[117] In any case, the largest classes should have a decline rate near ρ, while the smaller classes have a decline rate closer to γ.

From the data for these two industry groups,[118] we might estimate that the largest classes have a decline rate for correlations of about .98; thus we might estimate $\rho \approx .98$. The smaller classes might have a decline rate below .90; thus we might estimate $\gamma \approx .85$. Note that ρ corresponds to a half-life of 34 years, while γ corresponds to a half-life of about 4 years. There is clearly a great deal of uncertainty in these estimates.[119]

The midway point at which the decline in the correlations is between ρ and γ is even harder to estimate. As discussed in Section 5.2, we expect this midway point to be at about I. For illustrative purposes estimate this as \$100,000, so that $I \approx \$100{,}000$.

As discussed previously in Section 4.5, the maximum credibility in the absence of shifting risks parameters for one year of data is $1/(1 + J)$. Thus, if J were .1, the intercepts would ap-

[117]For the parameters selected in this section $(\Omega\rho + I\gamma)/(\Omega + I) = ((50{,}000)(.98) + (100{,}000)(.85))/(50{,}000 + 100{,}000) = .89$.

[118]We ignore here the real possibility that the covariance structure might differ significantly among different industry groups, since this data is well short of being able to distinguish if that is the case.

[119]Better estimates would require looking at similar data from a large number of individual states, each of reasonable size.

proach $1/1.1 = .909$ for large risk sizes. While it is unclear from this limited data precisely what that maximum intercept is, it is almost certainly greater than .85. Thus, J is probably .15 or less. In any case, for illustrative purposes $J = .10$ will be used.

7.5. *Estimating K*

The estimates of J and I, together with the intercepts by size of risk, can be used to estimate the value of K. In the absence of shifting risk parameters, the credibility for a single year of data is given by Equation 4.15:

$$Z = \frac{E+I}{E(1+J)+I+K}, \qquad E \geq \Omega.$$

Thus,

$$K = \left(\frac{1}{Z}-1\right)(E+I)-(JE). \tag{7.4}$$

Given an estimate of Z from the intercept, for a size E, and the previously estimated $I = \$100{,}000$ and $J = .10$, we can estimate K.

We get the following estimates:

Industry Group	Size[120] (000)	Intercept	Estimated K (\$000)
Manufacturing	20	.075	1,478
Manufacturing	65	.329	330
Manufacturing	200	.375	480
Manufacturing	650	.469	784
Manufacturing	2,000	.744	523
Manufacturing	6,500	.911	−5
Goods & Services	550	.605	369
Goods & Services	5,500	.837	541

Recall that for Manufacturing the smallest and largest size categories really do not contain enough classes to adequately

[120]Based on the midpoint of the size category.

quantify the intercept. In any case, for the largest size category, the estimate of K is extremely sensitive to the selection of J. For the smallest two size categories, Equation 5.9 rather than Equation 5.8 is likely to hold, since $E \leq \Omega$; thus the above estimate of K using the two smallest categories is likely to be invalid. Averaging the middle three size categories for Manufacturing plus the two size categories from Goods and Services, we get $K \approx \$500{,}000$. This value of K will be used for illustrative purposes.

7.6. *Selecting* Ω

Finally, we must select Ω, the value below which the classes are homogeneous; i.e., there is no significant impact from risk heterogeneity below size Ω. Conceptually, this is the size at which a class is likely to be made up of one significant sized employer.[121] On the other hand, it was seen before that choosing Ω somewhere close to I produces a smooth decline in credibilities.

For illustrative purposes choose $\Omega = \$50{,}000$. This corresponds for this data set to somewhere between 50 and 75 full-time employees.[122]

In the absence of shifting risk parameters over time, Equation 4.25 gives the credibility for one year of data as:

$$Z = \frac{E}{(1+J')E + K'}, \qquad E \leq \Omega = \$50{,}000$$

where

$$J' = J\left(\frac{\Omega}{I+\Omega}\right) = (.10)\left(\frac{50}{150}\right) = .033 \qquad \text{and}$$

[121] While situations where the data for a class comes from one significant employer are not common, they do occur.

[122] Assuming reported losses (at unit statistical plan level) of about 2.5% of payrolls and a State Average Weekly Wage of about \$600, 65 full-time employees have \$50,700 in expected annual losses.

$$K' = K\left(\frac{\Omega}{I+\Omega}\right) = 500{,}000\left(\frac{50}{150}\right) = \$166{,}667.$$

For $E = \$20{,}000$,

$$Z = \frac{20}{((1.033)(20) + 166.667)} = \frac{20}{187.3} = 10.7\%.$$

This compares to the estimated intercept of .075. Given the uncertainty of the estimated parameters, the uncertainty of the estimated intercept, and the approximate nature of the regression relation itself, these values of .107 and .075 are not inconsistent. Getting a somewhat more precise estimate of Ω would require analyzing data from many states over many years.

With all these caveats, we have estimated the essential features of the covariances. Equation 5.10 states for $\sqrt{E_iE_j} \geq \Omega$:

$$\mathrm{Cov}[X_i, X_j] = r^2\left\{\rho^{|i-j|} + I\gamma^{|i-j|}/\sqrt{E_iE_j} + \delta_{ij}\left(K/\sqrt{E_iE_j} + J\right)\right\}, \quad \sqrt{E_iE_j} \geq \Omega.$$

Similarly Equation 5.11 states that for $\sqrt{E_iE_j} \leq \Omega$:

$$\mathrm{Cov}[X_i, X_j] = r^2\left\{\rho^{|i-j|} + I\gamma^{|i-j|}/\Omega + \delta_{ij}\left(K/\sqrt{E_iE_j} + J\right)\right\}, \quad \sqrt{E_iE_j} \leq \Omega.$$

In both cases there is a factor of r^2 that multiplies the covariances that does not affect the credibilities.

7.7. *Illustrative Credibilities, No Weight to Overall Mean*

We can use Equations 5.10 and 5.11 together with the values of the parameters estimated in the previous section to estimate the covariances. These in turn can be used to estimate the credibilities using Equations 6.7 (for the case where no weight is being given to the mean).

The following illustrative values will be used to calculate credibilities:

$\rho = .98$	(rate of shifting parameters related to class homogeneity),
$\gamma = .85$	(rate of shifting parameters related to class heterogeneity),
$I = \$100{,}000$	(related to class heterogeneity),
$J = .10$	(related to parameter uncertainty),
$K = \$500{,}000$	(Bühlmann credibility parameter, related to process variance), and
$\Omega = \$50{,}000$	(size limit for class homogeneity).

For example, for years 1, 2, 3 and 4 being used to predict year 8, with each year of data having \$1 million in expected losses, Equations 6.7 become:

$$1.7Z_1 + 1.065Z_2 + 1.0327Z_3 + 1.0026Z_4 = .9002 + \lambda/2,$$
$$1.065Z_1 + 1.7Z_2 + 1.065Z_3 + 1.0327Z_4 = .9236 + \lambda/2,$$
$$1.0327Z_1 + 1.065Z_2 + 1.7Z_3 + 1.065Z_4 = .9483 + \lambda/2,$$
$$1.0026Z_1 + 1.0327Z_2 + 1.065Z_3 + 1.7Z_4 = .9746 + \lambda/2,$$
$$\text{and} \qquad Z_1 + Z_2 + Z_3 + Z_4 = 1,$$

with solution:[123]

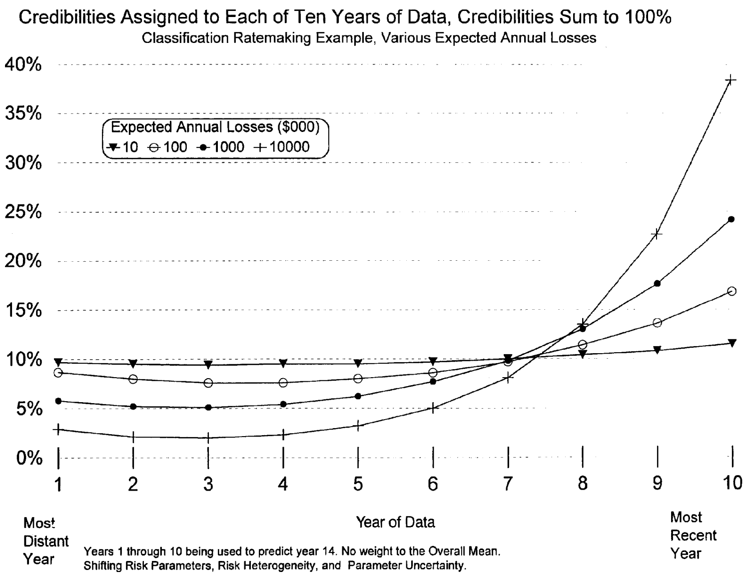
$$Z_1 = 21.08\%,$$
$$Z_2 = 21.98\%,$$
$$Z_3 = 25.34\%, \qquad \text{and}$$
$$Z_4 = 31.60\%.$$

Note that since no weight is given to the overall mean, the credibilities have been constrained to add up to 100%.

---

[123] The Lagrange Multiplier is .5416.

FIGURE 11



The credibilities assigned to ten individual years are shown in Figure 11 for various size classes, for years $1,2,\ldots,10$ being used to predict year 14. Note here it has been assumed that the credibilities are constrained to add to unity. Thus, the default weight is 10% to each of the ten years. However, as the classes get bigger and bigger we can make the estimation process more responsive and give more weight to the more recent data.[124] For \$10 million dollars in expected annual losses the most recent year gets about 38% of the weight. For small classes, we must use a

[124]The most distant year gets a slight amount of extra weight, due to the "edge effect." Year 1 contains valuable information about Year 0 due to the fact that they are correlated. Therefore, by giving a little more weight to Year 1, one gets some of the same benefit as if Year 0 were in the database. While, for example, Year 3 contains valuable information about Year 2 and Year 4, Years 2 and 4 are already in the database. In general, with shifting risk parameters over time, the most distant year(s) should receive somewhat more weight, due to this edge effect, than they would otherwise receive. In Figure 11 for \$10 million in Expected Annual Losses, as one goes to more distant years, at the edge the graph of credibilities bends slightly upwards rather than continuing to decline.

more stable method and give every available year of data significant weight. However, for small classes the parameters shift more quickly and thus there is a counter-balancing tendency to weight these older years less than more recent years. Nevertheless, for \$10,000 in annual expected losses the weights are all about 10%.

7.8. Illustrative Credibilities, Weight to Overall Mean

We can use Equations 5.10 and 5.11 together with the values of the parameters listed in Section 7.7 to estimate the covariances. These in turn can be used to estimate the credibilities using Equations 2.4, for the case where the complement of credibility is being given to the overall mean.

Assuming years $1,2,\ldots,10$ are being used to predict year 14, the credibilities assigned to the given years are shown in Figure 12. Larger sizes give more weight to recent years as well as more total credibility. Figure 13 shows the sums of the credibilities assigned to different classes. For ten years of data, the larger size classes are assigned up to 90% credibility.[125] The credibility goes to zero as the size of class goes to zero.[126] Also shown are the results for three years and one year of data.

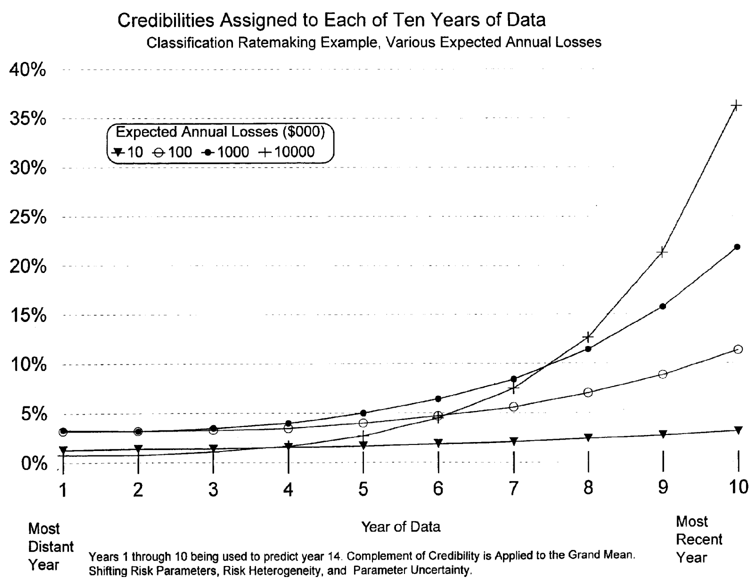
The class (expected) pure premiums within an industry group can easily vary by a factor of ten from lowest to highest. Thus, the average industry group pure premium, or equivalently a relativity of unity, is not a very good predictor for most classes. Therefore, the credibilities assigned to the classification data are relatively large. Assigning the complement of credibility to the average pure premium for the industry group, as illustrated here, is not generally done in practice.

---

[125]Without shifting risk parameters, the maximum credibility would be $Y/(Y+J) = 10/10.1 = 99\%$. With 10 years of data and $J = .1$, the effects of parameter uncertainty are not very significant.

[126]Since we've assumed no risk heterogeneity below size Ω.

FIGURE 12



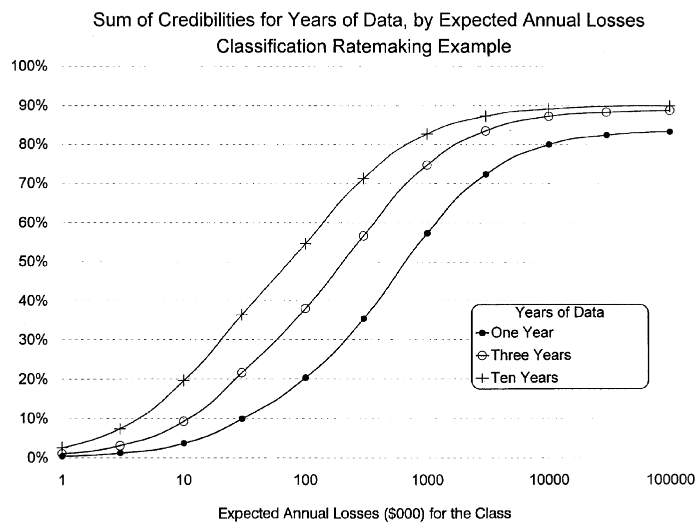
An alternative would be to work with loss ratios to premiums at current rates, as is done in Meyers [25]. Then the complement of credibility is given to the loss ratio for the industry group;[127] i.e., each class rate is changed by the industry group average rate change. This follows the general practice and is equivalent to giving the complement of credibility to the prior estimated relativity for each class.

7.9. Using Prior Estimates of Relativities

Assume that we have been estimating classification relativities for a long time. Then we might weight together the estimated relativity for each class based on the most recent data and the

[127] Meyers does not appear to divide the classes into industry groups. However, the technique presented could be applied equally well to industry groups. We would have to take a little more care in estimating the Bühlmann credibility parameters.

FIGURE 13



For example, Years 1 through 10 being used to predict year 14. Complement of Credibility is Applied to the Grand Mean.
Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty.

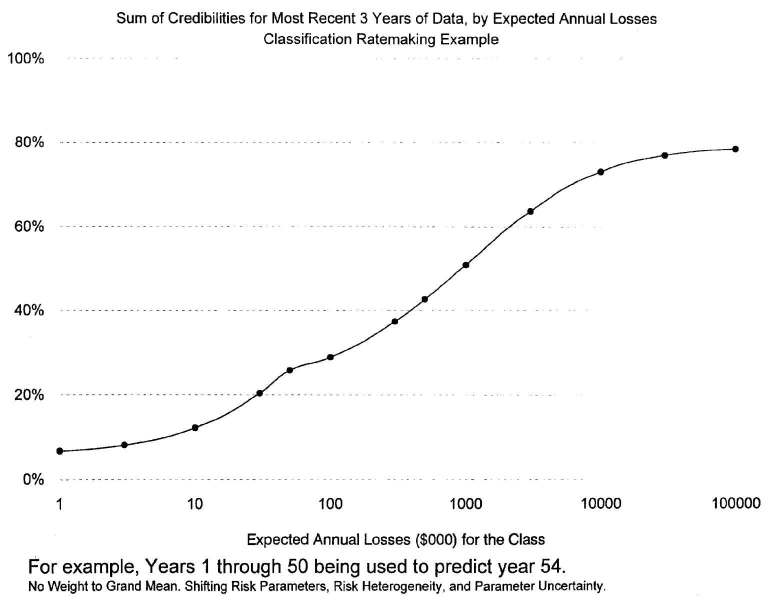
prior estimate of the relativity for that class. The issue is how much weight to apply to each of these estimates.

While there are other ways to think of this problem, we could fit it into the current framework by assuming some very long series of data, for example 50 years.[128] Then as in Section 7.7, we can compute the credibility to be assigned to each of these 50 years of data (with no weight to the overall mean). If three years of recent data are being used, then we can assign as the weight to the prior estimate the sum of the credibilities for the 47 less recent years.

For example, using the values from Section 7.7, for $1 million in expected annual losses, for years 1,2,...,50 being used to

[128]In the case of a workers compensation rating bureau, classification relativities have been estimated for about 80 years.

FIGURE 14



predict year 54, years 48, 49, and 50 have credibilities of: 11.8%, 16.3% and 22.8%. The prior estimate would be assigned a weight of 100% − (11.8% + 16.3% + 22.8%) = 49.1%.

Figure 14 shows the weight assigned to the most recent three years of data as the expected annual losses vary. The recent data for large classes gets less than 100% credibility; both the prior estimate and that from the recent data are assumed to be good estimators for large classes. The recent data for small classes gets considerable credibility; the prior estimate as well as that from the recent data are assumed to be poor estimators for small classes.

Note that the credibility curve in Figure 14 has a discontinuous derivative at the point $\Omega = 50{,}000$. This will be typical as we switch from Equations 5.10 for the covariances to Equations

5.11, as we go from a size where risks are generally heterogeneous to one where risks are generally homogeneous.[129]

7.10. General Effect of Differences in the Maturity of the Data

Conceptually the goal has been to estimate the expected future class relativity at *ultimate* report. Assume, as in Figure 11, we were predicting year 14 using data from years 1 to 10. Then we expect that year 1 at 10th report would be a better predictor of year 14 at ultimate than would year 1 at 5th report. Year 1 at 5th report is in turn a better predictor than year 1 at 1st report. Generally, the more mature the data from a single given year the better predictor of the future ultimate losses we expect it to be.[130]

Thus, actuaries will usually rely upon the latest *available* report for each year of data. In the case of the workers compensation classification example, we would have years 1 to 6 at fifth report,[131] year 7 at fourth report, year 8 at third report, year 9 at second report, and year 10 at first report.

In the example in Section 7.7, there is no weight to the overall mean; the credibilities assigned to the data sum to 100%. Thus in that situation, the credibilities reflect how good an estimator each year is *relative* to the others. If each of the ten years of data were at the same report, their relative value as estimators would be unaffected by maturity.

However, year 10 is only at first report while years 1 through 6 are at fifth report. Therefore, the tenth year of data is a poorer estimator relative to the other years than if it were available at fifth report. Thus, we should give year 10 somewhat less cred-

[129]In the model this switch is abrupt, leading to the discontinuous derivative of the credibility. While we could refine the model to make this derivative continuous, this would seem to be unlikely to have any practical significance.

[130]Thus, there is a dilemma. We prefer more recent years of data in order to minimize the impact of shifting risk parameters, but we also prefer more mature data. Section 9 discusses this from the point of view of an overall rate indication.

[131]Usually workers compensation classification data is only collected up to fifth report.

TABLE 4

CORRELATIONS BETWEEN REPORTS OF CLASSIFICATION RELATIVITIES

Various Size Classes, Based on Annual Expected Losses ($000)

	10 to 30 (19 classes)				30 to 100 (28 classes)			
	2nd	3rd	4th	5th	2nd	3rd	4th	5th
1st	.783	.757	.914	.834	.850	.754	.656	.624
2nd		.978	.849	.805		.860	.745	.730
3rd			.835	.783			.839	.809
4th				.949				.898
	100 to 300 (39 classes)				**300 to 1,000 (52 classes)**			
	2nd	3rd	4th	5th	2nd	3rd	4th	5th
1st	.863	.814	.822	.800	.879	.830	.809	.799
2nd		.935	.945	.917		.968	.945	.927
3rd			.957	.929			.980	.964
4th				.975				.975
	1,000 to 3,000 (49 classes)				**3,000 to 10,000 (22 classes)**			
	2nd	3rd	4th	5th	2nd	3rd	4th	5th
1st	.955	.924	.902	.884	.970	.964	.947	.939
2nd		.962	.946	.932		.980	.971	.965
3rd			.977	.965			.988	.977
4th				.986				.992

For each of five composite policy years, 84/85, 85/86, 86/87, 87/88 and 88/89, class relativities were calculated for the Manufacturing industry group. Then for each year, for classes in a given size category, correlations were calculated between the relativities at two different reports. The correlation matrices displayed here are an average of the five separate correlation matrices, one from each year.

ibility than was calculated in Section 7.7, while other years are assigned somewhat more credibility.

7.11. Correlations Between Differing Maturities

This effect of the differing maturities of data will be estimated by examining the correlations between class relative pure premiums from the same year of data but at different maturities. These correlations are calculated using Equation 7.2, where a difference in maturity is substituted for a difference in year. Table 4 displays correlation matrices for various size categories of classes from the Manufacturing industry group.[132] So, for example, for classes with expected annual losses between \$300,000 and \$1 million, the correlation between class relativities calculated from the same year of data at second report and fourth report is .945. In contrast, that between first and fifth report is .799. As expected, since more development occurs between first and fifth reports than between second and fourth report, the classification relativities are less highly correlated.[133]

In general, it is expected that the more loss development between two reports, the smaller the correlation of the relativities. The observed loss development factors (LDFs) were:[134]

1st to 2nd	1.249
2nd to 3rd	1.123
3rd to 4th	1.059
4th to 5th	1.040

Also, we expect that the relativities for smaller classes will be more affected by the random fluctuations caused by loss development. Therefore, the smaller the size category, the smaller the correlation of the relativities for different reports.

The simplest type of model would be one in which the correlation was some linear function of the class size and the loss development factor between reports. Since I was unable to find a useful model of that type, instead I first took the log of both the loss development factor and the correlation. Then I examined linear models involving the ln (correlation), ln (LDF), and size of class.

[132]Each correlation matrix is the average of five correlation matrices calculated for composite policy years 84/85, 85/86, 86/87, 87/88 and 88/89. A composite policy year includes July 1 to June 30.

[133]First and fifth report are further apart so that their correlation has more opportunity to decline from unity.

[134]For the Manufacturing industry group, for composite policy years 84/85, 85/86, 86/87, 87/88 and 88/89. All data was included independent of the size of class. Recall that the losses are paid plus case reserves.

One model of the relation of the correlations to the development that will have the desired properties is:[135]

$$\ln(\text{Correlation}) = -\ln(\text{LDF})/(\text{Linear Function of Size})$$
$$-\ln(\text{LDF})/\ln(\text{Correlation}) = \text{Linear Function of Size}. \qquad (7.5)$$

This model has the desired property that the correlation is 1 when the LDF is 1.[136] If the right hand side of Equation 7.5 is positive, then the correlation decreases as the amount of development increases. If the right hand side of the Equation 7.5 increases with size of class, then as desired the correlations will be closer to unity for larger classes.

For the data in Table 4, we can compute the ratio of the $-\ln(\text{LDF})/\ln(\text{correlation})$. For example, for the second to fourth report the LDF is (1.123)(1.059) = 1.189. For the size category \$300,000 to \$1 million in expected annual losses, the correlation between 2nd and 4th report is .945. Thus, $-\ln(\text{LDF})/\ln(\text{correlation}) = (-\ln(1.189)/\ln(.945)) = 3.06$.

Averaging over each correlation matrix we obtain by size of risk:[137]

Size ($ million)	$-\ln(\text{LDF})/\ln(\text{correlation})$
.02	1.76
.065	.76
.2	1.88
.65	2.38
2.0	3.35
6.5	6.20

---

135 For a given size of risk this model assumes the correlations decline as per a constant to the power of the "effective time" between reports. The "effective time" between reports is taken as the logarithm of the loss development factors.

136 In lines of insurance where salvage and subrogation are significant, the loss development factor can be less than unity. Equation 7.5 would not apply.

137 Expected Annual Losses for the midpoint of the size category.

A least squares linear regression would give

$$1.57 + .73\ (\text{Size}/1\ \text{million}).$$

Let,

$$c = -\ln(\text{LDF})/\ln(\text{correlation}).$$

Then,

$$\text{correlation} = \left(\frac{1}{\text{LDF}}\right)^{1/c}.$$

If we take for illustrative purposes:

$$c = 1.5 + .75\ (\text{Size}/1\ \text{million}),$$

then substituting into Equation 7.5 gives

$$\text{correlation} = 1/\text{LDF}^{1/(1.5+.75\ (\text{Size}/1\ \text{million}))}. \qquad (7.6)$$

For example, for the size category \$300,000 to \$1 million if we take a size of \$650,000 equivalent to the midpoint, then Equation 7.6 gives an estimated correlation of $1/\text{LDF}^{.5}$. For example, for the 2nd to 4th report the LDF is 1.189. Thus, for this size category the model correlation is about .91. (The observed correlation is .945.)

Table 5 displays the model correlations between reports for classes of various sizes. While the particular model represented by Equation 7.6 should be taken as solely for illustrative purposes, the general pattern of correlations in Table 5 is what we would expect. For a given report interval, the larger the class the higher the correlation. For a given size of class, the more development in a report interval, the lower the correlation. This pattern of correlations can be incorporated into the calculation of credibilities.

7.12. Credibilities Taking Into Account Differing Maturities

Returning to the example in Section 7.7, we can incorporate the impact of the differences in maturity. Given years 1 through 6 at fifth report, year 7 at fourth report, year 8 at third report, year

TABLE 5

CLASSIFICATION RATE RELATIVITIES
MODEL CORRELATIONS BETWEEN REPORTS

	Expected Annual Losses ($000)					
Reports	20	65	200	650	2,000	6,500
1 vs. 2	.864	.866	.874	.894	.929	.966
1 vs. 3	.800	.804	.815	.843	.893	.948
1 vs. 4	.770	.775	.787	.819	.876	.940
1 vs. 5	.750	.755	.768	.803	.865	.934
2 vs. 3	.926	.928	.932	.943	.962	.982
2 vs. 4	.892	.894	.900	.916	.944	.973
2 vs. 5	.869	.872	.879	.899	.932	.967
3 vs. 4	.963	.964	.966	.972	.981	.991
3 vs. 5	.938	.940	.943	.953	.968	.985
4 vs. 5	.974	.975	.977	.980	.987	.994

9 at second report and year 10 at first report, we try to predict year 14 at fifth report.

For a class with expected annual losses of $1 million, Equation 7.6 estimates the correlation between classification relativities at different reports as $1/\text{LDF}^{.444}$. For 2nd to 4th report, the estimated correlation is $(1.189)^{-.444} = .926$. Prior to taking into account the differences in maturity, the model covariance between year 7 and year 9 was 1.033.[138] It will be estimated that the covariance between year 7 at 4th report and year 9 at second report will be lower by a factor of the correlation[139] .926; (.926)(1.033) = .957.

The other model covariances involving at least one year of data at prior to 5th report are similarly adjusted.[140] (The vari-

[138]For $r^2 = 1$, $1 million in expected annual losses, and the parameters in Section 7.7.

[139]This is an approximation based on an assumption that the impact of maturity is largely independent of the other factors previously considered.

[140]For purposes of adjustment it was assumed Year 14, the year to be predicted, was at 5th report. If one assumed instead for example 20th report, all the covariances involving

ances along the diagonal of the variance-covariance matrix are unaffected.) The least squares credibilities differ from those obtained in Section 7.7:[141]

CREDIBILITY FOR CLASS WITH $1 MILLION IN EXPECTED ANNUAL LOSSES

Year, Report	Section 7.7	Taking into Account Differences in Maturity
1@5th	5.8%	6.7%
2@5th	5.2	6.2
3@5th	5.1	6.4
4@5th	5.4	7.1
5@5th	6.2	8.6
6@5th	7.7	10.8
7@4th	9.8	11.5
8@3rd	13.0	12.7
9@2nd	17.6	13.9
10@1st	24.2	16.0

As expected, more mature years of data are given more credibility than previously while less mature years receive less. For example, the data from year 10 at first report gets 16.0% credibility compared to the 24.2% credibility calculated in Section 7.7.

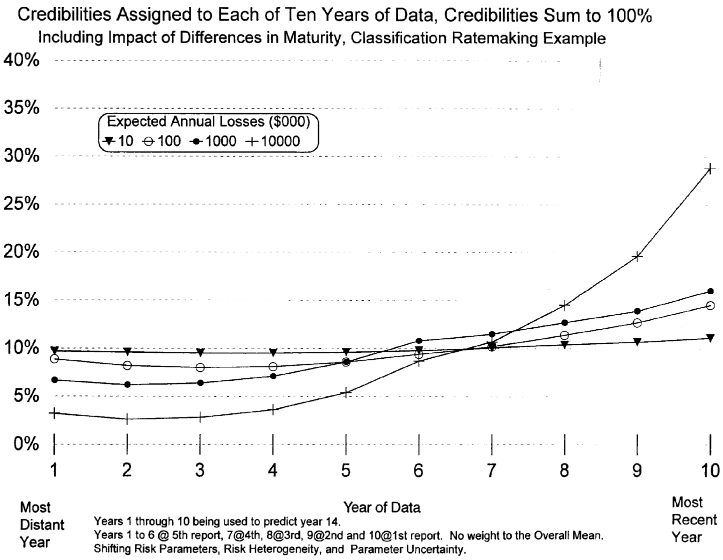
Figure 15 displays the credibilities for other size classes. The credibilities shown in Figure 15 that take into account differences in maturity can be compared to those in Figure 11, which ignore these differences. While the precise impact depends on the particular amount of loss development and the particular model used to estimate the correlations, the general pattern displayed here should occur in most situations.

The weights which would otherwise be given to immature years of data should decrease significantly for larger size classes.

Year 14 would be lower by the same factor but the resulting credibilities would all be the same, since they've been constrained to sum to 100%.

[141] As shown in Figure 11.

FIGURE 15



For smaller classes the weights assigned to recent years are already close to the default weight, in this case 10%, so taking into account their immaturity only produces a small decrease in the weight they would otherwise receive. In all cases, the more mature years of data receive more weight than when we ignored maturity. In this example, the largest increase in weight occurs for year 6, which is the most recent year which is available at "ultimate."[142]

So while taking into account shifting risk parameters over time tends to give more weight to recent years, taking into account the difference in maturity tends to counterbalance that tendency somewhat.[143] This will be true for overall ratemak-

[142]In this example, fifth report is the ultimate report actually received of Unit Statistical Plan data, even though there is loss development beyond fifth report.

[143]An example is given in Section 9 in which the most recent year of data is so immature it is given very little weight.

ing and experience rating as well as for classification ratemaking.

8. USE OF DATA FROM OTHER STATES

In estimating the classification relativities in a given state one may supplement the data from that state with data from other states, as in Harwayne [24].

8.1. Use of Data from One Other State

As a simple example, assume we are estimating Massachusetts relativities and will use New York experience in addition to that from Massachusetts. The key assumption is that the underlying expected class relativities in New York are similar to those in Massachusetts. Thus, observed relativities in New York are useful for predicting future relativities in Massachusetts. However, all other things being equal, a given volume of New York data is assumed to be less useful in predicting Massachusetts relativities than would be similar data from Massachusetts.[144] Thus, we expect that in this case the credibilities assigned to a given volume of data will be less for New York data than for Massachusetts data.

There are three steps to calculating the credibilities to assign to the years of data from Massachusetts and New York. First, we must model the covariance structures. Second, we must estimate the parameters in the covariance structures. Third, we must use these covariances together with the appropriate set of linear equations, in this case Equations 8.1, in order to solve for the credibilities. In this case, the first two steps will build on the results on classification relativities obtained in Section 7.

[144] Similarly, New York data would be more useful for predicting New York relativities than would data from Massachusetts.

8.2. Covariance Structure, Use of Data From One Other State

There are three types of variance-covariance matrices. The first type involves covariances between data from Massachusetts: S_{ij} = covariances within Massachusetts. The illustrative values from Section 7.7 will be used for these covariances. The second type involves covariances between data from New York: T_{ij} = covariances within New York. For a given volume of data, assume a similar covariance structure within New York to that estimated for Massachusetts; the illustrative values from Section 7.7 will therefore be used for the covariances within New York, T_{ij}.

The third type of covariance is that involving data from Massachusetts versus data from New York: U_{ij} = covariances between Massachusetts and New York. It is expected that for a given volume of data, the correlation of relativities between states is less than the correlation of relativities within states. This is what is observed.

8.3. Estimating Parameters, Between State Covariances

Classification data for Massachusetts, New York and several other large states was examined as discussed in Appendix F. Correlations of classification relativities between states were calculated for classes in various size categories for both the Manufacturing and the Goods and Services industry groups.

Based on the analysis discussed in Appendix F, with three exceptions the same parameters will be used for the interstate and intrastate covariances. The K parameter, related to the expected value of the process variance, will be zero for the interstate covariances. The J parameter, related to parameter uncertainty, will be selected for the interstate covariances as half of the intrastate J parameter.[145]

[145]The credibilities are relatively insensitive to this choice.

The r^2 parameter, setting the scale for the covariances, will be taken for the interstate covariances as 70% of its value for the intrastate covariances.[146] This will result in correlations of relativities between states that are lower than those within a state, all else being equal.

For the covariances the following inputs are used:[147]

	Intrastate	Interstate
r^2	1	.7
ρ	.98	.98
γ	.85	.85
I	100,000	100,000
J	.10	.05
K	500,000	0
Ω	50,000	50,000

8.4. *Equations for Credibilities, One Other State*

Assume we are trying to estimate class relativities in Massachusetts, without any weight to the overall mean. Let Z_i be the weight applied to the Massachusetts data and let W_i be the weight applied to the New York data. Then $\Sigma Z_i + \Sigma W_i = 1$, since there is no weight given to the overall mean. As shown in Appendix E, if we use Y years of data from each state, in order to predict year $Y + \Delta$, we obtain $2Y + 1$ equations in $2Y + 1$ unknowns:[148]

---

[146]The relative size of the interstate and intrastate covariances affects the calculation of credibilities. However, there is still an arbitrary choice of overall scale which does not affect the credibilities.

[147]The r^2 values contain an arbitrary scale factor. Since it is only their relationship that affects the credibilities, the actual r^2 values have not been estimated. Unlike Section 7.12, no adjustment is made for differing maturities here. Such an adjustment in the case of more than one state would parallel that for a single state as shown in Sections 7.11 and 7.12.

[148]There are YZ's, YW's, plus the Lagrange Multiplier λ. The equations would be somewhat different if the years for which we have Massachusetts data and New York data are not the same. Appendix E gives an example.

$$\sum_j Z_j S_{ij} + \sum_j W_j U_{ij} = \frac{\lambda}{2} + S_{i,Y+\Delta}, \qquad i = 1,2,\ldots Y,$$

$$\sum_i Z_i U_{ij} + \sum_i W_i T_{ij} = \frac{\lambda}{2} + U_{Y+\Delta,j}, \qquad j = 1,2,\ldots Y, \tag{8.1}$$

$$\sum_i Z_i + \sum_j W_j = 1,$$

where the covariance matrices are:

S_{ij} = covariances within Massachusetts,[149]

T_{ij} = covariances within New York,[150]

U_{ij} = covariances between Massachusetts and New York.[151]

8.5. Illustrative Credibilities, One Other State

For example, assume we are estimating year 54 class relativities in Massachusetts using data from years 1 to 50, with \$1 million of expected annual losses in Massachusetts and \$5 million of expected annual losses in New York. Then using Equations 8.1 the most recent three years of Massachusetts data would be given credibilities of 9.7%, 13.3% and 18.6%, while the three most recent years of New York data would be given credibilities of 2.5%, 7.0% and 15.3%. We could give the prior estimate the remaining weight of 33.6%.

8.6. Using Data From Several Other States

This example where the data from one outside state is used can be extended to one where data is used from several other states. Assume for simplicity that "countrywide data" is from

[149]More generally within the state for which we are trying to estimate class relativities.

[150]More generally within the supplementary data from outside the state of interest.

[151]More generally between the data from the state of interest and the data from outside the state of interest.

ten states, other than the state for which we are estimating class relativities.

Let C be the covariance matrix within states, while D is the covariance matrix between states.[152] Assume for simplicity that for a given volume of data, C is the same in each state and D is the same for each pair of states. Assume that the "countrywide data" is the average of data from ten states, each with a volume of data $\tilde{E}/10$. Let the covariance matrix between two non-Massachusetts states be D' and the covariance within a non-Massachusetts state be C'. Then the covariance between the countrywide data is the sum of 100 terms, 90 of which are between states, D', and 10 of which are within states, C'.[153] The covariance between the countrywide data is therefore $(90D' + 10C')/100 = .9D' + .1C'$. In general, if we had data from n other states each of the same size, the covariance between the countrywide data would be $((n-1)D' + C')/n$.

We have assumed $D' < C'$, so that $.9D' + .1C' < C'$. Due to the lack of homogeneity of the countrywide data, its covariance is less than that for an equivalent volume of data all from a single state.

The covariance of the countrywide data[154] with Massachusetts is just the average of ten similar terms all involving the covariance between the states.[155] Thus, the covariance between Massachusetts and the countrywide data is D.

In summary, for C and D calculated for the appropriate volumes of data for the state(s) involved:

S_{ij} = covariances within Massachusetts = C,

[152]Both C and D are a function of the volume of data in the state(s).

[153]$\mathrm{Cov}[\frac{1}{10}\{Y_1 + Y_2 + \cdots + Y_{10}\}, \frac{1}{10}\{Y_1 + Y_2 + \cdots + Y_{10}\}] = \frac{1}{100}\{\mathrm{Cov}[Y_1,Y_1] + \mathrm{Cov}[Y_1,Y_2] + \mathrm{Cov}[Y_1,Y_3] + \cdots \mathrm{Cov}[Y_{10},Y_{10}]\}$.

[154]The state of interest, in this case Massachusetts, is assumed to be excluded from the countrywide data.

[155]$\mathrm{Cov}[X,(Y_1 + Y_2 + \cdots + Y_{10})/10] = \frac{1}{10}\{\mathrm{Cov}[X,Y_1] + \mathrm{Cov}[X,Y_2] + \cdots + \mathrm{Cov}[X,Y_{10}]\}$.

T_{ij} = covariances within Countrywide[156] $= .1C' + .9D'$,

U_{ij} = covariances between Massachusetts and Countrywide $= D$.

8.7. Illustrative Credibilities, Data From Several Other States

These covariances can then be used in Equations 8.1, in order to solve for the credibilities. For example, assume we are estimating year 54 class relativities in Massachusetts using data from years 1 to 50, with \$1 million of expected annual losses in Massachusetts and \$1 million of expected annual losses in each of ten other states. Then using Equations 8.1, the most recent three years of Massachusetts data would be given credibilities of 8.5%, 11.0% and 14.9%. The most recent three years of countrywide data would be given credibilities of 1.8%, 9.8% and 28.5%. The remaining weight of 25.5% could be given the prior estimate of the class relativity.[157]

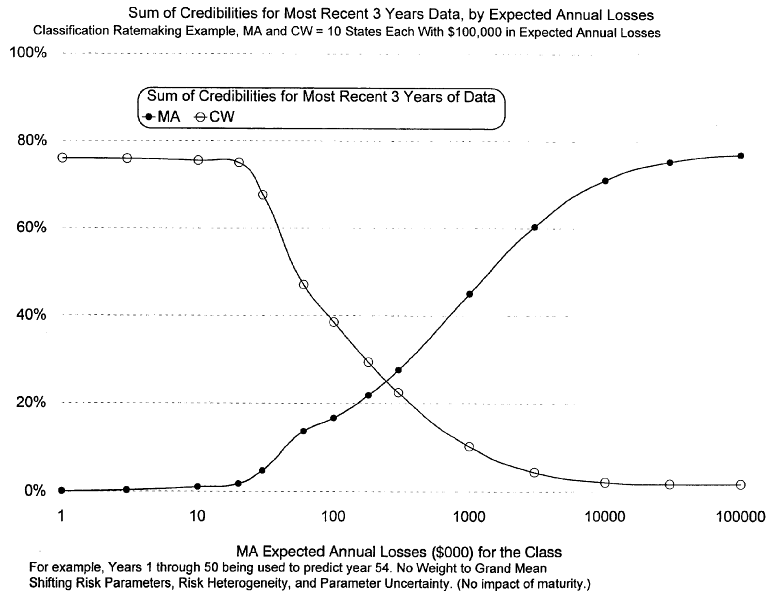
Figure 16 shows for a fixed amount of countrywide data, how the credibilities vary as the volume of data in Massachusetts changes. Since in Figure 16 there is assumed to be \$100,000 in expected annual losses in each of ten states other than Massachusetts, there is sufficient countrywide data to get a reasonable estimate of the class relativity. When there is very little Massachusetts data, for example \$3,000 in expected annual losses, then the most recent three years of Massachusetts data are given virtually no weight,[158] while the most recent three years of coun-

[156]This is for the case where "countrywide" data consists of 10 equal sized states. In general, the covariance of countrywide data will be some mixture of C and D covariance matrices.

[157]It should be noted that for this case, many older years of countrywide data are given negative weight. As a practical matter these weights could be set equal to zero and the weights given to more recent years of countrywide data could be reduced accordingly. This would increase the weight given to the prior estimate.

[158]This is in contrast to Figure 13 where, in the absence of the use of countrywide data, the Massachusetts data was given small but significant weight.

FIGURE 16

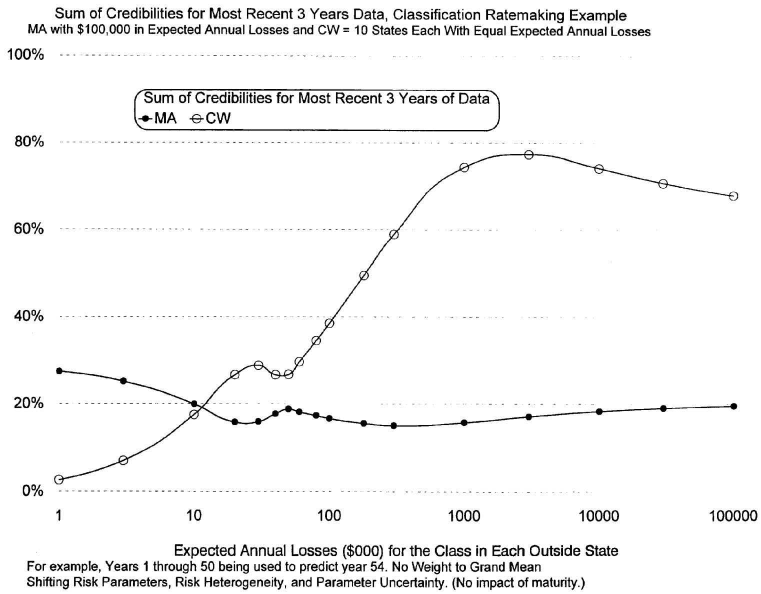


trywide data are given a weight of about 75%.[159] As the volume of Massachusetts data increases, while the volume of countrywide data remains the same, the weight assigned to the most recent three years of Massachusetts data increases up to about 75%, while that assigned to the countrywide data declines to zero.

Figure 17 displays the credibilities assigned to the most recent three years of data, for a fixed amount of Massachusetts data while the volume of countrywide data varies. As the volume of countrywide data increases, the credibility assigned to the most recent three years of countrywide data increases nonmonotonically to about 75%. The credibility assigned to the latest three years of Massachusetts data (with $100,000 in expected

[159] The remaining weight is given to the prior estimate.

FIGURE 17

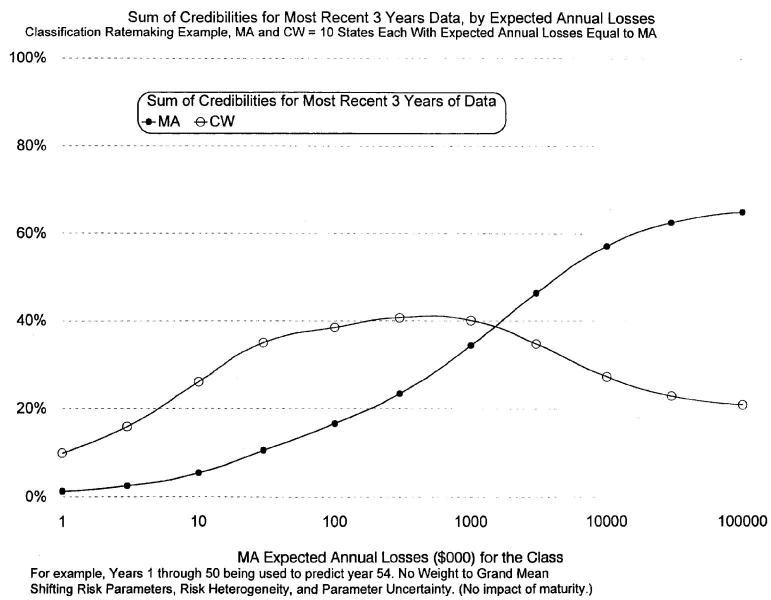


annual losses) varies between 27% and 15% as the volume of countrywide data varies.

Figure 18 displays the credibilities if Massachusetts and each of ten other states all have the same expected annual losses for a given class. As the size of class increases, the sum of the credibilities given to the most recent three years of Massachusetts data increases to about 65%. As the size of class increases, the sum of the credibilities given to the most recent three years of countrywide data increases and then decreases, as for very large classes the Massachusetts data is given more weight.

This behavior means that no simple formula for the amount of credibility given to the countrywide data will be appropriate. We must know how much data is available within the state of interest, before we know how much credibility to assign to the

FIGURE 18



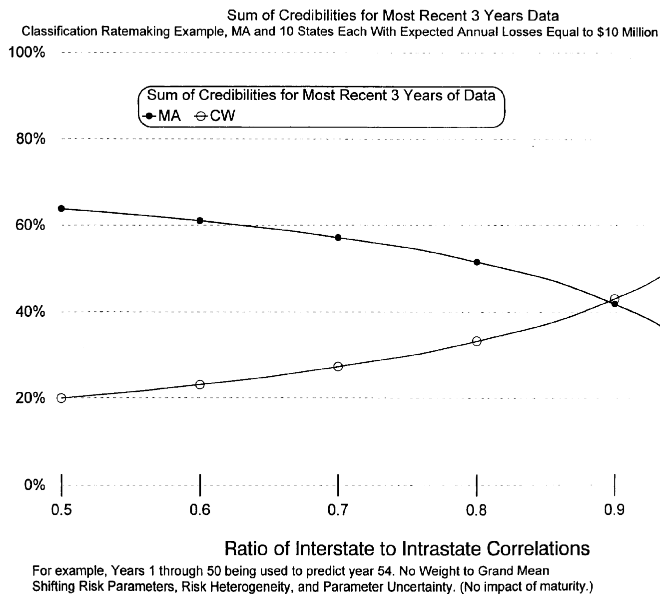
countrywide data.[160] If a simple formula such as the "square root rule" or the "Bühlmann credibility formula" were to be applied based solely on the volume of countrywide data, it would have to be supplemented by some other restriction on the credibility assigned to countrywide data. One commonly used rule of thumb is to restrict the credibility assigned to the countrywide data to be no more than:

$$\left(\tfrac{1}{2}\right)(1 - \text{credibility assigned to the state data}).$$

Figure 19 displays the sensitivity of the credibilities to the selected ratio of the interstate correlations to the intrastate correlations. For values of this ratio close to the selected value of 70%, the credibilities are relatively insensitive. Note that if the

[160]The reverse is also true, but the credibility of the Massachusetts data is less sensitive to the amount of countrywide data, as seen in Figure 16.

FIGURE 19



intrastate and interstate correlations were equal, then each outside state would get the same 7.7% credibility[161] as would Massachusetts.

9. A RATEMAKING EXAMPLE

This section will illustrate how the ideas in this paper might be applied to the calculation of an overall rate indication.[162] The issue explored here is how much weight should be given to different years of data. This example will illustrate how adjustments to the data for trend, development, etc. will affect the optimal weights.

Assume that for a given line of insurance the six most recent years of data are being combined in order to calculate a rate

[161] For a sum of 77% for ten outside states.

[162] This is an expansion of an example in Mahler [20].

indication. Specifically, assume we have loss ratios[163] from policy years[164] 1991, 1992, 1993, 1994, 1995 and 1996, all as of 12/31/96, which will be used to get a rate indication for policy year 1998.

9.1. Estimation Errors Due to Adjustments to Data

It is assumed that appropriate adjustments have been made to each year's data for development, trend, law changes, changes in deductibles, etc.[165] The necessity of these adjustments introduces estimation error into the process. For example, if we had policy year 1995 at ultimate rather than at first report, we could make a more precise estimate of policy year 1998 at ultimate.

The important consideration for this illustrative example is the pattern of errors for the different types of adjustments for the different years. For purposes of simplicity only two types of adjustments will be assumed. Development will be assumed to have larger estimation errors for recent years. In particular the "incomplete" policy year 1996 as of 12/31/96 will have an extremely large amount of development to ultimate. Trend[166] will be assumed to have larger estimation errors for more distant years.

For example, assume the reported Policy Year 1993 losses at 12/31/96 were $90 million. Further, assume that the point estimate[167] of Policy Year 1993 losses at ultimate is $96 million. This corresponds to a point estimate of the age to ultimate loss development factor of approximately 1.067. However, there is an error associated with this point estimate.

---

[163]The general ideas explored in this example would apply equally well to pure premiums.

[164]The general ideas explored in this example would apply equally well to calendar years or accident years of data.

[165]We assume that each of the adjusted loss ratios is intended to be an unbiased estimate of the Policy Year 1998 loss ratio.

[166]For illustrative purposes this can be thought of as trend, law amendment and other adjustments.

[167]Using data evaluated as of 12/31/96.

For example, an interval estimate of these ultimate losses might be \$92 million to \$100 million. This would correspond to an interval estimate of the age to ultimate loss development factor of approximately $1.067 \pm .044$.

A 95% confidence interval corresponds to about plus or minus two standard deviations. Therefore, this interval estimate of the loss development factor could result from a standard deviation of .022 or a variance of $.022^2 \approx .0005$. Any estimate is subject to error and in general one can estimate the variance of any estimator.[168]

Generally, estimation errors are quantified via variance-covariance matrices.[169] The covariances are introduced in order to capture the fact that the estimation errors for the years are usually positively correlated. If the development estimated for 1995 is too high, then it is likely that the development estimated for 1994 is too high as well. Similarly, if the trend applied to 1993 is too high, that applied to 1992 is likely to be too high as well.

Let $\boldsymbol{D}$ be the variance-covariance matrix quantifying the estimation errors related to development. An illustrative example of such a matrix is:

$$\boldsymbol{D} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 50 & 45 & 70 & 180 \\ 0 & 0 & 45 & 100 & 125 & 300 \\ 0 & 0 & 70 & 125 & 350 & 600 \\ 0 & 0 & 180 & 300 & 600 & 5{,}000 \end{pmatrix} \times 10^{-5}$$

[168] See, for example, Klugman, Panjer and Willmot [8].

[169] The diagonal elements are the variances quantifying the estimation errors. In this case, the element in the first row and second column is the covariance between the 1991 and 1992 errors. Readers may be familiar with the use of the inverse of the information matrix as a variance-covariance matrix when estimating parameters of loss distributions via the method of maximum likelihood. See, for example, Klugman, Panjer and Willmot [8].

The rows and columns correspond to the six years of data. For example, the variance of the estimated age to ultimate loss development factor of Policy Year 1993 is 50×10^{-5}.[170] The covariance between the estimated age to ultimate loss development factors for Policy Years 1993 and 1994 is 45×10^{-5}.

The particular values are chosen for illustrative purposes.[171] While the values would vary considerably depending on the particular application, the general pattern is expected to apply. The estimation errors for recent years are large,[172] and there is a positive correlation between the estimation errors for the different years.

Similarly, let $\boldsymbol{T}$ be the variance-covariance matrix quantifying the estimation errors related to trend. An illustrate example of such a matrix is:

$$\boldsymbol{T} = \begin{pmatrix} 350 & 292 & 240 & 192 & 150 & 110 \\ 292 & 300 & 247 & 198 & 155 & 114 \\ 240 & 247 & 250 & 201 & 157 & 115 \\ 192 & 198 & 201 & 200 & 156 & 115 \\ 150 & 155 & 157 & 156 & 150 & 110 \\ 110 & 114 & 115 & 115 & 110 & 100 \end{pmatrix} \times 10^{-5}$$

For example, the variance of the estimated trend factor from Policy Year 1994 to 1998 is 200×10^{-5}.[173] The covariance be-

---

[170] Thus, the standard deviation is $\sqrt{50 \times 10^{-5}} = .022$. If the point estimate of this loss development factor were, for example, 1.067, then using two standard deviations would result in an interval estimate of $1.067 \pm .044$.

[171] In particular, for longer tailed lines of insurance there would still be considerable development left for Policy Year 1991. In actual applications the actuary may have a good idea of how accurate an estimate is likely to be and thus could judgementally select a variance-covariance matrix.

[172] The error in developing the incomplete Policy Year 1996 is potentially extremely large.

[173] Thus, the standard deviation is $\sqrt{.002} = .045$. If the point estimate of this trend factor were, for example, 1.148, then using two standard deviations would result in an interval estimate of $1.148 \pm .090$.

tween the estimated trend factors to 1998 from 1994 and 1995 is 156×10^{-5}.

Again the particular values are chosen for illustrative purposes. The pattern was chosen such that the estimation error from trend is larger for more distant years and such that there is a large positive correlation[174] between the estimation errors for different years.[175]

9.2. *Covariance Structure for Years of Data*

Next we need to assume a variance-covariance structure for the year's loss ratios in the absence of any estimation error. Let this matrix be C. Then following the development in Mahler [1] of shifting risk parameters, assume that C has the form:[176]

$$C_{ij} = \delta_{ij} e^2 / \sqrt{E_i E_j} + r^2 \rho^{|i-j|}, \qquad \text{where} \quad \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}. \tag{9.1}$$

It is not necessary to know the source of e^2, r^2 and ρ in order to proceed. However, it may be helpful to think of ρ as the dominant eigenvalue (other than unity) of the transpose of a transition matrix of a Markov chain, r^2 as the variance of the hypothetical means, and $e^2/\sqrt{E_i E_j}$ as the expected value of the process variance.

In any case, ρ determines the rate of decline in the covariances as the separation between years increases.[177] So $\rho = .90$ would

[174]For example, the correlation between the estimated trend factors to 1998 from 1994 and 1995 is $156/\sqrt{(200)(150)} = .90$.

[175]A similar pattern would be expected for on-level factors to adjust for law amendments.

[176]This is the covariance structure in the presence of shifting risk parameters, equivalent to Equation 3.16. If appropriate we could instead use one of the more complicated covariance structures, for example Equations 5.10 and 5.11.

[177]For data from an individual insurer, one of the reasons that the covariances between years declines as the separation increases may be nonrenewals of insureds. The higher the lapse rate the faster the expected rate of decline in these covariances. As discussed in Busche [26], the higher the lapse rate, the lower the weight given to older years of data.

represent a more rapid decline than would $\rho = .99$; the former would correspond to more rapidly shifting parameters over time than the latter.

The relative magnitudes of r^2 and e^2 will control how much weight is given to distant versus recent years. The larger e^2, the more random noise there is in the data from any one year; when e^2 is large we must give each of the available years significant weight. The smaller e^2, the less random noise there is in the data and larger weight can be given to more recent years and insignificant weight to older years. When e^2 is small, we can use a more responsive method. When e^2 is large we have to use a more stable method.

If everything else is equal, the larger the volume[178] of data in a year, the smaller we expect the process variance of the loss ratios to be. We assume the process variance is inversely proportional to the volume of data.[179] Thus, how responsive our estimation method should be depends on the volume of data available per year. If more data is available per year, then the estimation method can be more responsive.

9.3. Credibilities

Assume we are estimating the year $Y+\Delta$ by weighting together years $1,2,\ldots,Y$. Then as shown in Appendix B, the least squares weights Z_i, $i = 1,2,\ldots,Y$, with $\sum_{i=1}^{N} Z_i = 1$, are the solution to the $Y+1$ Equations 6.7:

$$\sum_{i=1}^{Y} Z_i V_{ik} = V_{k,Y+\Delta} + \lambda/2, \qquad k = 1,2,\ldots,Y \qquad \text{and} \qquad \sum_{i=1}^{Y} Z_i = 1. \tag{9.2}$$

[178]The measurement of the volume of data would depend on the particular application. For example, it could be house-years, man-weeks, car-years, inflation adjusted sales, etc. See Bouska [27].

[179]For this example, it has been assumed e^2 is the same for each year.

where V is the variance-covariance matrix and λ is the Lagrange Multiplier.[180]

9.4. No Estimation Error

If we do not include any estimation error, then in our example $V = C$, $Y = 6$ and $\Delta = 2$. Thus, the Equations 9.2 become:

$$\sum_{i=1}^{6} Z_i C_{ik} = C_{k,8} + \lambda/2, \qquad k = 1,2,\ldots,6 \qquad \text{and}$$
$$\sum_{i=1}^{6} Z_i = 1. \tag{9.3}$$

Given values for E_i, e^2, r^2, and ρ we can use Equation 9.1 to calculate the matrix C and then solve these linear Equations 9.3 for the weights Z_i.

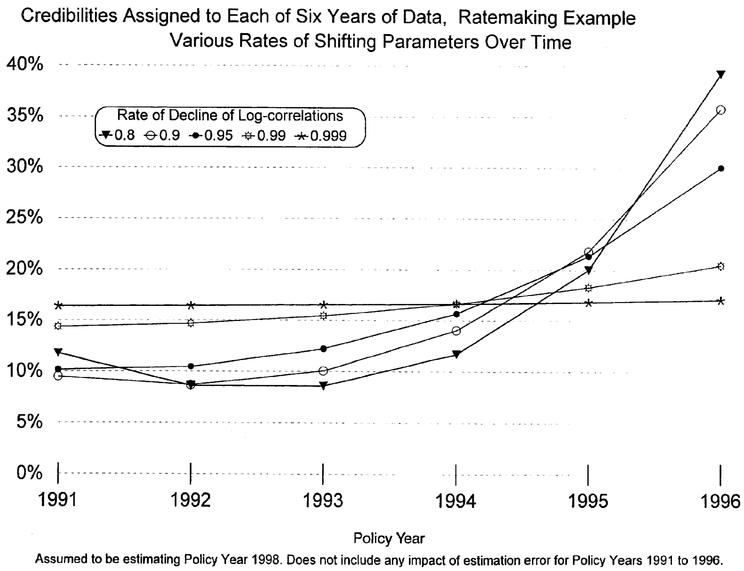
For example, with $E_i = 1$ for $i = 1$ to 8, $e^2 = .005$, $r^2 = .007$ and $\rho = .90$, we would get:

$$Z_1 = 9.5\%,$$
$$Z_2 = 8.7\%,$$
$$Z_3 = 10.1\%,$$
$$Z_4 = 14.0\%,$$
$$Z_5 = 21.8\%, \qquad \text{and}$$
$$Z_6 = 35.9\%.$$

Thus, as expected in the presence of shifting risk parameters, the more recent years 1996 and 1995 get more weight, while the earlier years 1991 and 1992 get less weight. Note that there is an "edge effect." The credibility assigned to 1991 is somewhat

[180] λ is an auxiliary variable, whose value will not be of particular interest.

FIGURE 20



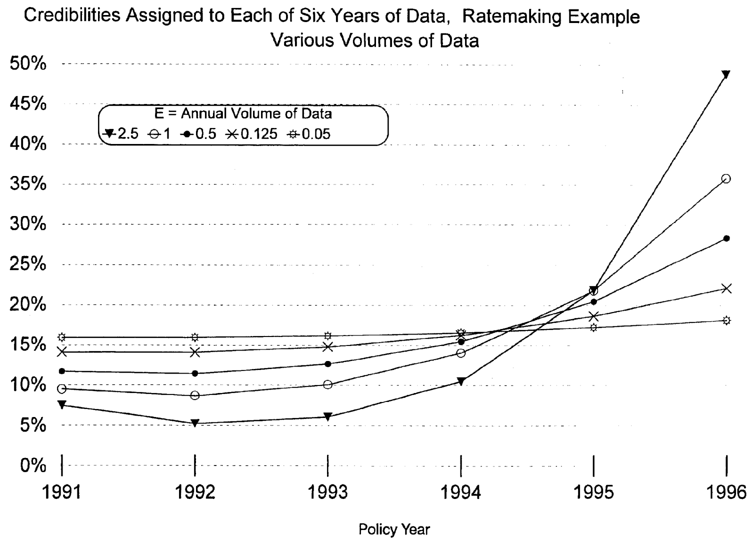
larger than it would otherwise be, since it is assumed to contain more unique information compared to 1992; the information content of 1992 is captured to some extent by the years 1991 and 1993 bracketing it on either side. The same "edge effect" applies to 1996, raising its credibility weight somewhat.

Figure 20 displays what happens as we vary ρ. As ρ approaches unity, parameters are shifting less rapidly, and therefore approximately equal weight is given to different years.[181] As ρ approaches zero, parameters are shifting more rapidly, and therefore less weight is given to the older years.

If we were to increase the expected value of the process variance, by taking $E = \frac{1}{2}$, keeping $e^2 = 0.005$, $r^2 = .007$ and $\rho = .90$,

[181]Recall that for this illustrative example the volume of data for each year is assumed to be the same.

FIGURE 21



Assumed to be estimating Policy Year 1998. Does not include any impact of estimation error for Policy Years 1991 to 1996.
Log-correlations declining at .90.

then the weights are:

$$Z_1 = 11.7\%,$$
$$Z_2 = 11.4\%,$$
$$Z_3 = 12.6\%,$$
$$Z_4 = 15.5\%,$$
$$Z_5 = 20.5\%, \qquad \text{and}$$
$$Z_6 = 28.4\%.$$

Compared to $E = 1$, with $E = \frac{1}{2}$ (a smaller volume of data) there is less weight given to more recent years and more weight given to more distant years. Figure 21 displays what happens as we vary E. As E (the volume of data) gets smaller, the weights become more equal. As E gets larger, more weight is given to recent years.

9.5. *Taking Into Account Estimation Error*

We can now introduce the impact of estimation error. First take the sum of the variance-covariance matrix discussed above for $E_i = 1$ for $i = 1$ to 8, $e^2 = .005$, $r^2 = .007$ and $\rho = .90$, and D, the assumed variance-covariance matrix for the estimation errors associated with development.[182]

$$V_{ij} = (.005)\delta_{ij} + (.007).90^{|i-j|} + D_{ij}, \qquad \text{for} \quad i,j \leq 6,$$

$$V_{ij} = (.005)\delta_{ij} + (.007).90^{|i-j|}, \qquad \text{for} \quad i \text{ or } j > 6.$$

For i or $j > 6$, the year is one whose losses we are trying to estimate. Since we are trying to estimate ultimate losses there is no additional development to be applied to those years. Thus, there is no D term, or alternately $D_{ij} = 0$ for i or $j > 6$.

Solving the Equations 9.2 for the weights we get:

$$Z_1 = 18.4\%,$$

$$Z_2 = 18.7\%,$$

$$Z_3 = 16.5\%,$$

$$Z_4 = 21.0\%,$$

$$Z_5 = 23.1\%, \qquad \text{and}$$

$$Z_6 = 2.3\%.$$

Taking into account the estimation errors due to development has decreased the weight given to recent years. In particular the weight given to incomplete policy year 1996 has declined very significantly. This is in line with the general practice of giving reduced or no weight to the incomplete policy year.

---

[182]We have assumed for simplicity that the estimation errors due to development are independent of the variance of the ultimate values for the years, so that the two variance-covariance matrices add. Also, we have for simplicity not had D depend on the volume of data E, even though in actual applications it is likely to be dependent.

Similarly we can include the impact of the estimation error due to trend using the previously selected variance-covariance matrix T

$$V_{ij} = (.005)\delta_{ij} + (.007).90^{|i-j|} + T_{ij}, \qquad \text{for} \quad i,j \leq 6,$$

$$V_{ij} = (.005)\delta_{ij} + (.007).90^{|i-j|}, \qquad \text{for} \quad i \text{ or } j > 6.$$

Solving the Equations 9.2 for the weights one gets:

$$Z_1 = 7.8\%,$$
$$Z_2 = 6.7\%,$$
$$Z_3 = 8.5\%,$$
$$Z_4 = 12.1\%,$$
$$Z_5 = 23.3\%, \qquad \text{and}$$
$$Z_6 = 41.6\%.$$

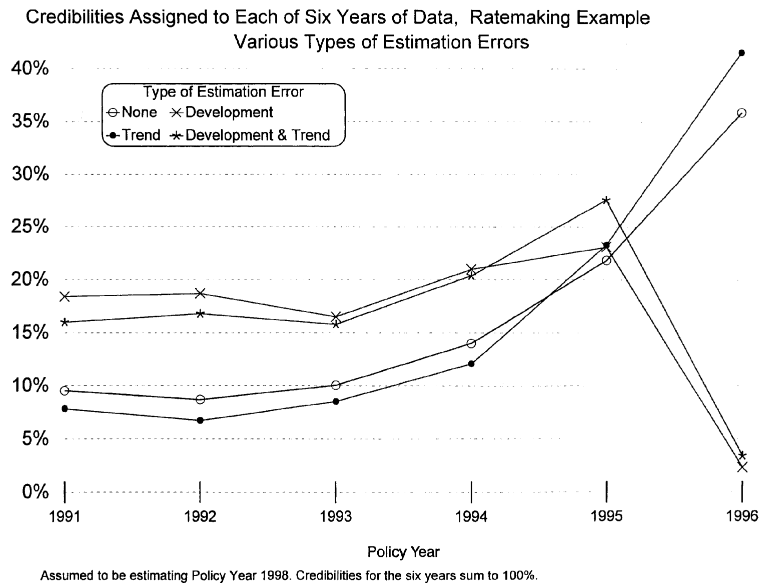
Taking into account the estimation errors due to trend has decreased the weight given to older years.

Finally, we can include the impact of both forms of estimation error by using the matrix $D+T$ in place of either D or T. (This assumes the estimation errors due to development and trend are independent.) The resulting weights are:

$$Z_1 = 16.0\%,$$
$$Z_2 = 16.8\%,$$
$$Z_3 = 15.8\%,$$
$$Z_4 = 20.4\%,$$
$$Z_5 = 27.6\%, \qquad \text{and}$$
$$Z_6 = 3.4\%.$$

Figure 22 compares the weights with and without the estimation errors.

FIGURE 22



9.6. *General Effects of Estimation Error*

The inputs used in the illustrative example can be varied. We can use more or less than six years of data. The gap between the latest year of data and the year to be estimated can differ. The volume of data and therefore the expected value of the process variance can vary by year. The relative size of the variance of the hypothetical means and the expected value of the process variance can differ. The rate at which parameters shift can be faster or slower. The pattern of estimation errors and their relative importance can differ.[183]

As any of these inputs vary, so do the calculated weights. Nevertheless, approximate values of the inputs can be used to estimate a pattern of weights that would be reasonable to use for a particular application.

[183] Also, in some cases the estimation errors would depend on the volume of data.

The general conclusions from analyzing this model all make sense. When we have a smaller volume of data per year we choose a more stable method.[184] Years with less data get less weight. When there is a lot of potential error from estimating loss development for a year, we give that year relatively less weight; this tends to affect more recent years. When there is a lot of potential error from estimating trend or on-level factors for a year, we give that year relatively less weight; this tends to affect more distant years. As there are more rapidly shifting parameters over time we choose a more responsive method.

Recall that in this illustrative example the weights always add to 100%. Thus the weight given to a particular year is a reflection of its value relative to the other years. Giving two years equal weight implies that they have the same value for purposes of estimation, but tells us nothing about what that value is in any absolute sense.

10. EXPERIENCE RATING

In this section, the previous results will be applied to a single split experience rating plan. While the values for the covariance structure used in this section were selected based on analyzing some workers compensation data from one state, they should be viewed as for illustrative purposes.

Section 10.1 describes the structure of a single-split experience rating plan. Section 10.2 describes the covariance structure. Section 10.3 displays the set of linear equations to be solved in order to get the credibilities. The parameters of the covariance structure are estimated and selected in Sections 10.4 to 10.8. Section 10.9 displays the credibilities that correspond to this covariance structure and parameters. Section 10.10 discusses the

[184]This customary practice is illustrated in Stern [28, p. 77]. The larger the premium volume, the more weight given to the latest year of data and the less weight given to the prior year of data. The smaller the premium volume, the more equal the weights given to the two years of data.

impact on the credibilities of taking into account the maturity of experience rating data.

10.1. Structure of the Experience Rating Plan

Assume we have split the losses into primary and excess portions, with the first \$5,000 of losses primary.[185] Assume only the first \$175,000 of any claim enters into experience rating.[186]

Assume we have Y years of data being used to predict year $Y + \Delta$.[187] We wish to determine credibilities to apply to the primary and excess data for each year.

Define the following quantities:

$$
\begin{aligned}
E_{Pi} &= \text{Expected Primary Losses for Year } i,\\
E_{Xi} &= \text{Expected Excess Losses for Year } i,\\
E_i &= E_{Pi} + E_{Xi} = \text{Expected Losses for Year } i,\\
A_{Pi} &= \text{Actual Primary Losses for Year } i,\\
A_{Xi} &= \text{Actual Excess Losses for Year } i,\\
D_i &= E_{Pi}/E_i = D\text{-ratio for Year } i,\\
P_i &= A_{Pi}/E_i,\\
X_i &= A_{Xi}/E_i,\\
\pi_i &= P_i - D_i = (A_{Pi} - E_{Pi})/E_i\\
&= \text{Primary ``Deviation Ratio'' for Year } i,\\
\xi_i &= X_i - (1 - D_i) = (A_{Xi} - E_{Xi})/E_i\\
&= \text{Excess ``Deviation Ratio'' for Year } i, \qquad \text{and}\\
M &= \text{Experience Modification.}
\end{aligned}
$$

---

[185] This is a single split experience rating plan. The \$5,000 split point is currently used for workers compensation. The general results illustrated here would be similar with a different split point.

[186] The \$175,000 limit is currently used in Massachusetts workers compensation. Other states use different limits.

[187] Typically $Y = 3$ and $\Delta = 2$. Years 1, 2 and 3 are predicting Year 5.

Then the experience modification will be of the form:

$$M = 1 + \sum_{i=1}^{Y} \pi_i Z_{Pi} + \sum_{i=1}^{Y} \xi_i Z_{Xi}. \tag{10.1}$$

The primary deviation ratio for year i, π_i, is given weight Z_{Pi}. The excess deviation ratio for year i, ξ_i, is given weight Z_{Xi}. The complement of credibility is given to unity, i.e., the average modification and the expected ratio of actual losses to expected losses.

If we were to introduce ballast and weighting values, as in the current experience rating plan,[188] then one could rewrite the credibilities as:

$$\begin{aligned} Z_{Pi} &= E_i/(E_i + B_i), \\ Z_{Xi} &= W_i Z_{Pi} = W_i E_i/(E_i + B_i). \end{aligned} \tag{10.2}$$

Note that there would be separate ballast and weighting values for each year in the treatment here. In the current experience rating plan there is a single B and W for a given insured.[189]

Then using the definitions of the deviation ratios:

$$\pi_i = (A_{Pi} - E_{Pi})/E_i \qquad \text{and} \qquad \xi_i = (A_{Xi} - E_{Xi})/E_i,$$

we can rewrite Equation 10.1 as:

$$M = 1 + \sum_{i=1}^{Y} \frac{A_{Pi} - E_{Pi} + W_i A_{Xi} - W_i E_{Xi}}{E_i + B_i}. \tag{10.3}$$

By giving each year its own weight, Equations 10.1 or 10.3 differ somewhat from the usual Equation 10.4.[190] If all the years of data were added together and assigned one combined primary credibility and one combined excess credibility, then Equation

[188] See Mahler [12] or Gillam and Snader [19].

[189] If an insured is interstate rated, the W and B values are a weighted average of those that would apply to that size risk if it were intrastate rated in each of the states involved.

[190] See for example Gillam and Snader [19] or Mahler [12].

10.1 would reduce to an equivalent of the usual equation for the experience modification for the Workers Compensation single split plan:

$$\begin{aligned} M &= 1 + Z_p(A_p/E - E_p/E) + Z_x(A_x/E - E_x/E) \\ &= 1 + \frac{A_p - E_p + WA_x - WE_x}{E+B} = \frac{A_p + WA_x + (1-W)E_x + B}{E+B} \end{aligned} \quad (10.4)$$

10.2. Variances and Covariances

The credibilities that appear in Equations 10.1 or 10.2 will be derived from the variance-covariance structure.[191]

There are three types of variances and covariances: those involving just primary deviation ratios, those involving just excess deviation ratios, and those involving both primary and excess deviation ratios. Each covariance will involve ratios from two (possibly different) years.

Define the relevant covariances as:

$$\begin{aligned} S_{ij} &= \mathrm{Cov}[\pi_i, \pi_j] = S_{ji}, \\ T_{ij} &= \mathrm{Cov}[\xi_i, \xi_j] = T_{ji}, \qquad \text{and} \\ U_{ij} &= \mathrm{Cov}[\pi_i, \xi_j]. \end{aligned} \quad (10.5)$$

Each of these three variance-covariance structures S, T and U would need to be modeled and/or estimated in a manner similar to that performed in previous sections of this paper. The covariances would differ by the amount of data and would be affected by risk heterogeneity, parameter uncertainty, and shifting risk parameters over time.

[191]The "best" credibilities will be taken as those that minimize the expected squared error. See Appendix D.

When the two years involved are the same, we obtain the total[192] variance or covariance:

$S_{1,1}$ = total variance of the primary deviation ratio,

$T_{1,1}$ = total variance of the excess deviation ratio, and

$U_{1,1}$ = total covariance of the primary and excess deviation ratios.

When the two years involved differ, we obtain in the *absence* of shifting risk parameters over time, the variance or covariance of the hypothetical means:

$S_{1,2}$ = variance of the hypothetical mean primary deviation ratios,

$T_{1,2}$ = variance of the hypothetical mean excess deviation ratios, and

$U_{1,2}$ = covariance of the hypothetical mean primary and excess deviation ratios.

In the presence of shifting risk parameters over time, it will be assumed that S, T and U each have a structure similar to that in Equations 5.10 and 5.11:

$$\mathrm{Cov}[X_i, X_j] = r^2\left\{\rho^{|i-j|} + I\gamma^{|i-j|}/\sqrt{E_iE_j} + \delta_{ij}\left(K/\sqrt{E_iE_j} + J\right)\right\}, \quad \sqrt{E_iE_j} \geq \Omega;$$

$$\mathrm{Cov}[X_i, X_j] = r^2\left\{\rho^{|i-j|} + I\gamma^{|i-j|}/\Omega + \delta_{ij}\left(K/\sqrt{E_iE_j} + J\right)\right\}, \quad \sqrt{E_iE_j} \leq \Omega.$$

The parameters r^2, I, J, K, ρ, γ and Ω in general may vary between the covariance structures for S, T and U. Thus, we will write each parameter with a subscript, p for primary, x for excess,

[192]"Total" means including both the variance (or covariance) of the hypothetical means and the expected value of the process variance (or covariance). See Mahler [11].

and m for mixed, resulting in the following equations:

$$S_{ij} = r_p^2 \left\{ \rho_p^{|i-j|} + I_p \gamma_p^{|i-j|} / \sqrt{E_i E_j} + \delta_{ij} \left(K_p / \sqrt{E_i E_j} + J_p \right) \right\}, \quad \sqrt{E_i E_j} \geq \Omega_p; \qquad (10.6)$$

$$S_{ij} = r_p^2 \left\{ \rho_p^{|i-j|} + I_p \gamma_p^{|i-j|} / \Omega_p + \delta_{ij} \left(K_p / \sqrt{E_i E_j} + J_p \right) \right\}, \quad \sqrt{E_i E_j} \leq \Omega_p; \qquad (10.7)$$

$$T_{ij} = r_x^2 \left\{ \rho_x^{|i-j|} + I_x \gamma_x^{|i-j|} / \sqrt{E_i E_j} + \delta_{ij} \left(K_x / \sqrt{E_i E_j} + J_x \right) \right\}, \quad \sqrt{E_i E_j} \geq \Omega_x; \qquad (10.8)$$

$$T_{ij} = r_x^2 \left\{ \rho_x^{|i-j|} + I_x \gamma_x^{|i-j|} / \Omega_x + \delta_{ij} \left(K_x / \sqrt{E_i E_j} + J_x \right) \right\}, \quad \sqrt{E_i E_j} \leq \Omega_x; \qquad (10.9)$$

$$U_{ij} = r_m^2 \left\{ \rho_m^{|i-j|} + I_m \gamma_m^{|i-j|} / \sqrt{E_i E_j} + \delta_{ij} \left(K_m / \sqrt{E_i E_j} + J_m \right) \right\}, \quad \sqrt{E_i E_j} \leq \Omega_m; \quad \text{and} \qquad (10.10)$$

$$U_{ij} = r_m^2 \left\{ \rho_m^{|i-j|} + I_m \gamma_m^{|i-j|} / \Omega_m + \delta_{ij} \left(K_m / \sqrt{E_i E_j} + J_m \right) \right\}, \quad \sqrt{E_i E_j} \leq \Omega_m. \qquad (10.11)$$

The covariance structure given by Equations 10.6 to 10.11 includes a total of 21 parameters. In theory, these parameters can be estimated using techniques similar to those used in the previous sections of this paper. As a practical matter, some of the parameters such as Ω, ρ and γ can be taken equal or approximately equal for S, T and U. So, for example, we could assume $\Omega_p = \Omega_x = \Omega_m$; in other words, we could assume that the transition from risk homogeneity to risk heterogeneity occurs at (approximately) the same size[193] for all three covariance structures.

[193] As applied here to experience rating, I have followed the current practice of using the total expected losses rather than the primary or excess losses to define the size of risk.

10.3. Equations for Credibilities

Set aside for now the difficult task of estimating the variance-covariance matrices: S, T and U. As shown in Appendix D, we can derive 2 Y linear equations for the 2 Y credibilities:

$$\sum_{i=1}^{Y}(Z_{Pi}S_{ik} + Z_{Xi}U_{ki}) = S_{k,Y+\Delta} + U_{k,Y+\Delta}, \qquad k = 1,2,\ldots Y, \quad \text{and} \tag{10.12}$$

$$\sum_{i=1}^{Y}(Z_{Pi}U_{ik} + Z_{Xi}T_{ki}) = U_{Y+\Delta,k} + T_{k,Y+\Delta}, \qquad k = 1,2,\ldots Y. \tag{10.13}$$

If the excess losses are set equal to zero; i.e., we have a no-split plan, then Equation 10.12 reduces to Equation 2.4. In the absence of shifting risk parameters over time, as shown in Appendix D, Equations 10.12 and 10.13 reduce to those derived in Mahler [11].[194]

10.4. Estimating the Parameters of the Covariance Structure

Prior sections have discussed how we might estimate some of the needed parameters. Also, the National Council on Compensation Insurance has estimated quantities which are similar to the I, J and K parameters here.[195] These NCCI estimates can aid in choosing the relative sizes of the I, J and K parameters.

The available data was insufficient to allow independent estimates of ρ_p, ρ_x and ρ_m, so it is assumed that $\rho_p \approx \rho_x \approx \rho_m$.

Thus, for a given insured, the size of risk to which we compare Ω_p, Ω_x or Ω_m is the same. In this case, I think it unlikely that Ω_p, Ω_x and Ω_m would differ. Nevertheless, for generality, I have labeled Ω with subscripts even though in the example $\Omega_p = \Omega_x = \Omega_m$.

[194]See Equations 5.3 and 5.4 in Mahler [11] for Z_p and Z_x for a split experience rating plan.

[195]See Gillam [13] and Mahler [12]. Note that the credibilities in the NCCI Revised Experience Rating Plan were derived without explicit recognition of the impact of what has been called herein U_{ij}, the covariance of the primary and excess losses.

Similarly, assume $\gamma_p = \gamma_x = \gamma_m$ and $\Omega_p = \Omega_x = \Omega_m$. So we have assumed that the rate of shifting parameters over time as it impacts S, T and U is similar and that risk homogeneity applies for risks of size less than Ω.

The primary losses are less subject to random fluctuations than the excess losses. Therefore, whenever possible the results of analyzing the primary deviation ratios will be relied upon.

The data analyzed was that used for intrastate experience rating in one state over a five year period.[196] The analysis was limited to risks that were experience rated over this whole period of time.[197] For each such risk, for each "rating year" the data consists of three separate years of actual primary losses, actual excess losses, expected primary losses, and expected excess losses, that were used to calculate the experience modification. The variance-covariance structure of this data was analyzed by size of risk.

For example, for risks with expected annual losses between \$10,000 and \$20,000 the correlations between primary deviation ratios, $(A_p - E_p)/(E_p + E_x)$, were computed for different separations and different reports.[198] For example, this primary correlation was .331 between the "rating year" 1991 data at first report[199] and the "rating year" 1992 data at first report. Table 6 displays the correlations.

There are 12 correlations corresponding to a separation of 1 year, 9 for 2 years, 6 for 3 years, and 3 for 4 years. Based

[196]For experience modifications applied to policies written during 1991, 1992, 1993, 1994 and 1995 in Massachusetts workers compensation.

[197]Employers who went out of business, left the state, became self-insured or became too small to be experience rated would therefore be excluded.

[198]This differs somewhat from Mahler [12] where correlations between $A_p/(E_p + E_x)$ were examined. The two sets of correlations are very similar.

[199]Generally data from a 1989 policy at first report, a 1988 policy at second report, and a 1987 policy at third report would be used to calculate the experience modification to apply to the 1991 policy. The data from the 1989 policy at first report is what is being referred to here.

TABLE 6

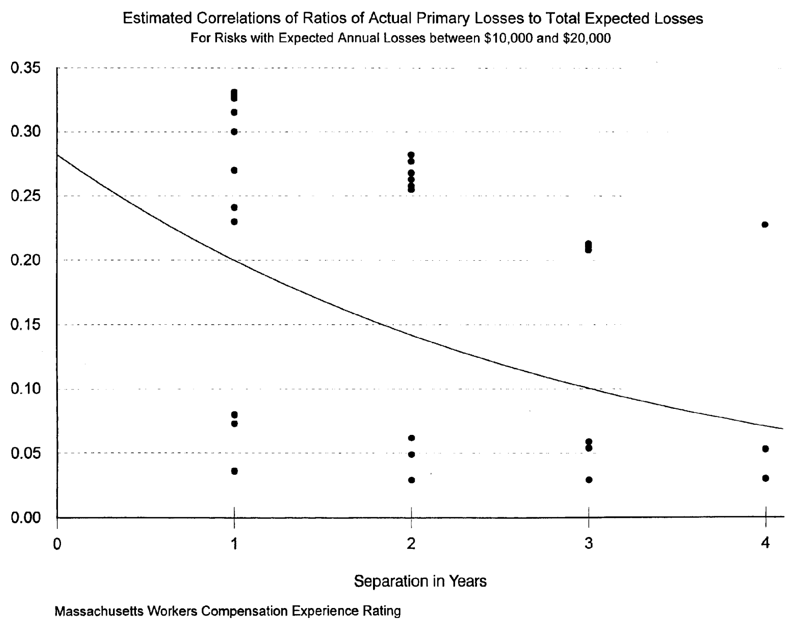
CORRELATIONS OF RATIOS OF ACTUAL PRIMARY LOSSES TO TOTAL EXPECTED LOSSES
Expected Annual Losses[1] Between $10,000 and $20,000
Massachusetts Workers Compensation Experience Rating

Rating Years[2]	Report	Separation	Correlation
91 92	1	1	.331
91 92	2	1	.230
91 92	3	1	.270
92 93	1	1	.328
92 93	2	1	.326
92 93	3	1	.241
93 94	1	1	.080
93 94	2	1	.300
93 94	3	1	.330
94 95	1	1	.036
94 95	2	1	.073
94 95	3	1	.315
91 93	1	2	.282
91 93	2	2	.255
91 93	3	2	.258
92 94	1	2	.062
92 94	2	2	.268
92 94	3	2	.263
93 95	1	2	.049
93 95	2	2	.029
93 95	3	2	.277
91 94	1	3	.059
91 94	2	3	.211
91 94	3	3	.208
92 95	1	3	.054
92 95	2	3	.029
92 95	3	3	.213
91 95	1	4	.053
91 95	2	4	.030
91 95	3	4	.228

[1]If E_1 and E_2 are the expected losses (primary plus excess) for the given report which are used for experience rating the two rating years, then $\sqrt{E_1 E_2}$ is between $10,000 and $20,000. There were an average of 3,060 such risks.

[2]91 refers to experience modifications applied to policies written in 1991.

FIGURE 23



on Mahler [1], it is expected that the logs of these correlations will decline linearly as the separation increases. A least squares regression was fit to these correlations, and the result was $c = (.282).709^s$, where c is the correlation and s is the separation. The value of .282 will be referred to as the "intercept" while the value of .709 will be referred to as the "slope" of this regression. This regression is illustrated in Figure 23.

Similar regressions were fit to the correlations for other size categories.[200] A similar analysis was performed for the correlations of excess deviation ratios and the correlations between primary and excess deviation ratios. The resulting slopes and intercepts are displayed in Table 7.

[200] A few estimated correlations were not positive and were excluded from the regressions.

TABLE 7

RESULTS OF EXPONENTIAL REGRESSIONS FIT TO CORRELATIONS OF RATIOS
Massachusetts Workers Compensation Experience Rating

Expected Annual Losses ($000)	Average Number of Risks	Primary		Excess		Mixed	
		Intercept	Slope	Intercept	Slope	Intercept	Slope
3 to 5	3,952	.224	.727	.034	.820	.089	.775
5 to 10	4,798	.169	.836	.086	.792	.102	.841
10 to 20	3,060	.282	.709	.098	.741	.126	.737
20 to 50	2,197	.380	.992	.146	.842	.167	.950
50 to 100	770	.579	.869	.260	.809	.272	.855
100 to 200	356	.717	.865	.442	.723	.408	.790
200 to 500	186	.661	.877	.471	.812	.355	.825
500 to 1,000	45	.869	.658	.693	.687	.397	.781
1,000 to 2,000	14	.882	.973	.776	.850	.583	.828

10.5. Estimating I_p, J_p, K_p, I_x, J_x and K_x

As discussed previously, the intercepts of the primary correlations are an estimate of the credibility to be assigned to a single year of data in the absence of shifting risk parameters.[201] Thus we expect a curve of the form:

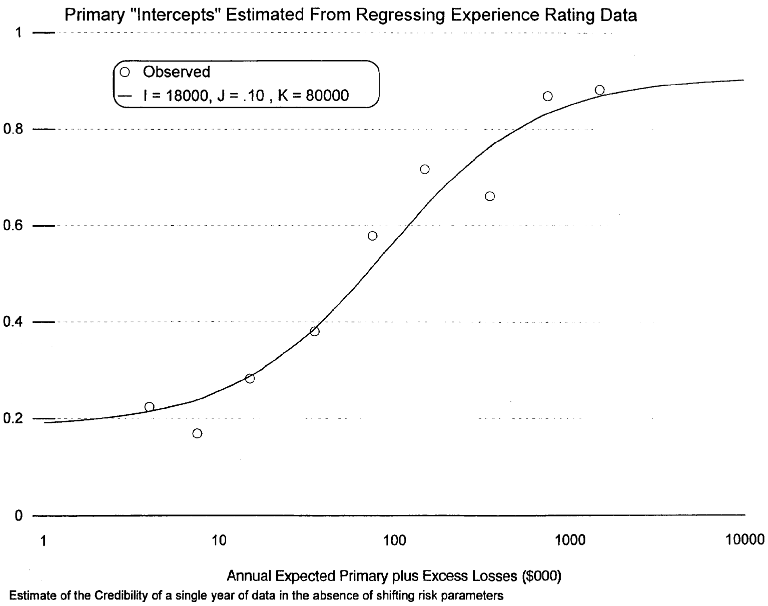
$$Z = (E + I_p)/\{(1 + J_p)E + I_p + K_p\}, \qquad \text{for} \quad E \geq \Omega_p.$$

As shown in Figure 24, the values $I_p = 18{,}000$, $J_p = .10$, and $K_p = 80{,}000$ do a reasonable job of approximating the estimated intercepts for the primary deviation ratios.[202]

Similarly, as seen in Figure 25, values of $I_x = 20{,}000$, $J_x = .15$, and $K_x = 315{,}000$ do a reasonable job of approximating the estimated intercepts for the excess deviation ratios.

[201] The correlation between primary and excess losses is also ignored.
[202] More data on extremely large risks would improve the estimate of J_p.

FIGURE 24



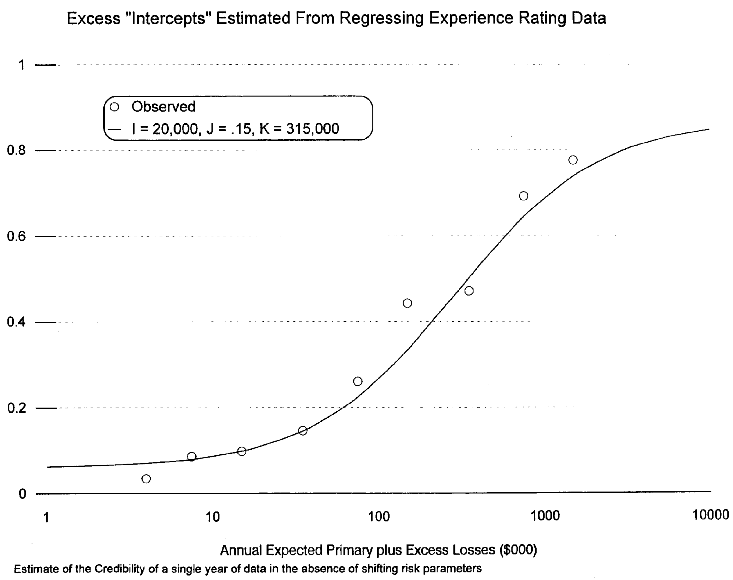
The intercepts of the regressions fit to the mixed correlations have a somewhat different interpretation. Using Equations 10.6, 10.8 and 10.10, for a primary deviation ratio π_i and excess deviation ratio ξ_j for different years $i \neq j$, we have for $\sqrt{E_i E_j} \geq \Omega_m$, $E_i \geq \Omega_p$ and $E_j \geq \Omega_x$:

$$\mathrm{Corr}[\pi_i, \xi_j] = \mathrm{Cov}[\pi_i, \xi_j] / \sqrt{\mathrm{Var}(\pi_i)\mathrm{Var}(\xi_j)} = U_{ij} / \sqrt{S_{ii} T_{jj}}, \qquad \text{and}$$

$$\mathrm{Corr}[\pi_i, \xi_j] = \frac{r_m^2 \left\{ \rho_m^{|i-j|} + I_m \gamma_m^{|i-j|} / \sqrt{E_i E_j} \right\}}{r_p r_x \sqrt{(1 + J_p + (I_p + K_p)/E_i)(1 + J_x + (I_x + K_x)/E_j)}}. \tag{10.14}$$

Note that the mixed correlation between different years does not involve J_m and K_m. Thus the regression fit to the mixed intercepts cannot be used to estimate these parameters. The intercept

FIGURE 25



of that regression should be:[203]

$$\frac{r_m^2\left\{1+I_m/\sqrt{E_iE_j}\right\}}{r_pr_x\sqrt{(1+J_p+(I_p+K_p)/E_i)(1+J_x+(I_x+K_x)/E_j)}}. \qquad (10.15)$$

These intercepts by size of risk will be used subsequently to check the reasonableness of selected parameter values.

10.6. Estimating γ and ρ

The slopes of the regressions fit to the correlations are displayed in Table 7. There is considerable random fluctuation, but generally the slopes are somewhere in the 75% to 90% range. As discussed previously, the slope for smaller sizes should be

[203]The result of substituting unity for ρ_m and γ_m in Equation 10.14.

approximately equal to γ, while that for larger sizes should be approximately equal to ρ. There is some tendency for the primary ratios for the slopes to be closer to unity for large sizes. For illustrative purposes, $\gamma_p = .80$ and $\rho_p = .85$ will be selected.

There is less evidence of a dependence on size of risk for the excess ratios; $\gamma_x = \rho_x = .80$ will be selected. The rates of shifting for the mixed correlations are similar to those for the primary and excess correlations; $\gamma_m = .80$ and $\rho_m = .83$ will be selected.

10.7. Estimating r^2, I_m, J_m and K_m

Besides analyzing correlations between data from different years, we need to analyze the variance of data from a single year. The variance is S_{ii} for primary deviation ratios:

$$S_{ii} = r_p^2(1 + J_p + (I_p + K_p)/E_i), \qquad E_i \geq \Omega. \qquad (10.16)$$

Similarly, for the excess deviation ratios the variance is

$$T_{ii} = r_x^2(1 + J_x + (I_x + K_x)/E_i), \qquad E_i \geq \Omega. \qquad (10.17)$$

For a given year, the covariance between the primary and excess deviation ratios is U_{ii}:

$$U_{ii} = r_m^2(1 + J_m + (I_m + K_m)/E_i), \qquad E_i \geq \Omega. \qquad (10.18)$$

The estimated variances and covariances for various sizes of risk are shown in Table 8.[204] Using the estimated primary variances and the previously selected values I_p, J_p and K_p we can estimate r_p^2. Similarly, we can estimate r_x^2. The covariances can be used to estimate r_m^2, I_m, J_m and K_m.

Table 9 shows the estimates of r_p^2 and r_x^2 that result from the estimated variances for the different sizes of risk and Equations

[204] In each case, the value shown is an average of 15 values from 5 years and 3 reports.

TABLE 8

VARIANCES AND COVARIANCES FOR A SINGLE YEAR Massachusetts Workers Compensation Experience Rating

Expected Annual Losses ($000)	Average Number of Risks	Primary Variance	Excess Variance	Mixed Covariance[1]
3 to 5	6,228	.271	15.444	1.221
5 to 10	6,619	.181	9.754	.791
10 to 20	4,081	.101	5.526	.442
20 to 30	1,569	.077	3.344	.301
30 to 50	1,318	.064	2.459	.222
50 to 100	1,034	.047	1.816	.177
100 to 200	506	.034	1.045	.112
200 to 500	262	.022	.637	.069
500 to 1,000	67	.019	.405	.055
1,000 to 2,000	23	.016	.213	.034

In each case the estimate shown is the average of 15 estimates from each of 5 years at 3 reports.
[1]Covariance of primary and excess deviation ratios for the same year.

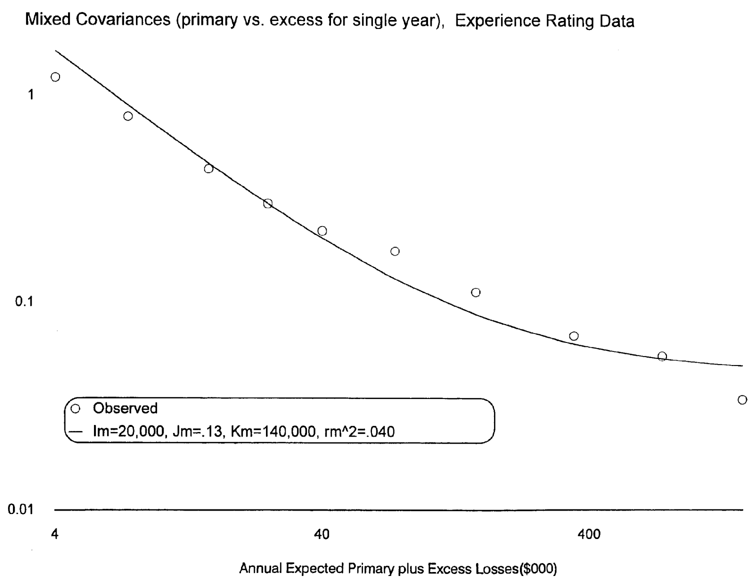
TABLE 9

ESTIMATES OF r^2 FROM OBSERVED VARIANCES

Expected Annual Losses ($000)	Primary Variance	r_p^2	Excess Variance	r_x^2
4	.271	.011	15.444	.182
7.5	.181	.013	9.754	.213
15	.101	.013	5.526	.235
25	.077	.015	3.344	.230
40	.064	.018	2.459	.258
75	.047	.020	1.816	.323
150	.034	.019	1.045	.309
350	.022	.016	.637	.302
750	.019	.015	.405	.254
1,500	.016	.014	.213	.155

$r^2 = \text{Variance}/(1 + J + (I + K)/E)$
$I_p = 18{,}000$, $J_p = .10$, $K_p = 80{,}000$
$I_x = 20{,}000$, $J_x = .15$, $K_x = 315{,}000$

FIGURE 26



10.16 and 10.17. The values of r_p^2 are all in the range of .015. The similarity of the estimates of r_p^2 that result from the different size categories tends to confirm the reasonableness of the previously selected values of I_p, J_p and K_p.[205]

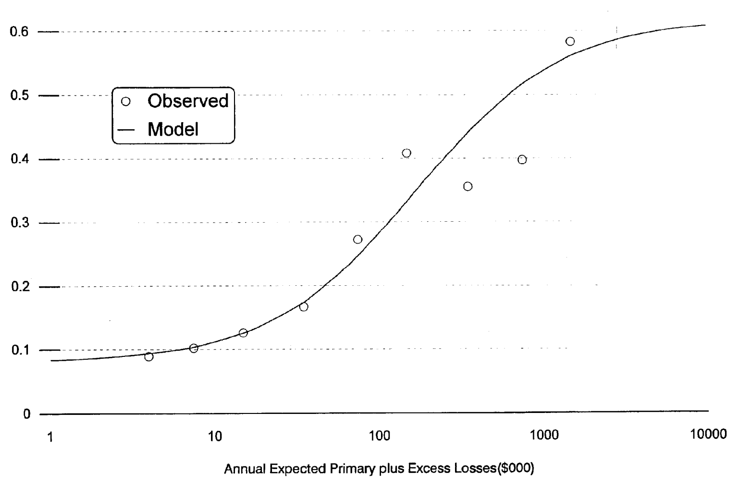
Similarly, Table 9 displays estimates for r_x^2 from the different size categories using Equation 10.17. The values of r_x^2 vary considerably. A value $r_x^2 \approx .26$ will be selected.

As seen in Figure 26, using Equation 10.18, the set of parameters: $I_m = 20{,}000$, $J_m = .13$, $K_m = 140{,}000$, and $r_m^2 = .040$, provides a reasonable fit to the estimated covariances by size of

[205]If the initially selected I_p, J_p and K_p did not seem to perform well here, then we could modify them somewhat so they performed better here. Then we would go back and check the performance in fitting the intercepts of the regressions fit to the correlations. We could iterate in this manner until we arrived at the best set of parameters.

FIGURE 27

Mixed "Intercepts" Estimated From Regressing Experience Rating Data



Model using values for the parameters of Ip = 18,000, Jp = .1, Kp = 80,000, Ix = 20,000, Jx = .15, Kx = 315,000, Im = 20,000, rp^2 =.015, rx^2 = .26, and rm^2 = .040.

risk.[206] Bear in mind that the largest size category has a limited number of risks and so the resulting estimate of the covariance is not very accurate.

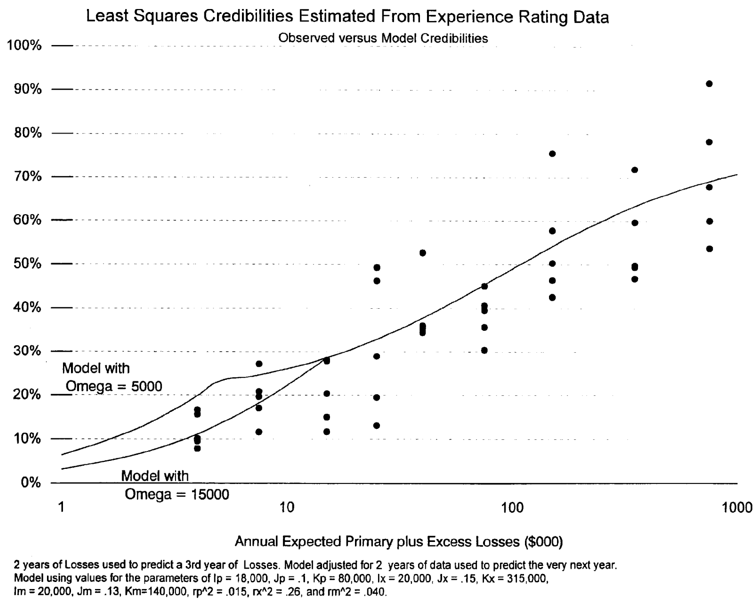
Using the selected set of parameters, we can compare the theoretical values from Equation 10.15 to the observed intercepts from the regressions fit to the mixed correlations. As seen in Figure 27, the fit is not unreasonable. Thus, the selected values of I_p, J_p, K_p, I_x, J_x, K_x, I_m, r_p^2, r_x^2, and r_m^2 seem consistent with the observed mixed intercepts.

10.8. *Selecting* Ω

The final parameters to be selected are Ω_p, Ω_x and Ω_m. Based on the reasonable fits obtained so far, Ω should be near the

[206]The value of J_m was selected to be between the selected J_p and J_x. More data on extremely large risks would improve the estimates of all the J parameters.

FIGURE 28



smaller sizes of risk examined or below the eligibility level for experience rating in Massachusetts.[207] Due to limited information, one value will be selected for all three parameters, $\Omega_p = \Omega_x = \Omega_m$.

Figure 28 displays least squares credibilities estimated from the 3 years of data used to experience rate policies. The credibilities are those that produced the smallest squared error when the first 2 years of data were used to predict the third.[208] These are compared to the model credibilities that result from the estimated parameters and the use of Equations 10.12 and 10.13.[209]

[207] If Ω were in the middle of the range of sizes examined, the observed covariance structure should have been affected.

[208] For each size category there are five estimates, one for each "rating year."

[209] The primary and excess credibilities were averaged: $Z = DZ_p + (1 - D)Z_x$, with $D = .22$.

The model credibilities are in the range indicated by the data.[210]

The credibilities for smaller size risks are shown for two values of Ω, $\Omega = \$5{,}000$ and $\Omega = \$15{,}000$. Based on Figure 28, $\Omega = \$15{,}000$ does a better job than $\Omega = \$5{,}000$. However, a better estimate of Ω would result from a more detailed analysis of data from risks barely eligible for or too small to be experience rated in Massachusetts.[211] While it is beyond the scope of this paper, a preliminary review of merit rating data for Massachusetts workers compensation indicates that $\Omega \approx \$5{,}000$ or perhaps even a little less. In any case, for illustrative purposes the selected values will be $\Omega_p = \Omega_x = \Omega_m = \$5{,}000$.[212]

10.9. Estimated Credibilities

The selected parameter values are:

$$\begin{array}{llll}
I_p = \$18{,}000 & J_p = .10 & K_p = \$80{,}000 & r_p^2 = .015 \\
I_x = \$20{,}000 & J_x = .15 & K_x = \$315{,}000 & r_x^2 = .26 \\
I_m = \$20{,}000 & J_m = .13 & K_m = \$140{,}000 & r_m^2 = .040
\end{array}$$

$$\begin{array}{lll}
\gamma_p = .80 & \rho_p = .85 & \Omega_p = \$5{,}000 \\
\gamma_x = .80 & \rho_x = .80 & \Omega_x = \$5{,}000 \\
\gamma_m = .80 & \rho_m = .83 & \Omega_m = \$5{,}000.
\end{array}$$

Using the above parameter values and Equations 10.12 and 10.13, credibilities were calculated for 3 years of data being used

[210]Since the parameters were estimated by a different analysis of this exact same data, this serves as a consistency check rather than an independent test of the results.

[211]For example, for a risk with \$1,000 in expected annual losses, with $\Omega = \$5{,}000$ $Z = 6.7\%$, while with $\Omega = \$15{,}000$ $Z = 3.3\%$. Thus an examination of the credibilities indicated by the data for smaller risks should help to determine the appropriate Ω.

[212]This would correspond to a minimum ballast value of $K_p\Omega_p/(I_p + \Omega_p) = (\$80{,}000)(5/(18+5)) \approx \$17.3$ thousand. Interestingly, for $g = 7$ as in Massachusetts, the NCCI minimum ballast value would be (7) (2,500) = \$17,500. The corresponding minimum weighting value would be $(K_p/K_x)(\Omega_p/\Omega_x)(I_x + \Omega_x)/(I_p + \Omega_p) = (80/300)(1)(25/23) = .29$. The NCCI minimum W is .07.

TABLE 10

EXPERIENCE RATING CREDIBILITIES[1]
Using Parameters Listed in Section 10.9

Expected Annual Losses ($000)	Primary[2]				Excess[3]				Combined[4]
	Z_1	Z_2	Z_3	Z_p	Z_1	Z_2	Z_3	Z_x	Z
.1	.9%	1.1%	1.3%	3.3%	.02%	.02%	.03%	.1%	.8%
.5	4.0	5.0	6.2	15.2	.09	.11	.14	.3	3.6
1	7.2	9.1	11.7	28.1	.2	.2	.3	.7	6.8
2	12.1	15.6	20.8	48.4	.5	.6	.7	1.7	12.0
3	15.3	20.3	28.1	63.7	.7	.9	1.1	2.7	16.2
4	17.5	23.8	34.1	75.4	1.0	1.3	1.5	3.8	19.6
5	19.0	26.4	39.2	84.5	1.3	1.7	2.0	5.0	22.5
7.5	19.9	27.8	41.6	89.3	1.5	1.8	2.2	5.4	23.9
10	20.6	29.0	43.9	93.5	1.6	2.0	2.4	5.9	25.2
25	22.7	33.5	54.5	110.6	2.3	2.9	3.6	8.8	31.2
50	22.0	35.8	65.4	123.3	3.4	4.4	5.5	13.3	37.5
100	17.3	34.7	77.3	129.3	5.0	6.6	8.7	20.3	44.3
250	6.9	27.0	90.6	124.4	6.8	10.0	15.3	32.1	52.4
500	.8	19.6	95.5	115.9	7.0	11.8	21.5	40.3	56.9
1,000	−1.3	14.5	94.8	108.1	6.1	12.4	27.9	46.4	59.9
2,500	−.2	12.6	89.1	101.5	4.3	11.6	35.1	51.0	62.1
5,000	1.0	13.1	84.6	98.8	3.3	10.8	38.8	52.9	63.0
10,000	1.9	14.0	81.4	97.3	2.7	10.1	41.1	53.8	63.4
∞	2.9	15.5	77.4	95.7	2.0	9.1	43.7	54.8	63.8

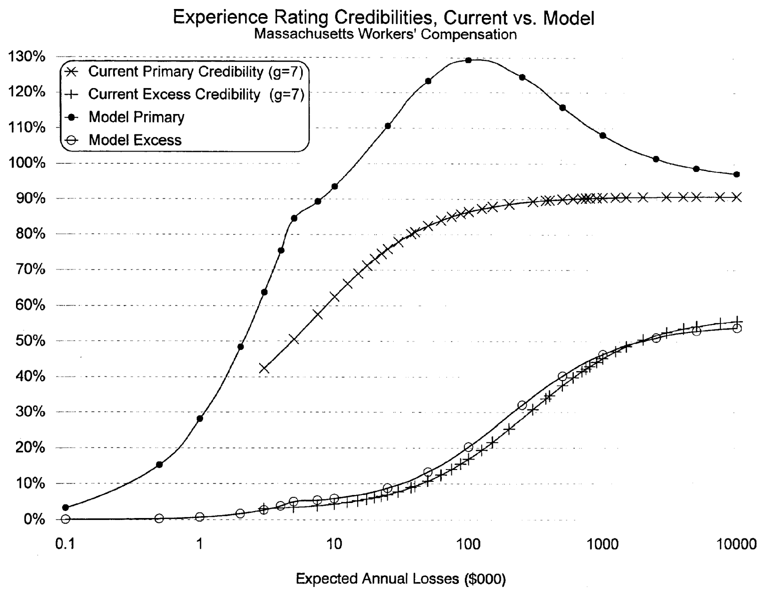
[1]Using data from years 1, 2 and 3 to predict year 5.
[2]Z_p is the sum of the primary credibilities for the three years.
[3]Z_x is the sum of the excess credibilities for the three years.
[4]$Z = DZ_p + (1 - D)Z_x$, for $D = .22$.

to predict the fifth year.[213] Table 10 displays the primary and excess credibilities assigned to each of the three years of data as well as the sum. Note that the primary credibilities can sum to greater than 100%. As pointed out in Mahler [11] and Mahler [12], this is not unusual when we take into account the covariance of the primary and excess losses.[214] In such circumstances

[213]For example, 1994, 1995 and 1996 data is being used to experience rate a policy written during 1998.

[214]Note that the numerator of π_i involves the primary losses while the denominator is the sum of the primary and excess expected losses. If instead the denominator had been just primary expected losses, then the ratio would be larger and the weight assigned to it would be smaller by a factor of the D-ratio. Then the primary weights would sum to less than 100%.

FIGURE 29

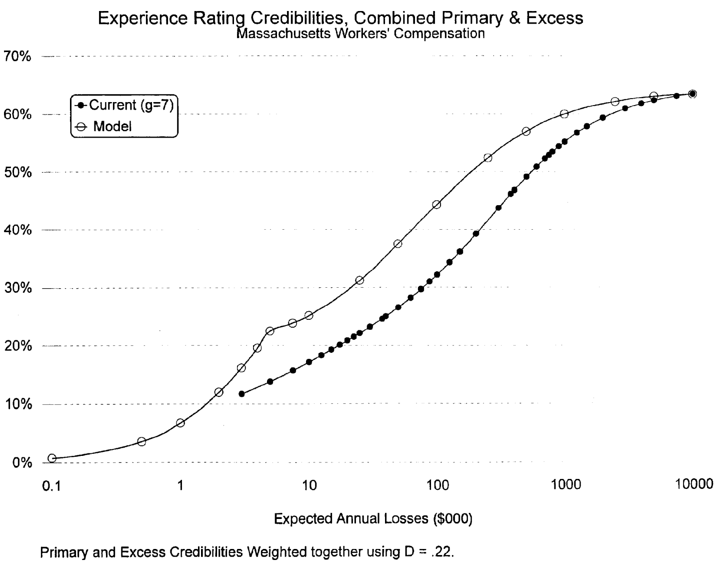


we could constrain the primary credibilities to be equal to unity, as shown in Mahler [11] and Mahler [12]. In any case, the combined credibility is between 0 and 1. It should also be noted that the uncertainty in the estimated J and Ω parameters produces uncertainty in the credibilities for large and small risks respectively.

As the size of risk increases, the weight assigned to the most recent year increases relative to that for the most distant year. For very large risks, we can rely almost solely on the latest year of data. For very small risks, it would be reasonable to rely on more than three years of data, since the older years would have credibilities close to that for the more recent years of data.

Figure 29 compares the primary and excess credibilities from Table 10 to those currently used in Massachusetts workers com-

FIGURE 30



pensation experience rating.[215] Figure 30 does the same comparison for the weighted average of the primary and excess credibilities. As in Mahler [12], the indicated primary and combined credibilities are generally higher than those from the NCCI plan.

At least part of this difference is due to the fact that Massachusetts average claim costs are higher than the national average. Using the same $5,000 split point between primary and excess in every state results in lower than average D-ratios in Massachusetts. Thus, the primary losses in Massachusetts are "very primary," while the excess losses are only "mildly excess." Thus, both the Massachusetts primary losses and excess losses contain more useful information and less random noise than in

[215]The NCCI Revised Experience Rating program, with $g = 7$. Here we have ignored the All Risk Rating Program (ARAP) which is currently applied on top of experience rating in Massachusetts and in combination produces more responsiveness to the insured's losses.

the average state. This would not be the case if the split point depended on the state specific parameter g.

On the other hand, due to the consideration here of the covariance between the primary and excess losses, the primary credibilities are higher and the excess credibilities are lower than they would otherwise be. The primary losses contain valuable information for predicting both the future primary and excess losses.

On balance, the excess credibilities for the current model are fairly close to those from the NCCI plan, while the primary credibilities are much greater. As stated before, the results would be expected to differ somewhat in low severity states.

In any case, the combined model credibilities are more similar to what would be obtained in other states.[216] The combined credibilities are between 0 and 1. In this case, they increase smoothly from zero to a maximum of about 63% for the largest risks.[217] Due to shifting risk parameters and parameter uncertainty, the maximum credibility is less than 100%.

The model combined credibilities are generally larger than those from the current NCCI plan. For example, for $100,000 in expected annual losses, the model has a combined credibility of 44.3%, while the current plan has 32.2%. While there are significant differences,[218] the overall magnitude and pattern of credibilities is very similar.

Note that model credibilities are also shown for risk sizes below the current eligibility level for experience rating.[219] Recall

---

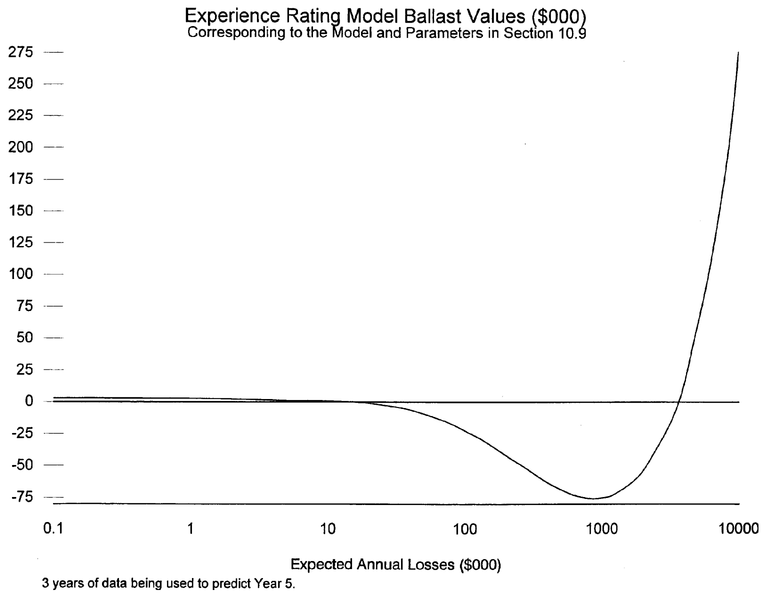
[216]The D-ratio is lower in Massachusetts than in the average state, so the primary credibilities receive less weight. This would result in lower combined credibilities, except that the primary credibilities are larger than average in Massachusetts.

[217]Due to the limited data for very large risks, the model parameters were chosen to some extent so that the maximum credibility would be close to that from the NCCI plan. In an average state the NCCI plan has a maximum credibility of about 67%, as shown in Mahler [12].

[218]For most insureds a 13% difference in credibilities would have a less than 3% effect on their experience modifications.

[219]The NCCI formulas for credibility are not intended to be applied to very small risks. As discussed in Mahler [12], minimum B values, etc., are used to deal with this problem. The NCCI credibilities graphed here are prior to any such refinements.

FIGURE 31

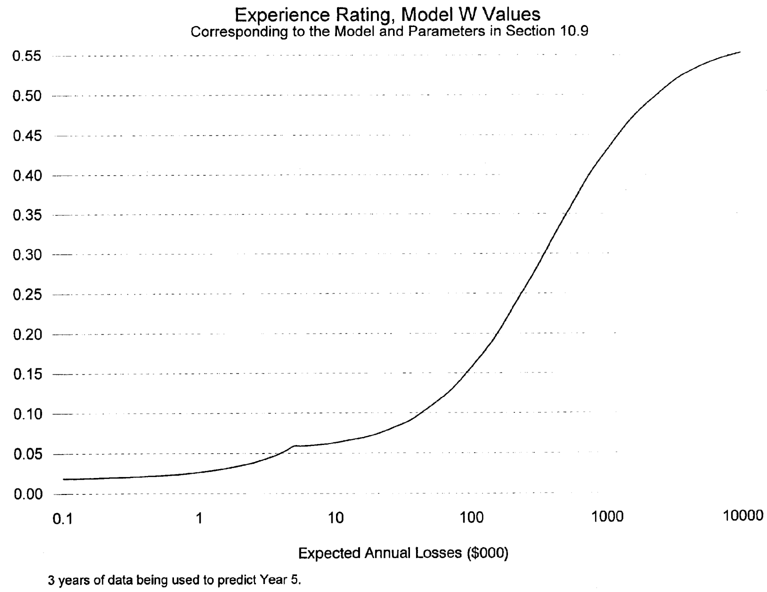


that these credibilities for very small risks depend very significantly on the estimated Ω parameter. An analysis of data from these very small risks would refine this estimate.

Figure 31 shows the ballast values corresponding to the model primary credibilities shown in Table 10. Since $B = E((1/Z_p) - 1)$, when $Z_p >> 100\%$, it follows that $B < 0$. While it is currently the case that B is positive, there is no mathematical reason why B cannot be negative.[220] Small risks have $B \approx 3{,}000$. B declines and becomes negative before increasing to very large positive values. Figure 32 shows W (weighting) values corresponding to the model credibilities shown in Table 10. Other than a discontinuity in the derivative of W that occurs at $\Omega = 5{,}000$, the W values increase smoothly with size of risk.

[220] $B = 0$ would correspond to $Z_p = 100\%$.

FIGURE 32

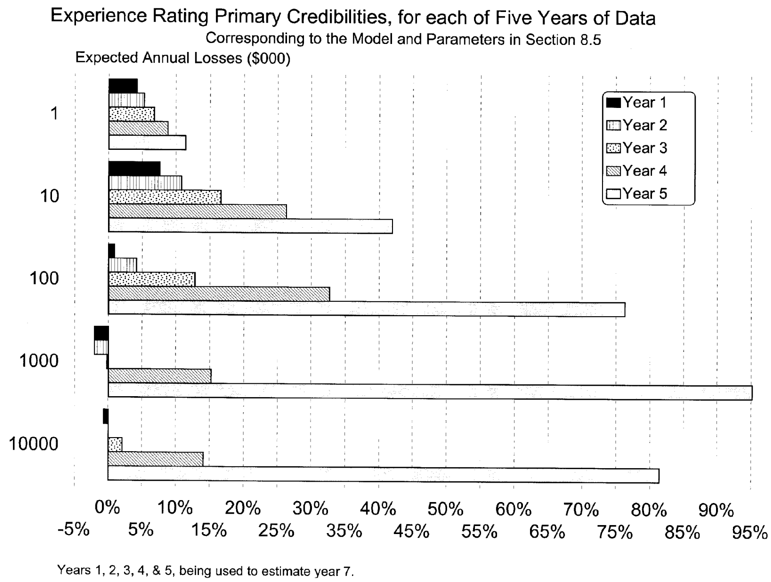


Currently each of three years used for experience rating is treated similarly. Instead each year could receive different credibilities. Figure 33 displays the model primary credibilities assigned to each of five years of data for various sizes of risks. Note that the weights assigned to an individual year of primary losses can be negative.[221] Figure 34 similarly displays the excess credibilities. The same pattern is observed in each figure, although for a given size of risk the weights given to different years are more similar to each other in the excess case than in the primary case.

It should be noted that for simplicity, equal expected losses have been assumed for each year. Equations 10.6 to 10.11 and Equations 10.12 and 10.13 apply equally well when the expected

[221] Also, the weights assigned to individual years of primary losses can in theory be greater than 100%.

FIGURE 33

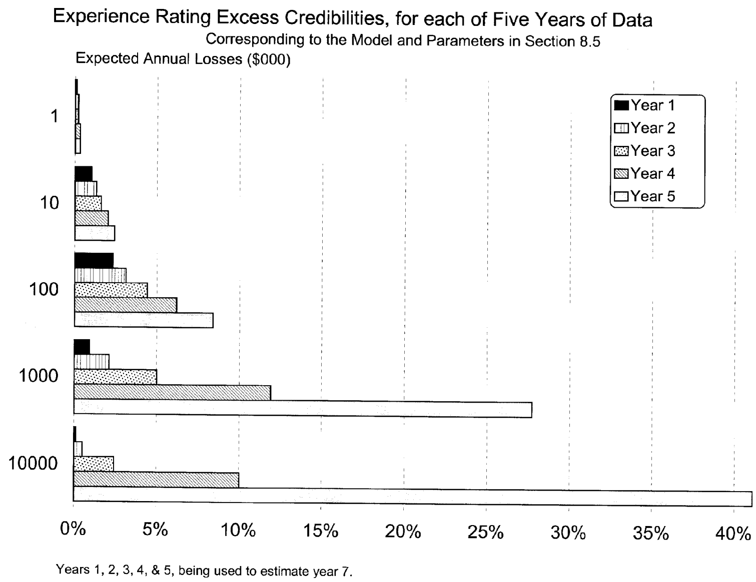


losses differ by year. In that case, years with more expected losses get more credibility than they would otherwise receive. The pattern due to varying volumes of data by year would be superimposed on that shown in Figures 33 and 34.

10.10. Taking Into Account Differences in the Maturity of the Experience Rating Data

Generally the data used for experience rating is at different reports. For workers compensation, generally three years of Unit Statistical Plan data is used for experience rating. For example, 1995 at first report, 1994 at second report, and 1993 at third report, might be used to experience rate a 1997 policy. The fact that the data are not at ultimate can affect the credibilities in two ways. First, as in Section 7.10, since the 1995 data is at an earlier report than the 1993 data, the 1995 data is a poorer estimator of 1997 ultimate losses compared to the 1993 data,

FIGURE 34



than if both 1993 and 1995 were at the same maturity.[222] Thus, the lack of maturity of the 1995 data reduces its value relative to the 1993 data and thus the credibility assigned to the 1995 data. In addition, all of the years of data are not at ultimate. Thus, they are all somewhat worse estimators than if they were available at ultimate. Thus, they all receive somewhat less credibility.[223]

As in Section 7.10, it will be assumed that the effect of loss development is to reduce the covariances between data at different reports. This refinement to the covariance structure will have the expected impact on the credibilities.

[222]Since 1995 is more recent than 1993, it is a better estimator of 1997, all other things being equal.
[223]Recall that unlike in Section 7.10, here the complement of credibility is given to the grand mean.

The first step is to estimate the correlations between the same experience rating data but at different reports. As before, various size categories will be used. Also, we can look at correlations between Primary Deviation Ratios, between Excess Deviation Ratios and between Primary and Excess Deviation Ratios.

For example, for risks with expected annual losses between $10,000 and $20,000, for the data at first report during Rating Year 91 and second report during Rating Year 92, the correlation of Primary Deviation Ratios is .937. Similar correlations can be calculated for Rating Year 92 vs. Rating Year 93, Rating Year 93 vs. Rating Year 94, and Rating Year 94 vs. Rating Year 95.[224] These four correlations have been averaged and are displayed in Table 11 as .942.

Table 11 displays similar correlations for other size categories as well as correlations for 2nd report vs. 3rd report and 1st report vs. 3rd report data.

The correlations between Primary Deviation Ratios and the correlations between Excess Deviation Ratios for different reports can be used directly, since for the same reports the correlation is one. However, for the mixed correlation between Primary Deviation Ratios and Excess Deviation Ratios, one would have to compare the correlation for different reports to the correlation for the same report, appropriately adjusted. Unfortunately this is not a practical solution,[225] therefore, the observed mixed correlations will not be used.

The primary and excess correlations in Table 11 do not display an obvious dependence on size of risk over the size categories examined.[226]

[224]These correlations are .952, .929, .949, illustrating the random fluctuation in the individual estimates for a given size category based on data from a single state.

[225]The actual correlations for a single report include a term involving the process variance. Unlike what was done in Section 7.11, we should not just totally remove this piece for comparison purposes since the different reports are not independent realizations of the same risk process, nor are the primary and excess losses independent.

[226]It is expected that the correlations will get closer to unity for very large risks, based on the analysis of classification data in Section 7.11.

TABLE 11

EXPERIENCE RATING, CORRELATIONS BETWEEN SAME DATA AT DIFFERENT REPORTS

Expected Annual Losses ($000)	1st to 2nd Corr.	1st to 2nd # Risks	2nd to 3rd Corr.	2nd to 3rd # Risks	1st to 3rd Corr.	1st to 3rd # Risks
	Primary Deviation Ratios					
3 to 5	.929	3404	.956	4052	.923	3447
5 to 10	.948	4212	.966	5247	.924	4229
10 to 20	.942	2742	.962	3458	.914	2764
20 to 30	.960	1078	.975	1359	.942	1073
30 to 50	.964	909	.972	1144	.941	900
50 to 100	.954	700	.975	900	.942	696
100 to 200	.955	330	.974	437	.940	332
200 to 500	.897	172	.958	222	.913	175
500 to 1,000	.890	46	.987	60	.877	52
1,000 to 2,000	.986	16	.987	19	.974	14
	Excess Deviation Ratios					
3 to 5	.833	3404	.893	4052	.758	3447
5 to 10	.836	4212	.910	5247	.785	4229
10 to 20	.861	2742	.897	3458	.769	2764
20 to 30	.852	1078	.930	1359	.804	1073
30 to 50	.877	909	.912	1144	.803	900
50 to 100	.858	700	.925	900	.798	696
100 to 200	.864	330	.919	437	.809	332
200 to 500	.779	172	.932	222	.720	175
500 to 1,000	.831	46	.930	60	.762	52
1,000 to 2,000	.924	16	.918	19	.857	14
	Mixed Deviation Ratios					
3 to 5	.561	3404	.574	4052	.552	3447
5 to 10	.564	4212	.584	5247	.554	4229
10 to 20	.546	2742	.569	3458	.546	2764
20 to 30	.585	1078	.610	1359	.605	1073
30 to 50	.529	909	.538	1144	.515	900
50 to 100	.585	700	.596	900	.576	696
100 to 200	.574	330	.607	437	.580	332
200 to 500	.530	172	.606	222	.540	175
500 to 1,000	.588	46	.648	60	.504	52
1,000 to 2,000	.551	16	.557	19	.504	14

Excluding the smallest and largest size intervals (with the least data), the average correlations between the different reports are:

	1–2	2–3	1–3
Primary	.939	.971	.924
Excess	.845	.919	.781

For illustrative purposes the following adjustments for loss development up to third report will be made to the primary, excess, and mixed covariances:[227]

	1–2	2–3	1–3
Primary Adjustment Factor	.94	.97	.92
Excess Adjustment Factor	.84	.92	.78
Mixed Adjustment Factor	.89	.94	.85

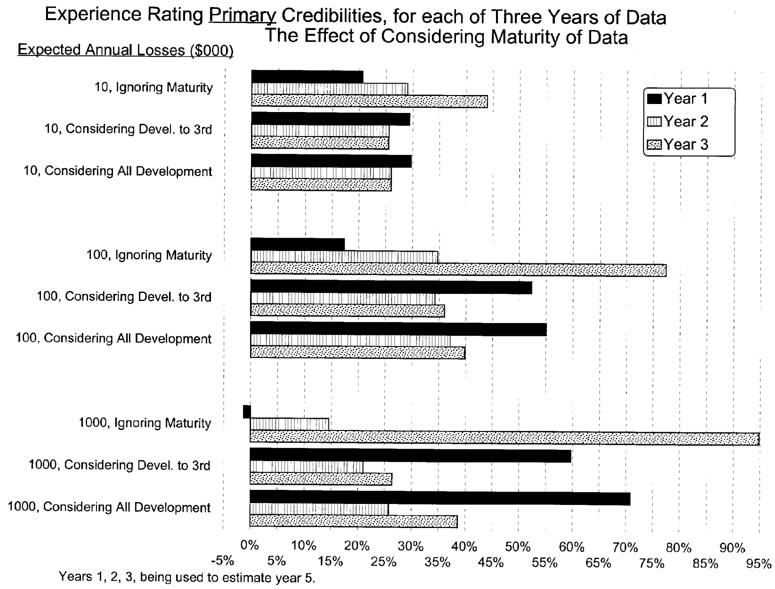
Using the parameters in Section 10.9, prior to any adjustment for differences in maturity, for $100,000 in expected annual losses we obtain the credibilities for Years 1, 2 and 3 predicting Year 5 shown both in Table 10 and in the first row below.

	Primary				Excess				Combined
	Year 1	Year 2	Year 3	Total	Year 1	Year 2	Year 3	Total	
No Adjustment for Maturity[228]	17.3%	34.7%	77.3%	129.3%	5.0%	6.6%	8.7%	20.3%	44.3%
Adjusting for Development to Third Report	25.3%	36.5%	67.9%	129.8%	5.7%	6.3%	5.8%	17.8%	42.4%
Adjusting for Development Both to Third Report and Beyond Third Report	27.4%	39.0%	69.7%	136.1%	4.1%	4.5%	3.7%	12.3%	39.5%

[227]These adjustment factors will only be applied to risks with expected annual losses between about $5,000 and $1 million. Risks outside that range would have adjustment factors that have not been estimated.

[228]See Table 10.

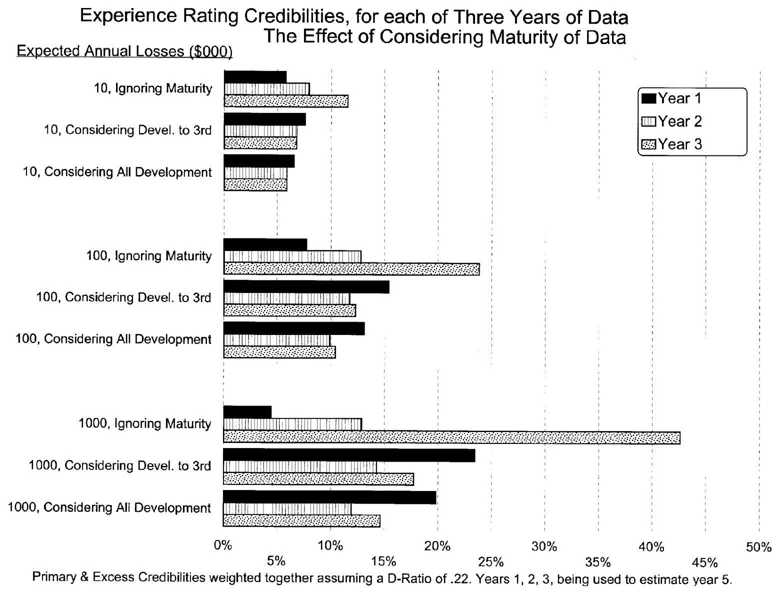
FIGURE 35



If we multiply all of the primary covariances between a year at first report and second report by an adjustment factor of .94, between second and third by .97, and between first and third by .92, with similar adjustments applied to the excess and mixed covariances, then the calculated credibilities are revised as shown in the second row above. The primary losses for Year 3 data at first report received less weight than when the maturity of the data was ignored. The primary losses for Year 1 at third report receive more weight. The overall credibility goes down somewhat.

Figure 35 displays the impact on primary credibilities for various sizes of risk, for each year separately. Figure 36 displays the impact on combined primary and excess credibilities for each year separately. Taking into account development up to third report alters the credibilities assigned to individual years; more

FIGURE 36



mature years get more weight while less mature years get less weight. Figure 37 shows the effect on the combined credibilities summed for the three years. The overall credibility is reduced by about 2% due to the consideration of development to third report.

The covariances are also affected by loss development beyond third report. The vast majority of such development affects excess losses rather than primary losses.[229] For illustrative purposes it will be assumed that development beyond third report reduces all the excess covariances between the data years and the year to be predicted (at ultimate) by .84, the adjustment factor

[229]In Massachusetts workers compensation, most claims of size less than $5,000 are closed by third report. Most claims open at third report have incurred amounts at third report that exceed $5,000 and also settle for more than $5,000.

for development from the first to the third report. The mixed covariances will be adjusted by a factor of .92, while the primary covariances are not adjusted at all.

The resulting credibilities were shown in the third row above for a risk with $100,000 in expected annual losses. Taking into account loss development beyond third report in this manner reduced the relative value of the excess losses as a predictor. Therefore, the credibilities assigned to the excess losses decreased, while those assigned to the primary losses increased.

Figures 35 to 37 compare the credibilities including the impacts of loss development to ultimate to those excluding any consideration of maturity as well as those including the impacts of loss development to third report. As expected, the inclusion of all loss development generally lowers the credibilities.[230]

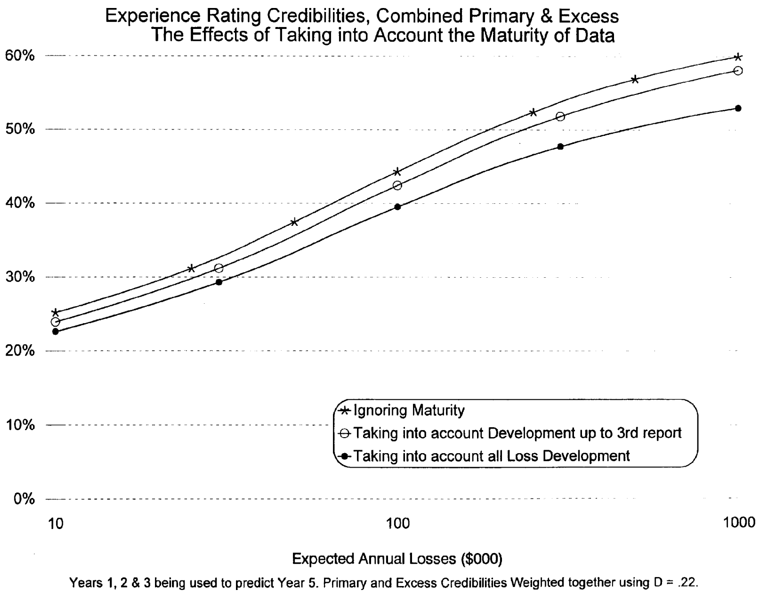
10.11. Conclusions-Experience Rating

While similar analyses of experience rating have been made in the past,[231] the present analysis incorporates shifting risk parameters, risk heterogeneity and parameter uncertainty in a comprehensive and integrated manner. While the example was for a single split experience rating plan for workers compensation, a similar analysis should be valuable for any experience rating type situation where the volume of data varies significantly between entities. For example, general liability experience rating or the assignment of towns to territories[232] would fall into this category.

On the other hand, situations such as private passenger automobile Safe Driver Insurance Plans or Bonus–Malus plans would allow a somewhat simpler analysis, since the size of the insured

[230]Recall that the adjustment factors were illustrative and not based on any specific experience rating data beyond third report.
[231]See, for example, Mahler [11] and Finger [29].
[232]See Conger [30].

FIGURE 37



is not significant.[233] The phenomenon of risk heterogeneity is not important in that case. Thus, the situation is the special case examined in Section 3.6, where parameter uncertainty and shifting risk parameters are important. In that case, we expect a covariance structure of the form:[234]

$$\mathrm{Cov}[X_i, X_j] = r^2 \rho^{|i-j|} + \delta_{ij} d^2. \qquad (10.19)$$

The size of risk E has been suppressed as not important in this case, and therefore the variance due to parameter uncertainty and that due to the expected value of the process variance can be combined into one term d^2. Equation 10.19 has the same form

[233]In medical malpractice, as discussed in Finger [29], the simpler situation is that of experience rating individual doctors, while the more general situation would involve experience rating groups of doctors.

[234]Compare to Equation 3.19.

as Equation 7.1 in Mahler [1]. Thus, the form of analysis in Mahler [1] should suffice in the case of frequency based private passenger automobile experience rating and similar situations.

11. MISCELLANEOUS

In this section the methods of Mahler [1] will be applied to the estimation of the market risk premium, the baseball models from Mahler [1] and Mahler [20] will be revisited, and the results in Boor [31] will be related to those herein.

11.1. Market Risk Premium

The market risk premium, an important economic concept used in the Capital Asset Pricing Model, is the excess return on stocks expected beyond the risk-free rate. A common estimate of the market risk premium is the difference between the return on large company stocks and the return on three-month U.S. Treasury Bills.[235] Table 12 shows this series from 1926 through 1995.

This series is very volatile. Ibbotson [32] recommends using a long-term (unweighted) average based on a belief that the expected real returns have been reasonably consistent over time. Using the currently available data from 1926 to 1995, the unweighted average is 8.76%.

While the risk parameters underlying this process are relatively stable, they are unlikely to have absolutely no shifting over time. The methods developed in Mahler [1] can be used to estimate the sensitivity of the estimated market risk premium to assumptions about the stability of the risk process.

Let X_i be the observed difference between the return on large company stocks and U.S. Treasury Bills for year i. Then one estimate of the market risk premium is to take all Y years of data

[235] See Chapter 8 of Ibbotson [32]. The market risk premium is referred to as the equity risk premium.

TABLE 12

PART 1

TOTAL RETURN

Year	Large Company Stocks	U.S. Treasury Bills	Difference
1926	11.62	3.27	8.35
1927	37.49	3.12	34.37
1928	43.61	3.56	40.05
1929	−8.42	4.75	−13.17
1930	−24.90	2.41	−27.31
1931	−43.34	1.07	−44.41
1932	−8.19	0.96	−9.15
1933	53.99	0.30	53.69
1934	−1.44	0.16	−1.60
1935	47.67	0.17	47.50
1936	33.92	0.18	33.74
1937	−35.03	0.31	−35.34
1938	31.12	−0.02	31.14
1939	−0.41	0.02	−0.43
1940	−9.78	0.00	−9.78
1941	−11.59	0.06	−11.65
1942	20.34	0.27	20.07
1943	25.90	0.35	25.55
1944	19.75	0.33	19.42
1945	36.44	0.33	36.11
1946	−8.07	0.35	−8.42
1947	5.71	0.50	5.21
1948	5.50	0.81	4.69
1949	18.79	1.10	17.69
1950	31.71	1.20	30.51
1951	24.02	1.49	22.53
1952	18.37	1.66	16.71
1953	−0.99	1.82	−2.81
1954	52.62	0.86	51.76
1955	31.56	1.57	29.99
1956	6.56	2.46	4.10
1957	−10.78	3.14	−13.92
1958	43.36	1.54	41.82
1959	11.96	2.95	9.01
1960	0.47	2.66	−2.19

TABLE 12

PART 2

TOTAL RETURN

Year	Large Company Stocks	U.S. Treasury Bills	Difference
1961	26.89	2.13	24.76
1962	−8.73	2.73	−11.46
1963	22.80	3.12	19.68
1964	16.48	3.54	12.94
1965	12.45	3.93	8.52
1966	−10.06	4.76	−14.82
1967	23.98	4.21	19.77
1968	11.06	5.21	5.85
1969	−8.50	6.58	−15.08
1970	4.01	6.52	−2.51
1971	14.31	4.39	9.92
1972	18.98	3.84	15.14
1973	−14.66	6.93	−21.59
1974	−26.47	8.00	−34.47
1975	37.20	5.80	31.40
1976	23.84	5.08	18.76
1977	−7.18	5.12	−12.30
1978	6.56	7.18	−0.62
1979	18.44	10.38	8.06
1980	32.42	11.24	21.18
1981	−4.91	14.71	−19.62
1982	21.41	10.54	10.87
1983	22.51	8.80	13.71
1984	6.27	9.85	−3.58
1985	32.16	7.72	24.44
1986	18.47	6.16	12.31
1987	5.23	5.47	−0.24
1988	16.81	6.35	10.46
1989	31.49	8.37	23.12
1990	−3.17	7.81	−10.98
1991	30.55	5.60	24.95
1992	7.67	3.51	4.16
1993	9.99	2.90	7.09
1994	1.31	3.90	−2.59
1995	37.43	5.60	31.83
Average	**12.52%**	**3.77%**	**8.76%**

Source: Ibbotson [23], Table 2–5.

and average:

$$\text{Estimate} = \sum_{i=1}^{Y} \frac{1}{Y} X_i.$$

More generally, we can weight together the X_i using weights Z_i such that $\sum Z_i = 1$:

$$\text{Estimate} = \sum_{i=1}^{Y} Z_i X_i.$$

The unweighted average is just a special case, with $Z_i = 1/Y$ for all years.

When parameters shift over time we would expect to have a covariance structure as per Equation 2.1:

$$\text{Cov}[X_i, X_j] = e^2 \delta_{ij} + r^2 \rho^{|i-j|}, \qquad \text{where} \quad \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}.$$

Equations 6.7 for the weights Z_i that minimize the expected squared error of the estimate of year $Y+1$ are:

$$\sum_{i=1}^{Y} Z_i \,\text{Cov}[X_i, X_k] = \text{Cov}[X_i, X_{Y+1}] + \lambda/2, \qquad k = 1, 2, \ldots Y,$$

where λ is the Lagrange multiplier. We can solve these Y linear equations plus the constraint equation for the desired weights Z_i. Given an assumed covariance structure, we can obtain weights and in turn use them to estimate the market risk premium.

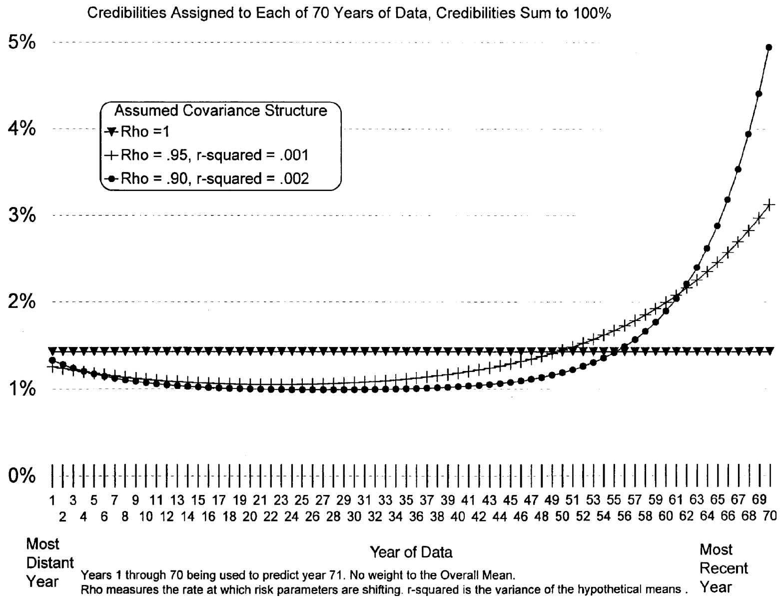
The variance of X is very large, about .0427.[236] The correlations are small.[237] Due to the large random fluctuations there is insufficient data to estimate the parameters of the covariance structure. However, we do have:

$$\text{Cov}[X, X] = \text{Var}[X] = e^2 + r^2 \approx .0427,$$

[236] The standard deviation is .207 compared to the mean of .0876.

[237] They are not statistically different from zero.

FIGURE 38



$$\begin{aligned}\mathrm{Corr}[X_i, X_{i+1}] &= \mathrm{Cov}[X_i, X_{i+1}]/\sqrt{\mathrm{Var}[X_i]\mathrm{Var}[X_{i+1}]}\\ &= r^2\rho/(e^2 + r^2)\\ &\approx r^2/(e^2 + r^2).\end{aligned}$$

Since the correlations between successive years are close to zero, r^2 is much smaller than e^2. For example, if $r^2 = .0005$ and $e^2 = .0422$ then $\mathrm{Corr}[X_1, X_2] \approx 1.2\%$. The effect of varying r^2 between .0005 and .0020 has been tested.

Ibbotson [32] believes that the parameters are *not* shifting rapidly. The parameter ρ measures the rate of shifting. For slow shifting, ρ is near 1. The effect of varying ρ between 1 and .90 has been tested. Figure 38 shows examples of the credibilities for various values of ρ and r^2.

TABLE 13

SENSITIVITY OF ESTIMATED MARKET RISK PREMIUM

$r^2(.0001)$	ρ	Estimated Market Risk Premium
5	1	8.76%[1]
5	.975	8.61
5	.95	8.68
5	.90	8.82
10	1	8.76[1]
10	.975	8.52
10	.95	8.67
10	.90	8.91
20	1	8.76[1]
20	.975	8.47
20	.95	8.75
20	.90	9.13

Based on seventy years of data from 1926–1995, (see Table 12).
Assuming total variance of .0427.
[1]Result of straight average.

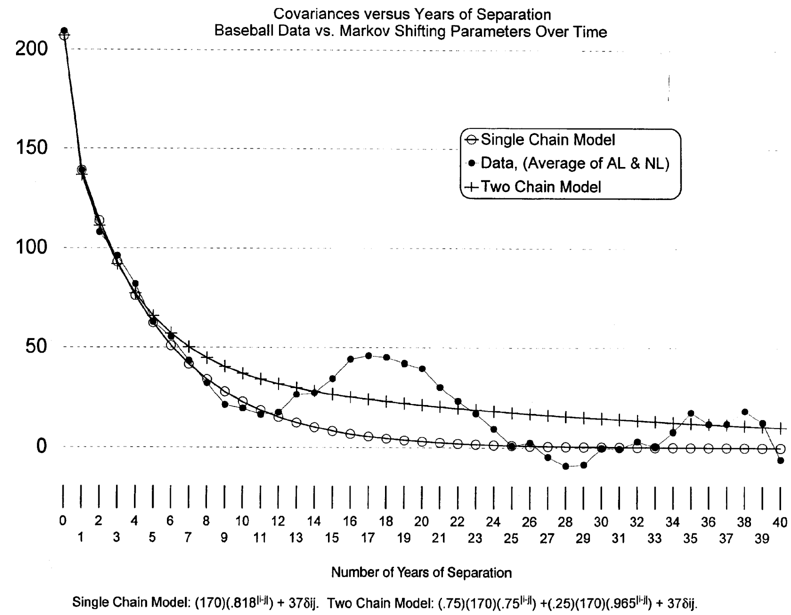
As shown in Table 13, the estimated market risk premium is relatively insensitive to the choice of the parameters of the covariance structure. Any reasonable set of inputs gives an answer in the same range. Bear in mind that just adding the 1995 data point changed the estimate of the market risk premium based on an unweighted average from 8.4% to 8.8%.

In conclusion, taking the straight unweighted average of the available data remains a reasonable method of estimating the market risk premium. While technical refinements could be included to take into account shifting risk parameters, they would not substantially improve or alter the estimate.

11.2. Baseball Example, Revisited

In Mahler [1], the data for the won-loss records of baseball teams was approximated by a model involving a single Markov chain with half-life of about $3\frac{1}{2}$ years. When expressed in terms

FIGURE 39



of games lost, the covariances between years of data X_i and X_j are:

$$\mathrm{Cov}[X_i, X_j] \approx (170)(.818^{|i-j|}) + (37)\,\delta_{ij}.$$

We could instead use a model involving two Markov chains with different half-lives. This would allow us to approximate the apparent longer term slower shifting as well as the shorter term rapid shifting. While the volume of data is not sufficient to allow us to fit a unique "two-chain" model, as seen in Figure 39 the following does a reasonable job:

$$\mathrm{Cov}[X_i, X_j] = (127.5)(.75^{|i-j|}) + (42.5)(.965^{|i-j|}) + 37\,\delta_{ij}.$$

The more swiftly shifting Markov chain has a dominant eigenvalue of .75 and a half-life[238] of about $2\frac{1}{2}$ years. The more slowly

[238] $(\ln 2) \div (\ln .75) = 2.4.$

shifting Markov chain has a dominant eigenvalue of .965 and a half-life[239] of about $19\frac{1}{2}$ years. The two Markov chains are given 75% weight and 25% weight, respectively.

This is an example of how two or more Markov chains of different half-lives could be used to attempt to model the different sources of shifting parameters over time.[240] Note that this data set does not lend itself to an examination of credibilities versus size of risk, since the seasons do not vary sufficiently in the number of games.

11.3. Boor, "Credibility Based on Accuracy"

As shown in Appendix B, linear Equations 6.7 for the credibilities with no weight to the mean are closely related to those in Boor. One difference is that Boor assumes only two estimators,[241] while Equations 6.7 assume two or more estimators.

A more fundamental difference is that Equations 6.7 assume that each of the estimators is unbiased. In Boor [31] no such assumption is made, so the results in Boor [31] apply in more general situations than Equations 6.7. Since the estimators in Boor [31] are possibly biased, the formulas for credibility involve terms such as $E[X_1 X_3]$, rather than $\mathrm{Cov}[X_1, X_3]$ such as in Equations 6.7.

12. SUMMARY AND CONCLUSIONS

In Sections 2 to 5, a general form of the covariances in the presence of shifting risk parameters, parameter uncertainty, and risk heterogeneity was developed. While a simple example using dice[242] was used to develop this covariance structure as shown

[239] $(\ln 2) \div (\ln .965) = 19.5$.

[240] Perhaps the chain with the shorter half-life relates to the baseball players while the chain with the longer half-life relates to shifts in management.

[241] The ideas in Boor [31] can be extended to more than two estimators. Boor presents the most common situation where two estimators are being combined.

[242] See Table 1.

in Equations 5.5, the model in Equations 5.10 and 5.11 is in a form appropriate for insurance applications.

Equation 3.10 in Section 3.4 shows that in the presence of parameter uncertainty there is a fundamentally different dependence of the credibility on the number of years of data and the size of risk in a single year. Section 3.7 discusses the fundamentally different dependence of the credibility on the number of years of data in the presence of shifting risk parameters versus parameter uncertainty. We can ameliorate the impact of parameter uncertainty by averaging over many years; in contrast, considering more than one year captures the effects of shifting risk parameters.

Section 5.2 includes a brief discussion of how we might estimate the parameters of the general covariance structure. Sections 7.3 to 7.6 and 7.11 illustrate how this might be done for classification data. Sections 10.4 to 10.8 and 10.10 illustrate how this might be done for experience rating data. While there are difficulties in estimating the required parameters, in general the results of applying credibility are relatively insensitive to the estimated parameters.[243]

Matrix equations are presented for calculating the (least squares) credibilities from the covariance structure. While Equations 2.4, 6.7, 8.1, 10.12 and 10.13 are similar, they each apply in a somewhat different situation. Each such set of linear equations depends on the covariance structure and can be solved for the credibilities using matrix methods.

Section 4.5 presents the credibilities in the important special case of stable (or very slowly shifting) risk parameters. Sections 4.3, 4.6, 4.7 and 5.3 explore the different behavior of credibilities expected for the smallest risks. As discussed in Section 4.8, the general covariance structure predicts the need for a minimum

[243]See Mahler [33].

ballast value in the revised Workers Compensation Experience Rating Plan.[244]

Sections 6.1 to 6.4, 7.7, 7.8, 7.12, 8.5, 8.7, 9.3 to 9.5, 10.9 and 10.10 contain illustrative calculations of credibilities. The general behaviors noted there should carry over to other similar situations.

Section 7 applies the ideas developed in this paper to an illustrative example of classification ratemaking for workers compensation. The parameters of the covariance structure were estimated in Sections 7.3 to 7.6. The behavior of the credibilities[245] when using data from one state was displayed by year[246] and by size of class in Sections 7.7 to 7.9.

Sections 7.10 to 7.12 illustrated the potential impact of the different maturity of the years of data on their credibilities. As expected, the most recent years of data at early reports get somewhat less weight than if we ignored the effects of different maturities.

Section 8 discusses how to incorporate data from outside the state. While the covariance structure has an extra layer of complication, it is still tractable. There are twice as many linear equations in twice as many unknowns, but they can still be easily solved for the credibilities. This general type of treatment should be useful whenever there is supplementary information analogous to the countrywide data.

Section 9 applies the ideas of this paper to an illustrative calculation of an overall rate indication. The effects on the weights

---

[244]The minimum ballast value was used based on practical considerations for almost a decade prior to the developments in this paper. It is pleasant to find an overall theoretical framework into which it fits.

[245]It is assumed in Section 7.9 that the complement of credibility is being given to the prior estimate of the class relativity. Section 7.7 assumed the weights assigned to the data sum to 100%, while Section 7.8 assumed the complement of credibility is given to the grand mean.

[246]The assignment of a separate weight to each year of data is an important refinement compared to the assignment of a single weight to the combined data for all years.

to be assigned to individual years of estimation errors, loss development and trend factors are discussed. Additional work would be required to adopt the general ideas presented to any particular situation. The general conclusions are far from surprising. When we have a smaller volume of data, we choose a more stable method; when we have a larger volume of data, we choose a more responsive method. Data subject to more estimation error is given less weight, all other things being equal.

Section 10 applies the ideas developed in this paper to workers compensation experience rating. This analysis should be useful for any commercial line in which the volume of data varies significantly from insured to insured. Sections 10.4 to 10.8 illustrate how we would estimate the parameters of the covariance structure in the case of a single split experience rating plan. Due to the effects of shifting parameters over time, the complicated behavior by size of risk, and the correlations of the primary and excess losses, the estimation of parameters is difficult and of necessity requires some judgment. Section 10.9 shows a sample calculation of the credibilities. The credibilities are displayed by year and by size of risk. Section 10.10 illustrates how to incorporate the impact of the different maturities of the data.

Section 11 contains miscellaneous results. The methodology is applied to an economic index, generalized to two Markov chains, and related to that in Boor [31].

In each of the various examples presented, there are three steps. First, we must specify the covariance structure.[247] Second, we must estimate the parameters of the covariance structure. Third, we must solve the appropriate set of linear equations for the credibilities.[248]

We live in a dynamic rather than stable environment. Therefore, the ideas presented in this paper on the covariance structure and resulting credibilities should have application in many

[247]See for example Table 2.
[248]See Table 3.

areas of actuarial work where risk parameters shift significantly over time. The methods presented can help answer fundamental questions about how many years of data to use in a particular situation and whether certain years of data should get significantly more weight than others. One needs to estimate how stable is the particular real world situation; how swiftly are risk parameters shifting over time?

REFERENCES

[1] Mahler, Howard C., "A Markov Chain Model for Shifting Risk Parameters," *PCAS* LXXXIV, 1997, pp. 581–659.

[2] Mayerson, Allen L., "A Bayesian View of Credibility," *PCAS* LI, 1964, pp. 85–104.

[3] Hewitt, Charles C., Jr., Discussion of Mayerson: "A Bayesian View of Credibility," *PCAS* LII, 1965, pp. 121–127.

[4] Hewitt, Charles C., Jr., "Credibility for Severity," *PCAS* LVII, 1970, pp. 148–171.

[5] Philbrick, Stephen W., "An Examination of Credibility Concepts," *PCAS* LXVIII, 1981, pp. 195–219.

[6] Herzog, Thomas N., *An Introduction to Credibility*, ACTEX, 1994.

[7] Venter, Gary G., "Credibility," Chapter 7 of *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990.

[8] Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot, *Loss Models: From Data to Decisions*, New York: John Wiley and Sons, 1998.

[9] Mahler, Howard C., "A Student's Guide to Bühlmann Credibility and Bayesian Analysis," *New England Actuarial Seminars*, 1997.

[10] Meyers, Glenn G., "An Analysis of Experience Rating," *PCAS* LXXII, 1985, pp. 278–317.

[11] Mahler, Howard C., Discussion of Meyers: "An Analysis of Experience Rating," *PCAS* LXXIV, 1987, pp. 119–189.

[12] Mahler, Howard C., Discussion of Gillam: "Parametrizing the Workers Compensation Experience Rating Plan," *PCAS* LXXXX, 1993, pp. 148–183.

[13] Gillam, William R., "Parametrizing the Workers Compensation Experience Rating Plan," *PCAS* LXXIX, 1992, pp. 21–56.

[14] Hewitt, Charles C., Jr., "Loss Ratio Distributions—A Model," *PCAS* LIV, 1967, pp. 70–88.

[15] Whitney, Albert W., "The Theory of Experience Rating," *PCAS* IV, 1918, pp. 274–292.

[16] Michelbacher, Gustav F., "The Practice of Experience Rating," *PCAS* IV, 1918, pp. 293–324.

[17] Mahler, Howard C., "Credibility: Practical Applications," Casualty Actuarial Society *Forum*, Fall 1989, pp. 187–199.

[18] Mahler, Howard C., "Introduction to Basic Credibility," Casualty Actuarial Society *Forum*, Winter 1997, pp. 67–103.

[19] Gillam, William R. and Richard H. Snader, "Fundamentals of Individual Risk Rating," 1992, available from the CAS.

[20] Mahler, Howard C., "An Example of Credibility and Shifting Risk Parameters," *PCAS* LXXVII, 1990, pp. 225–308.

[21] Finger, Robert J., "Risk Classification," Chapter 5 of *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990.

[22] Kallop, Roy H., "A Current Look at Workers Compensation Ratemaking," *PCAS* LXII, 1975, pp. 62–133.

[23] Feldblum, Sholom, "Workers Compensation Ratemaking (selected chapters)," Casualty Actuarial Society *Forum*, February 1993, pp. 241–312.

[24] Harwayne, Frank, "Use of National Experience Indicators in Workers Compensation Insurance Classification Ratemaking," *PCAS* LXIV, 1977, pp. 74–84.

[25] Meyers, Glenn G., "Empirical Bayesian Credibility for Workers Compensation Classification Ratemaking," *PCAS* LXXI, 1984, pp. 96–121.

[26] Busche, George R., "A Method to Include Multiple Years of Data in a Company's Rate Indication," Casualty Actuarial Society *Forum*, February 1993, pp. 401–420.

[27] Bouska, Amy S., "Exposure Bases Revisited," *PCAS* LXXVI, 1989, pp. 1–23.

[28] Stern, Phillip K., "Ratemaking Procedures for Automobile Liability Insurance," *PCAS* LII, 1965, pp. 139–202.

[29] Finger, Robert J., "Merit Rating for Doctor Professional Liability Insurance," *PCAS* LXXX, 1993, pp. 291–352.

[30] Conger, Robert F., "The Construction of Automobile Territories in Massachusetts," *PCAS* LXXIV, 1987, pp. 1–74.

[31] Boor, Joseph A., "Credibility Based on Accuracy," *PCAS* LXXIX, 1992, pp. 166–185.

[32] Stocks, Bonds, Bills and Inflation, 1996 Yearbook, Ibbotson Associates, Chicago.

[33] Mahler, Howard C., "An Actuarial Note on Credibility Parameters," *PCAS* LXXIII, 1986, pp. 1–26.

[34] Feller, William, *An Introduction to Probability Theory and its Applications*, New York: John Wiley and Sons, 1968.

[35] Resnick, Sidney I., *Adventures in Stochastic Processes*, Boston: Birkhauser, 1992.

[36] DeVlyder, F. Etienne, *Advanced Risk Theory—A Self-Contained Introduction*, University of Brussels, Belgium, 1996.

APPENDIX A

MARKOV CHAINS[247]

Assume that each year[248] an individual is in a "state." Each state could correspond to a different average claim frequency. Assume that there are a finite number of different states.

Assume with each new year that an individual in State i has a chance P_{ij} of going to State j. This chance is independent of which individual we have picked, what its past history was, or what year it is. The transition probability from State i to State j, P_{ij}, is only dependent on the two States, i and j.

Arrange these transition probabilities, P_{ij}, into a matrix P. This transition matrix P, together with the definition of the states, defines a (finite dimensional) Markov chain.

A vector containing the probability of finding an individual in each of the possible states is called a "distribution." If the distribution in Year 1 is β, then the expected distribution in Year 2 is βP, where βP is the matrix product of the (row vector) distribution β and the transition matrix P. The expected distribution in Year 3 is $(\beta P)P = \beta(PP) = \beta P^2$. The expected distribution in Year $1+g$ is βP^g.

Let P^T be the matrix transpose of P. Let Λ be the diagonal matrix with entries equal to the eigenvalues of P^T. Let V^T be the matrix each of whose columns are the eigenvectors of P^T. (V has as its rows the eigenvectors of P^T.) Then $(V^T)^{-1}P^TV^T = \Lambda$. Taking the transpose of both sides of this equation and noting that $\Lambda^T = \Lambda$, since Λ is symmetric: $VPV^{-1} = \Lambda$. So the matrix V can be used to diagonalize the transition matrix P:

$$V^{-1}\Lambda^2V = V^{-1}(VPV^{-1})^2V = V^{-1}VPV^{-1}VPV^{-1}V = P^2.$$

[247] See Feller [34], Resnick [35], and Appendix A in Mahler [1].

[248] Although in this paper the time interval is a year, in general, it can be anything else.

In general, $P^g = V^{-1}(VPV^{-1})^g V = V^{-1}\Lambda^g V$. So powers of P can be computed by taking powers of the diagonal matrix Λ and using the eigenvector matrix V to transform back. The elements of the diagonal matrix Λ^g are λ_i^g. $\lambda_1 = 1$ (the order of eigenvalues is arbitrary) and $|\lambda_i| < 1$ for $i > 1$ (ignoring the very unusual situation where $\lambda = 1$ is a multiple root of the characteristic equation).[249]

[249]If for any i, $|\lambda_i| > 1$, then there would be no limiting distribution, instead it would blow up. However, a finite dimensional Markov chain such that each state can be reached from every other state and such that no states are periodic has a unique stationary distribution, which is the limit as time goes to infinity. If for all i, $|\lambda_i| < 1$, then again there would be no non-zero limit, instead it would go to zero. Thus, we have all $|\lambda_i| \leq 1$ and at least one $|\lambda_i| = 1$.

APPENDIX B

MATRIX EQUATIONS FOR LEAST SQUARES CREDIBILITY WITH NO WEIGHT TO GRAND MEAN[250]

In this appendix, Equations 6.7 in the main text are derived by minimizing the squared error. The result is one constraint equation plus Y linear equations for the credibilities to be assigned to each of Y years of data. Thus the credibilities can be solved for in terms of the covariance structure. Also, the related result in Boor [31] is derived.

Let

$$C_{ij} = \mathrm{Cov}[X_i, X_j] = \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\mathrm{E}[X_j]$$
$$= \text{Covariance of Year } X_i \text{ and Year } X_j,$$
$$C_{ii} = \text{Variance of Year } X_i, \qquad \text{and}$$
$$\mu_i = \mathrm{E}[X_i] = \text{Expected value for Year } X_i.$$

Let Z_i be the credibility assigned to Year X_i. We wish to predict Year $X_{Y+\Delta}$ using Y years of data $X_1, X_2, \ldots, X_Y$. Assume $\sum_{i=1}^{Y} Z_i = 1$.

Then the estimate is:

$$F = \sum_{i=1}^{Y} Z_i X_i \qquad \text{and}$$

$$F - X_{Y+\Delta} = \left(\sum_{i=1}^{Y} Z_i X_i\right) - X_{Y+\Delta} = \sum_{i=1}^{Y} Z_i (X_i - X_{Y+\Delta})$$

since $\sum_{i=1}^{Y} Z_i = 1$.

[250] The derivation is along the same lines as those in Mahler [20] and Mahler [1].

Therefore,

$$(F - X_{Y+\Delta})^2 = \left(\sum_{i=1}^{Y} Z_i(X_i - X_{Y+\Delta})\right)\left(\sum_{j=1}^{Y} Z_j(X_j - X_{Y+\Delta})\right)$$
$$= \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j (X_i - X_{Y+\Delta})(X_j - X_{Y+\Delta}).$$

Then the expected value of the squared difference between the estimate F and $X_{Y+\Delta}$ is, as a function of the credibilities Z,

$$V(Z) = \mathrm{E}[(F - X_{Y+\Delta})^2]$$
$$= \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j \mathrm{E}[(X_i - X_{Y+\Delta})(X_j - X_{Y+\Delta})].$$

Now

$$\mathrm{E}[(X_i - X_{Y+\Delta})(X_j - X_{Y+\Delta})] = \mathrm{E}[X_i X_j] - \mathrm{E}[X_i X_{Y+\Delta}] - \mathrm{E}[X_j X_{Y+\Delta}] + \mathrm{E}[X_{Y+\Delta}^2]$$
$$\mathrm{E}[X_i X_j] = \mathrm{Cov}[X_i, X_j] + \mathrm{E}[X_i]E[X_j]$$
$$= C_{ij} + \mu_i \mu_j.$$

Thus,

$$\mathrm{E}[(X_i - X_{Y+\Delta})(X_j - X_{Y+\Delta})]$$
$$= C_{ij} - C_{i,Y+\Delta} - C_{j,Y+\Delta} + C_{Y+\Delta,Y+\Delta}$$
$$+ \mu_i\mu_j - \mu_i\mu_{Y+\Delta} - \mu_j\mu_{Y+\Delta} + \mu_{Y+\Delta}^2.$$

$$V(Z) = \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j$$
$$\times \{C_{ij} - C_{i,Y+\Delta} - C_{j,Y+\Delta} + C_{Y+\Delta,Y+\Delta} + \mu_i\mu_j$$
$$- \mu_i\mu_{Y+\Delta} - \mu_j\mu_{Y+\Delta} + \mu_{Y+\Delta}^2\}$$

$$V(Z) = \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j (C_{ij} + \mu_i\mu_j)$$
$$- \left(\sum_{i=1}^{Y}(C_{i,Y+\Delta} + \mu_i\mu_{Y+\Delta})Z_i\right)\left(\sum_{j=1}^{Y} Z_j\right)$$
$$- \left(\sum_{j=1}^{Y}(C_{j,Y+\Delta} + \mu_j\mu_{Y+\Delta})Z_j\right)\left(\sum_{i=1}^{Y} Z_i\right)$$
$$+ (C_{Y+\Delta,Y+\Delta} + \mu_{Y+\Delta}^2)\left(\sum_{i=1}^{Y} Z_i\right)\left(\sum_{j=1}^{Y} Z_j\right).$$

The last three terms all simplify since

$$\sum_{i=1}^{Y} Z_i = 1.$$

Therefore,

$$V(Z) = \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j (C_{ij} + \mu_i\mu_j) - 2\sum_{i=1}^{Y}(C_{i,Y+\Delta} + \mu_i\mu_{Y+\Delta})Z_i$$
$$+ C_{Y+\Delta,Y+\Delta} + \mu_{Y+\Delta}^2.$$

We can minimize $V(Z)$ given the constraint $\sum_{i=1}^{Y} Z_i - 1 = 0$ by using Lagrange multipliers. We set equal to zero the partial derivative with respect to Z_k of $V(Z) - \lambda(\sum_{i=1}^{Y} Z_i - 1)$:

$$2\sum_{i=1}^{Y} Z_i(C_{ik} + \mu_i\mu_k) - 2(C_{k,Y+\Delta} + \mu_k\mu_{Y+\Delta}) - \lambda = 0.$$

Therefore,

$$\sum_{i=1}^{Y} Z_i(C_{ik} + \mu_i\mu_k) = C_{k,Y+\Delta} + \mu_k\mu_{Y+\Delta} + \lambda/2 \qquad k = 1,2,\ldots,Y.$$

Also

$$\sum_{i=1}^{Y} Z_i = 1.$$

Thus we obtain $Y+1$ linear equations in $Y+1$ unknowns (the credibility assigned to each of Y years and the Lagrange multiplier λ).

If we assume each of the years X_i is an unbiased estimator of $X_{Y+\Delta}$, then $\mathrm{E}[X_i] = \mathrm{E}[X_{Y+\Delta}]$, or $\mu_i = \mu_{Y+\Delta}$. The above equations reduce to:

$$\sum_{i=1}^{Y} Z_i C_{ik} + \mu_{Y+\Delta}^2 \sum_{i=1}^{Y} Z_i = C_{k,Y+\Delta} + \mu_{Y+\Delta}^2 + \lambda/2.$$

Since $\sum_{i=1}^{Y} Z_i = 1$, this becomes Equations 6.7 in the main text:

$$\sum_{i=1}^{Y} Z_i C_{ik} = C_{k,Y+\Delta} + \lambda/2, \qquad k = 1,2,\ldots,Y$$

Equation 6.7 as well as Equations 2.4, 8.1 and 10.12 to 10.13, as shown in Table 2, are all variations on the so-called "normal equations" for credibilities. See, for example, De Vlyder [36] for an extensive discussion of the relation of the covariance structure to the credibilities.

Boor, "Credibility Based on Accuracy"

The result in Boor [31] can be obtained as a special case of the above development as follows, making no assumption concerning whether $\mathrm{E}[X_i]$ equals $\mathrm{E}[X_{Y+\Delta}]$. Assume we have two estimators X_1 and X_2 that we are using to estimate X_3. Then we get 2 linear equations plus the constraint equation:[251]

$$Z_1(C_{11} + \mu_1\mu_1) + Z_2(C_{12} + \mu_1\mu_2) = C_{13} + \mu_1\mu_3 + \lambda/2,$$

$$Z_1(C_{12} + \mu_1\mu_2) + Z_2(C_{22} + \mu_2\mu_2) = C_{23} + \mu_2\mu_3 + \lambda/2, \qquad \text{and}$$

$$Z_1 + Z_2 = 1.$$

[251] Note that $C_{21} = C_{12}$.

Subtracting the first two equations eliminates the Lagrange multiplier λ:

$$Z_1(C_{11} - C_{12} + \mu_1^2 - \mu_1\mu_2) + Z_2(C_{12} - C_{22} + \mu_1\mu_2 - \mu_2^2)$$
$$= C_{13} - C_{23} + \mu_1\mu_3 - \mu_2\mu_3.$$

Substituting $Z_2 = 1 - Z_1$ and solving for Z_1:

$$Z_1 = \frac{C_{13} - C_{12} - C_{23} + C_{22} + \mu_1\mu_3 - \mu_1\mu_2 - \mu_2\mu_3 + \mu_2^2}{C_{11} - 2C_{12} + C_{22} + \mu_1^2 - 2\mu_1\mu_2 + \mu_2^2}.$$

As in Boor [31], define the following quantities:

$$\tau_1^2 = \mathrm{E}[(X_1 - X_3)^2] = \mathrm{E}[X_1^2] - 2\mathrm{E}[X_1X_3] + \mathrm{E}[X_3^2]$$
$$= C_{11} + \mu_1^2 + C_{33} + \mu_3^2 - 2C_{13} - 2\mu_1\mu_3,$$
$$\tau_2^2 = \mathrm{E}[(X_2 - X_3)^2] = \mathrm{E}[X_2^2] - 2\mathrm{E}[X_2X_3] + \mathrm{E}[X_3^2]$$
$$= C_{22} + \mu_2^2 + C_{33} + \mu_3^2 - 2C_{23} - 2\mu_2\mu_3, \qquad \text{and}$$
$$\delta_{12}^2 = \mathrm{E}[(X_1 - X_2)^2] = \mathrm{E}[X_1^2] - 2\mathrm{E}[X_1X_2] + \mathrm{E}[X_2^2]$$
$$= C_{11} + \mu_1^2 + C_{22} + \mu_2^2 - 2C_{12} - 2\mu_1\mu_2.$$

Then we can verify that the numerator of Z_1 above is:

$$\tfrac{1}{2}(\tau_2^2 - \tau_1^2 + \delta_{12}^2).$$

The denominator of Z_1 above is δ_{12}^2. Therefore:

$$Z_1 = \frac{\tau_2^2 - \tau_1^2 + \delta_{12}^2}{2\delta_{12}^2}, \qquad \text{and}$$
$$Z_2 = 1 - Z_1 = \frac{\tau_1^2 - \tau_2^2 + \delta_{12}^2}{2\delta_{12}^2},$$

which is the result obtained in Boor [31].[252] We note that the key distinction is that Boor makes no assumption as to whether

[252]See page 169 of *PCAS* 1992.

the estimators are unbiased.[253] Thus his formulas involve terms like $\mathrm{E}[X_1 X_3]$ rather than the covariances such as in Equation 6.7 in the main text.

[253] Also, Boor only displays the result for combining two estimators. The development in this appendix works for any number of estimators; we get $Y + 1$ linear equations in $Y + 1$ unknowns.

APPENDIX C

CLASSIFICATION DATA

Unit Statistical Plan data for Massachusetts workers compensation insurance was examined.[254] A total of 13 composite policy years[255] of data were available at various reports.[256] For each year the latest available report was used: 80/81 to 88/89 @ 5th; 89/90 @ 4th; 90/91 @ 3rd; 91/92 @ 2nd; 92/93 @ 1st report.

For each classification, payrolls and losses were available. The losses were paid losses plus case reserves. Losses were broken down by injury kind and between medical and indemnity, but these splits were not used in the current analysis.

For example, for Class 2003, Bakeries, the experience in composite policy year 92/93 at first report was $68,928,691 in payroll and $1,477,837 in losses. This corresponds to a pure premium (per $100 of payroll) of 2.1440.

Class 2003 is one of 270 classes in the Manufacturing industry group. For composite policy year 92/93 at first report there was $3,896,021,286 in payroll and $67,944,193 in losses for the Manufacturing industry group. This corresponds to a pure premium of 1.7439. Thus the relative pure premium for Class 2003 for 92/93 @ 1st is 2.1440/1.7439 = 1.2294.

Performing similar calculations, we obtain the following relative pure premiums for two example classes:[257]

---

[254]Experience on all insureds in the state was included, except for large deductible policies. (Large deductibles were only available during the most recent three composite policy years.)

[255]A composite policy year runs from July to June. For example, composite policy year 92/93 includes experience from policies with policy effective dates from July 1, 1992 to June 30, 1993.

[256]First report is evaluated 18 months after policy inception. Subsequent reports are made at 12 month intervals, up to and including fifth report.

[257]The similar calculations were done for each class in the Manufacturing industry group. Similar but totally separate calculations were then done for the Goods & Services industry group.

	Relative Pure Premium	
Composite Policy Year	Class 2003	Class 3145[258]
92/93 @ 1st	1.2294	.7931
91/92 @ 2nd	1.2279	.3741
90/91 @ 3rd	1.5828	.5016
89/90 @ 4th	1.3713	.8561
88/89 @ 5th	1.2380	1.4134
87/88 @ 5th	1.7127	.5199
86/87 @ 5th	1.3507	1.0739
85/86 @ 5th	2.0721	1.1651
84/85 @ 5th	1.4784	.7649
83/84 @ 5th	1.6312	.9236
82/83 @ 5th	1.3711	1.6704
81/82 @ 5th	1.0365	1.5151
80/81 @ 5th	1.7196	.9415

The relative pure premiums show considerable fluctuation because these are medium-sized classes and the losses used are unlimited.[259]

In order to divide the classes into size categories, expected losses were calculated. Expected losses for a class for a composite policy year were obtained by multiplying the reported payrolls by three factors. The first factor was the ratio of the State Average Weekly Wage for Composite Policy Year 1992/1993[260] to that for the particular composite policy year. The second factor was the observed pure premium for the industry group for the particular composite policy year and report. The third and final factor was the ratio of the current rate[261] for the class to the average rate for the industry group.

[258]Class 3145 is Screw Manufacturing.

[259]For classification ratemaking individual claims are usually limited. Currently in Massachusetts workers compensation each claim is capped at $200,000 for classification ratemaking. (These excess losses are loaded back via factors which vary by hazard group and injury kind.)

[260]The most recent year of data used.

[261]Rates effective 5/1/96 were used.

For example, for Class 2003 for Composite Policy Year 91/92 @ 2nd report the payroll was \$88,136,418. The State Average Weekly Wage during 92/93 was \$580.73, while during 91/92 it was \$560.28. Thus the first adjustment factor is $\$580.73 \div \$560.28 = 1.036$. The observed pure premium for the Manufacturing industry group for 91/92 @ 2nd report is 2.361. The current manual rate for Class 2003 is \$5.77, while the average manual rate for Manufacturing is \$4.008. Thus the third adjustment factor is $\$5.77/\$4.008 = 1.43962$.

Thus the expected losses for Class 2003 for 91/92 @ 2nd are

$$(\$88{,}135{,}418 \div 100)(1.036)(2.361)(1.43962) = \$3{,}103{,}552.$$

A similar calculation of expected losses was made for each of the 13 years. Then the average expected annual losses were calculated for each class.[262] It is these average expected annual losses that were used to divide the classes into size categories for purposes of analysis.

[262]The average only included years in which the class had reported payrolls. Some classes were discontinued or newly erected during these 13 years.

APPENDIX D

Split Experience Rating Plan Matrix Equations for Least Squares Credibility

In this appendix, Equations 10.12 and 10.13 in the main text will be derived for the optimal primary and excess credibilities for a split experience rating plan.

Assume we have two well-defined portions of the total losses, which can be thought of as primary and excess.[263] Assume we have Y years of data being used to predict year $Y + \Delta$.[264] We wish to determine credibilities to apply to the primary and excess data for each year.

Define the following quantities:

$$E_{Pi} = \text{Expected Primary Losses for Year } i,$$

$$E_{Xi} = \text{Expected Excess Losses for Year } i,$$

$$E_i = E_{Pi} + E_{Xi} = \text{Expected Losses for Year } i,$$

$$A_{Pi} = \text{Actual Primary Losses for Year } i,$$

$$A_{Xi} = \text{Actual Excess Losses for Year } i,$$

$$D_i = E_{Pi}/E_i = \text{D-ratio for Year } i,$$

$$P_i = A_{Pi}/E_i,$$

$$X_i = A_{Xi}/E_i,$$

$$\pi_i = P_i - D_i = (A_{Pi} - E_{Pi})/E_i$$

$$= \text{Primary "Deviation Ratio" for Year } i, \text{ and}$$

$$\xi_i = X_i - (1 - D_i) = (A_{Xi} - E_{Xi})/E_i$$

$$= \text{Excess "Deviation Ratio" for Year } i.$$

---

[263]For workers compensation insurance, currently the first \$5,000 of each claim is primary, while the remainder up to a claim limit is excess. The claim limit for experience rating varies by state.

[264]Typically $Y = 3$ and $\Delta = 2$ currently. Years 1, 2 and 3 are predicting Year 5.

The quantity of interest in experience rating is how the insured's future losses will compare to the expected losses for the average insured in that class or mixture of classes. That estimate, the experience modification, can be written as:[265]

$$F = 1 + \sum_{i=1}^{Y} \pi_i Z_{Pi} + \sum_{i=1}^{Y} \xi_i Z_{Xi}.$$

This differs somewhat from the usual notation in, for example, Gillam and Snader [19] or Mahler [4], since each individual year of data will be assigned a separate credibility of each type, rather than adding the years of data together and having one overall Z_P and Z_X.

If we use the data from years 1 to Y in order to predict $P_{Y+\Delta} + X_{Y+\Delta}$, the ratio of actual to expected losses for year $Y+\Delta$, then the error is:

$$\begin{aligned} F - (P_{Y+\Delta} + X_{Y+\Delta}) &= F - (\pi_{Y+\Delta} + \xi_{Y+\Delta} + 1) \\ &= \sum_{i=1}^{Y} (\pi_i - \pi_{Y+\Delta}) Z_{Pi} + \sum_{i=0}^{Y} (\xi_i - \xi_{Y+\Delta}) Z_{Xi} \\ &\quad - \pi_{Y+\Delta}\left(1 - \sum_{i=1}^{Y} Z_{Pi}\right) - \xi_{Y+\Delta}\left(1 - \sum_{i=1}^{Y} Z_{Xi}\right). \end{aligned}$$

Define $\pi_0 = \xi_0 = 0$ and $Z_{P0} = 1 - \sum_{i=1}^{Y} Z_{Pi}$ and $Z_{X0} = 1 - \sum_{i=1}^{Y} Z_{Xi}$. Then the error is:

$$\sum_{i=0}^{Y} (\pi_i - \pi_{Y+\Delta}) Z_{Pi} + \sum_{i=0}^{Y} (\xi_i - \xi_{Y+\Delta}) Z_{Xi}.$$

[265]Equation 10.1 in the main text.

The squared error is:

$$\sum_{i=0}^{Y}\sum_{j=0}^{Y}(\pi_i - \pi_{Y+\Delta})(\pi_j - \pi_{Y+\Delta})Z_{Pi}Z_{Pj}$$
$$+ 2\sum_{i=0}^{Y}\sum_{j=0}^{Y}(\pi_i - \pi_{Y+\Delta})(\xi_j - \xi_{Y+\Delta})Z_{Pi}Z_{Xj}$$
$$+ \sum_{i=0}^{Y}\sum_{j=0}^{Y}(\xi_i - \xi_{Y+\Delta})(\xi_j - \xi_{Y+\Delta})Z_{Xi}Z_{Xj}.$$

Thus, the expected value of the squared error is:

$$\sum_{i=0}^{Y}\sum_{j=0}^{Y}Z_{Pi}Z_{Pj}\mathrm{E}[(\pi_i - \pi_{Y+\Delta})(\pi_j - \pi_{Y+\Delta})]$$
$$+ 2\sum_{i=0}^{Y}\sum_{j=0}^{Y}Z_{Pi}Z_{Xj}\mathrm{E}[(\pi_i - \pi_{Y+\Delta})(\xi_j - \xi_{Y+\Delta})]$$
$$+ \sum_{i=0}^{Y}\sum_{j=0}^{Y}Z_{Xi}Z_{Xj}\mathrm{E}[(\xi_i - \xi_{Y+\Delta})(\xi_j - \xi_{Y+\Delta})].$$

Define the following quantities in terms of covariances:

$$S_{ij} = \mathrm{Cov}[\pi_i, \pi_j] = S_{ji},$$
$$T_{ij} = \mathrm{Cov}[\xi_i, \xi_j] = T_{ji}, \qquad \text{and}$$
$$U_{ij} = \mathrm{Cov}[\pi_i, \xi_j].$$

Note that since $\mathrm{E}[\pi_i] = 0 = \mathrm{E}[\xi_j]$, $S_{ij} = \mathrm{E}[\pi_i\pi_j]$, $T_{ij} = \mathrm{E}[\xi_i\xi_j]$, and $U_{ij} = \mathrm{E}[\pi_i\xi_j]$. We can rewrite the expression for the expected value of the squared error in terms of S_{ij}, T_{ij} and U_{ij}.

For example, the "cross term" can be written as:

$$\mathrm{E}[(\pi_i - \pi_{Y+\Delta})(\xi_j - \xi_{Y+\Delta})]$$
$$= \mathrm{E}[\pi_i\xi_j] - \mathrm{E}[\pi_i\xi_{Y+\Delta}] - \mathrm{E}[\pi_{Y+\Delta}\xi_j] + \mathrm{E}[\pi_{Y+\Delta}\xi_{Y+\Delta}]$$
$$= U_{ij} - U_{i,Y+\Delta} - U_{Y+\Delta,j} + U_{Y+\Delta,Y+\Delta}.$$

Similarly,

$$E[(\pi_i - \pi_{Y+\Delta})(\pi_j - \pi_{Y+\Delta})] = S_{ij} - S_{i,Y+\Delta} - S_{Y+\Delta,j} + S_{Y+\Delta,Y+\Delta}$$

$$E[(\xi_i - \xi_{Y+\Delta})(\xi_j - \xi_{Y+\Delta})] = T_{ij} - T_{i,Y+\Delta} - T_{Y+\Delta,j} + T_{Y+\Delta,Y+\Delta}.$$

Note that:

$$\begin{aligned} E[(\pi_0 - \pi_{Y+\Delta})(\xi_j - \xi_{Y+\Delta})] &= -E[\pi_{Y+\Delta}(\xi_j - \xi_{Y+\Delta})] \\ &= -U_{Y+\Delta,j} + U_{Y+\Delta,Y+\Delta}, \end{aligned}$$

but $U_{0j} = \mathrm{Cov}[\pi_0, \xi_j] = \mathrm{Cov}[0, \xi_j] = 0$. Thus,

$$\begin{aligned} &E[(\pi_0 - \pi_{Y+\Delta})(\xi_j - \xi_{Y+\Delta})] \\ &\quad = U_{0,j} - U_{0,Y+\Delta} - U_{Y+\Delta,j} + U_{Y+\Delta,Y+\Delta}. \end{aligned}$$

Thus, the same notation works for index values of zero. Therefore, the expected value of the squared error is the following quadratic function of the primary and excess credibilities:

$$\begin{aligned} &\sum_{i=0}^{Y}\sum_{j=0}^{Y} Z_{Pi} Z_{Pj}(S_{ij} - S_{i,Y+\Delta} - S_{Y+\Delta,j} + S_{Y+\Delta,Y+\Delta}) \\ &\quad + 2\sum_{i=0}^{Y}\sum_{j=0}^{Y} Z_{Pi} Z_{Xj}(U_{ij} - U_{i,Y+\Delta} - U_{Y+\Delta,j} + U_{Y+\Delta,Y+\Delta}) \\ &\quad + \sum_{i=0}^{Y}\sum_{j=0}^{Y} Z_{Xi} Z_{Xj}(T_{ij} - T_{i,Y+\Delta} - T_{Y+\Delta,j} + T_{Y+\Delta,Y+\Delta}). \end{aligned}$$

Some simplification is possible using the facts that:

$$\sum_{i=0}^{Y} Z_{Pi} = 1 = \sum_{i=0}^{Y} Z_{Xi}.$$

Thus, the expected value of the squared error is:

$$\sum_{i=0}^{Y}\sum_{j=0}^{Y} Z_{Pi}Z_{Pj}S_{ij} - 2\sum_{i=0}^{Y} Z_{Pi}S_{i,Y+\Delta} + S_{Y+\Delta,Y+\Delta}$$
$$+ 2\sum_{i=0}^{Y}\sum_{j=0}^{Y} Z_{Pi}Z_{Xj}U_{ij} - 2\sum_{i=0}^{Y} Z_{Pi}U_{i,Y+\Delta}$$
$$- 2\sum_{i=0}^{Y} Z_{Xi}U_{Y+\Delta,i} + 2U_{Y+\Delta,Y+\Delta}$$
$$+ \sum_{i=0}^{Y}\sum_{j=0}^{Y} Z_{Xi}Z_{Xj}T_{ij} - 2\sum_{i=0}^{Y} Z_{Xi}T_{i,Y+\Delta} + T_{Y+\Delta,Y+\Delta}.$$

In order to minimize the expected value of the squared error we set each of the $2Y$ partial derivatives with respect to one of the credibilities equal to zero. We get $2Y$ linear equations in $2Y$ unknowns.

Taking the partial derivative of the expected squared error with respect to Z_{Pk} and setting it equal to zero yields:[266]

$$\sum_{i=0}^{Y} Z_{Pi}(S_{ik} - S_{k,Y+\Delta}) + \sum_{i=0}^{Y} Z_{Xi}(U_{ki} - U_{k,Y+\Delta}) = 0.$$

Taking the partial derivative of the expected squared error with respect to Z_{Xk} and setting it equal to zero yields:

$$\sum_{i=0}^{Y} Z_{Pi}(U_{ik} - U_{Y+\Delta,k}) + \sum_{i=0}^{Y} Z_{Xi}(T_{ki} - T_{k,Y+\Delta}) = 0.$$

Again some simplification is possible using the facts that:

$$\sum_{i=0}^{Y} Z_{Pi} = 1 = \sum_{i=0}^{Y} Z_{Xi}.$$

[266]Dividing each term by a factor of two.

The linear equations become:

$$\sum_{i=0}^{Y}(Z_{Pi}S_{ik} + Z_{Xi}U_{ki}) = S_{k,Y+\Delta} + U_{k,Y+\Delta}, \qquad \text{and}$$
$$\sum_{i=0}^{Y}(Z_{Pi}U_{ik} + Z_{Xi}T_{ki}) = U_{Y+\Delta,k} + T_{k,Y+\Delta}.$$

Since $S_{0k} = U_{0k} = T_{0k} = 0 = S_{k0} = U_{k0} = T_{k0}$, the summation on the left hand side can start at $i = 1$ rather than 0. The resulting $2Y$ linear equations are Equations 10.12 and 10.13 in the main text.

An important special case occurs in the absence of shifting parameters over time. Further assume we either use one year of data or combine several years of data together.

Then Equations 10.12 and 10.13 become:

$$Z_P S_{11} + Z_X U_{11} = S_{1,1+\Delta} + U_{1,1+\Delta}, \qquad \text{and}$$
$$Z_P U_{11} + Z_X T_{11} = U_{1+\Delta,1} + T_{1,1+\Delta}.$$

The solutions are:

$$Z_P = \frac{(S_{1,1+\Delta} + U_{1,1+\Delta})T_{11} - (U_{1+\Delta,1} + T_{1,1+\Delta})U_{11}}{S_{11}T_{11} - U_{11}^2}, \qquad \text{and}$$
$$Z_X = \frac{(T_{1,1+\Delta} + U_{1+\Delta,1})S_{11} - (U_{1,1+\Delta} + S_{1,1+\Delta})U_{11}}{S_{11}T_{11} - U_{11}^2}.$$

This matches the result in Mahler [11],[267] with:

S_{11} = Total variance of the primary losses,

T_{11} = Total variance of the excess losses,

[267] See Equations 5.3 and 5.4 in Mahler [11].

$S_{1,1+\Delta}$ = Variance of the hypothetical means of the primary losses,

$T_{1,1+\Delta}$ = Variance of the hypothetical means of the excess losses,

U_{11} = Total covariance of the primary and excess losses, and

$U_{1,1+\Delta} = U_{1+\Delta,1}$ = Covariance of the hypothetical means of the primary and excess losses.

In the notation in Mahler [11]:

a = Total variance of the primary losses,

b = Total variance of the excess losses,

c = Variance of the hypothetical means of the primary losses,

d = Variance of the hypothetical means of the excess losses,

r = Total covariance of the primary and excess losses, and

s = Covariance of hypothetical means of the primary and excess losses.

And in the absence of shifting risk parameters over time the optimum Z_P and Z_X are:

$$Z_P = \frac{(c+s)b-(d+s)r}{ab-r^2}, \qquad \text{and}$$

$$Z_X = \frac{(d+s)a-(c+s)r}{ab-r^2}.$$

APPENDIX E

USE OF COUNTRYWIDE CLASSIFICATION DATA, MATRIX EQUATIONS FOR LEAST SQUARES CREDIBILITY

This appendix will discuss Equations 8.1 in the main text for the optimal least squares credibility when combining classification data from more than one state.

Assume we have a series of observations of X_i, for example, the class relativities in Massachusetts for each of several years, $i = 1$ to Y. Assume we also have a related series of observations of A_i, for example, the relativities for the same class calculated from data from some other states.[268] Finally, assume we wish to predict $X_{Y+\Delta}$, the class relativity in Massachusetts in year $Y + \Delta$ in the example in the main text, using a weighted average of the X_i and A_i.

More specifically the predictor $F = \sum_{i=1}^{Y} Z_i X_i + \sum_{i=1}^{Y} W_i A_i$ and $\sum_{i=1}^{Y} Z_i + \sum_{i=1}^{Y} W_i = 1$.

Note that here the weights sum to 100%; there is no weight being given to the grand mean. Note that since we are predicting $X_{Y+\Delta}$, X and A will not enter into the matrix equations in a symmetric fashion.[269]

Let the covariances be:

$$S_{ij} = \mathrm{Cov}[X_i, X_j] = S_{ji},$$
$$T_{ij} = \mathrm{Cov}[A_i, A_j] = T_{ji}, \qquad \text{and}$$
$$U_{ij} = \mathrm{Cov}[X_i, A_j].$$

[268]It is not necessary that A_i be available for exactly the same years as X_i, $i = 1$ to Y, but the presentation is easier to follow if we assume that this is the case. (Years with no available data can be treated by giving them a weight of zero.)

[269]In contrast, the primary and excess losses did enter into the equations in Appendix D in a mathematically symmetric manner.

As in Appendices B and D, assume each X_i or A_i is an unbiased estimator of the quantity of interest, $X_{Y+\Delta}$. Then the expected value of the squared error is:

$$\begin{aligned}
V(Z,W) &= \mathrm{E}[(F - X_{Y+\Delta})^2] \\
&= \mathrm{E}\left[\left\{\sum_{i=1}^{Y} Z_i(X_i - X_{Y+\Delta}) + \sum_{i=1}^{Y} W_i(A_i - X_{Y+\Delta})\right\}^2\right] \\
&= \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j (S_{ij} - S_{i,Y+\Delta} - S_{Y+\Delta,j} + S_{Y+\Delta,Y+\Delta}) \\
&\quad + 2\sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i W_j (U_{ij} - S_{i,Y+\Delta} - U_{Y+\Delta,j} + S_{Y+\Delta,Y+\Delta}) \\
&\quad + \sum_{i=1}^{Y}\sum_{j=1}^{Y} W_i W_j (T_{ij} - U_{Y+\Delta,i} - U_{Y+\Delta,j} + S_{Y+\Delta,Y+\Delta}).
\end{aligned}$$

Some simplification of the expression for $V(Z,W)$ is possible. Since $S_{i,Y+\Delta} = S_{Y+\Delta,i}$

$$\sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j S_{i,Y+\Delta} = \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j S_{Y+\Delta,j}.$$

Therefore,

$$\begin{aligned}
&\sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j S_{i,Y+\Delta} + \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j S_{Y+\Delta,j} + 2\sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i W_j S_{i,Y+\Delta} \\
&\qquad = 2\sum_{i=1}^{Y}\sum_{j=1}^{Y} (Z_i Z_j + Z_i W_j) S_{i,Y+\Delta} \\
&\qquad = 2\left(\sum_{i=1}^{Y} Z_i S_{i,Y+\Delta}\right)\left(\sum_{j=1}^{Y} (Z_j + W_j)\right) = 2\sum_{i=1}^{Y} Z_i S_{i,Y+\Delta},
\end{aligned}$$

since $\sum_{j=1}^{Y}(Z_j + W_j) = 1$.

Similarly,

$$2\sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i W_j U_{Y+\Delta,j} + \sum_{i=1}^{Y}\sum_{j=1}^{Y} W_i W_j U_{Y+\Delta,i} + \sum_{i=1}^{Y}\sum_{j=1}^{Y} W_i W_j U_{Y+\Delta,j}$$

$$= 2\left(\sum_{i=1}^{Y}(Z_i + W_i)\right)\left(\sum_{j=1}^{Y} W_j U_{Y+\Delta,j}\right) = 2\sum_{i=1}^{Y} W_i U_{Y+\Delta,i}.$$

Also we have:

$$\sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j S_{Y+\Delta,Y+\Delta} + 2\sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i W_j S_{Y+\Delta,Y+\Delta}$$

$$+ \sum_{i=1}^{Y}\sum_{j=1}^{Y} W_i W_j S_{Y+\Delta,Y+\Delta}$$

$$= S_{Y+\Delta,Y+\Delta}\left(\sum_{i=1}^{Y}(Z_i + W_i)\right)\left(\sum_{j=1}^{Y}(Z_j + W_j)\right) = S_{Y+\Delta,Y+\Delta}.$$

Thus, $V(Z,W)$ simplifies to:

$$V(Z,W) = \sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i Z_j S_{ij} + 2\sum_{i=1}^{Y}\sum_{j=1}^{Y} Z_i W_j U_{ij}$$

$$+ \sum_{i=1}^{Y}\sum_{j=1}^{Y} W_i W_j T_{ij} - 2\sum_{i=1}^{Y} Z_i S_{i,Y+\Delta}$$

$$- 2\sum_{i=1}^{Y} W_i U_{Y+\Delta,i} + S_{Y+\Delta,Y+\Delta}.$$

The constraint equation $\sum_{i=1}^{Y} Z_i + \sum_{i=1}^{Y} W_i - 1 = 0$ is incorporated via the Lagrange multiplier λ. We minimize $V(Z,W)-$

$\lambda(\sum_{i=1}^{Y} Z_i + \sum_{i=1}^{Y} W_i - 1)$, by taking each of $2Y$ partial derivatives with respect to the Zs and Ws and setting them equal to zero. Setting the partial derivative with respect to Z_k equal to zero:

$$2\sum_{j=1}^{Y} Z_j S_{kj} + 2\sum_{j=1}^{Y} W_j U_{kj} - 2S_{k,Y+\Delta} - \lambda = 0.$$

This equation can be rewritten as:

$$\sum_{j=1}^{Y} Z_j S_{kj} + \sum_{j=1}^{Y} W_j U_{kj} = \lambda/2 + S_{k,Y+\Delta}.$$

Similarly, by setting the partial derivative with respect to W_k equal to zero we get:

$$\sum_{j=1}^{Y} Z_j U_{jk} + \sum_{j=1}^{Y} W_j T_{kj} = \lambda/2 + U_{Y+\Delta,k}.$$

The above $2Y$ linear equations (one from the partial derivative of each Z_k and each W_k) plus the constraint equation are Equations 8.1 in the main text. Note the similarities to the Equations 2.4, 6.7, 10.12 and 10.13 in the main text. Each set of equations applies to a somewhat different situation. However, each such set of linear equations depends on the covariance structure and can be solved for the credibilities using matrix methods.

In a situation where there were different years of Massachusetts and countrywide data, Equations 8.1 would be somewhat different in form. For example, assume Massachusetts data for years 1, 2, 3 and 4 plus countrywide data for years 2 and 3 were being used to predict Massachusetts relativities for year 8. The Equations 8.1 would become seven linear equations in seven unknowns:

$$Z_1S_{11} + Z_2S_{12} + Z_3S_{13} + Z_4S_{14} + W_2U_{12} + W_3U_{13} = \lambda/2 + S_{18},$$
$$Z_1S_{21} + Z_2S_{22} + Z_3S_{23} + Z_4S_{24} + W_2U_{22} + W_3U_{23} = \lambda/2 + S_{28},$$
$$Z_1S_{31} + Z_2S_{32} + Z_3S_{33} + Z_4S_{34} + W_2U_{32} + W_3U_{33} = \lambda/2 + S_{38},$$
$$Z_1S_{41} + Z_2S_{42} + Z_3S_{43} + Z_4S_{44} + W_2U_{42} + W_3U_{43} = \lambda/2 + S_{48},$$
$$Z_1U_{12} + Z_2U_{22} + Z_3U_{32} + Z_4U_{42} + W_2T_{22} + W_3T_{32} = \lambda/2 + U_{82},$$
$$Z_1U_{13} + Z_2U_{23} + Z_3U_{33} + Z_4U_{43} + W_2T_{23} + W_3T_{33} = \lambda/2 + U_{83},$$
$$\text{and} \qquad Z_1 + Z_2 + Z_3 + Z_4 + W_2 + W_3 = 1.$$

In this case, the equations each have four terms involving the four weights to each of the years of Massachusetts data, but only two terms involving the two weights to each of the years of countrywide data. There are $4 + 2 + 1 = 7$ unknowns, including the Lagrange multiplier.

APPENDIX F

ESTIMATING PARAMETERS OF BETWEEN STATE COVARIANCES

In order to calculate credibilities when using data from one or more outside states to calculate classification relativities, it is necessary to estimate the variance-covariance structure. This appendix will present an example of how to estimate the parameters of the between state covariances.

Assume as in Section 8 we are estimating Massachusetts class relativities and will use New York data in addition to that from Massachusetts.

Then there are three types of variance-covariance matrices. The first type involves covariances between data from Massachusetts. The second type involves covariances between data from New York. The third type of covariance is that involving data from Massachusetts versus New York. It is expected that for a given volume of data, the correlation of relativities between states is less than the correlation of relativities within states. This is what is observed.

For three years of data combined for each state, adjusted as it would be for classification ratemaking,[270] the relative pure premiums were calculated for the classes in the Manufacturing industry group. Then the correlations between the New York and Massachusetts relative pure premiums were calculated for various sizes of class.[271] The results are in the table below.

How does this compare to the correlation within a single state that we would expect if we could run the risk process twice

[270]Claims sizes are limited. Losses would be adjusted for law changes and loss development. See Kallop [12] and Feldblum [13]. In this case, these adjustments were performed by whichever rating bureau has responsibility for that state.

[271]The size of class was taken as the square root of the product of the payroll (summed over three years) for each state. In other words, the geometric average of the payroll for the two states was used for each class.

Manufacturing Industry Group				
Three Years of Payroll ($ million)			Correlations NY vs. MA	
Minimum	Maximum	Number of Classes	Observed	Capped[272]
1	3	12	.047	.041
3	10	35	.348	.382
10	30	44	.106	.127
30	100	45	.596	.592
100	300	29	.623	.623
300	1,000	5	.818	.823

and create two parallel universes?[273] In that case the portion of the covariance related to the expected value of the process variance, the term involving e^2 or K, would vanish.[274]

Ignoring shifting risk parameters,[275] the covariances between single years of data are given by Equation 4.13:

$$\mathrm{Cov}[X_i, X_j] = r^2\{1 + I/E + \delta_{ij}(K/E + J)\}, \qquad E \geq \Omega.$$

Then for the sum of three years of data, X_1, X_2 and X_3:

$$\mathrm{Cov}[X_1 + X_2 + X_3, X_1 + X_2 + X_3] = r^2\{9 + 9I/E + 3K/E + 3J\}, \qquad E \geq \Omega.$$

If we have data from two parallel versions of Massachusetts, we set $K = 0$, and J' to represent the possibly different J param-

[272]The relativity for each class was capped between 5 and 1/5, in order to limit the impact of any one class on the computed correlation.

[273]In the dice examples in Sections 3 and 4, one just rerolls the dice keeping everything else constant.

[274]For example, if X_1 and X_2 are each independent results of rolling 10 six-sided dice, then their covariance is zero, while the process variance of X_1 or X_2 is positive. The usual Bühlmann covariance structure is $\mathrm{Cov}[X_i, X_j] = \tau^2 + \delta_{ij}\eta^2$. For $i \neq j$, $\mathrm{Cov}[X_i, X_j] = \tau^2$; the term involving the expected value of the process variance, η^2, vanishes.

[275]Setting $\rho = 1$ and $\gamma = 1$.

eter when taking the covariance between two parallel versions of Massachusetts.

Then the correlation for three years of data from each of two parallel versions of Massachusetts is:

$$\frac{9+9I/E+3J'}{9+9I/E+3K/E+3J} = \frac{(3+J')E+3I}{(3+J)E+3I+K}, \qquad E \geq \Omega.$$

If we include shifting risk parameters, we get a slightly different expression for the correlations. The covariances are given by Equation 5.8:

$$\mathrm{Cov}[X_i, X_j] = r^2\{\rho^{|i-j|} + \gamma^{|i-j|} I/E + \delta_{ij}(K/E + J)\}, \qquad E \geq \Omega.$$

For the sum of three years of data, the terms involving ρ sum to $3+4\rho+2\rho^2$. Similarly, the terms involving γ sum to $(3+4\gamma+2\gamma^2)I/E$. Using the values from Section 7.7, for $\rho = .98$, $3+4\rho+2\rho^2 = 8.84$, while for $\gamma = .85$, $3+4\gamma+2\gamma^2 = 7.85$.

Thus, the correlations equal:

$$\frac{8.84+7.85I/E+3J'}{8.84+7.85I/E+3K/E+3J} = \frac{(2.95+J')E+2.62I}{(2.95+J)E+2.62I+K}, \qquad E \geq \Omega.$$

Similarly for $E \leq \Omega$ starting with Equation 5.9, we obtain a correlation of:

$$\frac{8.84+7.85I/\Omega+3J'}{8.84+7.85I/\Omega+3K/E+3J} = \frac{(2.95+J')E+2.62IE/\Omega}{(2.95+J)E+2.62IE/\Omega+K}, \qquad E \leq \Omega.$$

Depending on whether or not the effects that are responsible for parameter uncertainty are reproduced,[276] the term involving u^2 or J may or may not vanish. In the case of MA vs. NY, the two states would be affected by some of the same macroe-

[276] In the dice examples in Section 3, do we maintain the same coin flip or is the coin flipped again?

conomic and other forces that produce parameter uncertainty. Thus, for covariances between MA and NY, we would expect that a portion of the term involving J would remain. Since the two parallel versions of Massachusetts will be used to compare to the interstate situation, we will also assume that in that case a portion of the term involving J will remain. For illustrative purposes take $J' = .05$ for the covariances between two parallel versions of Massachusetts, one-half of the assumed value for the intrastate covariances.[277]

Using the estimated parameters from Section 7.7, with $J' = .05$, we get the correlations shown below as "Expected Intrastate."[278] These have been compared to the observed correlations between New York and Massachusetts.

Payroll 3 Years	Correlations			Ratio to Expected Intrastate	
($ million)	Expected Intrastate	NY vs. MA	NY vs. MA, Capped[279]	NY vs. MA	NY vs. MA, Capped
2	.215	.047	.041	.22	.19
6.5	.458	.348	.382	.76	.83
20	.600	.106	.127	.18	.21
65	.782	.596	.592	.76	.76
200	.900	.623	.623	.69	.69
650	.955	.818	.823	.86	.86

Similar comparisons were done for other large states and for the Goods and Services industry group. Here are the ratios of the observed interstate correlations to the expected intrastate corre-

[277]The credibilities are relatively insensitive to this choice.

[278]In order to translate payrolls into expected losses the payrolls were multiplied by the observed pure premium for the Manufacturing industry group of about $2.50 per $100 of payroll. Thus $1 million of payroll for 3 years corresponds to $8,333 of annual expected losses.

[279]The relativity for each class was capped between 5 and 1/5, in order to limit the impact of any one class on the computed correlation.

lations:

RATIO OF INTERSTATE TO EXPECTED INTRASTATE CORRELATIONS[280] BY THREE YEARS OF PAYROLL ($ MILLION)

	Manufacturing				Goods and Services	
State	3 to 10	10 to 30	30 to 100	100 to 300	10 to 100	100 to 1,000
Connecticut	−.07	.58	.39	.74	.90	.86
Florida	.08	.68	.81	.66	.23	.85
Georgia	.11	.34	.23	.40	.61	.84
Illinois	.50	.10	.60	.81	.58	.92
Michigan	—	.09	.59	.56	1.09	.82
Missouri	.12	.31	.45	.67	.94	.88
New Jersey	.72	.03	.53	.63	—	—
New York	.76	.18	.76	.69	1.09	.96
Oregon	.78	.88	.73	—	.95	.87
Wisconsin	.18	−.04	.69	.66	.86	.89
Average	**.35**	**.32**	**.58**	**.65**	**.81**	**.87**

Generally, the between state correlations are lower than the within state correlations, for a given volume of data. In this case, the between state correlations are perhaps 55% of the within state correlations for Manufacturing[281] and perhaps 85% for Goods and Services. A ratio of 70% would result if the r^2 factor multiplying the interstate covariances were 70% of the r^2 factor multiplying the intrastate covariances.[282]

This ratio of 70% will be used for illustrative purposes in Section 8. As is shown in Section 8.7, the credibilities are relatively insensitive to this choice for values within this general range.

[280]Only results for categories with 15 or more classes are displayed.

[281]The correlations for the smaller size categories for Manufacturing were affected by two classes whose observed Massachusetts relativities were vastly different than those observed in most other states.

[282]Recall that in the intrastate situation, the r^2 factor did not affect the calculated credibilities. In the interstate situation with two (or more) values for r^2, the relative size of the r^2 values will affect the credibilities.

For the interstate covariances the K parameter will be zero.[283] Also, the interstate covariances will use $J = .05$, one-half of the assumed value for the intrastate covariances.

[283]As discussed above, the portion of the covariance related to the expected value of the process variance would vanish when taking covariances between data from different states.

A GRAPHICAL ILLUSTRATION OF EXPERIENCE RATING CREDIBILITIES

HOWARD C. MAHLER

Abstract

This paper combines a simple experience rating example with a set of graphs in order to illustrate key credibility concepts as they relate to experience rating. As part of this graphical approach, credibility will be related to linear regression.

ACKNOWLEDGEMENT

I wish to thank Silvia Fitcher and Dotty Culleton for typing this paper.

1. INTRODUCTION

Philbrick [1] uses his excellent target shooting example to graphically illustrate some key concepts of credibility. Hewitt [2] uses a die/spinner example to illustrate important ideas of credibility. In this same spirit, this paper will combine a simple experience rating example with a set of graphs to illustrate key credibility ideas as they relate to experience rating. As part of the graphical approach, credibility ideas will be related to linear regression.

Prior and subsequent experience will be simulated for various sets of insureds for different sets of simple assumptions. This simulated data for the various examples will be used to illustrate that the slope of the regression line between prior and subsequent experience is one estimate of the Bühlmann credibility. Finally, these same examples will be used to illustrate that the expected squared error between the actual and predicted subsequent experience is minimized when the weight given to the observed experience is equal to the Bühlmann credibility.

2. EXPERIENCE RATING

The goal of experience rating is to use an individual insured's experience to help predict future loss costs.[1] If the individual risk's experience were observed to be worse than average, we would predict that his future experience would also likely be somewhat worse than average. Therefore, we would be likely to charge this insured somewhat more than average.

Credibility, as used in experience rating, quantifies how much worse or better an insured's future experience is expected to be based on a particular deviation from average observed in the past. In the simplest case:[2]

$$\begin{aligned}\text{New Estimate} &= (\text{Credibility})(\text{Observation})\\ &\quad + (1 - \text{Credibility})(\text{Overall Mean})\\ &= (\text{Overall Mean}) + (\text{Credibility})\\ &\quad \times (\text{Observation} - \text{Overall Mean}).\end{aligned}$$

In Appendix A, Bühlmann credibilities, Z, are calculated for various situations, using the formulas:

$$Z = N/(N + K)$$

$$K = EPV/VHM$$

[1]See, for example, Meyers [3], Mahler [4], Finger [5], Gillam and Snader [6], and Tiller [7].

[2]The actual applications have a number of complications beyond the scope of this paper.

where

Z = Bühlmann credibility,

N = Number of years of data (from a single insured),

K = Bühlmann credibility parameter,

EPV = Expected value of process variance for a single unit of the risk process (i.e., for one insured for one year), and

VHM = Variance of the hypothetical means for a single unit of the risk process (i.e., for one insured for one year).

3. SIMPLE EXAMPLE

The following very simplified assumptions will be used in various combinations to illustrate credibility ideas. See Table 1 for a summary of the different situations illustrated.

TABLE 1

SUMMARY OF DIFFERENT SITUATIONS

Situation Number*	Quantity of Interest	Types of Insureds	Figure Number(s)	Credibility Estimated	Credibility Theoretical
1	Frequency	50 Good, 50 Bad	1, 2	40%	33%
1	Frequency	3 Years of Prior Data 50 Good, 50 Bad	3	58%	50%
2	Frequency	50 Excellent, 50 Ugly	4, 5	78%	81.8%
3	Frequency	50 Excellent, 50 Good, 50 Bad, 50 Ugly	6, 7	72%	71.4%
4	Unlimited Losses	125 Excellent, 125 Ugly	8, 9	51.5%	52.9%
5	Limited Losses	125 Excellent 125 Ugly	10	71.4%	70.1%

*See Appendix A for more details.

Claim frequency for individual insureds is assumed to be Poisson.[3] Claim severity is assumed to be given by a Pareto distribution[4] with shape parameter 3 and scale parameter 20,000. Frequency and severity are assumed to be independent. There are four possible types of insureds with different Poisson parameters:

Type	Average Annual Claim Frequency
Excellent	5
Good	10
Bad	15
Ugly	20

In Appendix A, the usual Bühlmann credibility techniques have been applied to various situations involving these four types of insureds in order to quantify the credibility to be assigned to the past experience of an insured. A set of graphs has been constructed to illustrate these same situations.

These graphs illustrate the connection between Bühlmann credibility and least squares linear regression. For the simple situations dealt with here, the slope of the least squares regression line between the past and subsequent observations of insureds is an estimate of the Bühlmann credibility. Appendix B provides a mathematical demonstration of this relationship. Not only is this relationship approximate,[5] but the slope from the regression will vary in particular examples due to random fluctuations. Thus, the estimated credibility will not exactly equal the theoretical Bühlmann credibility.

4. GRAPHS OF FREQUENCY EXAMPLES

Assume we have 100 insureds all in the same risk classification, territory, etc. The first graph, Figure 1, shows simu-

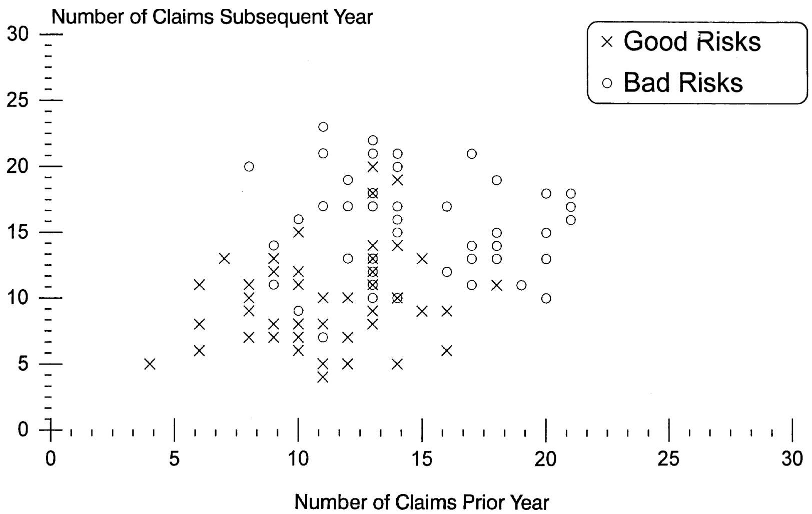
[3] The Poisson parameter for each insured stays the same over time.

[4] $F(x) = 1 - (20{,}000/(20{,}000 + x))^3$.

[5] As derived in Appendix B, one determines the expected value of a numerator and denominator separately and then assumes that $\mathrm{E}[A/B] \approx \mathrm{E}[A]/\mathrm{E}[B]$ in the sit-

FIGURE 1

SIMULATED CLAIMS EXPERIENCE



Situation 1: 50 Good Risks (Poisson 10) and 50 Bad Risks (Poisson 15)

lated claim counts for these 100 insureds divided into two equal groups. In this graph, the "Good Risks" are labeled with crosses and the "Bad Risks" with circles. In both the real world[6] and many of the subsequent graphs, the risks come without such labels attached. (If they did come with such labels, we would not need to use credibility.)

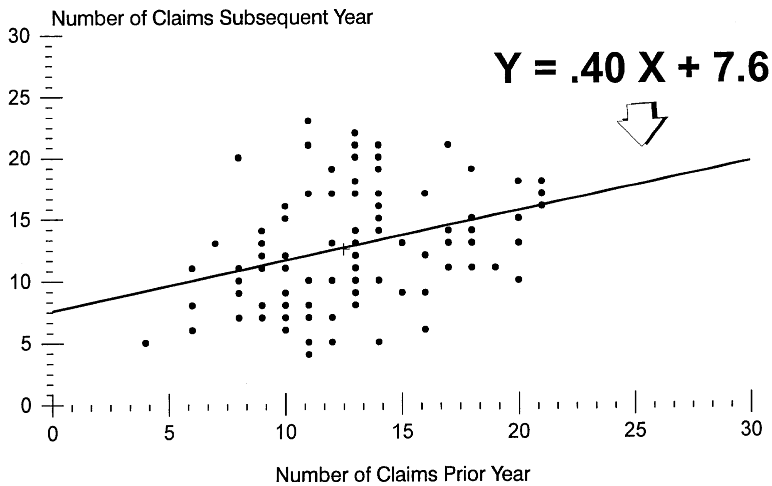
The 50 Bad Risks each have an expected claim frequency of 15 while the 50 Good Risks each have an expected claim frequency of 10. For each of the 100 insureds, a single prior year of simulated claim counts has been plotted against a single subsequent year of simulated claim counts. For example, one of

uations to which the result is being applied. In general, $E[A]/E[B]$ is not an unbiased estimator of A/B.

[6]In the real world, there is no way to precisely determine any individual's expected future frequency.

FIGURE 2

SIMULATED CLAIMS EXPERIENCE
GOOD AND BAD RISKS



Situation 1: 50 Good Risks (Poisson 10) and 50 Bad Risks (Poisson 15)

the Good Risks had 4 claims in the prior year and 5 claims in the subsequent year. This is indicated by a cross at the point (4,5). There is considerable overlap between the groups. Nevertheless, the Good Risks are more likely to be in the lower left while the Bad Risks are more likely to be in the upper right of the graph.

The next graph, Figure 2, shows the same 100 insureds without labels. In Figure 2 a least squares regression line has been fit to the points. One could use this fitted line to predict a future year's experience based on an observation. Since the line slopes upwards, a worse than average former year would lead one to predict a worse than average subsequent year.

So if one observed 20 claims in a year for an insured, one might predict about 15 claims for that insured next year, compared to the overall average of 12.5. The formula for this least squares line is approximately:

$$Y = .40X + 7.6.$$

The equation can be restated in the form of the "basic credibility formula:"

$$\text{Estimate} = Z(\text{Observation}) + (1 - Z)(\text{Overall Mean}),$$

with the credibility $Z = 40\%$ and

$$(1 - Z)(\text{Overall Mean}) = (60\%)(12.5) = 7.5 \approx 7.6.$$

With only 100 insureds, this result is subject to considerable random fluctuation. Thus, the estimated credibility of 40% is not equal to the theoretical Bühlmann credibility. The simulation with many more insureds would give a credibility of 1/3, the theoretical value as shown in Appendix A, Situation 1.

The credibility is just the slope of the straight line. It is the weight given to the observation.

Note the way that the fitted line passes through the point (12.5, 12.5), denoted by a plus. Average experience in the prior year yields an estimate of average experience in the subsequent year. This follows from rewriting the basic credibility formula as Estimate = Overall Mean + Z(Observation – Overall Mean).

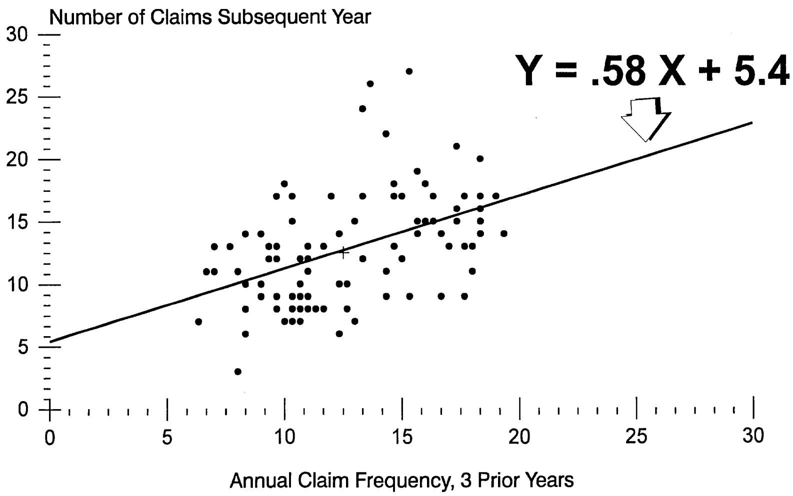
Note that the line $Y = X$, with a slope of unity, would correspond to 100% credibility, while the line $Y = 12.5$ with a slope of zero, would correspond to zero credibility. In general, the slope and the Bühlmann credibility will be between zero and one.

These general features displayed in Figure 2 will carry over to subsequent figures. The least squares line will slope upwards and pass through the point denoting average experience in the prior and subsequent period. The slope will be (approximately) equal to the credibility.

The next graph, Figure 3, is similar to Figure 2 but shows *three* years of prior experience rather than one. Note that the X-axis is

FIGURE 3

SIMULATED CLAIMS EXPERIENCE, 3 PRIOR YEARS
GOOD AND BAD RISKS



Situation 1: 50 Good Risks (Poisson 10) and 50 Bad Risks (Poisson 15)

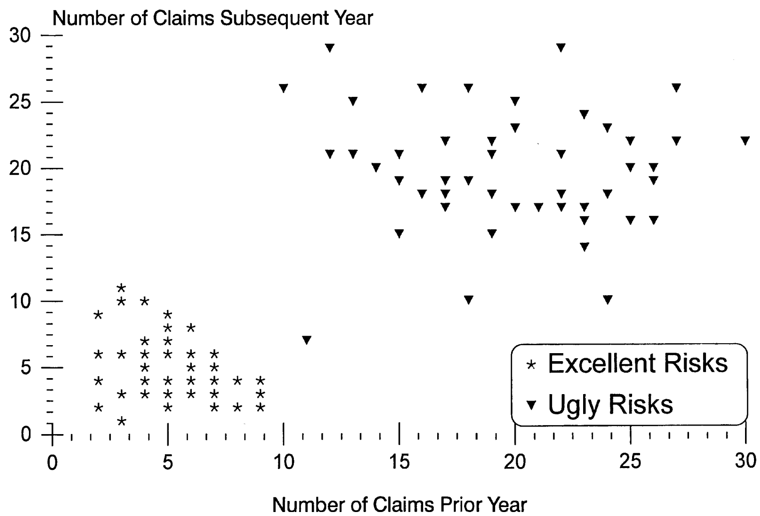
now the *annual* claim frequency observed over three years. We expect three years of data to contain more useful information and thus be given more weight than would one year. In fact, a fitted straight line has a larger slope of about 60% (actually 58%) corresponding to a credibility of 60%. One way to increase the credibility of data is to increase the volume of data.

In the case of Figures 2 and 3, the credibility is equal to $N/(N+K)$ where N is the number of years of data and $K = 2$. (See Appendix A, Situation 1.) This formula is used quite often, with the "Bühlmann credibility constant" K dependent on the statistical properties of the particular situation. Note that for Figure 2 with one year of prior data, $Z = 1/(1+2) = 33\%$, while in Figure 3 with three years of prior data, $Z = 3/(3+2) = 60\%$.

The next graph, Figure 4, shows 100 risks divided this time into Excellent Risks and Ugly Risks. The Excellent Risks are

FIGURE 4

SIMULATED CLAIMS EXPERIENCE



Situation 2: 50 Excellent Risks (Poisson 5) and 50 Ugly Risks (Poisson 20)

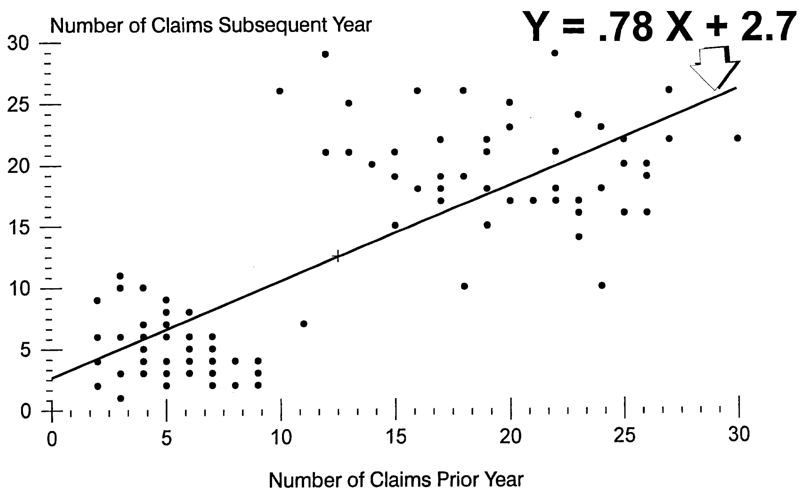
shown by asterisks and the Ugly Risks by wedges. The mean frequencies are 5 and 20 rather than 10 and 15 as in the previous exhibits. Therefore, the two groups are spread apart much more. Since there is more dispersion between risks,[7] each risk's data will be given more credibility than in the first graph.

This can be seen in the next graph, Figure 5, where a straight line has been fit to these points. The line has a much larger slope than the line in Figure 2, corresponding to higher credibility of about 82%. (The estimated credibility is 78%. Again the results of an experiment with only 100 risks differs from the theoreti-

[7]The experience is more likely to distinguish between excellent and ugly risks, than between good and bad risks. This is quantified via the variance of hypothetical means (*VHM*). As shown in Appendix A, the *VHM* in Situation 2 of 56.25 is much larger than that in Situation 1 of 6.25.

FIGURE 5

SIMULATED CLAIMS EXPERIENCE
EXCELLENT AND UGLY RISKS



Situation 2: 50 Excellent Risks (Poisson 5) and 50 Ugly Risks (Poisson 20)

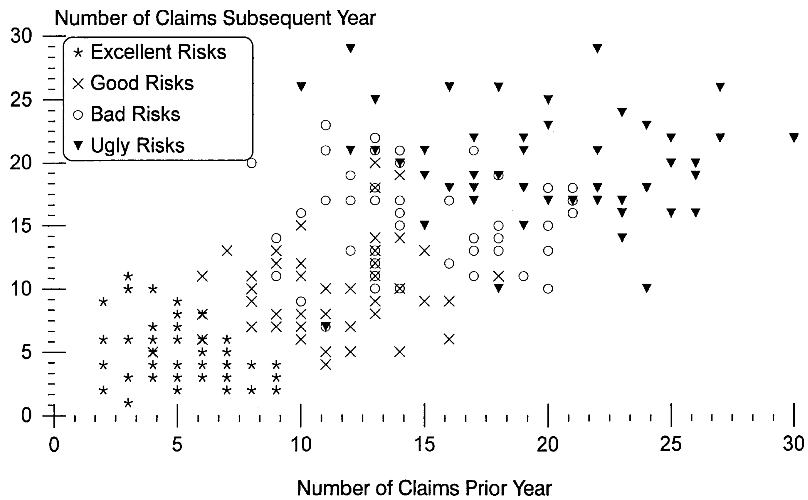
cal result of 81.8% in Appendix A, Situation 2, due to random fluctuation.) So due to the larger variation in hypothetical means (holding everything else equal) in Figure 5 versus Figure 2, the Bühlmann credibility increased from 33% to 82%. The value of the individual risk's information *increased relative* to the information contained in the overall mean. Conversely, the *relative* value of the information contained in the overall mean *decreased.*

The next graph, Figure 6, combines the four different types of insureds. This starts to approach the real world situations where risks' expected claim frequencies are assumed to be along a continuous spectrum, rather than being of discrete types.[8] We can see

[8]One could approach a continuous situation similar to the Gamma–Poisson frequency process. The Gamma–Poisson frequency process is explained, for example, in Hossack, Pollard and Zehnwirth [8], Herzog [9], or Mahler [10].

FIGURE 6

SIMULATED CLAIMS EXPERIENCE



Situation 3: 50 Excellent Risks (Poisson 5), 50 Good Risks (Poisson 10), 50 Bad Risks (Poisson 15), and 50 Ugly Risks (Poisson 20)

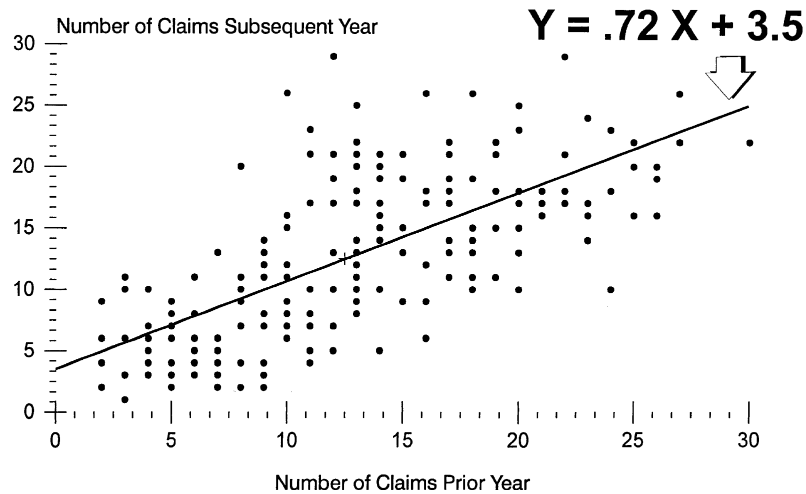
plenty of overlap between the four types of insureds, although since we labeled the insureds, we can discern the grouping of different types.

The next graph, Figure 7, shows a line fit to data from all four types. There the slope of 72% is between the slopes of 40% and 78% that we got when dealing with just two groups in Figures 2 and 5. All else being equal,[9] this makes sense since the variation of the hypothetical means is in between the variations of hypothetical means for those two situations. The theoretical credibility of 71% determined in Appendix A, Situation 3, is between the theoretical credibilities of 33% and 82% for Situations 1 and 2 which deal with only two groups.

[9] Specifically, the expected value of the process variance is the same in all three situations.

FIGURE 7

SIMULATED CLAIMS EXPERIENCE
EXCELLENT, GOOD, BAD, AND UGLY RISKS



Situation 3: 50 Excellent Risks (Poisson 5), 50 Good Risks (Poisson 10), 50 Bad Risks (Poisson 15), and 50 Ugly Risks (Poisson 20)

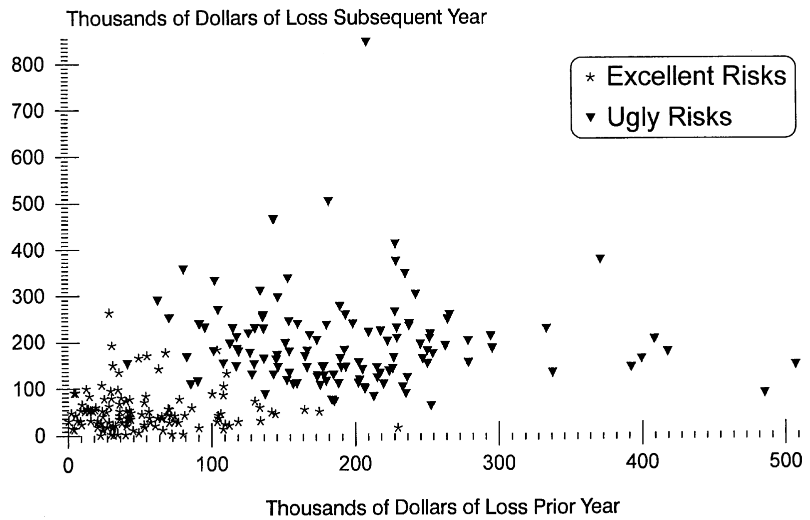
5. GRAPHS OF PURE PREMIUM EXAMPLES

The following graphs will all involve 125 Excellent and 125 Ugly Risks and not only deal with claim frequency, but with claim severity as well. By looking at dollars of loss rather than numbers of claims, as can be seen on the next graph, Figure 8, we introduce more random fluctuation.[10] Therefore, the relative value of the observation is less compared to the overall average; the credibility goes down. One way to *decrease* the credibility of data is to *increase* the variability of the data.

[10]In the absence of the labels, it would be somewhat easier to distinguish the Excellent and Ugly risks in Figure 4 dealing with frequency only than in Figure 8 dealing with dollars of loss.

FIGURE 8

SIMULATED LOSS EXPERIENCE



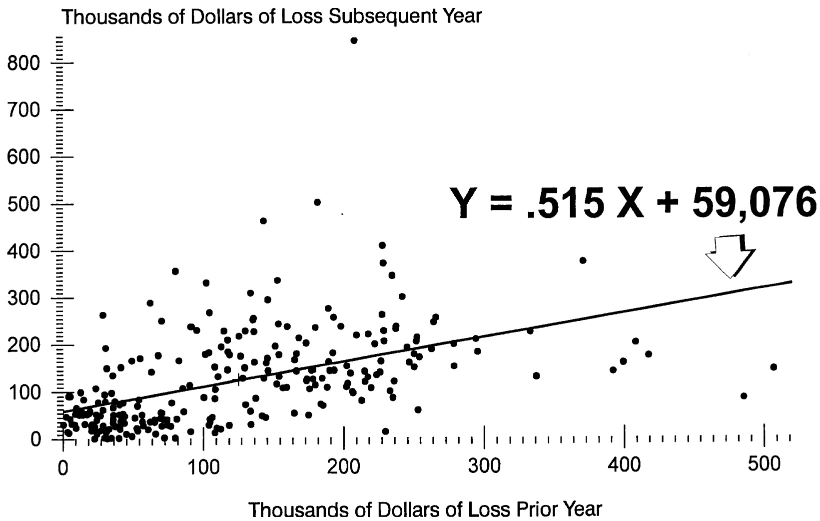
Situation 4: 125 Excellent Risks (Poisson 5) and 125 Ugly Risks (Poisson 20), Pareto Severity (3, $20,000)

As can be seen on the next graph, Figure 9, the slope of the fitted line is 51.5%. As shown in Appendix A, Situation 4, the theoretical credibility is 53% compared to 82% for the corresponding claim frequency Situation 2. The greater random fluctuation, which is quantified by the larger "process variance," has decreased the credibility assigned to the observations.

In practical applications, one often limits the size of claims entering into experience rating, since one way to decrease the variability of the data is to cap losses. The final graph in this series, Figure 10, shows the results of limiting each claim to $25,000. (This capping can be just for the purposes of experience rating or could involve an actual policy limit.) The slope of the fitted line between prior limited losses and subsequent limited losses is 71.4%. As determined in Appendix A, Situation 5,

FIGURE 9

SIMULATED LOSS EXPERIENCE



Situation 4: 125 Excellent Risks (Poisson 5) and 125 Ugly Risks (Poisson 20), Pareto Severity (3, $20,000)

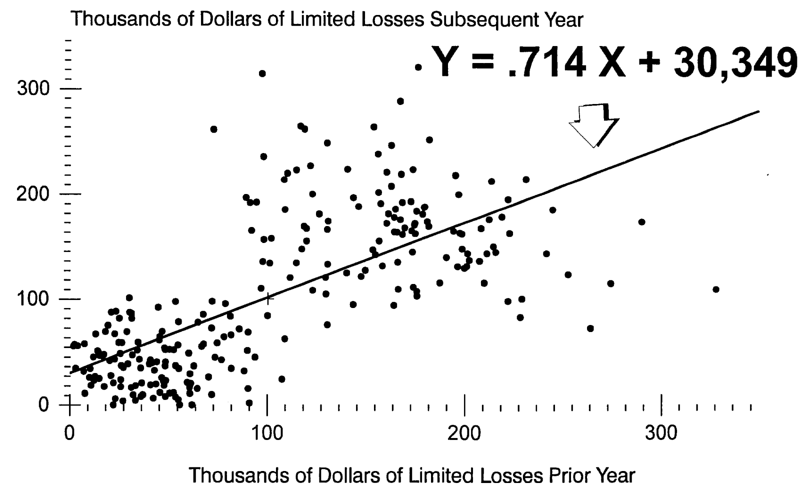
the theoretical credibility of 70% when using limited losses compares to 53% for total losses in Situation 4. Capping the losses has reduced the random fluctuations (i.e., has reduced the process variance) thereby increasing the credibility assigned to the experience. (Basic limit losses are less volatile than total limits losses.) For more on how to analyze experience rating plans, see for example Meyers [3] or Mahler [4].

6. EFFECT OF RANDOM FLUCTUATIONS ON ESTIMATED CREDIBILITIES

As mentioned above, the credibility estimated from regressing actual data sets will be affected by random fluctuations and, therefore, will not equal the theoretical Bühlmann credibility cal-

FIGURE 10

SIMULATED LOSS EXPERIENCE
Each Claim Limited to $25,000



Situation 5: 125 Excellent Risks (Poisson 5) and 125 Ugly Risks (Poisson 20), Pareto Severity (3, $20,000)

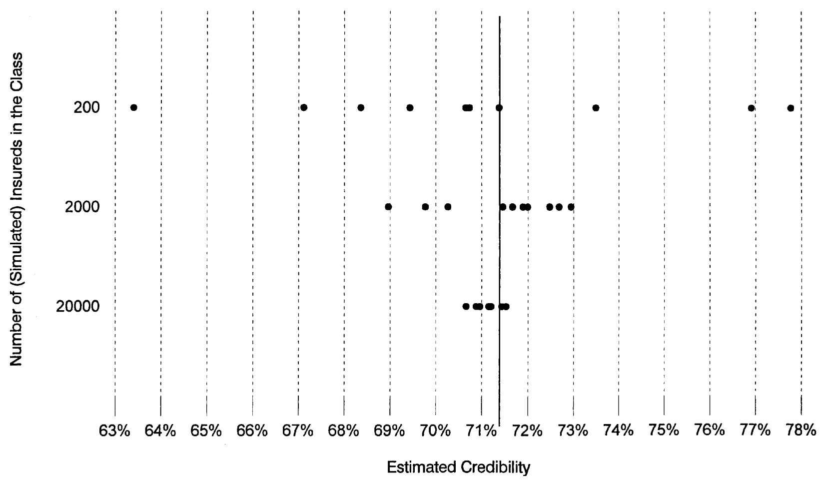
culated in Appendix A. The fewer insureds in the data set and/or the larger the process variance,[11] the larger is the impact from random fluctuations.

Figures 11 and 12 show the results of simulation experiments. Figure 11 deals with the frequency example with all four types of insureds as illustrated in Figures 6 and 7. The situation in Figure 7 with 200 insureds was simulated 10 separate times. This resulted in 10 different estimates of the credibility, ranging from 63.4% to 77.8%, as shown in Figure 11. Similar simulation ex-

[11] If the expected claim frequencies had been smaller, then the process variance would have been larger. For example, if the expected claim frequency for excellent risks were .05 rather than 5, one would need many more insureds to get as good an estimate of the credibility.

FIGURE 11

SIMULATION EXPERIMENTS
CREDIBILITIES ESTIMATED BY REGRESSION
CLAIM COUNTS
EXCELLENT, GOOD, BAD, AND UGLY RISKS



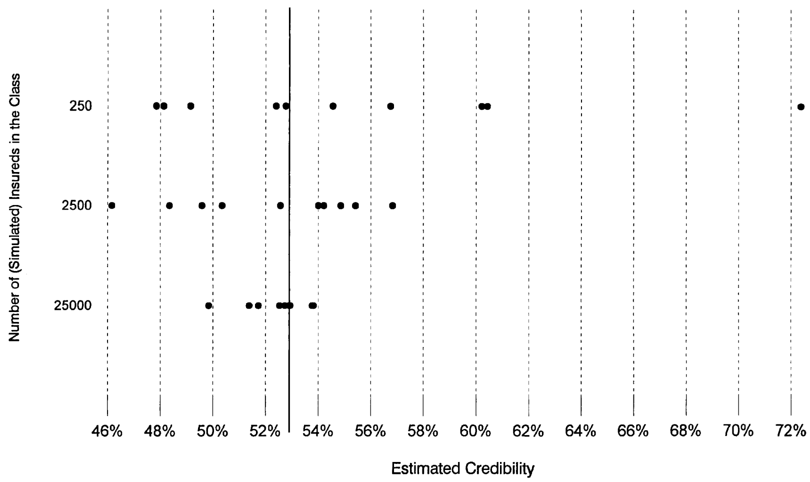
Credibilities are those to be applied to one observation of one insured. The theoretically correct value is 71.4%. Credibilities are estimated from the slope of the regression between one year of observations for the class and a subsequent year of observations for the class.

periments were performed for data sets of 2,000 and 20,000. As shown in Figure 11, with more insureds the credibility estimates are more tightly bunched and closer to the theoretically correct value.

Figure 12 is similar to Figure 11 but deals with the pure premiums rather than frequencies. With only 250 insureds there is considerable random fluctuation in the estimates. With 25,000 insureds the estimates are clustered between 50% and 54%. Due to the larger process variance, the estimates are less tightly clustered than they are in the examples involving frequency shown in Figure 11.

FIGURE 12

SIMULATION EXPERIMENTS
CREDIBILITIES ESTIMATED BY REGRESSION
LOSSES FOR EXCELLENT AND UGLY RISKS



Credibilities are those to be applied to one observation of one insured. The theoretically correct value is 52.9%. Credibilities are estimated from the slope of the regression between one year of observations for the class and a subsequent year of observations.

7. SQUARED ERRORS

Figures 13 through 17 illustrate the expected squared errors between the prediction and future observation for various weights applied to the observed data.

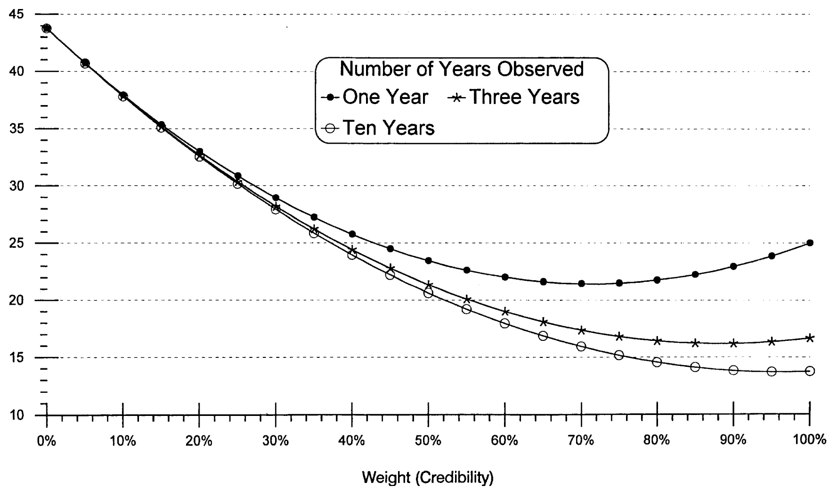
Figures 13 and 15 deal with the frequency example with all four types of insureds as illustrated in Figures 6 and 7. Figure 13 displays the expected squared error[12] as a function of the weight (credibility) given to the observed frequency. The expected squared error is a parabola as a function of the weight.[13]

[12]The expected value of the squared difference between the future observation and the prediction.

[13]This mathematical fact is demonstrated in Appendix C.

FIGURE 13

EXPECTED SQUARED PREDICTION ERRORS VS. WEIGHT GIVEN TO OBSERVED FREQUENCY
EXCELLENT, GOOD, BAD, AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5), Good Risks (Poisson 10), Bad Risks (Poisson 15), and Ugly Risks (Poisson 20). Expected value of process variance = 12.5, variance of the hypothetical means = 31.25. $K = 12.5/31.25 = 0.4$. Least squares credibilities are 71.4%, 88.2%, and 96.2%, for 1, 3, and 10 years of data, respectively.

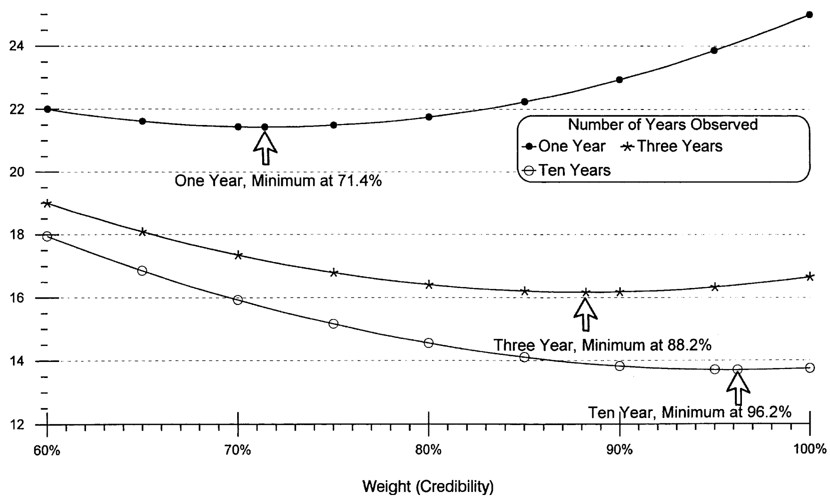
For one year of observed data, the expected squared error is minimized for a weight of 71.4%, the Bühlmann credibility for this situation. For three years of observed data, the minimum occurs for a weight of 88.2%. For ten years of observed data, the minimum occurs for a weight of 96.2%.[14]

As seen in Figure 13, as the number of years of observations increases, the prediction error from relying solely on the data (weight = 100%) declines, while the prediction error from relying solely on the a priori mean (weight = 0) remains the same. Thus, the place where the parabola reaches its minimum moves

[14]Note $10/(10 + .4) = 96.2\%$. Similarly $3/(3 + .4) = 88.2\%$ and $1/(1 + .4) = 71.4\%$.

FIGURE 14

EXPECTED SQUARED PREDICTION ERRORS VS. WEIGHT GIVEN TO OBSERVED FREQUENCY
EXCELLENT, GOOD, BAD, AND UGLY RISKS



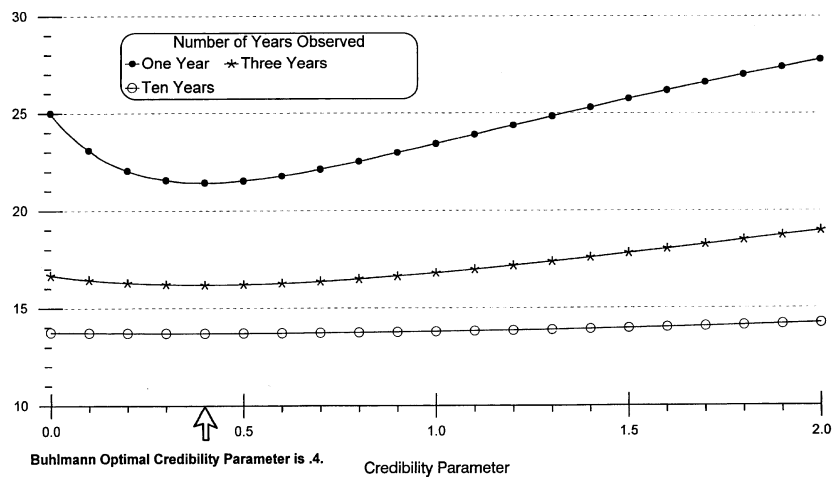
Equal numbers of: Excellent Risks (Poisson 5), Good Risks (Poisson 10), Bad Risks (Poisson 15), and Ugly Risks (Poisson 20). Expected value of process variance = 12.5, variance of the hypothetical means = 31.25. $K = 12.5/31.25 = 0.4$. Least squares credibilities are 71.4%, 88.2%, and 96.2%, for 1, 3, and 10 years of data, respectively.

to the right as the number of years of data increases; the credibility increases becoming 100% in the limit as the number of years increases. For example, for one year of data the parabola reaches its minimum at 71.4%, while for three years of data the corresponding parabola reaches its minimum at 88.2%. Figure 14 is a magnified version of Figure 13, which more clearly displays the minima.

Figure 15 is similar to Figure 13, but here the expected squared error is displayed as a function of the "credibility parameter." In other words, we give N years of data weight $Z = N/(N+K)$, using the Bühlmann credibility formula with credi-

FIGURE 15

EXPECTED SQUARED PREDICTION ERRORS VS. CREDIBILITY PARAMETER USED TO DETERMINE WEIGHT GIVEN TO OBSERVED FREQUENCY
EXCELLENT, GOOD, BAD, AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5), Good Risks (Poisson 10), Bad Risks (Poisson 15), and Ugly Risks (Poisson 20). Expected value of process variance = 12.5, variance of the hypothetical means = 31.25, $K = 12.5/31.25 = 0.4$.

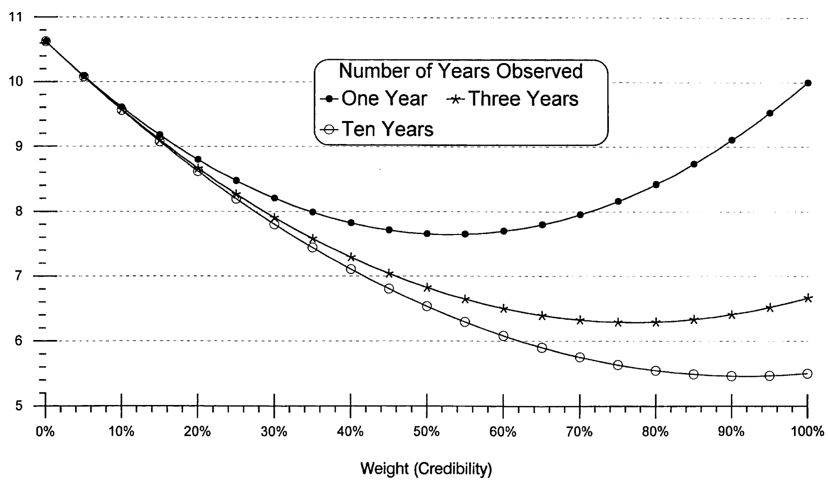
bility parameter K.[15] As shown in Appendix A, for Situation 3, the Bühlmann credibility parameter is 0.4; as seen in Figure 15, the expected squared error is indeed minimized for this value of the credibility parameter. Note the same credibility parameter of 0.4 is optimal regardless of the number of years of data observed.

Figures 16 and 17 are similar to Figures 13 and 15, but deal with the pure premiums rather than frequencies. Figure 16 shows the expected squared errors, which are parabolas as a function of

[15]In Figure 15 K is not necessarily the Bühlmann credibility parameter. Rather, we use a value of K to calculate a value of Z, which may not be the least squares Bühlmann credibility. In the case of Figure 15, 0.4 is the Bühlmann credibility parameter.

FIGURE 16

EXPECTED SQUARED PREDICTION ERRORS (BILLIONS) VS. WEIGHT GIVEN TO OBSERVED FREQUENCY
EXCELLENT AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5) and Ugly Risks (Poisson 20). Expected Value of Process Variance = 5,000 million, Variance of the Hypothetical Means = 5,625 million. $K = 0.8889$. Least Squares Credibilities are 52.9%, 77.1%, and 91.8%, for 1, 3, and 10 years of data respectively.

the weight applied to the observed losses. Again, the expected squared errors are minimized when the weight given to the observed losses corresponds to the Bühlmann credibility.

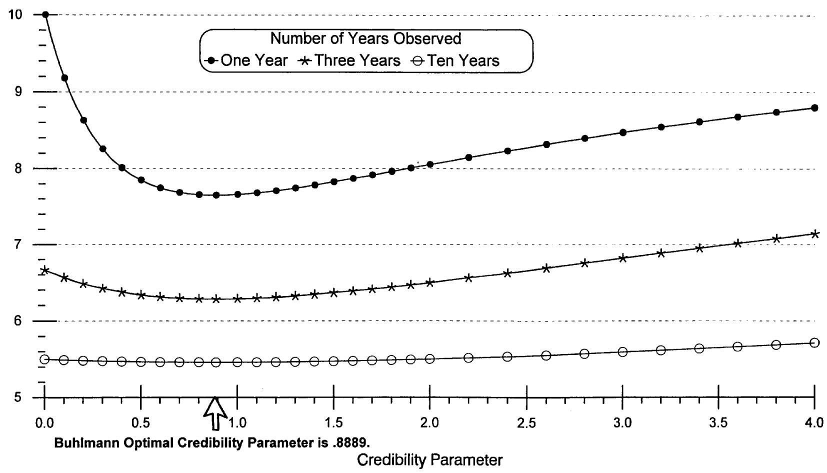
Figure 17 shows the expected squared error as a function of the credibility parameter. As shown in Appendix A, for Situation 4, the Bühlmann credibility parameter $K = .8889$. As seen in Figure 17, this value of the credibility parameter minimizes the expected squared errors.

8. CONCLUSIONS

Credibility, as used in experience rating, has been illustrated via graphs. The estimated credibility was equal to the slope of the line obtained from a least squares regression.

FIGURE 17

EXPECTED SQUARED PREDICTION ERRORS (BILLIONS) VS. CREDIBILITY PARAMETER USED TO DETERMINE WEIGHT GIVEN TO OBSERVED LOSSES EXCELLENT AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5) and Ugly Risks (Poisson 20). Expected value of process variance = 5,000 million, variance of the hypothetical means = 5,625 million. $K = 0.8889$.

Prior and subsequent experience has been simulated for various sets of insureds for different sets of simple assumptions. This simulated data for the various examples was used to illustrate that the slope of the regression line between prior and subsequent experience is one estimate of the Bühlmann credibility. Finally, these same examples were used to illustrate that the expected squared error between the actual and predicted subsequent experience is minimized when the weight given to the observed experience is equal to the Bühlmann credibility.

The regression technique shown here for illustrative purposes could be employed in simple situations. Where greater accuracy is desired or where the behavior is more complicated empirical Bayesian and other techniques have been developed to estimate credibilities from data.[16] In any case, the regression techniques applied to simulations of simple examples are another useful way to learn and understand the important basic ideas of credibility and experience rating.

[16]See for example ISO [11], Venter [12], Mahler [13], or Mahler [14].

REFERENCES

[1] Philbrick, Stephen W., "An Examination of Credibility Concepts," *PCAS* LXVIII, 1981, pp. 195–219.

[2] Hewitt, Charles C., "Credibility for Severity," *PCAS* LVII, 1970, pp. 148–171.

[3] Meyers, Glenn G., "An Analysis of Experience Rating," *PCAS* LXXII, 1985, pp. 278–317.

[4] Mahler, Howard C., Discussion of [3], *PCAS* LXXIV, 1987, pp. 119–189.

[5] Finger, Robert J., "Merit Rating for Doctor Professional Liability Insurance," *PCAS* LXXX, 1993, pp. 291–352.

[6] Gillam, William R. and Richard H. Snader, "Fundamentals of Individual Risk Rating," 1992, available from the CAS.

[7] Tiller, Margaret Wilkinson, "Individual Risk Rating," Chapter 3 of *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990.

[8] Hossack, I. B., J. H. Pollard and B. Zehnwirth, *Introductory Statistics with Applications to General Insurance*, New York, Cambridge University Press, 1983.

[9] Herzog, Thomas N., *An Introduction to Credibility*, ACTEX, 1994.

[10] Mahler, Howard C., "A Student's Guide to Conjugate Priors," *New England Actuarial Seminars*, 1997.

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

[12] Venter, Gary G., "Credibility," Chapter 7 of *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990.

[13] Mahler, Howard C., "A Markov Chain Model of Shifting Risk Parameters," *PCAS* LXXXIV, 1997, pp. 581–659.

[14] Mahler, Howard C., "Credibility with Shifting Risk Parameters, Risk Heterogeneity and Parameter Uncertainty," *PCAS* LXXXV, 1998, pp. 455–653.

[15] Mayerson, Allen L., "A Bayesian View of Credibility," *PCAS* LI, 1964, pp. 85–104.

[16] Hewitt, Charles C., Review of [15], *PCAS* LII, 1965, pp. 121–127.

[17] Mahler, Howard C., "A Student's Guide to Bühlmann Credibility and Bayesian Analysis," *New England Actuarial Seminars*, 1997.

[18] Mahler, Howard C., "A Student's Guide to Classical Credibility," *New England Actuarial Seminars*, 1997.

[19] Mahler, Howard C., "A Student's Guide to Loss Distributions," *New England Actuarial Seminars*, 1997.

[20] Klugman, Stuart A., Harry H. Panjer and Gordon E. Willmot, *Loss Models: From Data to Decisions*, Wiley, 1998.

[21] Hogg, Robert V. and Klugman, Stuart A., *Loss Distributions*, Wiley, 1984.

APPENDIX A

CREDIBILITY FOR THE EXAMPLES

The formulas to be used are:[17]

$$K = EPV/VHM,$$

$$Z = N/(N + K),$$

where

K = Bühlmann credibility parameter,

EPV = Expected value of process variance for a single unit of the risk process (i.e., for one insured for one year,

VHM = Variance of the hypothetical means for a single unit of the risk process (i.e., for one insured for one year,

Z = Bühlmann credibility, and

N = Number of years of data (from a single insured).

The following information will be used in various combinations to illustrate credibility ideas.

Claim frequency for individual insureds is assumed to be Poisson.[18] Claim severity is assumed to be given by a Pareto distribution[19] with shape parameter 3 and scale parameter 20,000.Frequency and severity are independent. There are four possible types of insureds with different Poisson parameters:

Type	Average Annual Claim Frequency
Excellent	5
Good	10
Bad	15
Ugly	20

[17]These formulas are explained or derived in, for example, Mayerson [15], Hewitt [16], Hewitt [2], Philbrick [1], Herzog [9], Venter [12], and Mahler [17].
[18]The Poisson parameter for each insured stays the same over time.
[19]$F(x) = 1 - (20{,}000/(20{,}000 + x))^3$.

Situation 1: Frequency with Good and Bad Risks

A risk is selected at random from a class made up equally of Good and Bad Risks.

Since the frequencies are Poisson, their variance is equal to their mean. Therefore, the expected value of the process variance is equal to the overall mean = 12.5.

$$VHM = \{(10 - 12.5)^2 + (15 - 12.5)^2\}/2 = 6.25.$$
$$K = EPV/VHM = 12.5/6.25 = 2.$$

For one year of data (as in Figure 2), $Z = 1/(1 + 2) = 33\%$.

For three years of data (as in Figure 3), $Z = 3/(3 + 2) = 60\%$.

Situation 2: Frequency with Excellent and Ugly Risks

A risk is selected at random from a class made up equally of Excellent and Ugly Risks.

$$EPV = \text{overall mean} = 12.5.$$
$$VHM = \{(5 - 12.5)^2 + (20 - 12.5)^2\}/2 = 56.25.$$
$$K = EPV/VHM = .222.$$

For one year of data (as in Figure 5), $Z = 1/(1 + .222) = 81.8\%$.

Situation 3: Frequency with Excellent, Good, Bad, and Ugly Risks

A risk is selected at random from a class made up equally of Excellent, Good, Bad, and Ugly Risks.

$$EPV = \text{overall mean} = 12.5.$$
$$VHM = \{(5 - 12.5)^2 + (10 - 12.5)^2 + (15 - 12.5)^2 + (20 - 12.5)^2\}/4 = 31.25.$$
$$K = 12.5/31.25 = .4.$$

For one year of data (as in Figure 7), $Z = 1/(1 + .4) = 71.4\%$.

Situation 4: Pure Premiums (Unlimited Losses) for Excellent and Ugly Risks

A risk is selected at random from a class made up equally of Excellent and Ugly risks.

With a Poisson frequency,[20] the process variance of the pure premiums = (mean frequency)(second moment of the severity). (See, for example, Mahler [18].) Since the severity distribution is assumed to be the same for all risks, the expected value of the process variance = (overall mean frequency)(second moment of the severity).

For a Pareto distribution, $F(x) = 1 - (\lambda/(\lambda + x))^{\alpha}$, the second moment of the severity is $2\lambda^2/\{(\alpha - 1)(\alpha - 2)\}$, which in this case is 400 million. Therefore, since the mean frequency is 12.5, *EPV* = (12.5) (400 million) = 5 billion.

For a Pareto distribution, the mean is $\lambda/(\alpha - 1) = 10{,}000$. Thus, the hypothetical mean pure premiums are 50,000 and 200,000. Thus, the *VHM* = 5.625 billion.

Thus, $K = EPV/VHM = 0.8889$.

For one year of data (as in Figure 9), $Z = 1/(1 + .8889) =$ 52.9%.

Situation 5: Limited Losses for Excellent and Ugly Risks

A risk is selected at random from a class made up equally of Excellent and Ugly Risks. One observes the losses limited to $25,000 per claim and attempts to predict the future limited losses for the same insured.

---

[20] In general, for cases where frequency and severity are independent, the process variance of the pure premium = (mean frequency)(variance of severity) + (mean severity)^2 (variance of frequency). For a Poisson, mean frequency = variance of the frequency. Thus, the process variance of the pure premiums = (mean frequency)(variance of severity + mean severity^2) = (mean frequency)(second moment of severity).

For a Pareto distribution, $F(x) = 1 - (\lambda/(\lambda + x))^{\alpha}$, the limited second moment is given by[21]

$$E[X^2; L] = E[X^2]\{1 - (1 + L/\lambda)^{1-\alpha}[1 + (\alpha - 1)L/\lambda]\}.$$

In this case, $E[X^2; 25{,}000] = 400$ million $\{1 - (1 + 1.25)^{-2}[1 + (2)(1.25)]\} = 123.5$ million. As in Situation 4, *EPV* = (overall mean frequency)(second moment of the severity) = (12.5)(123.5 million) = 1.544 billion.

For a Pareto distribution, the limited expected value is given by[22]

$$E[X; L] = E[X]\{1 - (1 + L/\lambda)^{1-\alpha}\}.$$

In this case, $E[X; 25{,}000] = (10{,}000)(1 - (1 + 1.25)^{-2}) = 8{,}025$. Thus, the hypothetical mean pure premiums are (5)(8,025) = 40,125 and (20)(8,025) = 160,500. Therefore, *VHM* = 3.623 billion. K = *EPV/VHM* = 0.426. Note that while both the *EPV* and *VHM* declined compared to Situation 4, the *EPV* declined more. Therefore, the Bühlmann credibility parameter K declined from 0.8889 to 0.426. Thus, for one year of data (as in Figure 10) $Z = 1/(1 + .426) = 70.1\%$.

[21] See Mahler [19] or Klugman, Panjer, and Willmot [20].

[22] See Hogg and Klugman [21], Mahler [19], or Klugman, Panjer, and Willmot [20].

APPENDIX B

REGRESSION AND CREDIBILITY

It turns out that in the example here,[23] the Bühlmann credibility is approximately the slope of the least squares line between prior and subsequent observations, as will be shown in this Appendix. Also, it will be shown that the regression line is expected to pass approximately through the point (M,M), where M is the overall mean.

Let X_i be the prior observations (for one year) for the insureds in the portfolio and let Y_i be the subsequent observations (for one year) for the insureds in the portfolio. A regression line $y = ax + b$, with slope a and intercept b, can be fit between the prior and subsequent observations.[24] Then the slope of the least squares line is given by:[25]

$$a = \frac{(\Sigma X_i Y_i/O) - (\Sigma X_i/O)(\Sigma Y_i/O)}{(\Sigma X_i^2/O) - (\Sigma X_i/O)^2}.$$

Where O is the number of insureds observed.

The numerator has an expected value equal to the covariance of X_i and Y_i, the observations in two separate years.[26] This is assumed to be τ^2, the variance of the hypothetical means.[27]

The denominator has an expected value equal to the variance of the observations in a single year.[28] It is assumed that this

---

[23]This result holds in the case of the covariance structure assumed in Appendices A and C. In particular, there are no shifting risk parameters over time. More general covariance structures are discussed, for example, in Meyers [3], Mahler [4], Mahler [13], and Mahler [14].

[24]As is done in Figures 2, 3, 5, 7, 9, and 10.

[25]For simplicity we have assumed that each insured is of the same size and gets the same weight. Thus, we perform an unweighted regression.

[26]Recall that $\mathrm{Cov}[A,B] = \mathrm{E}[AB] - \mathrm{E}[A]\mathrm{E}[B]$.

[27]One of the assumptions underlying Bühlmann's credibility formula is that the covariance between different years of data is the variance of the hypothetical means.

[28]Recall that $\mathrm{Var}[A] = \mathrm{E}[A^2] - \mathrm{E}[A]^2$.

expected value is $\tau^2 + \eta^2$, the sum of the variance of the hypothetical means and the expected value of the process variance.[29]

If one plugs in the expected values of both the numerator and denominator, then we expect:[30]

$$a \approx \tau^2/(\tau^2 + \eta^2) = \text{Bühlmann credibility for one year.}^{31}$$

If X had been an observation for N years rather than one year, then the expected value of the process variance (of the frequency or pure premiums) would have declined by a factor of $1/N$; it would have been η^2/N rather than η^2 as for one year.[32] On the other hand, the variance of the hypothetical means would have remained the same.[33] Thus, with N years of data rather than one, the expected value of the numerator would have been the same, but the expected value of the denominator would have been $\tau^2 + \eta^2/N$ and

$$a \approx \tau^2/(\tau^2 + \eta^2/N) = N/(N + \eta^2/\tau^2) = N/(N + K)$$
$$= \text{Bühlmann credibility for } N \text{ years.}$$

Thus, the slope of the regression line is approximately[34] equal to the Bühlmann credibility.

---

[29]The terms are each defined in terms of a single year of data. The total variance is equal to the *VHM* plus *EPV*.

[30]Note that this estimator which is a ratio of two unbiased estimators can be biased. This subject has been extensively discussed in relation to empirical Bayes credibility. See, for example, ISO [11] and Venter [12].

[31]$\tau^2/(\tau^2 + \eta^2) = VHM/(VHM + EPV) = 1/(1 + EPV/VHM) = N/(N + K)$, with $N = 1$ and $K = EPV/VHM$.

[32]See, for example, Mahler [17] or Mahler [14]. The process variance of the number of claims increases by a factor of N, since variances add for (independent) years. However, the claim frequency is the claim count divided by N, which introduces a factor of $1/N^2$ into the variances. The net result is a factor of $N/N^2 = 1/N$ for the process variance of the claim frequency.

[33]See, for example, Mahler [17] or Mahler [14]. The hypothetical annual means of the claim frequency are unchanged, thus, their variance is also unaffected. Alternately, the hypothetical mean claim counts are multiplied by N and, thus, their variance is multiplied by N^2. However, claim frequency is divided by N, which introduces a factor of $1/N^2$ into the variance. The *VHM* is, thus, multiplied by $N^2/N^2 = 1$.

[34]In general $E[A/B] \neq E[A]/E[B]$. Nevertheless, for the situations such as being dealt with here, $E[A/B] \approx E[A]/E[B]$.

Also, one can show that the regression line is expected to pass approximately through the point (M, M) where M is the overall mean. The intercept of the least squares line is

$$b = \frac{(\Sigma Y_i/O)(\Sigma X_i^2/O) - (\Sigma X_i Y_i/O)(\Sigma X_i/O)}{(\Sigma X_i^2/O) - (\Sigma X_i/O)^2},$$

where O is the number of insureds observed each year,

$\Sigma Y_i/O$ has an expected value equal to the overall mean M, and $\Sigma X_i^2/O$ has an expected value of the second moment of the average of N years of data. This is the sum of the variance of the average of N years of data plus the square of the overall mean. In turn, the variance of the average of N years of data[35] is equal to $\tau^2 + \eta^2/N$.

Thus, the expected value of $\Sigma X_i^2/O$ is equal to $\tau^2 + \eta^2/N + M^2$.

The numerator of a is equal to $\Sigma X_i Y_i/O - (\Sigma X_i/O)(Y_i/O)$. Thus, the expected value of $\Sigma X_i Y_i/O$ is equal to that of the numerator of a, τ^2, plus the expected value of $(\Sigma X_i/O)(\Sigma Y_i/O)$ which is (M) (M). Therefore, the expected value of $\Sigma X_i Y_i/O$ is $\tau^2 + M^2$.

$\Sigma X_i/O$ has an expected value equal to the overall mean M.

Thus, the expected value of the numerator of b with N years of data is

$$M(\tau^2 + \eta^2/N + M^2) - (\tau^2 + M^2)M = M\eta^2/N.$$

[35]Where $\mathbf{W}_j$ is the vector of data for year j

$$\begin{aligned}\mathrm{Var}[(1/N)\Sigma \mathbf{W}_j, (1/N)\Sigma \mathbf{W}_k] &= (1/N^2)\mathrm{Var}[\Sigma \mathbf{W}_j, \Sigma \mathbf{W}_k] \\ &= (1/N^2)\Sigma\,\mathrm{Var}[\mathbf{W}_j, \mathbf{W}_j] + (1/N^2)\Sigma_{j\neq k}\mathrm{Var}[\mathbf{W}_j, \mathbf{W}_k] \\ &= (1/N^2)N(\tau^2 + \eta^2) + (1/N^2)((N^2 - N)(\tau^2)) \\ &= \tau^2 + \eta^2/N.\end{aligned}$$

The denominator of b is the same as that of a, and has an expected value of $\tau^2 + \eta^2/N$ (for N years of data).

Thus, by substitution we expect

$$b \approx \frac{M\eta^2/N}{\tau^2 + \eta^2/N}.$$

Thus, since as was shown above,

$$a \approx \tau^2/(\tau^2 + \eta^2/N),$$

$$aM + b \approx M\tau^2/(\tau^2 + \eta^2/N) + (M\eta^2/N)/(\tau^2 + \eta^2/N) = M.$$

Thus, we indeed expect the regression line $y = ax + b$ to pass approximately through the point (M,M). Prior experience for an insured equal to the overall a priori expectation results in a prediction equal to the overall a priori expectation.

APPENDIX C

EXPECTED VALUE OF SQUARED ERRORS

This appendix discusses the expected value of the squared errors that result from the use of credibility to estimate an insured's future experience from the insured's past observed experience. The results of this appendix are illustrated in Figures 13 through 17. This appendix also shows how to calculate the Bühlmann credibility, which is the value that minimizes this expected squared error.[36]

Assume we have a time series, X_i, and we wish to estimate a future year of the same time series, $X_{N+\Delta}$, by weighting together observations X_i for $i = 1$ to N and the overall mean M. For example, the X_i could be the observed frequencies for a single insured over a series of individual years. If Z_i is the weight applied to year X_i, then

$$\text{Estimate} = \sum_{i=1}^{N} Z_i X_i + \left(1 - \sum_{i=1}^{N} Z_i\right) M.$$

Then the expected squared error comparing the estimate to the observation[37] is a quadratic function of the weights Z_i:[38]

$$V(Z) = \sum_{i=1}^{N}\sum_{j=1}^{N} Z_i Z_j C_{ij} - 2\sum_{i=1}^{N} Z_i C_{i,N+\Delta} + C_{N+\Delta,N+\Delta},$$

where $C_{ij} = \text{Cov}[X_i, X_j]$.

---

[36]It turns out that this value of credibility also minimizes the squared error between the predictions and the true/hypothetical means and between the predictions and the Bayesian estimates. See for example Mahler [17].

[37]The expected squared error compared to the observation, with respect to the hypothetical mean, or with respect to the Bayesian estimate are each minimized by the value for credibility calculated using the formula derived in this appendix.

[38]See for example Mahler [13].

In the examples in Appendix A, the covariance structure is that underlying the Bühlmann credibility formulation:

$$C_{ij} = \tau^2 + \delta_{ij}\eta^2,$$

where C_{ij} = Covariance of year i and year j,

η^2 = Expected value of the process variance,

τ^2 = Variance of the hypothetical means, and

$\delta_{ij} = 1$ if $i = j$ and 0 if $i \neq j$.

Due to symmetry in this case, it turns out that the expected squared errors are minimized for $Z_i = Z_j$. Let $Z = \sum_{i=1}^{N} Z_i$ = total weight to be applied to N years of data. Then if $Z_i = Z_j = Z/N$, substituting into the formula for the expected squared errors:

$$\begin{aligned} V(Z) &= \sum_{i=1}^{N}\sum_{j=1}^{N} Z_i Z_j(\tau^2 + \delta_{ij}\eta^2) - 2\sum_{i=1}^{N} Z_i\tau^2 + \tau^2 + \eta^2 \\ &= Z^2\tau^2 + \eta^2(Z/N)^2 N - 2\tau^2 Z + \tau^2 + \eta^2 \\ &= Z^2(\tau^2 + \eta^2/N) - 2\tau^2 Z + \tau^2 + \eta^2. \end{aligned}$$

For Situation 3 in Appendix A, $\eta^2 = 12.5$ and $\tau^2 = 31.25$. Thus, for ten years of data $V(Z) = 32.5Z^2 - 62.5Z + 43.75$. This is one of the parabolas shown in Figure 13.

In order to minimize the expected squared error, we set the derivative $V'(Z) = 0$. This results in $Z = \tau^2/(\tau^2 + \eta^2/N) = N/(N + \eta^2/\tau^2) = N/(N + K)$, where $K = \eta^2/\tau^2 = EPV/VHM$. For example, for 10 years of data in Figures 13 or 14, the parabola is minimized for $Z = 31.25/(31.25 + 12.5/10) = 0.962$. Alternatively, $K = 12.5/31.25 = 0.4$ and $Z = 10/(10 + .4) = 0.962$. As seen in Figure 15, this value of K minimizes the expected squared error.

THE MYERS–COHN PROFIT MODEL, A PRACTICAL APPLICATION

HOWARD C. MAHLER

Abstract

The Myers–Cohn Profit Model is presented via both a simple example and a practical application. The practical application is shown in considerable detail in order to illustrate some of the techniques required in applying theory in the real world. This should help actuaries understand the model as well as illustrate the importance of the inputs chosen and assumptions made. Since most of the inputs used in this profit model are required by other profit models, the illustrations of how to quantify these input values should be of more general applicability.

ACKNOWLEDGEMENT

I wish to thank Alexandra St. Onge for preparing the exhibits and Silvia Fitcher and Dotty Culleton for typing this paper.

1. INTRODUCTION

Beginning with Commissioner James M. Stone's automobile bodily injury liability decision for 1976 state set rates, explicit account has been taken of investment income in ratemaking for the major lines of automobile and workers compensation insurance in Massachusetts. Although the computational techniques have changed over the years, the common thread has been to attempt to allow insurers a fair return on their equity.

This paper will present one profit model that has been used. A simple example will be presented as well as a practical application.

2. THE MYERS–COHN MODEL

The Myers–Cohn net present value model was developed for the Massachusetts Rating Bureaus by Professors Stewart Myers and Richard Cohn. It was intended as an improvement of the Fairley model which was used previously.[1] The basic concepts underlying the Fairley model, the model shown in Mahler [2][2] and the Myers–Cohn model are all similar. Given similar inputs all three models give similar (but not identical) results. The Myers–Cohn model was first presented in the fall of 1981 at the 1982 automobile rate hearings. Then Commissioner Sabbagh used a modified version of this model to fix and establish the 1982 private passenger automobile rates. The Massachusetts Rating Bureaus used the Myers–Cohn model to derive its proposed workers compensation underwriting profit provision as well. It is currently used, with some technical refinements, to set profit provisions for both automobile and workers compensation insurance in Massachusetts.

The basic premise underlying the Myers–Cohn model can be stated this way: a fair premium must be equal to the expected losses and expenses discounted to present value at a risk-adjusted rate, plus the present value of the federal income taxes on underwriting and investment income discounted at an appropriate rate.[3] Premiums calculated this way should preserve the equity invested in the company and give the investor a fair return for the risk of underwriting by the company.

The Myers–Cohn model shares many features of other profit models. One estimates the length of time an insurer can invest

[1]The original Fairley Model, an improvement by Hill and Modigliani, and the Myers–Cohn Model, are all presented in *Fair Rate of Return in Property-Liability Insurance* [1].
[2]This model was first presented in the spring of 1981 and is described as "Model A" in Part III of the 1984 NAIC Study of Investment Income [3].
[3]As shown in Exhibits 3 and 5, and as discussed below, those underwriting taxes corresponding to the loss and expense payments are discounted at a risk-adjusted rate, while the other income taxes are discounted at the risk-free rate.

premium dollars prior to paying losses and expenses.[4] One estimates the investment income an insurer will earn on this cashflow and the necessary equity (surplus) backing up the policies. One takes into account the resulting income tax payments. Finally, one incorporates a reward to the insurer for taking the risk of writing insurance.

While this feature is shared with many other profit models,[5] the manner of doing so in the Myers–Cohn model is different. In the Myers–Cohn model selecting a risk-adjusted discount rate takes the place of selecting an appropriate rate of return on equity.

In the application of the Myers–Cohn model shown here, as well as the original paper by Myers and Cohn [1], the risk of writing insurance is quantified via the Capital Asset Pricing Model (CAPM). However, this is not a requirement. The Myers–Cohn model uses a risk-adjusted discount rate as an input. The difference between the risk-free and risk-adjusted rate determines the reward for taking the risk of underwriting. How this difference is selected is up to the person using the model. The CAPM is only one way to go about selecting this difference.

Once all the inputs have been determined, the Myers–Cohn equation yields the necessary premium as a ratio to losses and expenses. As shown in Exhibit 5,[6]

$$\frac{P}{L+E} = \frac{\kappa_1 - \tau_1\kappa_5}{\kappa_2 - \tau_2 r\kappa_3 - \tau_1\kappa_4 - \tau_1\alpha\kappa_6}.$$

Then one calculates the corresponding underwriting profit provision as 1 – (Losses + Expenses)/Premiums.

In order to illustrate the use of the Myers–Cohn model, a simplified example will be presented first. Later a practical application to Massachusetts workers compensation will be shown.

[4] Consideration of policyholder dividend payments may also be included in the model.
[5] See for example, Mahler [2].
[6] The terms in the equation are defined and discussed in Section 3 and in Exhibit 5.

It is neither the purpose nor intention of this paper to justify particular selections of inputs nor to determine the appropriate underwriting profit provision for use in any particular circumstance. All chosen inputs and calculated profit provisions are solely for illustrative purposes. As with all profit models, the profit provision calculated using the Myers–Cohn model is very sensitive to the inputs chosen and assumptions made. This sensitivity will be illustrated.

3. SIMPLE EXAMPLE

This section will illustrate the Myers–Cohn model via a simple example. The corresponding calculations are shown in Exhibits 1 through 5.

3.1. Simple Example, Inputs and Assumptions

For this simplified example, make the following assumptions:

- All premiums are collected in Quarter 1.
- All losses are paid in Quarter 5.
- Variable expenses are 20% of premiums, and are paid in Quarter 2.
- The ratio of fixed expenses to losses is 5%.
- Fixed expenses are paid in Quarter 2.
- Loss adjustment expenses are 10% of losses, and are paid when losses are in Quarter 5.
- The federal income tax rate on underwriting is 35%.
- Investments are made solely in risk-free Treasury securities.
- There are no investment expenses.
- The federal income tax rate on investment income is 35%.
- There is no discounting of reserves for tax purposes and no revenue offset feature of the tax code.

- There are no dividend payments.
- The risk-free rate is assumed to be 9%.
- A risk-adjusted rate of 7% is used. The important concept is that discounting "risky" loss and expense flows at the smaller risk-adjusted rate is intended to compensate insurers for the risk of underwriting insurance.[7]
- A 2-to-1 initial premium-to-surplus ratio is chosen.
- The surplus allocated to this policy is assumed to decline in proportion to the losses and expenses paid.

3.2. Simple Example, Result and Outputs

Using the Myers–Cohn profit model, the calculated underwriting profit provision is -3.0% as shown in Exhibit 1. The purpose of this example is to illustrate and help to understand the method of calculation, rather than concentrate on the answer itself. Exhibits 2, 3 and 4 show in detail how the cashflows are constructed and how the kappa values are determined. The kappa values are "timing parameters." They are calculated by discounting the various cashflows at either the risk-free or risk-adjusted rate. Exhibit 2 shows the cashflows for the initial set of weights.[8] However, as the profit provision varies so does the relative weight given to variable expenses, so that the profit model is solved via iteration.[9] The initial weights based on a profit provision of zero are used to calculate a profit provision which in turn yields a new relation of premiums to losses and a new set of

---

[7]One could combine a 9% risk-free rate with an assumed beta of liabilities of $-.2$ and a market risk premium of 10%, to get a risk adjusted rate of 7%. $7\% = 9\% - .2 \times 10\%$. This is the method used in Section 4.8. While this is based on the Capital Asset Pricing Model, some other means could be used to get the risk-adjusted rate.

[8]The cashflows are constructed for a single policy (or set of policies with the same effective date), with a policy effective period of Quarters 1, 2, 3, and 4. Thus, the policy effective date (time=0) is at the end of Quarter 0, and the beginning of Quarter 1.

[9]While one might attempt to solve for P in closed form, this would be very complicated and have little if any practical value. For an analogous situation, Mahler [2, p. 257] discusses an approximation which allows one to solve for P in closed form.

weights. This new set of weights is used to calculate a new profit provision. The process continues until the iteration converges to the "final weights" and profit provision.[10] Exhibit 4 shows the cashflows for the final weights.

3.3. *Simple Example, Details*

The top portion of Exhibit 1 shows the inputs and assumptions chosen for this example. Next are shown the various kappa values, which are defined as follows:[11]

κ_1 = Risk-adjusted discounted losses and expenses factor

κ_2 = Risk-free discounted premiums factor

κ_3 = Risk-free discounted investment balance tax factor

κ_4 = Risk-free discounted underwriting profit tax factor (contribution of premiums)

κ_5 = Risk-adjusted discounted underwriting profit tax factor (contribution of losses and expenses)

κ_6 = Risk-free discounted revenue offset tax factor.

The calculation of the kappa values is shown in Exhibit 3, for the initial weights. κ_1 is the risk-adjusted discounted loss and expense factor. It is calculated by discounting the loss and expense flows from Exhibit 2 at the risk-adjusted rate of 7%. The result is divided by the sum of losses and expenses, which has been selected as 1,000.

κ_2 is the result of discounting the premium flow at the 9% risk-free rate.

---

[10]Generally, this takes three or four iterations.

[11]The Myers–Cohn paper had only four kappas. One additional kappa was introduced in implementation to allow for the difference in timing between the payment of losses and expenses, and the timing of the tax consequences of incurring losses and expenses. κ_6 was introduced in order to take into account the "revenue offset" feature of the Tax Reform Act of 1986.

κ_3 is the result of discounting the investment balance for taxes at the risk-free rate. The investment balance for taxes shown on Exhibit 2 is the sum of the surplus plus the premium dollars collected that have yet to be paid out as losses, expenses, or dividends.

κ_4 is the discounted contribution of premiums to the underwriting profit tax. κ_5 is similar but for losses and expenses, and thus discounted at a risk-adjusted rate. It's assumed these take place evenly in the four policy quarters.

κ_6 is the discount factor used to take into account the revenue offset feature of the tax code.

The bottom portion of Exhibit 1 shows how the different factors are put together to calculate the ratio of premiums to losses and expenses and in turn the underwriting profit provision:

$$P/(L+E) = (\kappa_1 - \tau_1\kappa_5)/(\kappa_2 - \tau_2 r\kappa_3 - \tau_1\kappa_4 - \tau_1\alpha\kappa_6).$$

Those terms involving losses and expenses are in the numerator. The terms involving taxes include the tax rates, either τ_1, the underwriting tax rate, or τ_2, the investment income tax rate.

The term $\tau_2 r\kappa_3$ is the tax rate τ_2 times the investment income of $r\kappa_3$, which is the quarterly rate of return times the (discounted) investment balance.

Once the ratio of $P/(L+E)$ is calculated as 0.9712, the profit provision is $1-(1/0.9712) = -3.0\%$. This can be thought of as a target combined ratio of 103%.

4. PRACTICAL APPLICATION, MASSACHUSETTS WORKERS COMPENSATION

This section describes a practical application to Massachusetts workers compensation insurance. The calculations are shown in Exhibits 5 through 23.

Exhibit 5 shows the equations for the Myers–Cohn model.[12] As in the simple example in the previous section, the various inputs are brought together to calculate the profit provision shown in Exhibit 5.

In many cases, inputs have been taken from elsewhere in the ratemaking procedure.[13] The calculations that produced those inputs are beyond the scope of this paper. However, in general it is important to choose a set of consistent inputs to any underwriting profit model. The set of inputs should be consistent both internally and with other parts of the ratemaking process.

Certain complications present in recent rate filings have been removed to aid in exposition. Enough complications have been left to illustrate the usual types of difficulties that arise in practical situations. However, every application can have its own peculiar details that require special treatment. Many of those that have arisen in Massachusetts workers compensation are beyond the scope of this paper.

For completeness, the changes that were made from the filing for 1/1/98 rates to get the practical application shown here are listed in the Appendix.[14]

4.1. *Calculation of the Underwriting Profit Provision*

As in the simple example in the previous section, the various inputs are used to calculate six timing parameters, κ_1 through κ_6. These are then combined in Exhibit 5 using the formulas:

$$\frac{P}{L+E} = \frac{\kappa_1 - \tau_1\kappa_5}{\kappa_2 - \tau_2 r\kappa_3 - \tau_1\kappa_4 - \tau_1\alpha\kappa_6} \qquad \text{and}$$

$$\mu = 1 - (P/(L+E))^{-1},$$

[12]These were also used in the simple example in the previous section.

[13]For example, the estimate of loss flows employs estimates of ultimate losses by accident year.

[14]Among the complications not presented here, is the use of a simulation model (along the same general lines as in Venter and Gillam [4]) in order to estimate the impact on the indemnity loss flows of a major law change.

where r is the quarterly risk-free rate, τ_1 is the underwriting income tax rate, τ_2 is the investment income tax rate and α is a factor related to the revenue offset feature of the tax law.

As shown in Exhibit 5, this results in a model profit provision of -3.6% for the Massachusetts workers compensation example.[15] In order to apply this model profit provision in the usual Massachusetts workers compensation ratemaking procedure one final step is needed.[16]

Premium discounts are reductions in premiums for larger insureds to reflect their lower expense needs as a percent of Standard Premium. Standard Premium is prior to the impact of premium discounts.[17]

Ignoring the existence of Standard Premium and premium discounts in the profit model should have no economic impact since premium discounts merely represent money the insurer did not receive and never expected to receive.[18]

This idea is implemented as follows:

> The premium flow is net of premium discounts. (See Exhibit 6.)
>
> The expense flows do not include any weight for premium discounts. The initial weights are determined without the premium discount. Variable expenses are a percent of net premiums rather than Standard Premiums. (See Exhibit 9.)

---

[15]This result should be viewed as illustrative. Many of the input values (even if selected by the same individual) would vary considerably over time, state, line of insurance, etc.

[16]This step is needed because the rate indication is based on Standard Premiums (plus ARAP), prior to the impact of any premium discounts.

[17]Standard Premium can be thought of as the product of payroll, manual rate, and experience modification. (As reported on Financial Aggregate Data Calls, it includes expense constants as well.)

[18]The size of the premium discount is not uncertain.

Then as shown in Exhibit 5, one calculates the Underwriting Profit Allowance to load into the ratemaking procedure as:

$$\begin{aligned}&\text{Underwriting Profit Allowance}\\&\quad = (\text{Model Profit Provision})\\&\qquad \times (1 - \text{Premium Discount as \% of Standard Premium})\end{aligned}$$

In this case, the

$$\begin{aligned}&\text{Underwriting Profit Allowance}\\&\quad = (-3.6\%)(1 - 6.8\%) = -3.4\%.\end{aligned}$$[19]

4.2. Premium Cash Flow

The premium flow used in the profit model is shown in Exhibit 6, Part 1. It is estimated from a study conducted by the Rating Bureau and reported in the filing for rates to be effective 1/1/91.

Fourteen separate flows were calculated by combining the sample returns into categories formed by stock/non-stock, retro/non-retro, and size characteristics. Four premium size intervals, 0–4,999; 5,000–99,999; 100,000–499,999; and over 500,000 were used to distinguish among the premium flows for small, medium, and large risks.[20] The 14 flows were determined by calculating the time, in days, from the policy's effective date to the actual payment date. Summaries were then made for 90 day periods.

A raw combined flow was constructed by combining the fourteen individual flows with industrywide weights obtained from Unit Statistical Plan data and representing actual Mas-

[19]Thus, if expected losses, expenses including premium discounts, plus any provision for policyholder dividends, is equal to $1,000, then indicated Standard Premiums would be: $1,000/(1 − (−0.034)) = $967.12.

[20]These are the same size intervals that were used in the schedule of premium discounts at the time of the study.

sachusetts distributions of premium. The individual flows were first weighted by the stock/non-stock/retro/non-retro distribution. A final combination of those flows by size was accomplished using prospective Standard Premium size distributions at projected rates for each combination of stock/non-stock/retro/non-retro.

The raw weighted flow is shown in Exhibit 6, Part 2 as the "untrimmed" flow. Modifications are performed in order to arrive at a final company premium flow. First, the data for the quarter directly preceding the effective date is biased toward the end of that quarter. Most of that data represents deposit premiums which are made immediately prior to the effective date. Indeed, the average date in the sample for that quarter was only 6.5 days prior to the effective date. The use of this aggregate data valued as of the middle of the quarter (45 days) would produce erroneous results. In order to take this effect into account, the data was combined with the first quarter after the effective date for discounting purposes. The average date of the combined data should produce reasonably unbiased discounting results.[21]

Along with the above refinement, the tails of the "untrimmed" flows were truncated to eliminate the noise in the sample data and the remaining flow was normalized to unity. This result is shown in Exhibit 6 as the "trimmed" flow. It will serve as the paid premium flow, the flow pattern for commissions and as the net premium flow.

4.3. Policyholder Dividends

Historically, policyholder dividends have played an important part in a healthy workers compensation insurance market. Dividend plans have provided a means to reward those insureds with better experience. "Sliding scale" dividend plans, in which the payment of the dividend depends on the insured's loss ratio, have provided important incentives for safety and loss control.

[21] It is the discounted value of the flows that affects the underwriting profit provision calculated by the model.

Historically, substantial dividend payments have been made to Massachusetts workers compensation policyholders.

It is expected that companies will continue to pay dividends to policyholders to maintain their competitive position, particularly if the rates are adequate, as they are intended to be. Therefore, these anticipated policyholder dividend payments have been reflected in the cash flows used in the profit model, in the same way as all other flows are recognized. If these policyholder dividend flows were not recognized, imaginary investment income would be imputed to companies on funds they do not hold.

The payment of policyholder dividends has been estimated to occur at the seventh quarter. The magnitude of dividend payments is calculated in Exhibit 7 from the Massachusetts ratios of policyholder dividends to the earned premium from the previous year.

Since the proposed expenses and premium discounts elsewhere in the ratemaking process are based on all companies, the estimate of the level of policyholder dividends is based on the most recently available 11-year average ratio of dividends to net earned premiums for Massachusetts workers compensation for all companies.[22]

4.4. Expense Flows

The expense flows were derived using a weighted average of separately determined flows for commissions, premium and other taxes, general expenses, other acquisition expenses, allocated loss adjustment expenses, and unallocated loss adjustment expenses.

The magnitude of each of these flows is determined by the corresponding expense provisions determined elsewhere in the ratemaking process. The pattern of each of these flows is determined as described below.

[22]To the extent that in many of these years the rates were inadequate, this procedure may underestimate dividends expected in a healthy market with an adequate rate level.

In order to run the Myers–Cohn profit model, expenses that vary with premium are aggregated into one flow while expenses that do not vary with premiums are aggregated into another flow. Each of these combined expense flows is a weighted average that reflects the relative expense provisions in this filing. The weights are shown in Exhibit 9.

The resulting expense flows used in the profit model are displayed in Exhibit 8.

A study of general expense flow patterns was performed by the Rating Bureau and were reported in the July 13, 1977 filing for Massachusetts workers compensation rates. Briefly, general expenses were divided into general administration, audit, inspection, and Bureau expenses. A time line was constructed to indicate a particular type of expenditure's distance from the effective date of a typical policy. Expenses by cost center, including home and field office expenses, were analyzed to establish those patterns of expenditures relative to the effective date of the policy. The combination of all such expense patterns resulted in the overall general expense pattern listed in Exhibit 10.[23]

The distribution of other acquisition costs was estimated from the same time pattern study that was used for general expense. Marketing field offices and services, billing and collection, policy issuance, and advertising expenses were examined for their occurrence relative to the issuance of a policy. The combination of all such expense patterns resulted in the other acquisition expense pattern listed in Exhibit 10.

Premium taxes are estimated and paid quarterly based upon a flat percentage of a flat amount (the previous year's net written premium). An adjustment is made to this estimation process in the first quarter of the following year. For purposes of estimating the expense flows we assume that adjustment will be zero. Based on these statutory provisions, the premium tax liability for any

[23]The profit provision is insensitive to the precise timing of general expenses and other acquisition expenses.

individual policy is assumed to be incurred and paid as the policy is written. The other tax payment pattern was estimated from the same time pattern study that was used for general expense.

The commission flow pattern is assumed to coincide with the paid premium flow.[24]

The loss adjustment expense (LAE) flow patterns, both allocated and unallocated, are based on the loss flow. The allocated LAE flow is assumed to have the same pattern as the loss flow. This corresponds to an assumption that on average the allocated LAE payments occur at approximately the same time as claim payments.

The pattern of the unallocated LAE flow is assumed to be the same as the straight average of the loss flow and an earned premium flow. This corresponds to an assumption that on average half of the unallocated LAE payments are made as accidents occur over the course of the policy effective period and that the other half of the unallocated LAE payments are made as claims are paid.

The weights used to combine the various expense flow patterns into final expense flows are calculated using the expense provisions used elsewhere in the ratemaking process. Since the premium flow is constructed net of premium discounts, it is necessary to calculate the proportions of expenses to net premiums. The acquisition expense and premium taxes are treated as varying in proportion to net premium. Loss adjustment expense is treated as varying in proportion to losses. General expense and other taxes are assumed not to vary with premium levels. These shall be referred to as fixed expenses.

Since the underwriting profit provision is one factor that determines the premium, and since losses, loss adjustment expense, and fixed expenses are all treated as not varying with premium

[24]The commission flow is the same as the flow of premium payments (the trimmed flow). See Exhibit 6.

levels, their fraction of premiums depends on the underwriting profit provision. Thus, their weight relative to those items that vary in proportion to premiums will change as the profit provision does. However, their weight relative to each other will not change. Thus, it is important that the weights for loss adjustment expense and fixed expenses be in the proper ratio to losses. A set of such initial weights is calculated in Exhibit 9.

The profit model is run through several iterations until the weights and profit provision converge to their final values.[25] At each iteration the weights assigned to losses, loss adjustment expense, fixed expenses, and variable expenses are adjusted for the profit provision. These new weights are then used to calculate another profit provision which in turn leads to another set of weights. The final weights are shown in Exhibit 9.

4.5. Loss Cash Flow

A medical and an indemnity loss cash flow have been estimated from the most recent available Financial Aggregate data.[26] The combined loss flow used in the profit model reflects a weighted average of the medical and indemnity flows and is shown in Exhibit 11.[27]

As shown in Exhibit 12, the flow for medical losses is based on the paid losses combined with an estimate of ultimate medical losses for each accident year taken from elsewhere in the ratemaking process. The percent of these ultimate losses paid in each year is computed. (See Exhibit 12, Part 2.) The increment between reports for each accident year is then computed (see Exhibit 12, Part 3), and the latest three-year average has been calculated for each reporting interval until the 17th report.[28] Beyond that report, the selected percentage of paid to ultimate loss

[25]Usually convergence takes 3 or 4 iterations.

[26]These are the same data relied upon elsewhere in the ratemaking process in order to estimate ultimate losses.

[27]The loss flow in Exhibit 11 sums to 1,000 solely for convenience.

[28]A two year average was calculated for the 17th report.

has been extended judgmentally.[29] The resulting medical loss flow is shown in Part 4 of Exhibit 12.

The indemnity flow is estimated in a similar manner in Exhibit 13. The indemnity and medical flows are then weighted together using an estimate of the percentage of total losses represented by each type,[30] in order to get the loss flow shown in Exhibit 11.

4.6. *Determination of the Risk-Free Rate of Return*

The risk-free rate of return is calculated as an average, weighted by annual net cash flows of duration-matched Treasury yields. This calculation is displayed in Exhibit 14.[31] The yields are taken from Part 4 of Exhibit 14 and are calculated from the observed yields over the most recent 12 months for the different maturities of Treasury securities.[32] The weights are taken from Part 5 of Exhibit 14 and reflect the length of time between receipt of premiums and payment of losses and expenses for Massachusetts workers compensation estimated in prior exhibits.

4.7. *Federal Income Tax Rate*[33]

For the federal income tax rate on investments, the corporate 35% tax rate currently applicable to Treasury securities has been used. This corresponds to the so-called "statutory/regulatory company" assumption, which is used with the Myers–Cohn Model.

---

[29] The manner in which this is done has no significant impact on the net present value of the flow or the resulting profit provision.

[30] This estimate is taken from elsewhere in the ratemaking process. For this illustration the indemnity flow has been weighted 68%, while the medical flow has been weighted 32%.

[31] The risk-free rate resulting from this calculation (as applied to Massachusetts workers compensation) is approximately the yield available on seven-year treasury bonds. A more elaborate method of duration matching could be employed if desired. Any uncertainty in the timing of the cashflows (as well as their magnitude) is not taken into account here, but should be incorporated in the selection of the risk-adjustment.

[32] The yield on Treasury securities usually increases as the term increases.

[33] See Almagro and Ghezzi [5] for a discussion of federal income tax provisions affecting property/casualty insurers.

The model company is assumed to invest in risk-free U.S. government securities matched to the length of the expected cashflow. The statutory/regulatory company assumption was adopted by former Commissioner of Insurance Stone in 1976, and was used by both the Rating Bureau and former Commissioner of Insurance Sabbagh in the original implementations of the Myers–Cohn profit model. The assumption that the model insurer's entire portfolio is invested in taxable government securities has several important implications:

- The investment income to be imputed to insurers is to be determined by matching the maturities of taxable government securities with the investment cash flow.
- The tax rate to be applied to determine the after-tax investment income should be the tax rate applicable to taxable government securities.
- No adjustment for investment risk needs to be made, because the investment is "risk-free."
- A smaller allowance for investment expenses is appropriate, because such a model insurer would have smaller investment expenses than would an insurer investing in a variety of other assets.[34]
- No adjustment need be made to take into account the Alternate Minimum Tax.

In the author's opinion, this assumption has a number of advantages.[35] The assumption makes the measure of investment income relatively stable and predictable; it establishes an investment standard that real world companies can meet; and it insulates the policyholders from the fluctuations in the stock and bond market to which they might be exposed if an actual portfolio model were used. Using the statutory/regulatory assump-

[34]No provision for investment expenses has been included.

[35]At least when used by a Rating Bureau. Different considerations would apply in other situations.

tion, current purchasers of insurance are neither penalized by, nor credited with, past investment decisions of the insurer. Rather we assume that the insurer will invest the fresh funds supplied by the premium of the insured at the currently available rates of return. The policyholder is thus credited with investment income at the risk-free rate. The policyholder shares neither the risk nor the reward of any more risky investment strategy.[36]

A 35% tax rate for underwriting income (or losses) has also been used. Underwriting credits will be available at 35%, because this model insurer has investment income taxed at 35% which can be offset by an underwriting loss. Such a model insurer will also not be subject to the Alternate Minimum Tax. In any case, the investment income tax rate, investment strategy, investment return, reward for risk, etc., used in the profit model need to all be consistent.

How one might incorporate some assumed set of investment other than Treasury securities into the Myers–Cohn model has been a controversial subject from the model's inception. Other investments would have differing risk, return, and tax implications than Treasury securities. One requires a consistent set of inputs that properly takes into account all of the impacts on the operation of the model company, including its required rate of return.

While some calculations of profit provisions using the Myers–Cohn model assuming other investments have been presented, I am not convinced that the resulting profit provisions are reasonable. In my opinion, the structure of the Myers–Cohn model without an explicit rate of return on equity makes it very difficult to properly consider the impacts of investment choices on risk and the needed profit provision. In any case, this subject is beyond the scope of this paper.[37]

---

[36]Consistent with the model company, there is no loading for investment expenses. So the policyholder is not being asked to share the cost of any investment strategy.

[37]See for example, Derrig [6], which discusses the "Myers Theorem," which states that the present value of the tax on investment income does not depend on the risk of the securities held by the insurance company.

4.8. *Risk-Adjusted Rate of Return*

An input to the Myers–Cohn model is the risk-adjusted rate of return. In this implementation of the model, as well as the paper by Myers and Cohn, the risk-adjusted rate (r_L) is set equal to the risk-free rate (r_f) plus the product of the negative liability beta (β_L) and the long term market risk premium (M). As calculated in Exhibit 15:

$$r_L = r_f + \beta_L M.$$

Exhibit 16 displays the estimation of the market risk premium. For each available year the total return on large company stocks has subtracted from it the return on U.S. Treasury Bills.[38] Then, per the recommendation of Ibbotson Associates [7], the long term (unweighted arithmetic) average of these differences is taken as the estimate of the market risk premium.

The yearly points that form the basis for this average are an extremely volatile data series.[39] For example, Figure 1 shows the yearly points while Figure 2 shows the ten year moving average, which is still fairly volatile.

Figure 3 shows the average of the series starting in various years since 1926 through the present. Depending on when one starts, the average can range from about 6% to about 11%.

Thus, the years of data relied upon can have a significant impact on the estimated market risk premium. The use of a long term average is consistent with an assumption of a stable or relatively stable expected value over time, which Ibbotson believes is the case. Mahler [8] briefly discusses the insensitivity of the estimate to somewhat different weights rather than the long term unweighted average, provided one assumes a relatively slow rate of shifting parameters over time.

In any case, a value of the market risk premium between about 8.5% and 9% seems to be regarded as reasonable. There is noth-

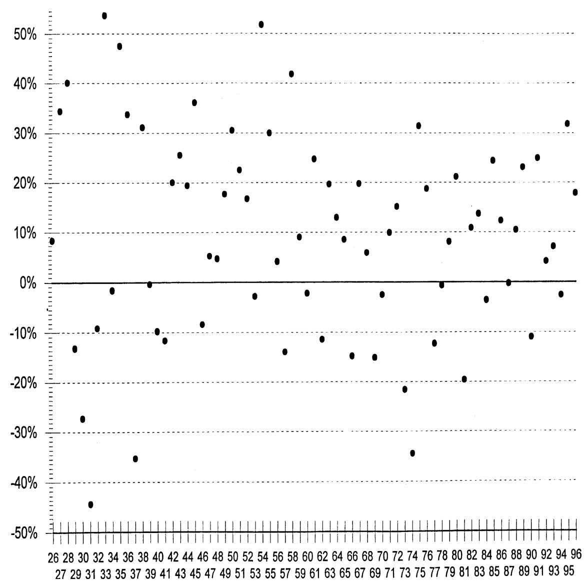
---

[38] See Ibbotson [7].

[39] While the mean is between 8% and 9%, the standard deviation is about 21%.

FIGURE 1

DIFFERENCE IN TOTAL RETURN ON LARGE COMPANY STOCKS AND U.S. TREASURY BILLS



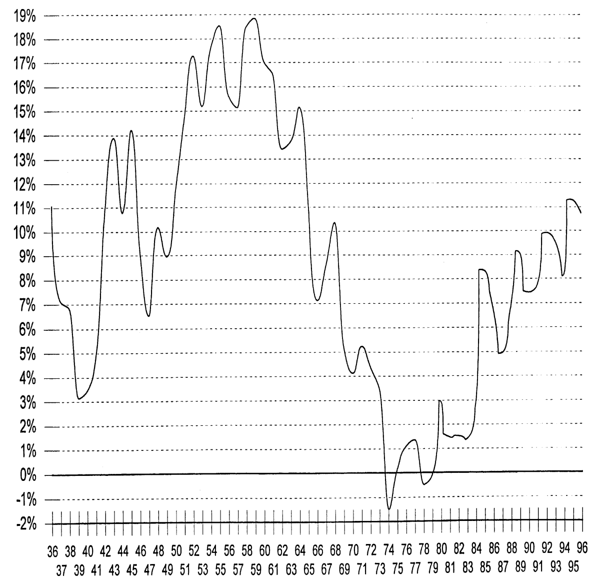
ing insurance specific about this value. This contrasts with the beta of liabilities which is insurance specific and for which there is no method of estimation generally regarded as reliable.

The beta of liabilities is intended to measure the covariance of insurance underwriting (as opposed to investing) with the stock market.[40] When combined with the market risk premium, it is intended to reward the insurer for the risk of underwriting insurance. Provided the beta of liabilities is negative, the risk-adjusted rate is smaller than the risk-free rate. Discounting the risky loss and expense flows at this smaller risk-adjusted rate results in a larger indicated premium than if these flows were discounted at

[40]This is based on the Capital Asset Pricing Model (CAPM).

FIGURE 2

DIFFERENCE IN TOTAL RETURN ON LARGE COMPANY STOCKS AND U.S. TREASURY BILLS
TEN YEAR MOVING AVERAGE OF THE SERIES



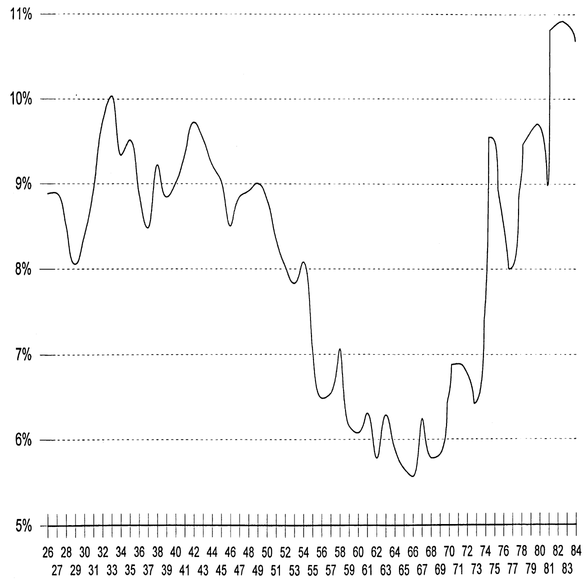
the risk-free rate. For this calculation a value of −0.21 used by the Massachusetts Commissioner of Insurance has been used for the beta of liabilities.

Unfortunately, as concluded by Kozik [9], "reliable estimates of the underwriting beta do not exist."[41] Thus, this is a major potential weakness of the Myers–Cohn model. Some technique must be employed to select or estimate the appropriate risk-adjustment. (The Capital Asset Pricing Model is the only

[41]As stated by Kozik, "Perhaps better methods of estimation may some day be developed." The discussion by Feldblum [10] is even more negative towards the whole idea of even considering something like a beta of liabilities.

FIGURE 3

DIFFERENCE IN TOTAL RETURN ON LARGE COMPANY STOCKS AND U.S. TREASURY BILLS
AVERAGE OF SERIES FROM THE GIVEN YEAR THROUGH 1996



technique the author has seen used for this piece of the Myers–Cohn model.) However, this is the same basic difficulty that one encounters in the use of other profit models that require the selection or estimation of the target rate of return or target internal rate of return. So while this presents a serious difficulty with the use of the Myers–Cohn model, it should be weighed against the similar difficulties in the use of other profit models.[42]

It should also be noted that since the reward for risk is based on using a risk-adjusted rate, the Myers–Cohn model would provide little risk return for a line of insurance that had a very quick

[42]This paper, in describing one profit model, is neither advocating for or against its use compared to some other profit model.

payment of losses such as hurricane insurance, unless one used a very large per-period risk-adjustment. Since it is very difficult to select an overall risk-adjustment to use on average for insurance, it would be extremely difficult, if not impossible, to come up with a risk-adjustment by line of insurance.

Exhibit 21, Part 1 displays the calculation of the "risk load." Let Z be the expected risk loading. Then $Z = 1 - (P^*/P)$, where P is the premium using a risk-adjusted discount rate while P^* is the premium (for the same cashflows) calculated with only the risk-free discount rate ($\beta_L = 0$). P^* is less than P, and the risk loading Z is positive.[43]

While in this illustrative calculation no specific use is made of Z,[44] it does quantify the effect of the risk-adjustment in the Myers–Cohn model.

4.9. Surplus

Initially surplus is assumed to be one-half of premiums and is assumed to decline in proportion to outstanding liabilities.[45] Thus, the surplus allocated to this policy or policy cohort is assumed to decline in proportion to the losses and expenses paid, as shown in Exhibit 17.

It should be noted that the Myers–Cohn model, as in the case with most profit models, is able to accommodate any magnitude or pattern of surplus flow selected by the user. However, for purposes of running the model, one does have to allocate surplus.

All of an insurer's surplus is in theory available to back up each policy, so in that sense one cannot allocate surplus to pol-

[43] For Massachusetts workers compensation, for $\beta_L = -0.21$, $Z \approx 5\%$ or 6% of premiums.

[44] As explained in the Appendix, in recent rate filings, Z has been used in a technical refinement that alters the investment balance for taxes.

[45] As explained in Mahler [2], the premium-to-surplus ratio one would observe for a given calendar year differs from the initial premium-to-surplus ratio selected here. Given the timing and magnitude of surplus flow selected here, one could compute what calendar year premium-to-surplus ratio would be observed for an assumed growth rate in premiums.

icy cohort, to line, or to state. On the other hand, each year, line, and state is expected to help contribute to the profitability of the insurer. The allocation of surplus for purposes of running the profit model allows one to allocate the needed *return* on surplus[46] among the different lines and states.[47] In that sense it is analogous to allocating by line and state certain expenses that have no direct relationship to any particular line or state. Expense allocation in the ratemaking process allows the collection of dollars needed to pay for these expenses. Similarly, surplus allocation allows the collection of dollars needed to achieve the desired return on surplus on an expected basis.

4.10. Construction of the Investment Balance for Tax Flow

The Investment Balance for Taxes is shown in Exhibit 18, Part 1.

The investment balance for any quarter is calculated as the sum of two components: assets available from the policy cash flow and those available from shareholders' equity. These two components are quantified each quarter as:

1. cumulative premiums minus cumulative losses, expenses, and dividends[48] (see Exhibit 18, Part 2) and

---

[46]While we have used the term "surplus" as per Myers and Cohn, a better term would be "equity." The concepts of surplus and equity are closely related but not identical. Surplus generally refers to statutory surplus while equity refers to economic net worth.

[47]Bingham [11], [12], and [13] discusses how one insurer uses allocation methods to measure returns and set targets by line of insurance. It is necessary to assign "benchmark" surplus to each line of insurance in order to apply the methodology used by Bingham. Bingham in his papers as well as Bender [14] discuss the relationship of risk, return, and required surplus. These issues apply when using the Myers–Cohn or most other profit models.

[48]It is important to note that in this computation the total premium equals the total losses, expenses, and dividends. In other words, this computation is performed using a profit provision of zero. This produces the appropriate estimate of investment income excluding any underwriting income (or loss). Mahler [2, Appendix VI] shows that the method used in Myers–Cohn to compute the investment income tax corresponds to a particular set of assumptions on the timing of income that is used in the model in that paper. Under these assumptions, the ratio of the present value of the income on the cashflows to the income on the cashflows is equal to the ratio of the present value of the outflows to the outflows.

2. surplus[49] (See Exhibit 17).

The tax on investment income can then be quantified by advancing the investment balance by one quarter (to the quarter in which the income is earned) and applying the quarterly investment rate and income tax rate.

4.11. Underwriting Tax Flows

The underwriting tax flows are shown in Exhibit 19, Part 1.

Premiums are earned equally throughout the year of the policy. This results in the premium portion of the underwriting tax flow shown in Exhibit 19, Part 1. This flow will be discounted to get κ_4 as shown in Exhibit 21.

The loss plus expense and dividend portion of the underwriting tax flow is shown in Exhibit 19, Part 1. This flow will be discounted to get κ_5 as shown in Exhibit 21.

The contribution of expenses (other than loss adjustment expense) and dividends to the underwriting tax flow is determined in Part 2 of Exhibit 19 by summing expenses and dividends paid in each quarter.

The contribution of losses and loss adjustment expense to this underwriting tax flow is determined in Part 3 of Exhibit 19, based on the reserve discount factors calculated in Exhibit 20. This follows the Tax Reform Act of 1986, which required insurers to discount loss reserves for tax purposes and specified how this was to be done.

The incurred loss plus LAE calculated in Part 3 of Exhibit 19 can be thought of as the sum of two pieces.[50] The first piece is the difference between the amount paid in a year and the discounted reserve previously held for those losses. The latter amount is

[49]As used in the Myers–Cohn model, "surplus" actually refers to shareholder equity rather than statutory surplus.
[50]Column 8 = Column 6 + Column 7, in Part 3 of Exhibit 19.

the paid losses in that year[51] times the appropriate reserve discount factor. Thus, the difference is the losses paid times unity minus the reserve discount factor.[52] The second piece is the change in discounted reserves on subsequent years. This is the product of the losses paid in subsequent years and the difference in reserve discount factors.[53]

The reserve discount factors are calculated in Exhibit 20 using a rolling sixty-month average of the mid-term "Applicable Federal Rate" (AFR) effective as of the beginning of each calendar month, and the reserve loss flow for workers compensation prescribed by the Internal Revenue Service.[54]

4.12. Discounted Flows

Each flow in the Myers–Cohn profit model has to be discounted at the appropriate risk-free or risk-adjusted rate. The risk-free discount rate is determined in Exhibit 14. The premium flow, the investment balance for tax, and the underwriting tax premium flow are discounted at the risk-free rate.

The risk-adjusted rate is determined in Exhibit 15. Discount factors based on the risk-adjusted rate are applied then to the total loss and expense flow and the underwriting tax loss flow.

Exhibit 21 shows the resulting values of the kappas. Also shown is the expected compensation to shareholders, i.e., the risk premium. This compensation for taking the risk of writing insurance is computed as unity minus the ratio of the premium

---

[51]For modeling purposes, the reserves and loss payments are assumed to be based on the same expected value. Also, the reserve discount factors are applied as if year one of the policy flow were accident year one in the Annual Statement, etc. This is only true for policies written January 1. This simplification has no significant impact on the calculation of the underwriting profit provision.

[52]Column 6 = 1 – Column 2 $\times$ Column 4, in Part 3 of Exhibit 19.

[53]Column 7 = Column 3 $\times$ Column 5, in Part 3 of Exhibit 19.

[54]For additional details, see Almagro and Ghezzi. [5, pp. 144, 145]. The reserve loss flow is updated by the IRS once every five years.

calculated using a risk-free discount rate ($\beta = 0$) and the premium calculated using a risk-adjusted rate.[55]

4.13. *Revenue Offset Provision of the Tax Reform Act of 1986*

Exhibit 22 contains the calculation of the unearned premium reserve "alpha" factor. When multiplied by κ_6 and the Federal Income Tax rate on underwriting, the alpha factor incorporates into the profit model the "revenue offset" provision of the 1986 Tax Reform Act.

This provision is explained in Almagro and Ghezzi [5]:

> Statutory income includes the change in unearned premium reserve during the tax year as a deduction. Insurers' acquisition expenses, however, are generally incurred and deducted near the time premiums are collected. Therefore, the statutory calculation does not accurately match recognition of premium income with recognition of related expenses.
>
> To approximately adjust for this mismatch, the IRS allows only 80% of the change in unearned premium reserve as a deduction. The limitation of the deduction is accomplished through an adjustment to statutory income, referred to as "revenue offset," whereby 20% of the unearned premium reserve change is added to statutory income for tax purposes.

This can be usefully thought of as accelerating the taxation of 20% of the premium income from a policy. Prior to this change, premium would be taxed as earned. Now 20% of premium is taxed as it is written or more precisely as an unearned premium reserve is set up. Then when the unearned premium reserve is

---

[55] This risk premium is calculated for informational purposes. While it is not used in the calculation of the underwriting profit provision, it is implicitly part of the profit provision calculated by the Myers–Cohn model. The value of the risk premium depends on the inputs chosen, most importantly the beta of liabilities, the market risk premium, and the timing of the cashflows.

taken down, 20% of the reduction in unearned premium reserve balances 20% of premium being earned at the same time. Thus, the timing of the reflection of this premium income has been moved from when it is earned to when the unearned premium reserve is set up.

Alpha is calculated in Exhibit 22 as 20% times the ratio of unearned premium reserves to premium times four times the quarterly risk-free rate.[56] κ_6 is calculated in Exhibit 23 based on the timing of the unearned premium reserves illustrated in the following example.[57] Table 1 shows how to specify the timing of the tax flows (due to the revenue offset) resulting from writing a new policy.

Assume $1,000 in written premiums and $120 in unearned premium reserves.[58] This 12.0% ratio of unearned premium reserves to premium approximates the current figure for workers compensation.

Continuing this example, let us assume a risk-free rate of 6% for illustrative purposes.[59] At 6%, the present value of the income tax due to the revenue offset is 0.4784.

Let κ_6 be the present value at 6% of a vector starting in Quarter 1[60] with the assumed pattern of unearned premium reserves:

$$\frac{180}{480}, \frac{140}{480}, \frac{100}{480}, \frac{60}{480}.$$

Then $\kappa_6 = 0.9702$.

---

[56]Thus alpha is approximately 20% of the unearned premium reserves times the annual risk-free rate.

[57]The final profit provision is insensitive to the particular choice of timing made.

[58]The sum over quarters is $4 \times \$120 = \$480$; this is $120 in unearned premium reserves on an annual basis.

[59]The actual calculation of κ_6 and alpha used in the calculation of the profit provision use the risk-free rate determined in Exhibit 14. The 6% value has been selected solely for illustrative purposes.

[60]The vector starts in Quarter 1 rather than Quarter 0 as per the unearned premium reserve. Advancing one quarter adds a factor of $(1 + r)^{-1}$ which is required in order to match the present value of the income tax due to the revenue offset.

TABLE 1

Quarter	Unearned Premium Reserve	Change in Unearned Premium Reserve	Income Tax Due to Revenue Offset[61]
0	180	180	12.60
1	140	−40	−2.80
2	100	−40	−2.80
3	60	−40	−2.80
4	0	−60	−4.20
Total	**480**	**0**	**0**

It is the case that:

$$0.4784 = 480 \times 35\% \times 20\% \times 0.01467 \times \kappa_6$$

where:

480 is the unearned premium reserve (UPR),

35% is the federal income tax rate on underwriting (FITU),

20% is the revenue offset factor, and

0.01467 is the quarterly risk-free rate of return (assuming a 6% annual rate).

Thus, the present value of the income tax due to the revenue offset is

$$\begin{aligned} & UPR \times FITU \times 20\% \times r \times \kappa_6 \\ &= 4 \times P \times UPRR \times FITU \times 20\% \times r \times \kappa_6 \\ &= P \times FITU \times \alpha \times \kappa_6, \end{aligned}$$

where

$$\alpha = 4 \times UPRR \times r \times 20,$$

$UPRR$ = unearned premium reserve ratio (to premiums),

r = quarterly risk-free rate, and

P = written premium.

[61] Change in unearned premium reserve × 35% × 20%.

This is the formula for alpha that is used in Exhibit 22. The impact of the revenue offset enters into the Myers–Cohn profit calculation via the term $FITU \times \alpha \times \kappa_6$, as shown in Exhibit 5.

5. SENSITIVITY ANALYSIS

For the practical example described in the previous section, the inputs combine to produce a model underwriting profit provision of -3.6%, as shown in Exhibit 5. As with any model, the result depends on both the structure of the model and particular inputs chosen.

Exhibit 24 shows the sensitivity of the Myers–Cohn model to the choice of different inputs. While individual inputs are varied one at a time for illustrative purposes, it is important to choose a consistent *set* of inputs for use in the profit model.

The risk-free rate of return can vary by several percentage points from one year to the next. Generally, in Massachusetts an average of the last year's rates available on a duration-matched portfolio of Treasury securities has been used to estimate the risk-free rate. For long-tailed lines like workers compensation, the profit provision is very sensitive to changes in interest rates. The higher the risk-free rate of return, the more investment income that can be earned, and therefore, the less premium is needed. Thus, all other things being equal, a higher risk-free rate of return corresponds to a more negative underwriting profit provision.

The more negative the beta of liabilities, the more positive the underwriting profit provision. If one assumed that underwriting was risk-free (beta of liabilities equal to zero), there would be a more negative profit provision. The difference between this profit provision and the calculated profit provision represents the reward for taking the risk of writing insurance. In recent workers compensation filings this risk premium has been about 5% or 6%. More generally, the further the risk-adjusted rate[62] is

[62]Whether one uses the CAPM or some other method to determine to the risk-adjusted rate.

from the risk-free rate the larger the risk load and the higher the profit provision.

Since the market risk premium and the beta of liabilities enter into the calculation only as a product,[63] their effect on the profit provision is similar. For a larger (magnitude) market risk premium, the profit provision is less negative, since the risk-adjustment is larger. The same effect is seen as for a similar increase in the magnitude of the beta of liabilities.

The investment income tax rate and premium-to-surplus ratio are other important and sometimes controversial inputs. The higher the assumed investment income tax rate, the more positive the profit provision. The insurer earns less investment income after taxes and thus needs more income from underwriting.

The higher the premium-to-surplus ratio, the more negative the profit provision. The more leveraged the insurance operation, the more important investment income considerations become. It should be noted that in this implementation of the Myers–Cohn model, the beta of liabilities is assumed to be independent of the premium-to-surplus ratio.

The sensitivity of the underwriting income tax rate depends on the profit provision. For profit provisions near zero, there is little underwriting income assumed and therefore little sensitivity to the tax rate. For substantially negative profit provisions, there is an assumed underwriting loss which is assumed to generate a credit against other taxable income. Thus, the higher the assumed underwriting tax rate, the more valuable is this tax credit. Therefore, the higher the underwriting tax rate the more negative the profit provision. The situation is reversed for a substantially positive underwriting profit provision. All other things being equal, the higher the underwriting tax rate, the further the underwriting profit provision is from zero. (A negative provision becomes

[63] The risk-adjusted rate is equal to the risk-free rate plus the product of the market risk premium and the beta of liabilities.

more negative, while a positive provision becomes more positive.)

It should be noted that in this example, the investments are assumed to be solely in Treasury securities taxed at the marginal corporate rate. Therefore, if the underwriting income tax rate were to change, one would also change the investment income tax rate. For example, in 1987 the marginal corporate tax rate declined from 46% to 34%. For a long-tailed line of insurance such as workers compensation, such a decline in *both* tax rates in the Myers–Cohn model would lead to a more negative profit provision. This is an example of why varying the inputs one at a time can only be for illustrative purposes.

The target underwriting profit provision calculated here includes the effect on investment income of the payment of expected policyholder dividends. In this case, policyholder dividends are paid out earlier than the average payment of losses plus expenses. Thus, dividend payments reduce expected investment income compared to the average payment for losses plus expenses. Therefore, the more that is assumed to be paid out in policyholder dividends (compared to losses and expenses) the more positive the underwriting profit provision.

The Tax Reform Act of 1986 introduced the discounting of loss reserves for tax purposes and the revenue offset feature. As expected, since each of these changes was intended to produce more taxes for the federal government, they each lead to a less negative underwriting profit provision. Insurers need more money to pay these taxes, all other things being equal.

Finally, the average timing of the loss payments is an extremely important input. The longer it takes to pay losses the more negative the profit provision. Investment income considerations are generally more important for long-tailed lines of insurance.

The risk-free rate, the size of the adjustment for risk, the investment income tax rate, the premium-to-surplus ratio, and the timing of the loss flow are usually the inputs to the Myers–Cohn model with the most significant impact on the underwriting profit provision. Of these, the investment income tax rate and the size of the adjustment for risk[64] have been the most intensely debated at rate hearings.

6. CONCLUSION

In Massachusetts, the Myers–Cohn model has been used to set many profit provisions over the last decade. As with any profit model, in any real world application, one must carefully examine the underlying assumptions and inputs to make sure that everything is consistent. It has proven very easy for two people to get extremely different profit provisions using the same model.[65]

The last two decades have demonstrated the impossibility of coming up with either a universally accepted profit model or profit provision. However, the possibility of differing answers no more makes profit models useless than would the inability to agree on future loss levels make trending and loss development techniques useless. Profit models provide a framework for a rational discussion and allow the testing of the effect of changes to the tax law, investment policy, claims payment patterns, economic conditions, etc. The Myers–Cohn model provides one framework in which to attempt to quantify these effects.

[64]In the CAPM implementation, the adjustment for risk is the product of the beta of liabilities and the market risk premium.

[65]Disagreements about the risk-free rate, the risk-adjusted rate, the investment income tax rate, the amount of surplus, etc., can quickly add up to a substantial disagreement on the overall profit provision. Even when using the same profit model for workers compensation insurance, disagreements of 10% or more in proposed profit provisions are not unheard of. These disagreements parallel those that can occur at contested rate hearings with respect to the indicated rate change, where expert witnesses can have very significant disagreements with respect to loss development, trend, law impacts, etc.

REFERENCES

[1] Cummins, J. David and Scott E. Harrington, eds., *Fair Rate of Return in Property-Liability Insurance*, Hingham, MA, Kluwer–Nijhoff, 1986.

[2] Mahler, Howard C., "An Introduction to Underwriting Profit Models," *PCAS* LXXII, 1985, pp. 239–277.

[3] National Association of Insurance Commissioners. *Report of the Investment Income Task Force*, June 1984.

[4] Venter, Gary G., and William R. Gillam, "Simulating Serious Workers' Compensation Claims," *Reinsurance*, Casualty Actuarial Society Discussion Paper Program, 1986, pp. 226–258.

[5] Almagro, Manuel and Thomas L. Ghezzi, "Federal Income Taxes—Provisions Affecting Property/Casualty Insurers," *PCAS* LXXV, 1988, pp. 95–162.

[6] Derrig, Richard A., "Theoretical Considerations of the Effect of Federal Income Taxes on Investment Income in Property–Liability Ratemaking," *The Journal of Risk and Insurance*, 61, 4, December 1994, pp. 691–709.

[7] *Stocks, Bonds, Bills, and Inflation Yearbook*, Ibbotson Associates, Chicago, IL. (Published yearly.)

[8] Mahler, Howard C., "Credibility with Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty," *PCAS* LXXXX, 1998, pp. 455–653.

[9] Kozik, Thomas J., "Underwriting Betas—The Shadow of Ghosts," *PCAS* LXXXI, 1994, pp. 303–329.

[10] Feldblum, Sholom, Discussion of [8], *PCAS* LXXXIII, 1996, pp. 648–656.

[11] Bingham, Russell E., "Discounted Return—Measuring Profitability and Setting Targets," *PCAS* LXXVII, 1990, pp. 124–159.

[12] Bingham, Russell E., "Surplus—Concepts, Measures of Return, and Determination," *PCAS* LXXX, 1993, pp. 55–109.

[13] Bingham, Russell E., "Rate of Return—Policyholder, Company, and Shareholder Perspectives," *PCAS* LXXX, 1993, pp. 110–147.

[14] Bender, Robert K., Discussion of [12], *PCAS* LXXXIV, 1997, pp. 44–105.

EXHIBIT 1

MYERS–COHN PROFIT MODEL
EXAMPLE OF CALCULATION OF UNDERWRITING PROFIT PROVISION

Inputs

Risk-Free Rate = 9%
Beta of Liabilities = −.20
Market Risk Premium = 10%
Risk-Adjusted Rate = 9% − .20 × 10% = 7%
Premium-to-Surplus Ratio = 2
Federal Income Tax Rate on Underwriting = 35%
Federal Income Tax Rate on Investment = 35%
Expenses (other than loss adjustment expense) are all paid in Quarter 2.
Variable Expenses are 20% of Premium.
Fixed Expenses are 5% of Losses.
Loss Adjustment Expenses are 10% of Losses.
Premiums are collected in Quarter 1.
Losses and loss adjustment expense are paid in Quarter 5.
There are no Policyholder Dividends paid.
There is no discounting of reserves (for tax purposes).
There is no revenue offset provision; alpha = 0.

Kappas	Initial Weights	Final Weights	
κ_1	.9380	.9378	Risk-adjusted discounted losses and expenses factor.
κ_2	.9893	.9893	Risk-free discounted premiums factor.
κ_3	4.8935	4.9165	Risk-free discounted investment balance tax factor.
κ_4	.9478	.9478	Risk-free discounted underwriting profit tax factor.
κ_5	.9588	.9588	Risk-adjusted discounted underwriting profit tax factor.
κ_6	N.A.	N.A.	Risk-free discounted revenue offset tax factor.

Profit Provision

$$\frac{P}{L+E} = \frac{\kappa_1 - \tau_1\kappa_5}{\kappa_2 - \tau_2 r\kappa_3 - \tau_1\kappa_4 - \tau_1\alpha\kappa_6}$$

$$= \frac{.9378 - .35(.9588)}{.9893 - (.35 \times .021778 \times 4.9165) - (.35 \times .9478)}$$

$$= .9712$$

$$\mu = 1 - (P/(L+E))^{-1} = -3.0\%$$

EXHIBIT 2

EXAMPLE CASHFLOWS
(Initial Weights)

Quarter	Premiums	Losses	Expenses[1]	Cumulative Difference	Surplus[2]	Investment Balance[3]
0	0	0	0	0	250.00	250.00
1	1,000.00	0	0	1,000.00	500.00	1,500.00
2	0	0	234.78	765.22	382.61	1,147.83
3	0	0	0	765.22	382.61	1,147.83
4	0	0	0	765.22	382.61	1,147.83
5	0	695.65	69.57	0	0	0
Total	1,000.00	695.65	304.35			

[1]Expenses are the sum of $200 (20% of premium) representing variable expense in Quarter 2, 34.78 (5% of losses) representing fixed expense in Quarter 2, and 69.57 (10% of losses) representing LAE in Quarter 5. Note that for the initial weights, losses plus expenses = 1,000 = premiums.
[2]Initially, surplus is taken as half of premiums at policy inception. (This is approximated by having $250 of surplus flow in during Quarter 0, prior to policy inception and an additional $250 of surplus flow in during Quarter 1.) The suplus allocated to this policy is assumed to decline in proportion to the payment of losses and expenses.
[3]Investment Balance is the sum of the surplus and the cumulative difference of premiums and losses plus expenses.

The policy inception date is at the end of Quarter 0 and the beginning of Quarter 1.

EXHIBIT 3

Example Calculation of Kappas (Initial Weights)

κ_1 = Risk-adjusted discounted losses and expenses factor
$.76522 \times (1.07)^{-4.5/4} + .23478 \times (1.07)^{-1.5/4}$
= .9380

Note: Losses and loss adjustment expenses discounted to the middle of the fifth quarter. Expenses discounted to the middle of the second quarter.

κ_2 = Risk-free discounted premiums factor
= Discounted Value of Premium Flow
= .9893

Note: Discounting to the middle of the first quarter $.9893 = (1.09)^{-.5/4}$.

κ_3 = Risk-free discounted investment balance tax factor
= Discounted Investment Balance for Taxes
= $\{(250 \times .9893) + (1500 \times .9682) + (1147.83 \times .9476) + (1147.83 \times .9274) + (1147.83 \times .9076)\}/1000$
= 4.8935

κ_4 = Risk-free underwriting profit tax factor (contribution of premiums)
= $(.25 \times .9787) + (.25 \times .9578) + (.25 \times .9374) + (.25 \times .9174)$
= .9478

Note: Discounting to the end of the first, second, third, and fourth quarters.

κ_5 = Risk-adjusted discounted underwriting profit tax factor (contribution of losses and expenses)
= $(.25 \times .9832) + (.25 \times .9667) + (.25 \times .9505) + (.25 \times .9346)$
= .9588

Note: Discounting to the end of the first, second, third, and fourth quarters.

κ_6 = Not applicable since no revenue offset provision is assumed.

EXHIBIT 4

EXAMPLE CASHFLOWS
(Final Weights)[1]

Quarter	Premiums[2]	Losses	Expenses[3]	Cumulative Difference	Surplus	Investment Balance[4]
0	0	0	0	0	250.00	250.00
1	1,000.00	0	0	1,000.00	500.00	1,500.00
2	0	0	229.27	770.73	385.37	1,156.10
3	0	0	0	770.73	385.37	1,156.10
4	0	0	0	770.73	385.37	1,156.10
5	0	700.66	70.07	0	0	0
Total	1,000.00	700.66	299.34			

[1]As the profit provision varies so does the relative weight given to variable expenses, so that the profit model is solved via iteration.
[2]Premiums shown are prior to the profit loading. The premium loaded for profit is 971.18.
[3]Expenses are the sum of 194.24 (20% of premiums loaded for profit of 971.18) representing variable expense in Quarter 2, 35.03 (5% of losses) representing fixed expense in Quarter 2, and 70.07 (10% of losses) representing LAE in Quarter 5. Note that losses plus expenses = 1000.
[4]Investment Balance is the sum of the surplus and the cumulative difference of premiums and losses plus expenses.

The policy inception date is at the end of Quarter 0 and the beginning of Quarter 1.

EXHIBIT 5

PART 1

THE MYERS–COHN COST OF CAPITAL UNDERWRITING PROFIT PROVISION MODEL[1]

Let

Flows		Capital Market Rates	
P	= Premium	r	= Risk-Free Rate
L	= Losses	r_L	= Risk-Adjusted Rate (Adjusted for Risk of Underwriting by Line)
E	= Expenses	τ_1	= Federal Underwriting Income Tax Rate
IVB	= Investment Balance	τ_2	= Federal Investment Income Tax Rate
IVBT	= Investment Balance for Tax	μ	= Underwriting Profit Margin
UWP	= Underwriting Profit	α	= Revenue Offset Factor for Taxes

Then, given the basic valuation equations of the Myers–Cohn model,

Present Value of Premium	=	Present Value of Losses and Expenses plus Present Value of Federal Tax Liabilities on Underwriting Profits and Investment Income on the Investment Balance,

or

(1) $$\mathrm{PV}(P) = \mathrm{PV}(L+E) + \mathrm{PV}(\mathrm{UWP}\tau_1) + \mathrm{PV}(\mathrm{IVBT}r\tau_2)$$

where

> UWP is Underwriting Profit and IVBT is the Investment Balance for Taxes. The investment balance flow, IVB, is defined as the funds available for investment from the policy cash flow, cumulative premium minus cumulative losses, plus those funds available from other supporting assets. IVBT is IVB advanced one quarter to the time period when the income is earned and the tax liability is incurred.

Then, if premiums and investment income are valued at the risk-free rate r, losses and expenses valued at a risk-adjusted rate r_L; underwriting and investment income taxed at

[1] Chapter 3 of *Fair Rate of Return on Property-Liability Insurance* [1].

rates τ_1 and τ_2; and underwriting profits taxed using discounted loss reserves:

$$\text{(2)} \qquad \mathrm{PV}_r(\mathrm{P}) = \mathrm{PV}_{r_L}(L+E) + \mathrm{PV}_r(P\tau_1\mathrm{UWP}/(P-(L+E))) - \mathrm{PV}_{r_L}((L+E)\tau_1\mathrm{UWP}/(P-(L+E)) + \mathrm{PV}_r(r\tau_2(\mathrm{IVBT})).$$

The various discounted values can be rewritten in terms of the kappas defined below. Note that the term involving κ_6 relates to the revenue offset provision, which as explained in Section 4.13 adjusts the timing for income tax purposes of the premium portion of the underwriting profit.

$$P\kappa_2 = (L+E)\kappa_1 + P\tau_1\kappa_4 + P\alpha\tau_1\kappa_6 - (L+E)\tau_1\kappa_5 + Pr\tau_2\kappa_3$$

$$P(\kappa_2 - \tau_2 r\kappa_3 - \tau_1\kappa_4 - \alpha\tau_1\kappa_6) = (L+E)(\kappa_1 - \tau_1\kappa_5)$$

or

$$\text{(3)} \qquad \frac{P}{L+E} = \frac{\kappa_1 - \tau_1\kappa_5}{\kappa_2 - \tau_2 r\kappa_3 - \tau_1\kappa_4 - \alpha\tau_1\kappa_6}$$

and

$$\mu = 1 - (P/(L+E))^{-1}$$

where	κ_1 = Risk-adjusted discounted losses and expenses factor
	κ_2 = Risk-free discounted premiums factor
	κ_3 = Risk-free discounted investment balance tax factor
	κ_4 = Risk-free discounted underwriting profit tax factor
	κ_5 = Risk-adjusted discounted underwriting profit tax factor
	κ_6 = Risk-free discounted revenue offset tax factor

EXHIBIT 5

PART 2

MASSACHUSETTS WORKERS COMPENSATION

(1)		Model Profit Allowance (Part 4)	−3.6%
(2)		Average Premium Discount as a Percent of Standard Premium plus ARAP[2]	6.8%
(3)		Adjustment for Investment Expenses	0.0%
		Underwriting Profit Allowance	
(4)	=	$[(1) \times (1 - (2))] + (3)$	−3.4%

Parameters

1.	Cash Flows		
	a.	Premium	Exhibit 6
	b.	Expenses	Exhibit 8
	c.	Losses	Exhibit 11
	d.	Expense/Loss Weights	Exhibit 9
	e.	Policyholder Dividends	Exhibit 7
	f.	Surplus	Exhibit 17
	g.	Underwriting Tax Flow	Exhibit 19
2.	Capital Market Rates		
	a.	Risk-Free Rate	6.60%
	b.	Risk-Adjusted Rate (Beta = −.21, Market Risk Premium 8.9%)	4.73%
3.	Federal Income Tax Rates		
	a.	Underwriting	35%
	b.	Investment	35%
4.	Initial Premium/Surplus Ratio		2 to 1

[2]From elsewhere in the ratemaking process. ARAP (All Risk Adjustment Program) is applied in Massachusetts workers compensation as a surcharge on top of experience rating. The rate indication is calculated in terms of Standard Premium plus ARAP = payrolls × manual rates × experience modification × ARAP surcharge, if any.

EXHIBIT 5

Part 3

Massachusetts Workers Compensation
Calculation of Underwriting Profit Provisions Using Myers–Cohn Cost of Capital Model

$$\frac{P}{L+E} = \frac{\kappa_1 - \tau_1\kappa_5}{\kappa_2 - \tau_2 r\kappa_3 - \tau_1\kappa_4 - \tau_1\alpha\kappa_6}$$

$$\mu = 1 - (P/(L+E))^{-1}$$

$r = 0.016107$ $\quad r_L = 0.011623$ $\quad \tau_1 = 0.35$ $\quad \tau_2 = 0.35$

$\beta = -0.21$ $\quad r_M - r = 0.089$ $\quad \alpha = 0.00155$

Discounting Factors

$\kappa_1 = .856659$

$\kappa_2 = .962190$

$\kappa_3 = 14.558852$

$\kappa_4 = .960994$

$\kappa_5 = .949474$

$\kappa_6 = .967392$

$$\frac{P}{L+E} = \frac{0.856659 - 0.35(0.949474)}{0.962190 - 0.35(0.016107)(14.558852) - 0.35(0.960994) - 0.35(0.00155)(0.967392)}$$

$$= 0.965210$$

$$\mu = 1 - (.965210)^{-1} = -0.0360$$

Model Provision = −3.6%

EXHIBIT 6

Part 1

Massachusetts Workers Compensation Premium Flow

Quarter	Premium Flow
1	0.2397
2	0.2120
3	0.2355
4	0.1948
5	0.0462
6	0.0159
7	0.0271
8	0.0043
9	0.0060
10	0.0148
11	0.0043
12	0.0007
13	0.0001
14	−0.0007
15	0.0000
16	−0.0002
17	−0.0006
18	0.0000
19	0.0000
20	−0.0002
21	−0.0001
22	0.0004
Sum	1.0000

From Exhibit 6, Part 2, selected net premium flow.

EXHIBIT 6

PART 2

DETERMINATION OF SELECTED PREMIUM FLOW FROM PREMIUM CALL

Days From Effective Date		(1) Untrimmed Flow	(2) Trimmed Flow*	(3) Selected Net Premium
−89–	0	0.0082		
1–	90	0.2316	0.2397	0.2397
91–	180	0.2120	0.2120	0.2120
181–	270	0.2355	0.2355	0.2355
271–	360	0.1948	0.1948	0.1948
361–	450	0.0462	0.0462	0.0462
451–	540	0.0159	0.0159	0.0159
541–	630	0.0271	0.0271	0.0271
631–	720	0.0043	0.0043	0.0043
721–	810	0.0060	0.0060	0.0060
811–	900	0.0148	0.0148	0.0148
901–	990	0.0043	0.0043	0.0043
991–	1080	0.0007	0.0007	0.0007
1081–	1170	0.0001	0.0001	0.0001
1171–	1260	−0.0007	−0.0007	−0.0007
1261–	1350	0.0000	0.0000	0.0000
1351–	1440	−0.0002	−0.0002	−0.0002
1441–	1530	−0.0006	−0.0006	−0.0006
1531–	1620	0.0000	0.0000	0.0000
1621–	1710	0.0000	0.0000	0.0000
1711–	1800	−0.0002	−0.0002	−0.0002
1801–	1890	−0.0001	−0.0001	−0.0001
1891–	1980	0.0003	0.0004	0.0004
1981–	2070	0.0000		
2071–	2160	0.0000		
2161–	2250	0.0001		
2251–	2340	0.0000		
2341–	2430	0.0000		
2431–	2520	0.0000		
2521–	2610	0.0000		
2611–	2700	0.0000		
2701–	2790	0.0000		
2791–	2880	0.0000		
2881–	2970	0.0000		
2971–	3060	0.0000		
3061–	3150	0.0000		
3151–	3240	0.0000		
3241–	3330	0.0000		
3331–	3420	0.0000		
		1.0000	1.0000	1.0000

*The quarter preceding the effective date in (1) was biased toward the end of the quarter (average time from effective date = −6.5 days). Therefore, that percentage of premium from quarter zero was added into the first quarter. The combined first quarter in (2) has a resulting average effective date at 56 days.

EXHIBIT 7

MASSACHUSETTS WORKERS COMPENSATION RATIO OF POLICYHOLDER DIVIDENDS TO THE EARNED PREMIUM FROM THE PREVIOUS YEAR*

Massachusetts				Countrywide			
Year	Total Stock	Non-Stock	All Companies	Year	Total Stock	Non-Stock	All Companies
				85	9.72	15.32	11.34
86	6.36	12.04	7.85	86	8.26	13.09	9.79
87	5.05	9.57	6.24	87	7.41	11.11	8.70
88	3.55	6.43	4.32	88	6.52	10.17	7.88
89	2.70	2.37	2.60	89	6.30	8.55	7.08
90	2.62	2.10	2.44	90	5.66	8.88	6.74
91	1.76	1.81	1.77	91	4.82	7.69	5.78
92	1.19	2.19	1.57	92	3.91	6.72	4.81
93	1.04	2.30	1.52	93	4.19	6.44	4.96
94	0.87	1.79	1.21	94	5.25	9.02	6.54
95	1.14	2.44	1.57	95	4.72	9.48	6.38
96	1.84	3.47	2.25	96		Not Available	
Average	2.56	4.23	3.03	Average	6.07	9.68	7.27

*Computed using the data compiled from Annual Statements.

Policyholder Dividends Net Premium[1]	Policyholder Dividends Premium Tax Rate[2]	Net of Premium Tax
3.03%	2.30%	3.0%

Policyholder dividends are assumed on average to be paid in quarter 7.[3]

[1] Average for all insurers in Massachusetts.
[2] From elsewhere in the ratemaking process.
[3] This corresponds to 19.5 months on average from policy inception.

EXHIBIT 8

MASSACHUSETTS WORKERS COMPENSATION EXPENSE FLOWS

	Fixed Expense Flow[1]		Variable Expense Flow[2] Plus Dividend Flow			Fixed Expense Flow[1]		Variable Expense Flow[2] Plus Dividend Flow	
Quarter	Initial Weights	Final Weights	Initial Weights	Final Weights	Quarter	Initial Weights	Final Weights	Initial Weights	Final Weights
−3	0.690	0.694	0.580	0.560	36	0.374	0.376	0.000	0.000
−2	0.740	0.744	1.740	1.679	37	0.266	0.268	0.000	0.000
−1	10.300	10.355	3.480	3.359	38	0.266	0.268	0.000	0.000
0	16.170	16.256	12.760	12.316	39	0.266	0.268	0.000	0.000
1	27.883	28.032	43.925	42.397	40	0.266	0.268	0.000	0.000
2	20.093	20.200	11.392	10.996	41	0.118	0.118	0.000	0.000
3	19.403	19.506	12.301	11.872	42	0.118	0.118	0.000	0.000
4	19.403	19.506	10.225	9.869	43	0.118	0.118	0.000	0.000
5	13.850	13.924	2.936	2.834	44	0.118	0.118	0.000	0.000
6	7.740	7.782	0.811	0.783	45	0.223	0.224	0.000	0.000
7	7.050	7.088	31.382	30.290	46	0.223	0.224	0.000	0.000
8	7.050	7.088	0.219	0.212	47	0.223	0.224	0.000	0.000
9	4.018	4.039	0.306	0.295	48	0.223	0.224	0.000	0.000
10	4.018	4.039	0.755	0.729	49	0.230	0.231	0.000	0.000
11	4.018	4.039	0.219	0.212	50	0.230	0.231	0.000	0.000
12	4.018	4.039	0.036	0.034	51	0.230	0.231	0.000	0.000
13	2.274	2.286	0.005	0.005	52	0.230	0.231	0.000	0.000
14	2.274	2.286	−0.036	−0.034	53	0.262	0.263	0.000	0.000
15	2.274	2.286	0.000	0.000	54	0.262	0.263	0.000	0.000
16	2.274	2.286	−0.010	−0.010	55	0.262	0.263	0.000	0.000
17	1.361	1.368	−0.031	−0.030	56	0.262	0.263	0.000	0.000
18	1.361	1.368	0.000	0.000	57	0.197	0.198	0.000	0.000
19	1.361	1.368	0.000	0.000	58	0.197	0.198	0.000	0.000
20	1.361	1.368	−0.010	−0.010	59	0.197	0.198	0.000	0.000
21	0.928	0.933	−0.005	−0.005	60	0.197	0.198	0.000	0.000
22	0.928	0.933	0.020	0.020	61	0.248	0.249	0.000	0.000
23	0.928	0.933	0.000	0.000	62	0.248	0.249	0.000	0.000
24	0.928	0.933	0.000	0.000	63	0.248	0.249	0.000	0.000
25	0.632	0.636	0.000	0.000	64	0.248	0.249	0.000	0.000
26	0.632	0.636	0.000	0.000	65	0.171	0.171	0.000	0.000
27	0.632	0.636	0.000	0.000	66	0.171	0.171	0.000	0.000
28	0.632	0.636	0.000	0.000	67	0.171	0.171	0.000	0.000
29	0.452	0.455	0.000	0.000	68	0.171	0.171	0.000	0.000
30	0.452	0.455	0.000	0.000	69	0.278	0.279	0.000	0.000
31	0.452	0.455	0.000	0.000	70	0.278	0.279	0.000	0.000
32	0.452	0.455	0.000	0.000	71	0.278	0.279	0.000	0.000
33	0.374	0.376	0.000	0.000	72	0.278	0.279	0.000	0.000
34	0.374	0.376	0.000	0.000	73	0.169	0.170	0.000	0.000
35	0.374	0.376	0.000	0.000	74	0.169	0.170	0.000	0.000
					etc.[3]				

[1] General expense, other tax, allocated loss adjustment expense, and unallocated loss adjustment expense flows combined using the weights in Exhibit 9.
[2] Commissions, other acquisition expense and premium tax combined using the weights in Exhibit 9.
[3] Flow continues out to the same quarter as the loss flow.

EXHIBIT 9

PART 1

MASSACHUSETTS WORKERS COMPENSATION CASH FLOW WEIGHTS

	Initial Weights	Final Weights	
Item of Expense Allowance	Prem(%)[1]	Prem(%)	Loss + Exp +Div(%)
Premium (Net of Premium Discounts)	100.0%	100.00%	96.52%
Expected Losses	65.9%	68.64%	66.25%
Total Expenses Plus Dividends	34.1%	34.96%	33.75%
Fixed Expenses: (Total Expenses not varying with Premium)	20.8%	21.66%	20.91%
Loss Adjustment Expense[1]	13.9%	14.48%	13.97%
Allocated[2]	7.4%	7.71%	7.44%
Unallocated[2]	6.5%	6.77%	6.53%
General Expenses[3]	6.4%	6.67%	6.43%
Other Tax[3]	0.5%	0.52%	0.50%
Variable Expenses Plus Dividends: (Varying with Premium)	13.3%	13.30%	12.84%
Total Acquisition	8.0%	8.00%	7.72%
Commissions[3]	5.1%	5.10%	4.92%
Other Acquisition[3]	2.9%	2.90%	2.80%
Premium Tax	2.3%	2.30%	2.22%
Policyholder Dividends	3.0%	3.00%	2.90%

[1] From Part 2.
[2] The loss adjustment expense split between allocated and unallocated is 53.4% & 46.6% based on a two-year average of Annual Statement data for thirteen major writers in Massachusetts.
[3] Weighted based on calculations underlying other portions of the ratemaking process.

EXHIBIT 9

PART 2

MASSACHUSETTS WORKERS COMPENSATION DETERMINATION OF INITIAL CASH FLOW WEIGHTS

(1)	Acquisition and Field Supervision (as a Percent of Net Premium)	8.0%
(2)	Premium Taxes (as a Percent of Net Premium)	2.3%
(3)	Policyholder Dividends (as a Percent of Net Premium, Net of Taxes)	3.0%
(4)	Variable Expenses (excluding profit provision) plus Policyholder Dividends	13.3%
(5)	Loss, Loss Adjustment Expense Ratio, and Fixed Expense Ratio	86.7%
(6)	Loss Adjustment Expense as a Percent of Losses	21.0%
(7)	Ratio of Fixed Expense as a Percent of Losses	10.5%
(8)	Loss Ratio to Net of Premium Discount (if there were no loading for profits)	65.9%
(9)	Loss Adjustment Expense as a Percent of Premiums Net of Premium Discount	13.9%
(10)	Fixed Expenses as a Percent of Premiums Net of Premium Discount	6.9%
(11)	Expenses (excluding profit provision) plus Policyholder Dividends	34.1%

(1) From elsewhere in the ratemaking process.
(2) From elsewhere in the ratemaking process.
(3) From elsewhere in the ratemaking process.
(4) $= (1) + (2) + (3)$
(5) $= 1 - (4)$
(6) From elsewhere in the ratemaking process.
(7) From elsewhere in the ratemaking process.
(8) $= (5)/[1 + (6) + (7)]$
(9) $= (6) \times (8)$
(10) $= (7) \times (8)$
(11) $= (4) + (9) + (10)$

Values may differ somewhat due to rounding and the desire to have the weights add up to exactly 100% for illustrative purposes.

EXHIBIT 10

Massachusetts Workers Compensation Percentage Distributions of General, Other Acquistion, and Taxes

(1) Time from Eff. Date (Days)	(2) Distribution (%) General Exp.	(3) Distribution (%) Other Acquisition	(4) Distribution (%) Premium Tax	(5) Distribution (%) Other Tax
−359 to −270	1	2	0	1
−269 to −180	1	6	0	2
−179 to − 90	15	12	0	14
−89 to 0	23	44	0	29
1 to 90	21	30	100	23
91 to 180	10	2	0	8
181 to 270	9	1	0	7
271 to 360	9	1	0	7
361 to 450	10	2	0	8
451 to 540	1	0	0	1
Total	100	100	100	100

Source: (2) from filing for 1977 Massachusetts workers compensation rates, Exhibit 20.
(3) & (5) from filing for 1977 Massachusetts workers compensation rates, Exhibit 21.

EXHIBIT 11

MASSACHUSETTS WORKERS COMPENSATION COMBINED LOSS FLOW

Quarter	Losses	Quarter	Losses	Quarter	Losses
0	0.000	46	2.096	92	1.590
1	48.524	47	2.096	93	1.590
2	48.524	48	2.096	94	1.590
3	48.524	49	2.159	95	1.590
4	48.524	50	2.159	96	1.590
5	66.200	51	2.159	97	1.590
6	66.200	52	2.159	98	1.590
7	66.200	53	2.457	99	1.590
8	66.200	54	2.457	100	1.590
9	37.726	55	2.457	101	1.590
10	37.726	56	2.457	102	1.590
11	37.726	57	1.853	103	1.590
12	37.726	58	1.853	104	1.590
13	21.349	59	1.853	105	1.590
14	21.349	60	1.853	106	1.590
15	21.349	61	2.324	107	1.590
16	21.349	62	2.324	108	1.590
17	12.779	63	2.324	109	1.590
18	12.779	64	2.324	110	1.590
19	12.779	65	1.601	111	1.590
20	12.779	66	1.601	112	1.590
21	8.713	67	1.601	113	1.590
22	8.713	68	1.601	114	1.590
23	8.713	69	2.610	115	1.590
24	8.713	70	2.610	116	1.590
25	5.936	71	2.610	117	1.228
26	5.936	72	2.610	118	1.228
27	5.936	73	1.590	119	1.228
28	5.936	74	1.590	120	1.228
29	4.248	75	1.590	121	1.190
30	4.248	76	1.590	122	1.190
31	4.248	77	1.590	123	1.190
32	4.248	78	1.590	124	1.190
33	3.512	79	1.590	125	1.190
34	3.512	80	1.590	126	1.190
35	3.512	81	1.590	127	1.190
36	3.512	82	1.590	128	1.190
37	2.500	83	1.590	129	1.190
38	2.500	84	1.590	130	1.190
39	2.500	85	1.590	131	1.190
40	2.500	86	1.590	132	1.190
41	1.104	87	1.590	133	0.024
42	1.104	88	1.590	134	0.024
43	1.104	89	1.590	135	0.024
44	1.104	90	1.590		
45	2.096	91	1.590		

EXHIBIT 12

PART 1

MASSACHUSETTS WORKERS COMPENSATION
ACCIDENT YEAR FINANCIAL AGGREGATE DATA
Paid Medical Losses ($000)

Accident Year	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Report																		
1										53,868	57,773	69,032	73,437	71,349	59,620	56,209	47,474	46,756
2									100,050	125,246	143,539	163,308	168,043	157,921	125,313	111,694	98,415	
3								101,006	120,685	156,030	179,265	202,461	206,221	189,375	144,506	129,386		
4							99,464	111,294	134,064	171,728	198,412	224,423	224,206	202,641	154,817			
5						85,250	103,981	118,337	141,457	181,742	212,212	235,114	233,222	209,992				
6					84,128	87,491	106,993	123,617	146,424	189,738	218,820	241,023	238,891					
7				80,364	85,169	89,279	109,980	126,548	150,553	194,425	223,554	244,963						
8			64,825	81,159	86,736	90,931	111,508	129,785	153,184	197,840	226,442							
9		61,458	65,473	82,235	88,021	92,239	113,482	134,600	155,011	200,696								
10	48,866	62,020	66,200	83,335	88,691	94,155	114,027	135,889	156,469									
11	49,373	62,220	66,408	84,521	90,436	94,869	114,545	135,572										
12	49,986	62,915	66,701	86,405	90,784	96,394	114,737											
13	50,157	63,424	68,101	87,640	91,257	96,521												
14	50,780	64,781	68,370	88,671	92,472													
15	52,900	65,188	68,557	89,697														
16	53,271	65,745	69,034															
17	53,289	66,584																
18	55,017																	
Estimated Ultimate Losses*																		
	59,496	73,415	75,107	105,024	104,280	110,334	131,378	158,061	182,265	241,389	270,238	294,102	291,717	262,854	200,627	179,797	155,381	149,490

*From elsewhere in the ratemaking process.

EXHIBIT 12

PART 2

MASSACHUSETTS WORKERS COMPENSATION RATIO OF CUMULATIVE PAID LOSSES TO ULTIMATE LOSSES[1]

									MEDICAL Accident Year Data									
Accident Year	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Report																		
1										0.223	0.214	0.235	0.252	0.271	0.297	0.313	0.306	0.313
2									0.549	0.519	0.531	0.555	0.576	0.601	0.625	0.621	0.633	
3								0.639	0.662	0.646	0.663	0.688	0.707	0.720	0.720	0.720		
4							0.757	0.704	0.736	0.711	0.734	0.763	0.769	0.771	0.772			
5						0.773	0.791	0.749	0.776	0.753	0.785	0.799	0.799	0.799				
6					0.807	0.793	0.814	0.782	0.803	0.786	0.810	0.820	0.819					
7				0.765	0.817	0.809	0.837	0.801	0.826	0.805	0.827	0.833						
8			0.863	0.773	0.832	0.824	0.849	0.821	0.840	0.820	0.838							
9		0.837	0.872	0.783	0.844	0.836	0.864	0.852	0.850	0.831								
10	0.821	0.845	0.881	0.793	0.851	0.853	0.868	0.860	0.858									
11	0.830	0.848	0.884	0.805	0.867	0.860	0.872	0.858										
12	0.840	0.857	0.888	0.823	0.871	0.874	0.873											
13	0.843	0.864	0.907	0.834	0.875	0.875												
14	0.853	0.882	0.910	0.844	0.887													
15	0.889	0.888	0.913	0.854														
16	0.895	0.896	0.919															
17	0.896	0.907																
18	0.925																	

[1]Ratio calculated from the data in Exhibit 12, Part 1.

EXHIBIT 12

PART 3

MASSACHUSETTS WORKERS COMPENSATION RATIO OF INCREMENTAL PAID LOSSES TO ULTIMATE LOSSES[1]

MEDICAL
Accident Year Data

Accident Year	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Report																		
1										0.223	0.214	0.235	0.252	0.271	0.297	0.313	0.306	0.313
2										0.296	0.317	0.321	0.324	0.329	0.327	0.309	0.328	
3									0.113	0.128	0.132	0.133	0.131	0.120	0.096	0.098		
4								0.065	0.073	0.065	0.071	0.075	0.062	0.050	0.051			
5							0.034	0.045	0.041	0.041	0.051	0.036	0.031	0.028				
6						0.020	0.023	0.033	0.027	0.033	0.024	0.020	0.019					
7					0.010	0.016	0.023	0.019	0.023	0.019	0.018	0.013						
8				0.008	0.015	0.015	0.012	0.020	0.014	0.014	0.011							
9			0.009	0.010	0.012	0.012	0.015	0.030	0.010	0.012								
10		0.008	0.010	0.010	0.006	0.017	0.004	0.008	0.008									
11	0.009	0.003	0.003	0.011	0.017	0.006	0.004	(0.002)										
12	0.010	0.009	0.004	0.018	0.003	0.014	0.001											
13	0.003	0.007	0.019	0.012	0.005	0.001												
14	0.010	0.018	0.004	0.010	0.012													
15	0.036	0.006	0.002	0.010														
16	0.006	0.008	0.006															
17	0.000	0.011																
18	0.029																	

[1]Increments of the ratios in Exhibit 12, Part 2.

EXHIBIT 12

PART 4

MASSACHUSETTS WORKERS COMPENSATION MEDICAL LOSS FLOW ESTIMATED FROM FINANCIAL AGGREGATE DATA

Annual Flow			Quarterly Flow					
Accident Year Report	Selected % Paid In Year[1]	Cumulative Paid	Quarter	% Paid in Quarter	Quarter	% Paid in Quarter	Quarter	% Paid in Quarter
1	0.310309	0.310309	1	0.077577	46	0.001552	91	0.001250
2	0.321294	0.631603	2	0.077577	47	0.001552	92	0.001250
3	0.104578	0.736181	3	0.077577	48	0.001552	93	0.001250
4	0.054505	0.790686	4	0.077577	49	0.001454	94	0.001250
5	0.031742	0.822428	5	0.080324	50	0.001454	95	0.001250
6	0.021326	0.843754	6	0.080324	51	0.001454	96	0.001250
7	0.016776	0.860530	7	0.080324	52	0.001454	97	0.001250
8	0.013091	0.873621	8	0.080324	53	0.002088	98	0.001250
9	0.017439	0.891060	9	0.026145	54	0.002088	99	0.001250
10	0.006767	0.897827	10	0.026145	55	0.002088	100	0.001250
11	0.002804	0.900631	11	0.026145	56	0.002088	101	0.001250
12	0.006206	0.906837	12	0.026145	57	0.001484	102	0.001250
13	0.005815	0.912652	13	0.013626	58	0.001484	103	0.001250
14	0.008350	0.921002	14	0.013626	59	0.001484	104	0.001250
15	0.005935	0.926937	15	0.013626	60	0.001484	105	0.001250
16	0.006725	0.933662	16	0.013626	61	0.001681	106	0.001250
17	0.005863	0.939525	17	0.007936	62	0.001681	107	0.001250
18	0.005000	0.944525	18	0.007936	63	0.001681	108	0.001250
19	0.005000	0.949525	19	0.007936	64	0.001681	109	0.001250
20	0.005000	0.954525	20	0.007936	65	0.001466	110	0.001250
21	0.005000	0.959525	21	0.005332	66	0.001466	111	0.001250
22	0.005000	0.964525	22	0.005332	67	0.001466	112	0.001250
23	0.005000	0.969525	23	0.005332	68	0.001466	113	0.001250
24	0.005000	0.974525	24	0.005332	69	0.001250	114	0.001250
25	0.005000	0.979525	25	0.004194	70	0.001250	115	0.001250
26	0.005000	0.984525	26	0.004194	71	0.001250	116	0.001250
27	0.005000	0.989525	27	0.004194	72	0.001250	117	0.000119
28	0.005000	0.994525	28	0.004194	73	0.001250	118	0.000119
29	0.005000	0.999525	29	0.003273	74	0.001250	119	0.000119
30	0.000475	1.000000	30	0.003273	75	0.001250	120	0.000119
			31	0.003273	76	0.001250		
			32	0.003273	77	0.001250		
			33	0.004360	78	0.001250		
			34	0.004360	79	0.001250		
			35	0.004360	80	0.001250		
			36	0.004360	81	0.001250		
			37	0.001692	82	0.001250		
			38	0.001692	83	0.001250		
			39	0.001692	84	0.001250		
			40	0.001692	85	0.001250		
			41	0.000701	86	0.001250		
			42	0.000701	87	0.001250		
			43	0.000701	88	0.001250		
			44	0.000701	89	0.001250		
			45	0.001552	90	0.001250		

[1] Latest three-year average of increments in Exhibit 12, Part 3.

EXHIBIT 13

PART 1

MASSACHUSETTS WORKERS COMPENSATION ACCIDENT YEAR FINANCIAL AGGREGATE DATA Paid Indemnity Losses ($000)

Accident Year	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Report																		
1										120,247	137,690	142,992	140,914	117,052	68,573	56,731	48,624	43,578
2									236,187	319,315	381,688	421,680	414,181	324,040	182,154	151,241	134,800	
3								285,961	344,820	469,125	571,781	633,427	617,310	478,981	260,130	223,780		
4							305,843	348,226	433,480	574,344	698,980	758,832	737,862	565,118	305,445			
5						282,023	339,136	395,469	493,370	646,570	774,682	828,605	799,948	623,872				
6					274,255	299,399	365,637	431,819	532,881	698,477	823,743	872,214	847,431					
7				256,162	286,337	314,986	389,072	451,884	560,608	728,175	851,545	900,818						
8			217,492	261,645	298,092	327,759	401,139	464,595	574,907	745,590	870,900							
9		213,574	220,787	267,421	307,369	336,265	409,779	471,651	582,627	758,099								
10	157,572	215,640	225,003	272,338	312,842	340,707	416,787	475,324	591,827									
11	160,492	218,918	227,921	276,273	316,244	344,828	420,620	473,875										
12	163,310	222,244	230,692	278,138	319,978	350,551	422,424											
13	164,222	225,142	232,076	281,285	322,427	355,902												
14	166,056	226,086	234,416	284,195	327,349													
15	167,521	227,993	236,156	287,414														
16	170,118	229,065	239,648															
17	170,660	231,683																
18	173,122																	
Estimated Ultimate Losses*	186,627	251,361	261,737	319,256	364,175	402,259	482,020	542,046	694,626	946,041	1,107,184	1,164,583	1,133,705	890,031	468,719	408,273	357,818	303,942

*From elsewhere in the ratemaking process.

EXHIBIT 13

PART 2

MASSACHUSETTS WORKERS COMPENSATION
RATIO OF CUMULATIVE PAID LOSSES TO ULTIMATE LOSSES[1]

INDEMNITY
Accident Year Data

Accident Year	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Report																		
1										0.127	0.124	0.123	0.124	0.132	0.146	0.139	0.136	0.143
2									0.340	0.338	0.345	0.362	0.365	0.364	0.389	0.370	0.377	
3								0.528	0.496	0.496	0.516	0.544	0.545	0.538	0.555	0.548		
4							0.635	0.642	0.624	0.607	0.631	0.652	0.651	0.635	0.652			
5						0.701	0.704	0.730	0.710	0.683	0.700	0.712	0.706	0.701				
6					0.753	0.744	0.759	0.797	0.767	0.738	0.744	0.749	0.747					
7				0.802	0.786	0.783	0.807	0.834	0.807	0.770	0.769	0.774						
8			0.831	0.820	0.819	0.815	0.832	0.857	0.828	0.788	0.787							
9		0.850	0.844	0.838	0.844	0.836	0.850	0.870	0.839	0.801								
10	0.844	0.858	0.860	0.853	0.859	0.847	0.865	0.877	0.852									
11	0.860	0.871	0.871	0.865	0.868	0.857	0.873	0.874										
12	0.875	0.884	0.881	0.871	0.879	0.871	0.876											
13	0.880	0.896	0.887	0.881	0.885	0.885												
14	0.890	0.899	0.896	0.890	0.899													
15	0.898	0.907	0.902	0.900														
16	0.912	0.911	0.916															
17	0.914	0.922																
18	0.928																	

[1] Ratio calculated from the data in Exhibit 13, Part 1.

EXHIBIT 13

PART 3

MASSACHUSETTS WORKERS COMPENSATION
RATIO OF INCREMENTAL PAID LOSSES TO ULTIMATE LOSSES[1]

INDEMNITY
Accident Year Data

Accident Year	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Report																		
1										0.127	0.124	0.123	0.124	0.132	0.146	0.139	0.136	0.143
2									0.210	0.210	0.220	0.239	0.241	0.233	0.242	0.231	0.241	
3								0.167	0.156	0.158	0.172	0.182	0.179	0.174	0.166	0.178		
4							0.105	0.115	0.128	0.111	0.115	0.108	0.106	0.097	0.097			
5						0.063	0.069	0.087	0.086	0.076	0.068	0.060	0.055	0.066				
6					0.043	0.043	0.055	0.067	0.057	0.055	0.044	0.037	0.042					
7				0.029	0.033	0.039	0.049	0.037	0.040	0.031	0.025	0.025						
8			0.020	0.017	0.032	0.032	0.025	0.023	0.021	0.018	0.017							
9		0.014	0.013	0.018	0.025	0.021	0.018	0.013	0.011	0.013								
10	0.013	0.008	0.016	0.015	0.015	0.011	0.015	0.007	0.013									
11	0.016	0.013	0.011	0.012	0.009	0.010	0.008	−0.003										
12	0.015	0.013	0.011	0.006	0.010	0.014	0.004											
13	0.005	0.012	0.005	0.010	0.007	0.013												
14	0.010	0.004	0.009	0.009	0.014													
15	0.008	0.008	0.007	0.010														
16	0.014	0.004	0.013															
17	0.003	0.010																
18	0.013																	

[1]Increments of the ratios in Exhibit 13, Part 2.

EXHIBIT 13

PART 4

MASSACHUSETTS WORKERS COMPENSATION INDEMNITY LOSS FLOW ESTIMATED FROM FINANCIAL AGGREGATE DATA

Annual Flow			Quarterly Flow					
Accident Year Report	Selected % Paid In Year[1]	Cumulative Paid	Quarter	% Paid in Quarter	Quarter	% Paid in Quarter	Quarter	% Paid in Quarter
1	0.139406	0.139406	1	0.034852	49	0.002490	97	0.001750
2	0.238216	0.377622	2	0.034852	50	0.002490	98	0.001750
3	0.172706	0.550328	3	0.034852	51	0.002490	99	0.001750
4	0.099931	0.650259	4	0.034852	52	0.002490	100	0.001750
5	0.060230	0.710489	5	0.059554	53	0.002631	101	0.001750
6	0.041214	0.751703	6	0.059554	54	0.002631	102	0.001750
7	0.027021	0.778724	7	0.059554	55	0.002631	103	0.001750
8	0.018825	0.797549	8	0.059554	56	0.002631	104	0.001750
9	0.012450	0.809999	9	0.043177	57	0.002027	105	0.001750
10	0.011521	0.821520	10	0.043177	58	0.002027	106	0.001750
11	0.005174	0.826694	11	0.043177	59	0.002027	107	0.001750
12	0.009407	0.836101	12	0.043177	60	0.002027	108	0.001750
13	0.009961	0.846062	13	0.024983	61	0.002627	109	0.001750
14	0.010523	0.856585	14	0.024983	62	0.002627	110	0.001750
15	0.008106	0.864691	15	0.024983	63	0.002627	111	0.001750
16	0.010508	0.875199	16	0.024983	64	0.002627	112	0.001750
17	0.006660	0.881859	17	0.015058	65	0.001665	113	0.001750
18	0.013000	0.894859	18	0.015058	66	0.001665	114	0.001750
19	0.007000	0.901859	19	0.015058	67	0.001665	115	0.001750
20	0.007000	0.908859	20	0.015058	68	0.001665	116	0.001750
21	0.007000	0.915859	21	0.010304	69	0.003250	117	0.001750
22	0.007000	0.922859	22	0.010304	70	0.003250	118	0.001750
23	0.007000	0.929859	23	0.010304	71	0.003250	119	0.001750
24	0.007000	0.936859	24	0.010304	72	0.003250	120	0.001750
25	0.007000	0.943859	25	0.006755	73	0.001750	121	0.001750
26	0.007000	0.950859	26	0.006755	74	0.001750	122	0.001750
27	0.007000	0.957859	27	0.006755	75	0.001750	123	0.001750
28	0.007000	0.964859	28	0.006755	76	0.001750	124	0.001750
29	0.007000	0.971859	29	0.004706	77	0.001750	125	0.001750
30	0.007000	0.978859	30	0.004706	78	0.001750	126	0.001750
31	0.007000	0.985859	31	0.004706	79	0.001750	127	0.001750
32	0.007000	0.992859	32	0.004706	80	0.001750	128	0.001750
33	0.007000	0.999859	33	0.003113	81	0.001750	129	0.001750
34	0.000141	1.000000	34	0.003113	82	0.001750	130	0.001750
			35	0.003113	83	0.001750	131	0.001750
			36	0.003113	84	0.001750	132	0.001750
			37	0.002880	85	0.001750	133	0.000035
			38	0.002880	86	0.001750	134	0.000035
			39	0.002880	87	0.001750	135	0.000035
			40	0.002880	88	0.001750	136	0.000035
			41	0.001294	89	0.001750	137	0.000000
			42	0.001294	90	0.001750	138	0.000000
			43	0.001294	91	0.001750	139	0.000000
			44	0.001294	92	0.001750	140	0.000000
			45	0.002352	93	0.001750	141	0.000000
			46	0.002352	94	0.001750	142	0.000000
			47	0.002352	95	0.001750	143	0.000000
			48	0.002352	96	0.001750	144	0.000000

[1]Latest three-year average of increments in Exhibit 14.

EXHIBIT 14

PART 1

CALCULATION OF RISK-FREE RATE OF RETURN

Duration	Yield*	Weight**
1	5.70%	1× 152.045
2	6.09%	2× 91.034
3	6.25%	3× 66.128
4	6.35%	4× 39.887
5	6.45%	5× 26.394
6	6.52%	6× 18.176
7	6.57%	7× 13.008
8	6.62%	8× 10.752
9	6.66%	9× 7.656
10	6.70%	10× 3.380
11	6.74%	11× 6.416
12	6.78%	12× 6.608
13	6.82%	13× 7.524
14+	6.83%	14× 93.996
Weighted Average	6.60%	

*Yield from Exhibit 14, Part 4.
**Weight is the product of the duration and the corresponding values from Exhibit 14, Part 5.

EXHIBIT 14

PART 2

TREASURY BOND YIELD RATES

Month	1 Year	2 Year	3 Year	5 Year	7 Year	10 Year	30 Year
Jul 96	5.85	6.27	6.45	6.64	6.76	6.87	7.03
Aug 96	5.67	6.03	6.21	6.39	6.52	6.64	6.84
Sep 96	5.83	6.23	6.41	6.60	6.73	6.83	7.03
Oct 96	5.55	5.91	6.08	6.27	6.42	6.53	6.81
Nov 96	5.42	5.70	5.82	5.97	6.10	6.20	6.48
Dec 96	5.47	5.78	5.91	6.07	6.20	6.30	6.55
Jan 97	5.61	6.01	6.16	6.33	6.47	6.58	6.83
Feb 97	5.53	5.90	6.03	6.20	6.32	6.42	6.69
Mar 97	5.80	6.22	6.38	6.54	6.65	6.69	6.93
Apr 97	5.99	6.45	6.61	6.76	6.86	6.89	7.09
May 97	5.87	6.28	6.42	6.57	6.66	6.71	6.94
Jun 97	5.69	6.09	6.24	6.38	6.46	6.49	6.77
Average	5.69	6.07	6.23	6.39	6.51	6.60	6.83

(July 1996–June 1997)
Source: Federal Reserve Board (Statistical Release G-13)

EXHIBIT 14

PART 3

DURATIONS OF TREASURY SECURITIES

Maturity	Yield*	Duration
1 Year	5.69%	0.99
2 Year	6.07%	1.91
3 Year	6.23%	2.79
5 Year	6.39%	4.37
7 Year	6.51%	5.75
10 Year	6.60%	7.51
30 Year	6.83%	13.26

Note: Duration is a weighted average term to maturity, where the years are weighted by the present value of the related cash flow.
For bonds with semiannual coupons, duration in years is:

$$[(1+Y)-(1+Y)\wedge(1-2M)]/2Y$$

where Y is the semi-annual coupon yield = $[(1+\text{yield})\wedge .5]-1$ and M is the maturity.
*From Exhibit 14, Part 2.

EXHIBIT 14

PART 4

INTERPOLATED YIELDS BY DURATION

Duration		Yield	
0.99	*	5.69%	*
1.00		5.70%	**
1.91	*	6.07%	*
2.00		6.09%	**
2.79	*	6.23%	*
3.00		6.25%	**
4.00		6.35%	**
4.37	*	6.39%	*
5.00		6.45%	**
5.75	*	6.51%	*
6.00		6.52%	**
7.00		6.57%	**
7.51	*	6.60%	*
8.00		6.62%	**
9.00		6.66%	**
10.00		6.70%	**
11.00		6.74%	**
12.00		6.78%	**
13.00		6.82%	**
13.26	*	6.83%	*
14.00		6.83%	***

*From Exhibit 14, Part 3.
**Interpolated.
***Taken equal to last observed value.

EXHIBIT 14

PART 5

MASSACHUSETTS WORKERS COMPENSATION CALCULATION OF NET CASH FLOWS BY YEAR

Quarter	Net Cash Flow	Duration Period	Sum for Duration Period	Quarter	Net Cash Flow	Duration Period	Sum for Duration Period
−3	−1.270			30	−3.252	7 Yr	
−2	−2.480			31	−3.252	7 Yr	
−1	−13.780			32	−3.252	7 Yr	13.008
0	−28.930			33	−2.688	8 Yr	
1	135.915			34	−2.688	8 Yr	
2	148.538			35	−2.688	8 Yr	
3	171.820			36	−2.688	8 Yr	10.752
4	133.195			37	−1.914	9 Yr	
5	−14.213	1 Yr		38	−1.914	9 Yr	
6	−36.277	1 Yr		39	−1.914	9 Yr	
7	−54.959	1 Yr		40	−1.914	9 Yr	7.656
8	−46.596	1 Yr	152.045	41	−0.845	10 Yr	
9	−23.186	2 Yr		42	−0.845	10 Yr	
10	−14.834	2 Yr		43	−0.845	10 Yr	
11	−24.799	2 Yr		44	−0.845	10 Yr	3.38
12	−28.215	2 Yr	91.034	45	−1.604	11 Yr	
13	−16.247	3 Yr		46	−1.604	11 Yr	
14	−17.007	3 Yr		47	−1.604	11 Yr	
15	−16.342	3 Yr		48	−1.604	11 Yr	6.416
16	−16.532	3 Yr	66.128	49	−1.652	12 Yr	
17	−10.351	4 Yr		50	−1.652	12 Yr	
18	−9.782	4 Yr		51	−1.652	12 Yr	
19	−9.782	4 Yr		52	−1.652	12 Yr	6.608
20	−9.972	4 Yr	39.887	53	−1.881	13 Yr	
21	−6.764	5 Yr		54	−1.881	13 Yr	
22	−6.290	5 Yr		55	−1.881	13 Yr	
23	−6.670	5 Yr		56	−1.881	13 Yr	7.524
24	−6.670	5 Yr	26.394	57	−1.418	14 Yr	
25	−4.544	6 Yr		58	−1.418	14 Yr	
26	−4.544	6 Yr		59	−1.418	14 Yr	
27	−4.544	6 Yr		60	−1.418	14 Yr	5.672
28	−4.544	6 Yr	18.176	61–400	−88.324	15+ Yr	88.324
29	−3.252	7 Yr					

Note: Net Cash Flow = Premium − Total Losses & Expenses (including dividends).

EXHIBIT 15

CALCULATION OF THE RISK-ADJUSTED RATE OF RETURN

(1)	Risk-Free Rate of Return	6.60%
(2)	Beta of Liabilities	−0.21
(3)	Market Risk Premium	8.9%
(4)	Risk-Adjusted Rate of Return = (1) + [(2) × (3)]	4.73%

(1) From Exhibit 14.
(2) The Beta of Liabilities is the same as used by the Massachusetts Commissioner of Insurance in the past to set private passenger automobile rates.
(3) From Exhibit 16.

EXHIBIT 16

MARKET RISK PREMIUM

Year	Total Return Large Company Stocks	Total Return U.S. Treasury Bills	Difference
1926	11.62	3.27	8.35
1927	37.49	3.12	34.37
1928	43.61	3.56	40.05
1929	−8.42	4.75	−13.17
1930	−24.90	2.41	−27.31
1931	−43.34	1.07	−44.41
1932	−8.19	0.96	−9.15
1933	53.99	0.30	53.69
1934	−1.44	0.16	−1.60
1935	47.67	0.17	47.50
1936	33.92	0.18	33.74
1937	−35.03	0.31	−35.34
1938	31.12	−0.02	31.14
1939	−0.41	0.02	−0.43
1940	−9.78	0.00	−9.78
1941	−11.59	0.06	−11.65
1942	20.34	0.27	20.07
1943	25.90	0.35	25.55
1944	19.75	0.33	19.42
1945	36.44	0.33	36.11
1946	−8.07	0.35	−8.42
1947	5.71	0.50	5.21
1948	5.50	0.81	4.69
1949	18.79	1.10	17.69
1950	31.71	1.20	30.51
1951	24.02	1.49	22.53
1952	18.37	1.66	16.71
1953	−0.99	1.82	−2.81
1954	52.62	0.86	51.76
1955	31.56	1.57	29.99
1956	6.56	2.46	4.10
1957	−10.78	3.14	−13.92
1958	43.36	1.54	41.82
1959	11.96	2.95	9.01
1960	0.47	2.66	−2.19
1961	26.89	2.13	24.76
1962	−8.73	2.73	−11.46
1963	22.80	3.12	19.68
1964	16.48	3.54	12.94
1965	12.45	3.93	8.52

EXHIBIT 16

MARKET RISK PREMIUM (Continued)

Year	Total Return Large Company Stocks	Total Return U.S. Treasury Bills	Difference
1966	−10.06	4.76	−14.82
1967	23.98	4.21	19.77
1968	11.06	5.21	5.85
1969	−8.50	6.58	−15.08
1970	4.01	6.52	−2.51
1971	14.31	4.39	9.92
1972	18.98	3.84	15.14
1973	−14.66	6.93	−21.59
1974	−26.47	8.00	−34.47
1975	37.20	5.80	31.40
1976	23.84	5.08	18.76
1977	−7.18	5.12	−12.30
1978	6.56	7.18	−0.62
1979	18.44	10.38	8.06
1980	32.42	11.24	21.18
1981	−4.91	14.71	−19.62
1982	21.41	10.54	10.87
1983	22.51	8.80	13.71
1984	6.27	9.85	−3.58
1985	32.16	7.72	24.44
1986	18.47	6.16	12.31
1987	5.23	5.47	−0.24
1988	16.81	6.35	10.46
1989	31.49	8.37	23.12
1990	−3.17	7.81	−10.98
1991	30.55	5.60	24.95
1992	7.67	3.51	4.16
1993	9.99	2.90	7.09
1994	1.31	3.90	−2.59
1995	37.43	5.60	31.83
1996	23.07	5.21	17.86
Average	12.67	3.79	8.88

Selected Market Risk Premium is 8.9.
Source: SBBI, 1997 Year Book from Ibbotson Associates, Table 2-5.

EXHIBIT 17

SURPLUS FLOW
(Computed Using Final Weights)

Quarter	Proportion of Loss + Expense and Dividends Remaining to be Paid	Surplus[1]
0	0.954041	238.510
1	0.851465	425.733
2	0.788122	394.061
3	0.724595	362.298
4	0.663072	331.536
5	0.602455	301.227
6	0.550032	275.016
7	0.468795	234.397
8	0.417636	208.818
9	0.388307	194.154
10	0.358545	179.272
50	0.105393	52.697
100	0.034422	17.211

Note: Quarters 11, 12, etc. have not been displayed solely in the interests of space.

[1]Equal to the premium times the proportion of loss, expenses, and dividends remaining to be paid, divided by the premium-to-surplus ratio. For example, in Quarter 50, $(1000) \times (.105393)/2 = 52.697$. The premium-to-surplus ratio has been selected as 2 for all quarters. In Quarter 0, only one-half of the calculated surplus is included to represent the surplus flow starting at policy inception, which occurs at the end of Quarter 0.

EXHIBIT 18

PART 1

INVESTMENT BALANCE FOR TAX FLOW

Quarter	Using Initial Weights	Using Final Weights
−3	0.000	0.000
−2	−1.270	−1.254
−1	−3.750	−3.677
0	−17.530	−17.391
1	191.926	192.547
2	514.335	516.894
3	631.142	633.879
4	771.122	774.089
5	873.515	876.604
6	829.096	831.878
7	766.730	769.144
8	670.742	674.388
9	598.699	601.950
10	560.920	563.957
50	159.725	160.578
100	53.181	53.465

Note: Quarters 11, 12, etc. have not been displayed solely in the interests of space.
The Investment Balance for Taxes is the Investment Balance advanced one quarter.
The Investment Balance is the sum of the Surplus Flow (Exhibit 17) and the Cumulative Premiums minus Losses, Expenses, and Dividends (Exhibit 18, Part 2).

EXHIBIT 18

PART 2

CUMULATIVE PREMIUM MINUS LOSSES, EXPENSES, AND DIVIDENDS

Quarter	Using Initial Weights	Using Final Weights
−3	−1.270	−1.254
−2	−3.750	−3.677
−1	−17.530	−17.391
0	−46.460	−45.963
1	89.455	91.161
2	237.994	239.818
3	409.813	411.791
4	543.009	545.068
5	528.796	530.651
6	492.519	494.128
7	437.560	439.991
8	390.964	393.132
9	367.779	369.803
10	352.944	354.841
50	104.830	105.389
100	34.236	34.418

Note: Quarters 11, 12, etc. have not been displayed solely in the interests of space.

EXHIBIT 19

PART 1

UNDERWRITING TAX FLOWS

Contribution to the Underwriting Tax Flow of:			
		Losses and Expenses (plus Dividends)*	
Quarter	Premiums	Final Weights	Initial Weights
−3	0.00	1.254	1.270
−2	0.00	2.423	2.480
−1	0.00	13.714	13.780
0	0.00	28.572	28.930
1	250.00	231.876	232.398
2	250.00	192.643	192.075
3	250.00	192.826	192.294
4	250.00	190.823	190.218
5	0.00	14.936	14.974
6	0.00	6.743	6.739
7	0.00	35.556	36.620
8	0.00	5.478	5.458
9	0.00	3.712	3.704

*Loss and LAE contribution from Exhibit 19, Part 3, converted to a quarterly flow. (Exhibit 19, Part 3 only displays the result for the initial weights.) Dividends plus Expenses Other than LAE from Exhibit 19, Part 2. For example, for the initial weights for Quarter 6, $(20.953/4) + 1.501 = 6.739$. Quarters 10, 11, etc. have not been displayed solely in the interests of space.

EXHIBIT 19

PART 2

UNDERWRITING TAX FLOW FOR EXPENSES (INCLUDING DIVIDENDS) OTHER THAN LAE

Quarter	Final Weights	Initial Weights
−3	1.254	1.270
−2	2.423	2.480
−1	13.714	13.780
0	28.572	28.930
1	57.065	58.515
2	17.832	18.192
3	18.015	18.411
4	16.012	16.335
5	9.670	9.736
6	1.476	1.501
7	30.290	31.382
8	0.212	0.219
9	0.295	0.306
10	0.729	0.755
11	0.212	0.219
12	0.034	0.036
13	0.005	0.005
14	−0.034	−0.036
15	0.000	0.000
16	−0.010	−0.010
17	−0.030	−0.031

Note: Quarters 18, 19, etc. have not been displayed solely in the interests of space.

EXHIBIT 19

PART 3

CONTRIBUTION TO THE UNDERWRITING TAX FLOW OF LOSS AND LOSS ADJUSTMENT EXPENSE–INITIAL WEIGHTS

(1) Year	(2) Reserve Discount Factors[1]	(3) Difference in Reserve Discount Factors	(4) Loss + LAE Paid During Year[2]	(5) Loss + LAE Paid Subsequent to the Year	(6)[3]	(7) Change in Discounted Reserves[4]	(8) Loss + LAE Contribution to Underwriting Tax Flow
1	0.0000	0.8339	181.079	616.924	181.079	514.453	695.533
2	0.8339	−0.0307	202.706	414.219	33.669	−12.717	20.953
3	0.8032	−0.0306	115.518	298.701	22.734	−9.140	13.594
4	0.7726	−0.0188	65.369	233.331	14.865	−4.387	10.478
5	0.7538	−0.0104	39.128	194.203	9.633	−2.020	7.614
6	0.7434	−0.0146	26.678	167.525	6.846	−2.446	4.400
7	0.7288	0.0014	18.175	149.350	4.929	0.209	5.138
8	0.7302	−0.0051	13.006	136.344	3.509	−0.695	2.814
9	0.7251	0.0335	10.753	125.591	2.956	4.207	7.163
10	0.7586	0.0263	7.655	117.936	1.848	3.102	4.950
11	0.7849	0.0287	3.380	114.556	0.727	3.288	4.015
12	0.8136	0.0316	6.417	108.138	1.196	3.417	4.613
13	0.8452	0.0354	6.610	101.529	1.023	3.594	4.617
14	0.8806	0.0407	7.524	94.005	0.898	3.826	4.724
15	0.9213	0.0486	5.674	88.332	0.447	4.293	4.739
16	0.9699	0.0000	7.117	81.215	0.214	0.000	0.214
17	0.9699	0.0000	4.903	76.311	0.148	0.000	0.148
18	0.9699	0.0000	7.992	68.319	0.241	0.000	0.241

[1]Exhibit 20.
[2]Sum of quarterly paid losses from Exhibit 11 plus paid LAE.
[3]Losses paid in the year minus previously held discounted reserve for those losses.
[4]On losses for subsequent year.
(6) = [1 − (2)] × (4)
(7) = (3) × (5)
(8) = (6) + (7)
Note: Years beyond 18 are not displayed solely in the interest of space. The contribution to the underwriting tax flow declines slowly to zero.

EXHIBIT 20

Part 1

Summary of Discount Reserve Factors

Year	Discount Reserve Factor
1	0.8339
2	0.8032
3	0.7726
4	0.7538
5	0.7434
6	0.7288
7	0.7302
8	0.7251
9	0.7586
10	0.7849
11	0.8136
12	0.8452
13	0.8806
14	0.9213
15	0.9699
16	0.9699

Calculated using the reserve flow from Exhibit 20, Part 3 and the interest rate (average mid-term AFR) from Exhibit 20, Part 4. The calculation of the values for the first two years are shown on Exhibit 20, Part 2; the remaining values are calculated similarly.

EXHIBIT 20

PART 2

CALCULATION OF COUNTRYWIDE LIABILITY RESERVE DISCOUNT FACTORS

Year	Reserve Flow*	Discount Factors**	Discounted Reserve Flow
1	28.3575	0.9699	27.5043
2	15.4945	0.9124	14.1377
3	8.2342	0.8584	7.0679
4	5.1434	0.8075	4.1532
5	4.1564	0.7596	3.1573
6	2.4089	0.7146	1.7214
7	2.3136	0.6723	1.5553
8	0.5173	0.6324	0.3271
9	0.9641	0.5949	0.5736
10	0.9641	0.5597	0.5396
11	0.9641	0.5265	0.5076
12	0.9641	0.4953	0.4775
13	0.9641	0.4659	0.4492
14	0.9641	0.4383	0.4226
15	5.2530	0.4124	2.1661
16	0.0000	0.3879	0.0000
Total	77.6634		64.7605
Total Discounted Reserve/Total Reserve =			0.8339

Year	Reserve Flow*	Discount Factors**	Discounted Reserve Flow
1	15.4945	0.9699	15.0283
2	8.2342	0.9124	7.5131
3	5.1434	0.8584	4.4149
4	4.1564	0.8075	3.3562
5	2.4089	0.7596	1.8299
6	2.3136	0.7146	1.6533
7	0.5173	0.6723	0.3478
8	0.9641	0.6324	0.6097
9	0.9641	0.5949	0.5736
10	0.9641	0.5597	0.5396
11	0.9641	0.5265	0.5076
12	0.9641	0.4953	0.4775
13	0.9641	0.4659	0.4492
14	5.2530	0.4383	2.3025
15	0.0000	0.4124	0.0000
Total	49.3059		39.6032
Total Discounted Reserve/Total Reserve =			0.8032

*From Exhibit 20, Part 3.
**Based on the average mid-term AFR (see Exhibit 20, Part 4) of 6.30%.

EXHIBIT 20

PART 3

WORKERS COMPENSATION COUNTRYWIDE RESERVE FLOW

Year	Annual Liability Loss Flow
1	22.3366
2	28.3575
3	15.4945
4	8.2342
5	5.1434
6	4.1564
7	2.4089
8	2.3136
9	0.5173
10	0.9641
11	0.9641
12	0.9641
13	0.9641
14	0.9641
15	0.9641
16	5.2530

Source: Revenue Procedure 92-47 (Tables of Discount Factors).

EXHIBIT 20

PART 4

CALCULATION OF INTEREST RATE FOR RESERVE DISCOUNT FACTORS

Month		Midterm AFR*	Month		Midterm AFR*
Jun.	1992	7.04%	Jun.	1995	6.83%
Jul.	1992	6.85%	Jul.	1995	6.28%
Aug.	1992	6.49%	Aug.	1995	6.04%
Sep.	1992	5.98%	Sep.	1995	6.38%
Oct.	1992	5.78%	Oct.	1995	6.31%
Nov.	1992	5.68%	Nov.	1995	6.11%
Dec.	1992	6.15%	Dec.	1995	5.91%
Jan.	1993	6.34%	Jan.	1996	5.73%
Feb.	1993	6.22%	Feb.	1996	5.61%
Mar.	1993	5.88%	Mar.	1996	5.45%
Apr.	1993	5.45%	Apr.	1996	5.88%
May	1993	5.46%	May	1996	6.36%
12 Month Average		**6.11%**	**48 Month Average**		**6.24%**
Jun.	1993	5.33%	Jun.	1996	6.58%
Jul.	1993	5.54%	Jul.	1996	6.74%
Aug.	1993	5.32%	Aug.	1996	6.84%
Sep.	1993	5.35%	Sep.	1996	6.64%
Oct.	1993	5.00%	Oct.	1996	6.72%
Nov.	1993	4.92%	Nov.	1996	6.60%
Dec.	1993	5.07%	Dec.	1996	6.31%
Jan.	1994	5.32%	Jan.	1997	6.10%
Feb.	1994	5.34%	Feb.	1997	6.38%
Mar.	1994	5.36%	Mar.	1997	6.42%
Apr.	1994	5.88%	Apr.	1997	6.49%
May	1994	6.43%	May	1997	6.85%
24 Month Average		**5.76%**	**60 Month Average**		**6.30%**
Jun.	1994	6.92%			
Jul.	1994	6.83%			
Aug.	1994	7.05%			
Sep.	1994	7.05%			
Oct.	1994	7.10%			
Nov.	1994	7.45%			
Dec.	1994	7.74%			
Jan.	1995	7.92%			
Feb.	1995	7.96%			
Mar.	1995	7.75%			
Apr.	1995	7.34%			
May	1995	7.12%			
36 Month Average		**6.29%**			

*Midterm "Applicable Federal Rate" published monthly by the Internal Revenue Sevice.

EXHIBIT 21

PART 1

VALUES OF KAPPAS

	Initial Weights	Final Weights
κ_1 = Risk-adjusted discounted loss, expense and dividend factor	.8573	.8567
κ_2 = Risk-free discounted premiums	.9622	.9622
κ_3 = Risk-free discounted investment value tax	14.4878	14.5589
κ_4 = Risk-free discounted underwriting profit tax factor (contribution of premiums)	.9610	.9610
κ_5 = Risk-adjusted discounted underwriting profit tax factor (contribution of losses, expenses, and dividends)	.9496	.9495
κ_6 = Risk-free discounted revenue offset tax factor	.9674	.9674

EXHIBIT 21

PART 2

VALUES USED SOLELY TO COMPUTE THE RISK PREMIUM

	Initial Weights	Final Weights
Risk-Free κ_1	.8195	.8187
Risk-Free κ_5	.9331	.9330

EXHIBIT 21

PART 3

CALCULATION OF THE RISK PREMIUM

(1)	Premium calculated using a risk-adjusted discount rate	965
(2)	Premium calculated using a risk-free discount rate (Beta = 0)	906
(3)	Risk Load 1 – (2)/(1)	6.1%

EXHIBIT 21

PART 4

CALCULATION OF KAPPAS

	Flow Discounted	Discount Rate	Discounted to Time Zero From Middle or End of Quarter
κ_1	Losses, Expenses, and Dividends (Exhibits 10 and 11)	Risk-Adjusted	Middle
κ_2	Premiums (Exhibit 6)	Risk-Free	Middle
κ_3	Investment Balance for Taxes (Exhibit 18)	Risk-Free	Middle
κ_4	Premium Contribution to U/W Tax Flow (Exhibit 19)	Risk-Free	End
κ_5	Loss, Expense, and Dividends Contribution to U/W Tax Flow (Exhibit 19)	Risk-Adjusted	End
κ_6	Unearned Premium Reserve Contribution to Revenue Offset Tax Provision (Exhibit 22)	Risk-Free	End

EXHIBIT 22

PART 1

MASSACHUSETTS WORKERS COMPENSATION CALCULATION OF ALPHA— THE REVENUE OFFSET TAX FACTOR

(1)	Unearned Premium Reserve/Premium (Exhibit 22, Part 2)	0.120
(2)	Risk-Free Rate (See Exhibit 14)	6.60%
(3)	Quarterly Risk-Free Rate = $[1 + (2)]^{0.25} - 1$	1.61%
(4)	Proportion of Unearned Premium Reserve change brought into income (TRA 1986)	20%
(5)	Alpha = $4 \times (1) \times (3) \times (4)$	0.00155

EXHIBIT 22

PART 2

MASSACHUSETTS WORKERS COMPENSATION CALCULATION OF UNEARNED PREMIUM RESERVE RATIO

(1)	Countrywide Net Written Premium—1995	26,188,620
(2)	Unearned Premium (Prior Year—1994)	3,323,798
(3)	Unearned Premium (Current Year—1995)	3,506,306
(4)	Average Unearned Premium = $[(2) + (3)]/2$	3,415,052
(5)	Ratio Unearned/Written Premium (Prior Year) = $(2)/(1)$	0.127
(6)	Ratio Unearned/Written Premium (Current Year) = $(3)/(1)$	0.134
(7)	Average Ratio = $(4)/(1)$	0.130
(8)	Ratio Underlying Current Rates	0.110
(9)	Selected Unearned Premium Reserve Ratio	0.120

Source: "1996 Best's Aggregates & Averages" ($000)
Annual Statement & Insurance Expense Exhibit
"Underwriting & Investment Exhibit, Part 2—Premium Earned".

EXHIBIT 23

MASSACHUSETTS WORKERS COMPENSATION RISK-FREE DISCOUNTED UNEARNED PREMIUM TAX FACTOR

(1) Quarter	(2) Unearned Premium Reserve	(3) Unearned Premium Reserve Lagged One Quarter
0	180	0
1	140	180
2	100	140
3	60	100
4	0	60
Total	480	480

(4)	Annual Risk-Free Rate (Exhibit 14)	6.60%
(5)	Present value of Column (3) at interest rate in (4)	464.3483
(6)	κ_6 = (5)/Sum of (3)	0.967392

(2) = Selected Relative Values (see Text).

EXHIBIT 24

SENSITIVITY ANALYSIS BASED ON PRACTICAL EXAMPLE IN SECTION 4

Risk-Free Rate	Model Profit Provision	Difference
12.6%	−11.5%	−7.9%
10.6%	−9.2%	−5.6%
8.6%	−6.6%	−3.0%
6.6%	−3.6%	Base
4.6%	−0.1%	+3.5%
Beta of Liabilities		
−.11	−6.9%	−3.3%
−.21	−3.6%	Base
−.31	−0.2%	+3.4%
Investment Income Tax Rate		
25%	−8.2%	−4.6%
30%	−5.9%	−2.3%
35%	−3.6%	Base
40%	−1.4%	+2.2%
Underwriting Income Tax Rate[1]		
25%	−3.3%	+0.3%
30%	−3.4%	+0.2%
35%	−3.6%	Base
40%	−3.8%	−0.2%
(Initial) Premium-to-Surplus Ratio		
3	−5.7%	−2.1%
2	−3.6%	Base
1	2.5%	+6.1%

[1]The sensitivity exhibited here is not typical. This type of sensitivity will be present when a small negative underwriting profit provision has been calculated. The magnitude and direction of sensitivity to the underwriting income tax rate depends on whether there is an indicated underwriting loss or gain and the magnitude of that loss or gain.

EXHIBIT 24

Sensitivity Analysis
Based on Practical Example in Section 4
(Continued)

Market Risk Premium		
8%	−4.3%	−0.7%
8.9%	−3.6%	Base
10%	−2.7%	+0.9%
Policyholder Dividends[2]		
0	−3.9%	−0.3%
3%	−3.6%	Base
5%	−3.4%	+0.2%
10%	−3.0%	+0.6%
Reserves for Tax Purposes		
No Discounting	−5.2%	−1.6%
Discounting as per TRA 1986	−3.6%	Base
Revenue Offset Feature[3]		
None	−3.7%	−0.1%
As per TRA 1986	−3.6%	Base
Timing of Loss Payments[4]		
Two Quarters Later	−4.4%	−0.8%
One Quarter Later	−4.0%	−0.4%
As per Exhibit 9	−3.6%	Base
One Quarter Earlier	−3.3%	+0.3%

[2]The observed sensitivity is due to the profit provision taking into account the effect of policyholder dividends on investment income. It does not include any change in rates due to any loading of a provision for dividends themselves.

[3]The impact would be greater for lines of insurance with a larger ratio of unearned premium reserves to premium. Also the impact is greater the higher the risk-free rate.

[4]Includes the impact of the corresponding changes in the LAE flows. The impact is greater with a higher risk-free rate. (A higher risk-free rate enhances the impact of time on the value of money.)

APPENDIX

Section 4 contains a practical application of the Myers–Cohn model to Massachusetts workers compensation insurance. Input values were selected to be reasonable at the time this calculation was first prepared in 1997. Most of the inputs will change over time and thus should be updated on a regular basis. While the particular input values shown will not be up to date, the means of getting these values should still be applicable.

In many cases, inputs have been taken from elsewhere in the ratemaking procedure. The calculations that produced those inputs are beyond the scope of this paper. However, in general it is important to choose a set of consistent inputs to any underwriting profit model. The set of inputs should be consistent both internally and with other parts of the ratemaking process.

In this application of the Myers–Cohn model, time has been divided into quarters of a year. While this has been found to be a very useful choice in practical applications, there is no reason why some other choice could not be made.[66] Claim payments in workers compensation insurance can extend for 70 years or more from the date of accident. Therefore, in the rate filing the loss flows extend out about 300 quarters.[67]

Certain complications present in recent rate filings have been removed to aid in exposition. Enough complications have been left to illustrate some of the difficulties that arise in practical situations. However, every application can have its own peculiar details that require special treatment. Many of those that have arisen in Massachusetts workers compensation are beyond the scope of this paper.

---

[66] In that case the risk-free and risk-adjusted rates used in the Myers–Cohn formula should be adjusted to be appropriate for the selected periods of time.

[67] The detailed behavior in the extreme tail of the loss flow has little impact on the profit provision. The fact that the loss flow is very long does have a significant impact.

For completeness, the changes that were made from the filing for 1/1/98 rates to get the practical application shown here are:

1. Massachusetts imposes a 1% tax on the investment income of domestic insurers. The final tax rate for investment income was 0.2% higher in the rate filing to reflect the pro-rated impact of this tax after federal income taxes.

2. In the rate filing at the suggestion of the Insurance Department, the risk-adjusted rate increases linearly to the risk-free rate from Quarter 5 to the end of the loss and expense flow. Equivalently, the absolute value of beta decreases linearly to zero. Consistent with the change in risk-adjusted rate, the surplus/liabilities ratio used in the rate filing decreases linearly to zero from Quarter 5 to the end of the loss and expense flow. No adjustment was made in the surplus ratio or the risk-adjustment by quarter in the practical application presented here.

3. In the rate filing at the suggestion of the Insurance Department, the expected compensation to shareholders[68] contained in the investment balance is reduced such that only 25% of expected shareholder compensation remains in the investment balance after Quarter 5. No such adjustment was made in the practical application presented here.

4. Massachusetts has had two major reform laws within the last fifteen years. Chapter 572, effective 10/1/86 introduced escalation of benefits and increased the maximum durations of benefits, among other changes. This lengthened the indemnity loss flow considerably. Chapter 398

[68]This expected compensation to shareholders for the risk of writing insurance can be calculated by comparing the profit provision with a beta of zero, i.e. with the risk-adjusted rate equal to the risk-free rate, to that calculated with the selected beta. In recent Massachusetts workers compensation rate filings, the expected compensation to shareholders has been about 5% or 6%.

effective 12/23/91 cuts back on the escalation of benefits and the maximum durations of benefits, among other changes. Recent rate filings have included these impacts on the indemnity loss flows. However, for simplicity, neither impact is presented here. More generally, estimates of the loss flows in particular applications could include estimates of the effects of changes in the law or other changes in payment patterns.

5. Recent rate filings have contained no provisions for policyholder dividends, due to changes in the law governing rate filings in Massachusetts. As calculated herein, we have assumed a policyholder dividend provision has been included in the proposed rates.[69] Dividends have been included in the calculation of the profit provision both here and in the rate filing in order to show the impact on the cash flows. It should be noted that as calculated herein, the profit provision takes into account the loss of opportunity to earn investment income but not the money paid out in dividends itself. One could add the dividend provision to the calculated profit provision to get a "profit and dividends provision." However, to the extent dividend payments have been explicitly allowed for elsewhere in the ratemaking process, one would only need to reflect the loss of opportunity to earn investment income in the calculation of a profit provision, as is done here.

[69] The loading for policyholder dividends is assumed to be 3% of net premiums. As with all inputs, this should be viewed as illustrative only. In those circumstances in which policyholder dividends should not be considered, the weight to the policyholder dividend flow can be set equal to zero.

STUDYING POLICY RETENTION RATES USING MARKOV CHAINS

JOSEPH O. MARKER

Abstract

How does one measure the effect of improved policy retention on such key variables as market share and profitability?

This paper will analyze this problem by:

- *using the theory of Markov chains to model policy retention and to determine key values such as steady-state probabilities;*
- *using current spreadsheet technology to solve the key matrix equations from Markov chain theory; and*
- *applying these results to determine key business variables such as effects on profitability and market share.*

1. INTRODUCTION AND PROBLEM STATEMENT

You run an insurance company. You know that retaining policies is good business, but you want to quantify its value.[1] To simplify the analysis, you assume that all policies are written for a fixed policy term, expire at the same time, and have no mid-term activity. It turns out that the theory of Markov chains provides help with the analysis.

Markov chains assume discrete time periods and a system with "states" and "transition probabilities," the probabilities of moving from one state to another in one time period. For example, a physical system may consist of particles that move from a state to state in each discrete time period.

[1] D'Arcy and Doherty [1] discuss the "aging phenomenon." Their paper looks at this phenomenon relative to the profitability of insuring a policyholder for several periods. This paper views the same phenomenon from the aggregate financial viewpoint of an entire corporation.

The number of successes in a sequence of independent Bernoulli trials with probability of success p is a Markov chain. The system is defined to be in state k at time n if there have been exactly k successes in the first n trials. The transition probability of going from state k to state $k+1$ is p and the transition probability of staying at state k is $q = 1 - p$. In this paper, the term "Markov chain"[2] refers to a system with stationary transition probabilities. This means that if a particle is in state j at time t, then the conditional probability of going to state k at time $t+1$ does not depend on t, nor does it depend on any of the states that the particle was in prior to time t.

For the policy retention problem of this paper we replace the term "particle" by the term "customer." We say that the customer is in state k for $k = 0,1,2\ldots$ if the customer has been insured with the company for k consecutive time periods (one time period is equal to one policy term).

$k = 0$ refers to a person not currently insured with the company.

$k = 1$ refers to a policyholder in his/her first policy term.

$k = 2$ refers to a policyholder who has renewed once.

To study retention we define retention probabilities $\{r_k,\ k = 0,1,2,\ldots\}$ such that

r_k is the probability of renewing a policy that has been with the company for k time periods (that is, r_k is the probability that a customer currently in state k will pass to state $k+1$ in the next time period), and

r_0 is the probability of writing a customer who is not currently insured with the company.

We need an initial distribution $\{p_k^{(0)},\ k = 0,1,2\ldots\}$ where $p_k^{(0)}$ is the proportion of the entire population that has been insured

[2]For discussion of Markov chains see Feller [2, p. 372] or Resnick [4, p. 60]. The notation in this paper more closely follows that of Feller.

with the company for k years. Note that $1 - p_0^{(0)}$ is the company's initial "market share."

With this notation the *matrix of transition probabilities*[3] is:

$$\mathbf{A} = \begin{pmatrix} 1-r_0 & r_0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1-r_1 & 0 & r_1 & 0 & 0 & \cdots & 0 & 0 \\ 1-r_2 & 0 & 0 & r_2 & 0 & \cdots & 0 & 0 \\ 1-r_3 & 0 & 0 & 0 & r_3 & \cdots & 0 & 0 \\ 1-r_4 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1-r_{N-1} & 0 & 0 & 0 & 0 & \cdots & 0 & r_{N-1} \\ 1-r_N & 0 & 0 & 0 & 0 & \cdots & 0 & r_N \end{pmatrix}.$$

Here a_{ij} is the conditional probability of going from state i to state j in one time period. The indices i and j range from 0 to N. (See Appendix C for a discussion of chains with an infinite number of states.) The maximum value N may be set at a number of policy periods after which the retention is essentially constant.

For this retention problem only the first column (more correctly called the zeroth column) and the superdiagonal are non-zero, along with a_{NN}. This is because a customer in state j either moves to state $j+1$ (if the policy renews) or to state 0 (if the customer takes his/her business elsewhere). The retention rate is simply the probability that the policy will renew at its next expiration.

Notation Conventions

Superscripts within parentheses, such as $^{(n)}$, refer to time periods, or in the case of matrix elements, refer to n-step transi-

[3]Feller [2] contains an example that is mathematically equivalent to this Markov chain, except that, in his example, the number of states is infinite. He refers to state k as the "age," and says that at the next time period the system will either pass to age $k+1$ or will go back to age 0 and start afresh. See [2, pp. 382, 390, 398, 403].

tion probabilities. The n-step transition probability of going from state i to state j is the conditional probability that a customer in state i will, in n periods, be in state j.

Plain superscripts refer to exponents, except ^{t} refers to matrix transpose.

Subscripts refer to states of the system.

Vectors and matrices are in boldface.

2. GENERAL RESULTS ABOUT MARKOV CHAINS AND THE STEADY-STATE DISTRIBUTION

Given an initial distribution of states $\mathbf{p}^{(0)}$, the distribution of states in the next period is given by $\mathbf{p}^{(1)} = \mathbf{A}^t\mathbf{p}^{(0)}$, where $\mathbf{A}^t$ is the transpose of $\mathbf{A}$. This follows immediately, since $p_k^{(1)} = \sum_j a_{jk} p_j^{(0)}$. Each term on the right represents the probability that the system is in state j at time 0 and passes to state k at time 1. The summation over j then is the total probability of being in state k at time 1. The kth element of $\mathbf{p}^{(1)}$ is thus the inner product of the vector $\mathbf{p}^{(0)}$ and the kth column of matrix $\mathbf{A}$, which is the definition of multiplication on the left by the transpose.

Similarly, the conditional probability $a_{jk}^{(2)}$ of moving to state k in two steps given initial state j is given by

$$a_{jk}^{(2)} = \sum_m a_{jm} a_{mk},$$

which means that the two-step transition matrix is given by $\mathbf{A}^{(2)} = \mathbf{A}^2$. This is intuitively obvious by observing that, in order to get from j to k in two steps, one must stop at some state m at the first step. By induction, the n-step transition matrix[4] is given by $\mathbf{A}^n$. By definition, the element $a_{jk}^{(n)}$ is the probability, given state j, of being in state k n-periods later.

[4]Feller [2, pp. 382, 383].

Let $\mathbf{B}$ = transpose of $\mathbf{A} = \mathbf{A}^t$. Then the distribution at time n is given by

$$\mathbf{p}^{(n)} = (\mathbf{A}^t)\mathbf{p}^{(n-1)} = \mathbf{B}\mathbf{p}^{(n-1)} = (\mathbf{B}^n)\mathbf{p}^{(0)} \qquad \text{for any} \quad n \geq 1.$$

The *steady-state (or invariant) probability distribution* is defined as that solution of the equation $\mathbf{p}^* = \mathbf{B}\mathbf{p}^*$ for which $\sum p_j^* = 1$. It turns out that the steady-state probabilities are very important to our original business retention problem. We will discuss later how to calculate $\mathbf{p}^*$.

A key result is that

$$\mathbf{B}^n\mathbf{p}^{(0)} \to \mathbf{p}^* \quad \text{as} \quad n \to \infty \quad \text{for any initial distribution } \mathbf{p}^{(0)}.$$

The proof is in Appendix B. This limiting result says that the ultimate distribution of customers by state (remember: "state" is the number of consecutive renewals) is independent of the initial distribution but depends only on the steady state probabilities associated with the retentions.

3. CALCULATING THE STEADY-STATE (INVARIANT) DISTRIBUTION

There are several approaches to calculating the invariant distribution for our retention problem.

3.1. Use the Definition Directly

Recall that the matrix $\mathbf{B}$ for the retention problem is given by

$$\mathbf{B} = \mathbf{A}^t = \begin{pmatrix} 1-r_0 & 1-r_1 & 1-r_2 & 1-r_3 & 1-r_4 & \cdots & 1-r_{N-1} & 1-r_N \\ r_0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & r_1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & r_2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & r_3 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & r_{N-1} & r_N \end{pmatrix},$$

where all the r_k are strictly between 0 and 1.

The defining equations for invariance are:

$$p_k = r_{k-1}p_{k-1} \qquad \text{for} \quad k = 1,2,3\cdots N-1, \tag{3.1}$$

$$p_N = r_{N-1}p_{N-1} + r_N p_N, \qquad \text{and} \tag{3.2}$$

$$p_0 = (1-r_0)p_0 + (1-r_1)p_1 + \cdots + (1-r_N)p_N. \tag{3.3}$$

From Equation 3.1 we obtain

$$p_k = r_0 r_1 r_2 \cdots r_{k-1} p_0 \qquad \text{for} \quad k = 1,2,3,\ldots N-1. \tag{3.4}$$

From Equation 3.4 we can see that the terms 0 through $N-1$ on the right-hand side of Equation 3.3 add to $p_0 - r_{N-1}p_{N-1}$. From Equation 3.2 the last term on the right-hand side of Equation 3.3 equals $r_{N-1}p_{N-1}$. Thus, we can choose an arbitrary value for p_0, define the remaining p_k by Equations 3.1 and 3.2, and Equation 3.3 will be automatically satisfied. Once all the p_k are calculated, just rescale them so they add to 1 and these values are the invariant probabilities.

Thus the retention problem has a particularly simple form of transition matrix that allows the steady-state probabilities to be calculated directly from the definition.

3.2. A Simple Machine-Oriented Approach[5]

The vector $\mathbf{p}^*$, whose transpose is defined by

$$(\mathbf{p}^*)^t = (1,1,\ldots 1,1)(\mathbf{I} - \mathbf{A} + \mathbf{ONE})^{-1},$$

defines an invariant distribution. Here **I** is the identity matrix and **ONE** is the square matrix all of whose entries are 1. Resnick [4] proves this handy proposition. This result requires that **A** be irreducible, which we prove in Appendix C.

[5]Resnick [4, p. 138].

3.3. Use a Spreadsheet "Solver"

A spreadsheet "solver"[6] can solve for the steady-state probabilities. A typical spreadsheet solver

(a) maximizes, minimizes, or sets a target cell to a specific value

(b) subject to constraint equations or inequalities

(c) by changing a set of "decision cells."

The use of the target cell is optional. The solver can be used to simply produce values of the decision cells that satisfy the given constraints.

Recall that steady state probability vector is simply the solution $\mathbf{p}^*$ of the matrix equation $\mathbf{Bx} = \mathbf{x}$, for which $\sum x_j = 1$, where $\mathbf{B} = \mathbf{A}^t$. This equation can be rewritten as $\mathbf{Cx} = 0$, where $\mathbf{C} = \mathbf{B} - \mathbf{I}$ and $\mathbf{I}$ is the identity matrix.

Now setting up the solver is simple:

1. Set up the matrix $\mathbf{C}$, which is a function of the transition probabilities $\mathbf{A}$.
2. Set up a vector $\mathbf{x}$, the vector of decision variables that are allowed to change when the solver is run.
3. Set up a vector $\mathbf{z}$ as the matrix product $\mathbf{Cx}$.
4. Run the solver with the following constraints:

$$\mathbf{z} = 0 \qquad \text{and} \qquad \sum x_j = 1.$$

The resultant vector $\mathbf{x}$ is the steady-state probability vector $\mathbf{p}^*$.

We present this solution using the solver because solvers are being commonly used to handle problems involving maximizing, minimizing, and satisfying constraints, and a solver for our

[6] The particular solver used in this paper is that from the Microsoft Excel spreadsheet.

retention problem does not require the same linear algebra skills that other solutions entail.

4. SPREADSHEET EXAMPLE TO MODEL THE RETENTION PROBLEM

Recall that we have translated the retention problem into Markov chain terms and have reviewed some characteristics of Markov chains. Appendix A displays printouts from a spreadsheet set up to analyze retentions. The spreadsheet is documented, but here are some of the highlights.

The Basic Data section asks us to input the retention probabilities $\{r_k,\ k = 0,1,2,\ldots N\}$ and initial probability distribution $\{p_k^{(0)}, k = 0,1,2\ldots N\}$. Recall that r_k is the probability that a policyholder that has been insured for k policy periods will renew when his/her policy expires, and r_0 is the probability that the company will capture as new business a customer not currently insured with the company. The end of the Basic Data section translates these retention probabilities into the matrix $\mathbf{A}$ of one-step transition probabilities.

For example, in Appendix A the company's initial market share is 10%, since the proportion $p_0^{(0)}$ of the market not insured by the company is 90%. Since $N = 9$ and $p_N^{(0)} = .043$, 4.3% of the market has been insured with the company nine or more policy periods. At the next renewal cycle $r_0 = 1.0\%$ of the population not insured by the company will be captured as new business.

The section labeled "Distribution At Time n" shows how the distribution $\mathbf{p}^{(n)}$ changes after n time periods. Recall that $\mathbf{p}^{(n)} = \mathbf{B}\mathbf{p}^{(n-1)}$, and that $\mathbf{p}^{(n)} = \mathbf{B}^n\mathbf{p}^{(0)}$, where $\mathbf{B} = \mathbf{A}^t$. We have shown that $\mathbf{p}^{(n)} \to \mathbf{p}^*$ (the steady-state probability) as $n \to \infty$.

The value of these calculations is that they allow us to get a feel for how fast the limit is approached. In the real world, a company does not have an infinite time horizon to wait for the limiting behavior to be realized. The "Distribution At Time n"

explanation also allows us to restrict the model to a finite planning horizon. Too many managerial changes during the convergence period could invalidate the Markov chain assumption that the transition matrix is stationary over time. The n-step transition matrices in Appendix A converge to a matrix which has the steady-state distribution vector in each of its columns.

5. RETURN TO THE ORIGINAL RETENTION PROBLEM

When any action affects retention, it changes the transition matrix **A**. Improved retention means larger superdiagonal elements (probabilities of renewal) and smaller elements in the first column (probabilities of non-renewal). In this section, we will study our original set of retention assumptions and their effect on key business variables. The spreadsheet with those results is shown as Appendix A. Then we will see how a shift in retention (Appendix B) may change the results.

We have used the theory of Markov chains, along with spreadsheet tools, to compute steady-state probabilities for a given set of retention rates. We have shown that the distribution of states of the system (recall that the "state" of an individual customer is the number of consecutive policy renewals for that customer) approaches the steady-state probabilities, as time goes on, regardless of the initial distribution.

The spreadsheet in Appendix A gives us a sense for how quickly this convergence takes place. It is easy to calculate the distribution at time n, because the n-step transition matrix is just the nth power of the one-step transition matrix. Mahler's paper, "A Markov Chain Model of Shifting Risk Parameters," provides a mathematical treatment of the rate of convergence [3].

Thus we have a wealth of tools that give us information about the probability distribution of states throughout time. These probabilities are not in themselves of much interest to management. However, there are functions of these probabilities that are of great interest. For example, the projected market share is of keen interest. Because we have included state 0 in our definition of states (the customer is in state 0 if he/she is not currently insured with the company), the market share at time n is given by $1 - p_0^{(n)}$.

Loss ratios, expense ratios, and combined ratios greatly interest management. Most observers would agree that renewing an existing policy is much less expensive than writing a new policy. It follows that increasing the retention rate will improve the expense ratio. Most would also agree that the loss ratio for a customer who has been on the company books for a long period of time will be lower than for a new or recent customer. Actions that improve retention should improve the loss ratio.

The last page of Appendix B, Combined Ratio Differential, illustrates how to estimate this effect. To estimate the effect of retention on combined ratio one needs a sense of how loss/expense ratios vary by state (the number of consecutive policy renewals). The phrase "needs a sense of" is intentionally vague. It could mean that we have data on loss or expense ratios by state. More likely it means that we have some information that would enable us to make an assumption about how the loss or expense ratio varies by state. For instance we may be able to say that a new policyholder has a 10% worse loss ratio than a long-standing policyholder. Or it could mean that we accept a management estimate of this differential and use the model to check the effect of retention under different estimates.

Once we have made a reasonable assumption about these differentials by state, we are ready to estimate the effect of improved retention. This is simply a matter of:

1. entering the initial distribution and retention probabilities in to the spreadsheet;
2. running the spreadsheet to determine the steady-state probabilities and how quickly the system approaches those limiting probabilities; and

TABLE 1

ORIGINAL RETENTION ASSUMPTIONS

	State k									
	0	1	2	3	4	5	6	7	8	9
Retention	.0100	.8500	.9000	.9000	.9000	.9000	.9000	.9000	.9000	.9000
Steady-State Probability	.9132	.0091	.0077	.0069	.0063	.0057	.0051	.0046	.0041	.0371
Combined Ratio Differential	.2000	.1000	.0800	.0600	.0400	.0200	0	0	0	0

3. applying the differentials in loss/expense ratios to the various probabilities to arrive at an "average differential" or "average loss/expense ratio."

Then we make the same calculation using the "improved retentions" in Appendix B and compare the results to estimate the effect of the change in retention.

The "Combined Ratio Differential" section in Appendix A shows a calculation of this nature for the original retention probabilities. Here we externally determined (or hypothesized) various combined ratio differentials by state relative to the combined ratio for a long-standing (i.e., seven term or longer) policyholder. The results are summarized in Table 1. The retention and combined ratio differentials are inputs to the calculation. The steady-state probabilities and the average differential are calculated. From Appendix A the average differential is +.0446 using the steady-state probabilities as weights. That is, on average the book of business will have a 4.5% higher average combined ratio than if the book consisted entirely of long-term customers.

Now suppose that the company takes some action that improves retention. Such an action might be a new billing option, more advertising, etc. The number of such actions is limited only

TABLE 2

IMPROVED RETENTION ASSUMPTIONS

	State k									
	0	1	2	3	4	5	6	7	8	9
Retention	.0120	.8700	.9200	.9200	.9200	.9200	.9200	.9200	.9200	.9200
Steady-State Probability	.8753	.0105	.0091	.0084	.0077	.0071	0066	.0060	.0055	.0637
Combined Ratio Differential	.2000	.1000	.0800	.0600	.0400	.0200	0	0	0	0

by the creativity of the sales or marketing manager. In the example in Appendix B, our improved retention assumption is that r_k increases by .02 for $k >= 1$ and r_0 increases by .002 (recall r_0 is the probability that the company writes a new customer). We can use the same spreadsheet with the revised retention and obtain the results shown in Table 2, assuming that the differentials have not changed.

Now what has been the effect of the management action to improve the retention? The ultimate market share increases from 8.7% to 12.5%. The ultimate loss ratio decreases by 0.8% (i.e., the combined ratio differential drops from 4.4% to 3.6%). Now remember that these are "ultimate" results and we know that, for Markov chains, it may take quite a few renewal cycles to approach these limiting results!

The insurer must weigh these benefits against the costs. For example, if an improved billing system produces the increased retention, then the improved market share and loss ratio must overcome the cost of maintaining and building the billing system. If instead, a rate decrease is used to improve retention then it is likely that the overall combined ratio itself will increase and wipe out the benefits from the retention improvement. The exact effect will depend on the price elasticity of demand for the product.

5. COMPARISON TO SINGLE POLICYHOLDER APPROACH

The approach used in this paper examines the financial effects of retaining policies on the entire company's book of business. Starting with an initial distribution of business by policy age and a set of transition probabilities, we use Markov chain theory to model the distribution over time. Because one of the states of the system (i.e., state 0) consists of potential customers not insured by the company, the model produces estimates of total growth as well as distribution by policy age. We then hypothesize differences in loss ratio by policy age to examine changes in profitability over time. The Markov chain approach enables us to examine the effects on growth and profitability of changes in the transition probabilities.

This entire approach is an aggregate approach in that it looks at the growth and profitability of a company's entire book of business over time. In contrast, D'Arcy and Doherty [1] approach the "aging phenomenon" by tracing the profitability of a single insured over time. They start with a new customer (corresponding to state 1 in this paper) and calculate the profitability of that customer's policies from the initial date through the last renewal, discounting all calculations to the initial policy inception date. D'Arcy and Doherty hypothesize differing levels of profitability by policy age. In their model the probability of renewal is constant over time.

D'Arcy and Doherty study the price that will optimize present and future profits from a customer added to the books. How do the approach of this paper and D'Arcy-Doherty relate? We can express the D'Arcy-Doherty models in Markov chain terms as follows:

The initial distribution $\mathbf{p}^0$ consists of a probability of 1.0 of being a new policyholder.

The retention probabilities r_k are constant (called W in [1]).

The state 0 becomes an "absorbing state." That is, there is no more action for the individual policyholder once he/she non-renews.

Using Markov chains to study the aging phenomenon in [1] is not useful because the transition probabilities are so simple that Markov chain theory is not needed. D'Arcy and Doherty concentrate on a single policyholder and the transition matrix does not satisfy the criteria for using the theorems about invariant distributions.

D'Arcy and Doherty concentrate on the single policyholder and are sophisticated in treating differing loss ratios and the time value of money in arriving at proper prices. Their analysis could be used as an input to this paper's aggregate model. We could use the models in [1] to enable us to calculate the expected present value of profit for each policy renewal (i.e. for each state k). This gives us a set of expected profits corresponding to the various states in the retention model. There is no need to sum these discounted present values for all the renewals of a single customer as is done in [1]. We can then hypothesize an initial distribution and use the transition matrix as was done in this paper. Our Markov chain model determines the distribution of states of the system over time. With this information and the expected profits by state, we can determine the company's expected profit over time. The Markov chain model allows us to easily vary the retention rate by state of the system.[7]

Both papers refer to optimizing profitability over time. In [1] this is done by calculating the present value of expected profits over the life of an individual policyholder as a function of price. The renewal rate W is adversely affected by increasing price, so that there is an optimum price above which the profits begin decreasing.

---

[7] The possibility of renewal rates changing by policy age is mentioned briefly in [1, p. 38].

In this paper, the expected profits for the entire corporation are calculated using Markov chains. Increasing the price for policies increases the profit at each renewal. However, price increases lower the renewal probabilities r_k. This decreases both the market share and the number of customers in the higher states (i.e., long-term policyholders). At some point raising the price adversely affects renewal probabilities so much that total profit is adversely affected.

Both papers mention elasticity of demand (with respect to price) as critical values. Basically, the more elastic the demand, the more difficult it is to increase overall profit through price increases.

In general, we could use D'Arcy and Doherty [1] to establish the expected profit by age of policy. Then we could plug this information into the Markov chain model to determine aggregate profitability over time and growth for the company.

6. OBSERVATIONS AND CAVEATS

Many companies have neither very good retention information nor very good ideas of loss ratio differentials by retention. The Markov chain model is useful even in these circumstances. To illustrate, some company managements have wildly inflated ideas of the benefits of improved retention on market share and profitability. Let's assume that the actuary can persuade management to "guess" the improvement both in retention rate and in combined ratio differentials by state. The company can then use the model to produce profitability and market share change estimates that are more realistic than management's original "feeling." As the company obtains better data, some of the hypothesizing can be replaced with observations. There is a high probability that retention data will improve because it is of universal interest among top management.

In using this type of modeling one must be careful not to compound too many assumed improvements. For example, suppose the retention on long-standing business is 95%. A new billing plan claims to increase this by 2% (additively). A few months later the ability to "account sell" increases the number again by 2%. Then a fancy new endorsement produces another 2% increase. Now the implied retention rate is 101%, which is absurd. This sounds ridiculous, but companies do act this way when the actions are separated in time and the company loses its memory due to management changes.

A better way to view this is to express these increases as reductions in the non-renewal or lapse rate and then compound them properly. For example, we might say that each of the three actions above reduces the lapse rate by 40% (i.e., reduces it from .05 to .03), so that the resultant retention from this series of actions becomes:

$$1 - (.05 \times .60 \times .60 \times .60) = .989.$$

The assumption that the policy renewal process is a Markov chain is a simplification of the real world. Recall that the Markov property says that the probability of passing to a given future state depends on the current state but does not depend on any prior history. This implies, for example, that the probability r_0 of capturing a new customer is the same whether or not that customer has ever been previously insured with the company. This is probably not an accurate assumption.

We can attempt to get around this assumption by defining two "0" states: state "0a" for potential customers who have never been with the company, and state "0b" for potential policyholders who had been previously insured. With this formulation the transition matrix is such that current policyholders (state 1 or higher) can never get to state 0a. It turns out that the invariant distribution assigns probability 0 to state 0a (that is, everyone eventually becomes a policyholder or former policyholder of the company).

In this situation the distribution of states approaches the invariant distribution very slowly. In one reasonable example (where the probability of capturing a new customer is a high 5%) the limiting distribution was not approached even after sixty-four time periods. Thus, in this situation the Markov chain model is useful for finite time periods, but the study of the invariant distribution is somewhat academic.

The model in this paper assumes that time is discrete, that all customers have policies with inception dates at these discrete time periods, and that the only possible actions are renewal or non-renewal. Of course, we know that customers can cancel or purchase policies at any time, and that endorsement activity is probably more frequent than renewal activity. This would require a continuous time Markov process with a richer set of options.

In selecting actions that improve financial results through "improved retention," we must verify that the action itself does not adversely affect the profitability for each state. A classic action that violates this condition is a rate decrease. Obviously, this action would decrease the profitability of each state, even though it improves retention.

7. CONCLUSION

This paper uses the theory of Markov chains to analyze retention rates and how they affect key insurance variables. In the paper, the Markov chain state for a customer is the number of consecutive policy periods the customer has been insured with the company. Determining the ultimate, or limiting, distribution for Markov chains involves solving matrix equations of the form $\mathbf{Bx} = \mathbf{x}$.

The paper shows how to do this using spreadsheets. Finally, the paper illustrates how changing the retention rates (i.e. the transition probabilities in the Markov chain) might change key business variables such as profitability and market share. There is also a discussion of how the model interrelates with an earlier "policy age" model by D'Arcy and Doherty.

REFERENCES

[1] D'Arcy, Stephen and Neil Doherty, "The Aging Phenomenon and Insurance Prices," *PCAS* LXXVI, 1989, pp. 24–44.

[2] Feller, William, *An Introduction to Probability Theory and its Applications*, Volume I, John Wiley & Sons, Inc., 1968.

[3] Mahler, Howard, "A Markov Chain Model of Shifting Risk Parameters," *PCAS* LXXXIV, 1997, pp. 581–659.

[4] Resnick, Sidney, *Adventures in Stochastic Processes*, Birkhauser: Boston, 1992.

APPENDIX A

INITIAL RETENTION ASSUMPTIONS

INTRODUCTION

This spreadsheet carries out the calculations for the insurance retention problem. The retention problem is set up as as a Markov Chain, where a customer is in state k if he/she has been insured with the company for k consecutive periods. State 0 refers to a potential customer not currently insured with the company. Each state k has an associated "retention probability" r_k, where r_k is the probability that a customer in state k renews his/her policy. The customer non-renews, i.e. moves to state 0, with probability $1 - r_k$.

The retention problem translates to a Markov chain as follows:

The states of the Markov chain are defined as in the retention problem.

The matrix of transition probabilities $\mathbf{A} = (a_{i,j})$ is defined as follows:

$$a_{k,k+1} = r_k,$$
$$a_{k,0} = 1 - r_k, \qquad \text{and}$$
$$a_{k,j} = 0 \qquad \text{for all other } j.$$

William Feller [2, p. 382] discusses this Markov chain problem. Sidney Resnick [4] describes this Markov chain as the "Success Run Chain."

The retention problem also requires a vector $\mathbf{p}^{(0)}$, the initial probability distribution of states.

With the transition probabilities $\mathbf{A}$ and the initial distribution $\mathbf{p}^{(0)}$ specified, the spreadsheet calculates the "steady-state," or invariant, distribution of states, to which the system converges; the probability distributions at various points in time, to check for rate of convergence; and changes in market share and profitability over time.

BASIC DATA

Section 1: Calculate the distribution given the matrix of transition probabilities **A**, where $a_{i,j}$ is probability of going from state i to state j in one step. The initial distribution is $\mathbf{p}^{(0)}$. This particular example is an effort to model insurance retention. State i is the number of years the customer has been insured with the company. The first state (zero) refers to a potential customer not currently insured. The next state (one) refers to a first-year insured, etc.

Input Section

Input the *retention probabilities* of going from state i to state $i+1$. That is, the input for state 0 is the probability that someone currently insured elsewhere will be written as new business. The input for state $i>0$ is the probability of renewing a policy of someone that the company has insured for i years.

Then input the *initial distribution* $\mathbf{p}^{(0)}$ of insureds. For $i = 0$, this is the proportion of the population not currently insured with the company. For $i > 0$, this is the proportion of the population insured with the company for i consectutive policy terms. The last column is the proportion insured with the company for 9 or more consecutive terms.

	State i									
	0	1	2	3	4	5	6	7	8	9
Retention Probabilities	0.01	0.85	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
Initial Distribution $\mathbf{p}^{(0)}$	0.9	0.01	0.009	0.008	0.007	0.007	0.006	0.005	0.005	0.043

Following is the resultant matrix **A** of one-step transition probabilities:

0.99	0.01	0	0	0	0	0	0	0	0
0.15	0	0.85	0	0	0	0	0	0	0
0.10	0	0	0.90	0	0	0	0	0	0
0.10	0	0	0	0.90	0	0	0	0	0
0.10	0	0	0	0	0.90	0	0	0	0
0.10	0	0	0	0	0	0.90	0	0	0
0.10	0	0	0	0	0	0	0.90	0	0
0.10	0	0	0	0	0	0	0	0.90	0
0.10	0	0	0	0	0	0	0	0	0.90
0.10	0	0	0	0	0	0	0	0	0.90

DISTRIBUTION AT TIME *n*

This section shows how to calculate the distribution at time n, given the initial distribution $\mathbf{p}^{(0)}$ and the matrix **A**.

Note that if **p** is the distribution of states at any time, then $\mathbf{A}^t$ (the transpose of **A**) times **p** is the distribution in the next time period. That is, the probability that the system is in state m in the next time period is the mth column of matrix **A** times the distribution **p**.

Matrix $\mathbf{A}^t$

0.99	0.15	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
0.01	0	0	0	0	0	0	0	0	0
0	0.85	0	0	0	0	0	0	0	0
0	0	0.90	0	0	0	0	0	0	0
0	0	0	0.90	0	0	0	0	0	0
0	0	0	0	0.90	0	0	0	0	0
0	0	0	0	0	0.90	0	0	0	0
0	0	0	0	0	0	0.90	0	0	0
0	0	0	0	0	0	0	0.90	0	0
0	0	0	0	0	0	0	0	0.90	0.90

Probability Distributions $\mathbf{p}^{(n)}$ at time period n, for $n = 1, 2, 3\ldots$

$\mathbf{p}^{(0)}$	State	1	2	3	4	5	6	7	8	9
0.900	0	0.9015	0.9028	0.9039	0.9049	0.9059	0.9067	0.9074	0.9080	0.9086
0.010	1	0.0090	0.0090	0.0090	0.0090	0.0090	0.0091	0.0091	0.0091	0.0091
0.009	2	0.0085	0.0077	0.0077	0.0077	0.0077	0.0077	0.0077	0.0077	0.0077
0.008	3	0.0081	0.0077	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069	0.0069
0.007	4	0.0072	0.0073	0.0069	0.0062	0.0062	0.0062	0.0062	0.0062	0.0062
0.007	5	0.0063	0.0065	0.0066	0.0062	0.0056	0.0056	0.0056	0.0056	0.0056
0.006	6	0.0063	0.0057	0.0058	0.0059	0.0056	0.0050	0.0050	0.0050	0.0050
0.005	7	0.0054	0.0057	0.0051	0.0052	0.0053	0.0050	0.0045	0.0045	0.0045
0.005	8	0.0045	0.0049	0.0051	0.0046	0.0047	0.0048	0.0045	0.0041	0.0041
0.043	9	0.0432	0.0429	0.0430	0.0433	0.0431	0.0430	0.0430	0.0428	0.0422

Note that the n-step transition probability is given by raising matrix $\mathbf{A}$ to the nth power. The distribution at time n is given by $(\mathbf{A}^t)^n$ times $\mathbf{p}^{(0)}$.

Shown below are the transposes of some n-step transition matrices:

Two-step									
0.9816	0.2335	0.1890	0.1890	0.1890	0.1890	0.1890	0.1890	0.1890	0.1890
0.0099	0.0015	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010
0.0085	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.7650	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.8100	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.8100	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.8100	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.8100	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8100	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8100	0.8100	0.8100

Four-step									
0.9675	0.3741	0.3388	0.3388	0.3388	0.3388	0.3388	0.3388	0.3388	0.3388
0.0097	0.0031	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027	0.0027
0.0083	0.0020	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016
0.0076	0.0011	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008
0.0069	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.6197	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.6561	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.6561	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.6561	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.6561	0.6561	0.6561	0.6561	0.6561

STEADY-STATE PROBABILITIES

This section shows the steady-state probabilities found using the solver.

The steady-state probability $\mathbf{p}^*$ is characterized by $\mathbf{A}^t \times \mathbf{p}^* = \mathbf{p}^*$, or $(\mathbf{A}^t - \mathbf{I}) \times \mathbf{p}^* = 0$, where $\mathbf{I}$ is the identity matrix.

Use the solver to find the solution. Let $\mathbf{C} = \mathbf{A}^t - \mathbf{I}$. The steady-state probability $\mathbf{p}^*$ is the solution of the linear system $\mathbf{Cx} = 0$ for which the elements of $\mathbf{x}$ sum to 1.0. After using the solver, the vector $\mathbf{x}$ contains the steady-state probabilities, and the vector $\mathbf{z} = \mathbf{Cx}$ contains all zeros.

Matrix $\mathbf{C} = \mathbf{A}^t - \mathbf{I}$										$\mathbf{x}$	$\mathbf{z} = \mathbf{Cx}$
−0.01	0.15	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.91324	0
0.01	−1.00	0	0	0	0	0	0	0	0	0.00913	0
0	0.85	−1.00	0	0	0	0	0	0	0	0.00776	0
0	0	0.90	−1.00	0	0	0	0	0	0	0.00699	0
0	0	0	0.90	−1.00	0	0	0	0	0	0.00629	0
0	0	0	0	0.90	−1.00	0	0	0	0	0.00566	0
0	0	0	0	0	0.90	−1.00	0	0	0	0.00509	0
0	0	0	0	0	0	0.90	−1.00	0	0	0.00458	0
0	0	0	0	0	0	0	0.90	−1.00	0	0.00413	0
0	0	0	0	0	0	0	0	0.90	−0.10	0.03713	0
										1.00000	

In this application of the solver, the "target cell" for the solver is undefined, since there is no objective function to maximize or minimize.

COMBINED RATIO DIFFERENTIAL

This section illustrates how studying the retention problem helps businesses evaluate profitability and market share. Generally combined ratios are better for customers who have been retained longer, due to lower expenses and/or better loss ratios. By comparing the average combined ratios before and after improving retention, one can measure the financial effects of changing policy retention.

State i	Steady State Probability **x**	Assumed Combined Ratio Differential	Assumed Base Combined Ratio	Combined Ratio
0	0.913242	N/A		
1	0.009132	20.00%		115.54%
2	0.007763	10.00%		105.54%
3	0.006986	8.00%		103.54%
4	0.006288	6.00%		101.54%
5	0.005659	4.00%		99.54%
6	0.005093	2.00%		97.54%
7	0.004584	0.00%		95.54%
8	0.004125	0.00%		95.54%
9	0.037128	0.00%	95.54%	95.54%
Market share		8.68%		
Average combined ratio differential		4.46%		
Average combined ratio		100.00%		

APPENDIX B

IMPROVED RETENTION ASSUMPTIONS

BASIC DATA

Section 1: Calculate the distribution given the matrix of transition probabilities **A**, where $a_{i,j}$ is the probability of going from state i to state j in one step. The initial distribution is $\mathbf{p}^{(0)}$. This particular example is an effort to model insurance retention. State i is the number of years a customer has been insured with the company. The first state (zero) refers to a potential customer not currently insured. The next state (one) refers to a first-year insured, etc.

Input Section

Input the retention probabilities of going from state i to state $i+1$. That is, the input for state 0 is the probability that someone currently insured elsewhere will be written as new business. The input for state $i > 0$ is the probability of renewing a policy of someone that the company has insured for i years.

Then input the initial distribution $\mathbf{p}^{(0)}$ of insureds. For $i = 0$, this is the proportion of the population not currently insured with the company. For $i > 0$, this is the proportion of the population insured with the company for i consectutive policy terms. The last column is the proportion insured with the company for nine or more consecutive terms.

	State i									
	0	1	2	3	4	5	6	7	8	9
Retention Probabilities	0.012	0.87	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
Initial Distribution $\mathbf{p}^{(0)}$	0.90	0.01	0.009	0.008	0.007	0.007	0.006	0.005	0.005	0.043

Following is the resultant matrix **A** of one-step transition probabilities:

0.988	0.012	0	0	0	0	0	0	0	0
0.130	0	0.87	0	0	0	0	0	0	0
0.080	0	0	0.92	0	0	0	0	0	0
0.080	0	0	0	0.92	0	0	0	0	0
0.080	0	0	0	0	0.92	0	0	0	0
0.080	0	0	0	0	0	0.92	0	0	0
0.080	0	0	0	0	0	0	0.92	0	0
0.080	0	0	0	0	0	0	0	0.92	0
0.080	0	0	0	0	0	0	0	0	0.92
0.080	0	0	0	0	0	0	0	0	0.92

DISTRIBUTION AT TIME n

This section shows how to calculate the distribution at time n, given the initial distribution $\mathbf{p}^{(0)}$ and the matrix **A**.

Note that if **p** is the distribution of states at any time, then $\mathbf{A}^t$ (the transpose of **A**) times **p** is the distribution in the next time period. That is, the probability that the system is in state m in the next time period is the mth column of matrix **A** times the distribution **p**.

Matrix $\mathbf{A}^t$									
0.988	0.13	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
0.012	0	0	0	0	0	0	0	0	0
0	0.87	0	0	0	0	0	0	0	0
0	0	0.92	0	0	0	0	0	0	0
0	0	0	0.92	0	0	0	0	0	0
0	0	0	0	0.92	0	0	0	0	0
0	0	0	0	0	0.92	0	0	0	0
0	0	0	0	0	0	0.92	0	0	0
0	0	0	0	0	0	0	0.92	0	0
0	0	0	0	0	0	0	0	0.92	0.92

Probability Distributions $\mathbf{p}^{(n)}$ at time period n, for $n = 1, 2, 3\ldots$

$\mathbf{p}^{(0)}$	State	1	2	3	4	5	6	7	8	9
0.900	0	0.8977	0.8957	0.8938	0.8921	0.8906	0.8892	0.8879	0.8867	0.8857
0.010	1	0.0108	0.0108	0.0107	0.0107	0.0107	0.0107	0.0107	0.0107	0.0106
0.009	2	0.0087	0.0094	0.0094	0.0094	0.0093	0.0093	0.0093	0.0093	0.0093
0.008	3	0.0083	0.0080	0.0086	0.0086	0.0086	0.0086	0.0086	0.0086	0.0085
0.007	4	0.0074	0.0076	0.0074	0.0080	0.0079	0.0079	0.0079	0.0079	0.0079
0.007	5	0.0064	0.0068	0.0070	0.0068	0.0073	0.0073	0.0073	0.0073	0.0073
0.006	6	0.0064	0.0059	0.0062	0.0064	0.0062	0.0067	0.0067	0.0067	0.0067
0.005	7	0.0055	0.0059	0.0055	0.0057	0.0059	0.0057	0.0062	0.0062	0.0062
0.005	8	0.0046	0.0051	0.0055	0.0050	0.0053	0.0055	0.0053	0.0057	0.0057
0.043	9	0.0442	0.0449	0.0459	0.0473	0.0481	0.0491	0.0502	0.0510	0.0522

Note that the n-step transition probability is given by raising matrix $\mathbf{A}$ to the nth power. The distribution at time n is given by $(\mathbf{A}^t)^n$ times $\mathbf{p}^{(0)}$.

Shown below are the transposes of some n-step transition matrices:

Two-step									
0.9777	0.1980	0.1526	0.1526	0.1526	0.1526	0.1526	0.1526	0.1526	0.1526
0.0119	0.0016	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010
0.0104	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.8004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.8464	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.8464	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.8464	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.8464	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8464	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8464	0.8464	0.8464

Four-step									
0.9598	0.3161	0.2786	0.2786	0.2786	0.2786	0.2786	0.2786	0.2786	0.2786
0.0116	0.0031	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026
0.0102	0.0021	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016
0.0095	0.0012	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008
0.0088	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.6775	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.7164	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.7164	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.7164	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.7164	0.7164	0.7164	0.7164	0.7164

STEADY-STATE PROBABILITIES

This section shows the steady-state probabilities found using the solver.

The steady-state probability $\mathbf{p}^*$ is characterized by $\mathbf{A}^t \times \mathbf{p}^* = \mathbf{p}^*$, or $(\mathbf{A}^t - \mathbf{I}) \times \mathbf{p}^* = 0$, where $\mathbf{I}$ is the identity matrix.

Use the solver to find the solution. Let $\mathbf{C} = \mathbf{A}^t - \mathbf{I}$. The steady-state probability $\mathbf{p}^*$ is the solution of the linear system $\mathbf{Cx} = 0$ for which the elements of $\mathbf{x}$ sum to 1.0. After using the solver, the vector $\mathbf{x}$ contains the steady-state probabilities, and the vector $\mathbf{z} = \mathbf{Cx}$ contains all zeros.

Matrix $\mathbf{C} = \mathbf{A}^t - \mathbf{I}$										$\mathbf{x}$	$\mathbf{z} = \mathbf{Cx}$
−0.012	0.13	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.87527	0
0.012	−1.00	0	0	0	0	0	0	0	0	0.01050	0
0	0.87	−1.00	0	0	0	0	0	0	0	0.00914	0
0	0	0.92	−1.00	0	0	0	0	0	0	0.00841	0
0	0	0	0.92	−1.00	0	0	0	0	0	0.00773	0
0	0	0	0	0.92	−1.00	0	0	0	0	0.00712	0
0	0	0	0	0	0.92	−1.00	0	0	0	0.00655	0
0	0	0	0	0	0	0.92	−1.00	0	0	0.00602	0
0	0	0	0	0	0	0	0.92	−1.00	0	0.00554	0
0	0	0	0	0	0	0	0	0.92	−0.08	0.06372	0
										1.00000	

In this application of the solver, the "target cell" for the solver is undefined, since there is no objective function to maximize or minimize.

COMBINED RATIO DIFFERENTIAL

This section illustrates how changing the retention assumptions affects profitability. Generally combined ratios are better for customers who have been retained longer, due to lower expenses and/or better loss ratios. By comparing the average combined ratios before (100.0%) and after improving retention (99.2%), one can measure the financial effects of changing policy retention.

State i	Steady State Probability **x**	Assumed Combined Ratio Differential	Assumed Base Combined Ratio	Combined Ratio
0	0.875274	N/A		
1	0.010503	20.00%		115.54%
2	0.009138	10.00%		105.54%
3	0.008407	8.00%		103.54%
4	0.007734	6.00%		101.54%
5	0.007116	4.00%		99.54%
6	0.006546	2.00%		97.54%
7	0.006023	0.00%		95.54%
8	0.005541	0.00%		95.54%
9	0.063719	0.00%	95.54%	95.54%
	Market share	12.47%		
	Average combined ratio differential	3.66%		
	Average combined ratio	99.20%		

APPENDIX C

PROOF THAT **A** IS IRREDUCIBLE

In this appendix we prove that an invariant distribution exists for the Markov chain formulation of the retention problem. Recall that the transition matrix for this problem is given by:

$$\mathbf{A} = \begin{pmatrix} 1-r_0 & r_0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1-r_1 & 0 & r_1 & 0 & 0 & \cdots & 0 & 0 \\ 1-r_2 & 0 & 0 & r_2 & 0 & \cdots & 0 & 0 \\ 1-r_3 & 0 & 0 & 0 & r_3 & \cdots & 0 & 0 \\ 1-r_4 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1-r_{N-1} & 0 & 0 & 0 & 0 & \cdots & 0 & r_{N-1} \\ 1-r_N & 0 & 0 & 0 & 0 & \cdots & 0 & r_N \end{pmatrix},$$

where all the r_k are strictly between 0 and 1.

To prove the result we need to define the terms "aperiodic" and "irreducible."

State j is defined to be "periodic" if there exists an integer $t > 1$ such that $a_{jj}^{(n)} = 0$ unless n is an integer multiple of t. Here $a_{jj}^{(n)}$ is the n-step probability of returning to state j. The matrix **A** is aperiodic if no states are periodic.

We show that this system is aperiodic. Consider any state j. For any $k > 0$, with $0 < k \leq N - j$ the system can return to state j in $j + k + 1$ steps through the sequence

$$j \to j+1 \to j+2 \to \cdots \to j+k \to 0 \to 1 \to \cdots \to j \quad \text{for} \quad k \leq N - j. \tag{C.1}$$

For $k > N - j$, the system can return to state j in $j + k + 1$ steps through the same sequence except that it "parks" at state N for

$k-(N-j)$ steps before going to state 0. For example, if $j=1$, $N=4$, and $k=6$, then the system returns to state j in $k+j+1$ (= 8) steps through the sequence of states:

$$1 \to 2 \to 3 \to 4 \to 4 \to 4 \to 4 \to 0 \to 1.$$

This last complication only comes about because we set N as the highest state. If we had allowed an infinite number of states then Equation C.1 holds for all $k > 0$.

Thus we have shown that a system in state j can return to state j in n steps for all $n = \{j+2, j+3, j+4, \ldots\}$.[8] This means **A** has no period; i.e., **A** is aperiodic.

A chain is defined to be "irreducible" if and only if every state can be reached from every other state. This means that, given any two states j and k, there exists an integer n such that the system can move from j to k in n steps.

The chain **A** is clearly irreducible since the system can move from state j to state k through the sequence of states:

$$j \to 0 \to 1 \to \cdots \to k$$

We have now established that **A** is aperiodic and irreducible.

We now show directly that an invariant distribution **p** exists for **A** by calculating it.

The defining equations for invariance are

$$p_k = r_{k-1} p_{k-1} \qquad \text{for} \quad k = 1, 2, 3 \cdots N-1, \tag{C.2}$$

$$p_N = r_{N-1} p_{N-1} + r_N p_N, \qquad \text{and} \tag{C.3}$$

$$p_0 = (1-r_0) p_0 + (1-r_1) p_1 + \cdots + (1-r_N) p_N. \tag{C.4}$$

From Equation C.2 we get

$$p_k = r_0 r_1 r_2 \cdots r_{k-1} p_0 \qquad \text{for} \quad k = 1, 2, 3 \cdots N-1. \tag{C.5}$$

[8]This is true for $n = j+1$ also, but this is not needed for the proof.

From Equation C.5 we can see that the terms 0 through $N-1$ on the right-hand side of Equation C.4 add to $p_0 - r_{N-1}p_{N-1}$. From Equation C.3 the last term on the right-hand side of Equation C.4 equals $r_{N-1}p_{N-1}$. Thus, we can choose an arbitrary value for p_0, define the remaining p_k by Equation C.2 and C.3, and Equation C.4 will be automatically satisfied. Once all the p_k are calculated, just rescale them so they add to 1 and these values of **p** are the invariant probabilities.

The following theorem[9] will now enable us to say that the n-step distributions converge to the invariant distribution, regardless of the initial distribution.

Suppose a chain is irreducible and aperiodic and that there exist probabilities $\{p_k,\ k = 0,1,2,\ldots\}$ with all $p_k \geq 0$ that satisfy the invariant distribution conditions:

$$\mathbf{p} = \mathbf{A}^t\mathbf{p}.$$

Then

$$a_{jk}^{(n)} \to p_k \qquad \text{as} \quad n \to \infty$$

independently of the initial state j, and the chain is ergodic.

We have already shown that **A** satisfies all the conditions of the theorem. (Note the term "ergodic" means that the mean recurrence time to revisit any state j is finite). What the conclusion means is that the n-step transition matrix $\mathbf{A}^n$ ultimately approaches the matrix for which every column is the invariant distribution.

[9]Feller, [2, p. 393]. Actually, the theorem in Feller is more powerful in that it provides a converse which states that if the limits exist, then they form the invariant distribution.

TESTING THE ASSUMPTIONS OF AGE-TO-AGE FACTORS

GARY G. VENTER

Abstract

The use of age-to-age factors applied to cumulative losses has been shown to produce least-squares optimal reserve estimates when certain assumptions are met. Tests of these assumptions are introduced, most of which derive from regression diagnostic methods. Failures of various tests lead to specific alternative methods of loss development.

ACKNOWLEDGEMENT

I would like to thank the Committee on Review of Papers for comments leading to numerous expository improvements over previous drafts. Any obscurity that remains is of course my own.

INTRODUCTION

In his paper "Measuring the Variability of Chain Ladder Reserve Estimates" Thomas Mack presented the assumptions needed for least-squares optimality to be achieved by the typical age-to-age factor method of loss development (often called "chain ladder"). Mack also introduced several tests of those assumptions. His results are summarized below, and then other tests of the assumptions are introduced. Also addressed is what to do when the assumptions fail. Most of the assumptions, if they fail in a particular way, imply least-squares optimality for some alternative method.

The organization of the paper is to first show Mack's three assumptions and their result, then to introduce six testable implications of those assumptions, and finally to go through the testing of each implication in detail.

PRELIMINARIES

Losses for accident year w evaluated at the end of that year will be denoted as being as of age 0, and the first accident year in the triangle is year 0. The notation below will be used to specify the models. Losses could be either paid or incurred. Only development that fills out the triangle is considered. Loss development beyond the observed data is often significant but is not addressed here. Thus age ∞ will denote the oldest possible age in the data triangle.

Notation

$c(w,d)$:	cumulative loss from accident year w as of age d
$c(w,\infty)$:	total loss from accident year w when end of triangle reached
$q(w,d)$:	incremental loss for accident year w from $d-1$ to d
$f(d)$:	factor applied to $c(w,d)$ to estimate $q(w,d+1)$
$F(d)$:	factor applied to $c(w,d)$ to estimate $c(w,\infty)$

Assumptions

Mack showed that some specific assumptions on the process of loss generation are needed for the chain ladder method to be optimal. Thus if actuaries find themselves in disagreement with one or another of these assumptions, they should look for some other method of development that is more in harmony with their intuition about the loss generation process. Reserving methods more consistent with other loss generation processes will be discussed below. Mack's three original assumptions are slightly restated here to emphasize the task as one of predicting future incremental losses. Note that the losses $c(w,d)$ have an evaluation date of $w+d$.

1. $\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)c(w,d)$.

   In words, the expected value of the incremental losses to emerge in the next period is proportional to the total losses emerged to date, by accident year. Note that in Mack's definition of the chain ladder, $f(d)$ does not depend on w, so the factor for a given age is constant across accident years. Note also that this formula is a linear relationship with no constant term. As opposed to other models discussed below, the factor applies directly to the cumulative data, not to an estimated parameter, like ultimate losses. For instance, the Bornhuetter-Ferguson method assumes that the expected incremental losses are proportional to the ultimate for the accident year, not the emerged to date.

2. Unless $v = w$, $c(w,d)$ and $c(v,g)$ are independent for all v, w, d and g.

   This would be violated, for instance, if there were a strong diagonal, when all years' reserves were revised upwards. In this case, instead of just using the chain ladder method, most actuaries would recommend eliminating these diagonals or adjusting them. Some model-based methods for formally recognizing diagonal effects are discussed below.

3. $\mathrm{Var}[q(w,d+1) \mid \text{data to } w+d] = a[d,c(w,d)]$.

   That is, the variance of the next increment observation is a function of the age and the cumulative losses to date. Note that $a(\cdot,\cdot)$ can be any function but does not vary by accident year. An assumption on the variance of the next incremental losses is needed to find a least-squares optimal method of estimating the development factors. Different assumptions, e.g., different functions $a(\cdot,\cdot)$ will lead to optimality for different methods of estimating the factor f. The form of $a(\cdot,\cdot)$ can be tested by trying different forms, estimating the f's, and seeing if the variance formula holds. There will almost always be some function $a(\cdot,\cdot)$ that reasonably accords with the observations, so the issue with this assumption is not its validity but its implications for the estimation procedure.

Results (Mack)

In essence what Mack showed is that under the above assumptions the chain ladder method gives the minimum variance unbiased linear estimator of future emergence. This gives a good justification for using the chain ladder in that case, but the assumptions need to be tested. Mack assumed that $a[d,c(w,d)] = k(d)c(w,d)$, that is, he assumed that the variance is proportional to the previous cumulative loss, with possibly a different proportionality factor for each age. In this case, the minimum variance unbiased estimator of $c(w,\infty)$ from the triangle of data to date $w+d$ is $F(d)c(w,d)$, where the age-to-ultimate factor $F(d) = [1+f(d)][1+f(d+1)]\cdots$, and $f(d)$ is calculated as:

$$f(d) = \sum_w q(w,d+1) \Big/ \sum_w c(w,d),$$

where the sum is over the w's mutually available in both columns (assuming accident years are on separate rows and ages are in separate columns). Actuaries often use a modified chain ladder that uses only the last n diagonals. This will be one of the alternative methods to test if Mack's assumptions fail. Using only part of the data when all the assumptions hold will reduce the accuracy of the estimation, however.

Extension

In general, the minimum variance unbiased $f(d)$ is found by minimizing

$$\sum_w [f(d)c(w,d) - q(w,d+1)]^2 k(d)/a[d,c(w,d)].$$

This is the usual weighted least-squares result, where the weights are inversely proportional to the variance of the quantity being estimated. Because only proportionality, not equality, to the variance is required, $k(d)$ can be any convenient function of d—usually chosen to simplify the minimization.

For example, suppose $a[d,c(w,d)] = k(d)c(w,d)^2$. Then the $f(d)$ produced by the weighted least-squares procedure is the average of the individual accident year d to $d+1$ ratios, $q(w,d+1)/c(w,d)$. For $a[d,c(w,d)] = k(d)$, each $f(d)$ regression above is then just standard unweighted least squares, so $f(d)$ is the regression coefficient $\sum_w c(w,d)q(w,d+1)/\sum_w c(w,d)^2$. (See Murphy [8].) In all these cases, $f(d)$ is fit by a weighted regression, and so regression diagnostics can be used to evaluate the estimation. In the tests below just standard least-squares will be used, but in application the variance assumption should be reviewed.

Discussion

Without going into Mack's derivation, the optimality of the chain ladder method is fairly intuitive from the assumptions. In particular, the first assumption is that the expected emergence in the next period is proportional to the losses emerged to date. If that were so, then a development factor applied to the emerged to date would seem highly appropriate. Testing this assumption will be critical to exploring the optimality of the chain ladder. For instance, if the emergence were found to be a constant plus a percent of emergence to date, then a different method would be indicated—namely, a factor plus constant development method. On the other hand, if the next incremental emergence were proportional to ultimate rather than to emerged to date, a Bornhuetter-Ferguson type approach would be more appropriate.

To test this assumption against its alternatives, the development method that leads from each alternative needs to be fit, and then a goodness-of-fit measure applied. This is similar to trying a lot of methods and seeing which one you like best, but it is different in two respects: (1) each method tested derives from an alternative assumption on the process of loss emergence; (2) there is a specific goodness-of-fit test applied. Thus the fitting is a test of the emergence patterns that the losses are subject to, and not just a test of estimation methods.

TESTABLE IMPLICATIONS OF ASSUMPTIONS

Verifying a hypothesis involves finding as many testable implications of that hypothesis as possible, and verifying that the tests are passed. In fact a hypothesis can never be fully verified, as there could always be some other test you haven't thought of. Thus the process of verification is sometimes conceived as being really a process of attempted falsification, with the current tentatively-accepted hypothesis being the strongest (i.e., most easily testable) one not yet falsified. (See Popper [9].) The assumptions (1)–(3) are not directly testable, but they have testable implications. Thus they can be falsified if any of the implications are found not to hold, which would mean that the optimality of the chain ladder method could not be shown for the data in question. Holding up under all of these tests would increase the actuary's confidence in the hypothesis, still recognizing that no hypothesis can ever be fully verified. Some of the testable implications are:

1. Significance of factor $f(d)$.
2. Superiority of factor assumption to alternative emergence patterns such as:
    (a) linear with constant: $E[q(w,d+1) \mid \text{data to } w+d] = f(d)c(w,d) + g(d)$;
    (b) factor times parameter: $E[q(w,d+1) \mid \text{data to } w+d] = f(d)h(w)$;
    (c) including calendar year effect: $E[q(w,d+1) \mid \text{data to } w+d] = f(d)h(w)g(w+d)$.

Note that in these examples the notation has changed slightly so that $f(d)$ is a factor used to estimate $q(w, d+1)$, but not necessarily applied to $c(w,d)$. These alternative emergence models can be tested by goodness of fit, controlling for number of parameters.

3. Linearity of model: look at residuals as a function of $c(w,d)$.
4. Stability of factor: look at residuals as a function of time.
5. No correlation among columns.
6. No particularly high or low diagonals.

The remainder of this paper consists of tests of these implications.

TESTING LOSS EMERGENCE—IMPLICATIONS 1 & 2

The first four of these implications are tests of assumption (1). Standard diagnostic tests for weighted least-squares regression can be used as measures.

Implication 1: Significance of Factors

Regression analysis produces estimates for the standard deviation of each parameter estimated. Usually the absolute value of a factor is required to be at least twice its standard deviation for the factor to be regarded as significantly different from zero. This is a test failed by many development triangles, which means that the chain ladder method is not optimal for those triangles.

The requirement that the factor be twice the standard deviation is not a strict statistical test, but more like a level of comfort. For the normal distribution this requirement provides that there is only a probability of about 4.5% of getting a factor of this absolute value or greater when the true factor is zero. Many analysts are comfortable with a factor with absolute value 1.65 times its standard deviation, which could happen about 10% of the time by chance alone. For heavier-tailed distributions, the same ratio of factor to standard deviation will usually be more likely to occur by chance. Thus, if a factor were to be considered not significant for the normal distribution, it would probably be even less significant for other distributions. This approach could be made into a formal statistical test by finding the distribution that the factors follow. The normal distribution is often satisfactory, but it is not unusual to see some degree of positive skewness, which would suggest the lognormal. Some of the alternative models discussed below are easier to estimate in log form, so that is not an unhappy finding.

It may be tempting to do the regression of cumulative on previous cumulative and test the significance of that factor in order to justify the use of the chain ladder. However it is only the incrementals that are being predicted, so this would have to be carefully interpreted. In a cumulative-to-cumulative regression, the significance of the difference of the factor from unity is what needs to be tested. This can be done by comparing that difference to the standard deviation of the factor, which is equivalent to testing the significance of the factor in the incremental-to-cumulative regression. Some alternative methods to try when this assumption fails are discussed below.

Implication 2: Superiority to Alternative Emergence Patterns

If alternative emergence patterns give a better explanation of the data triangle observed to date, then assumption (1) of the chain ladder model is also suspect. In these cases development based on the best-fitting emergence pattern would be a natural option to consider. The sum of the squared errors (SSE) would be a way to compare models (the lower the better) but this should be adjusted to take into account the number of parameters used. Unfortunately it appears that there is no generally accepted method to make this adjustment. One possible adjustment is to compare fits by using the SSE divided by $(n-p)^2$, where n is the number of observations and p is the number of parameters. More parameters give an advantage in fitting but a disadvantage in prediction, so such a penalty in adjusting the residuals may be appropriate. A more popular adjustment in recent years is to base goodness of fit on the Akaike Information Criterion, or AIC (see Lütkepohl [5]). For a fixed set of observations, multiplying the SSE by $e^{2p/n}$ can approximate the effect of the AIC. The AIC has been criticized as being too permissive of over-parameterization for large data sets, and the Bayesian Information Criterion, or BIC, has been suggested as an alternative. Multiplying the SSE by $n^{p/n}$ would rank models the same as the BIC. As a comparison, if you have 45 observations, the improvement in SSE needed to justify adding a 5th parameter to a 4 parameter model is about 5%, $4\frac{1}{2}$%, and almost 9%, respectively, for these three adjustments. In the model testing below the sum of squared residuals divided by $(n-p)^2$ will be the test statistic, but in general the AIC and BIC should be regarded as good alternatives.

Note again that this is not just a test of development methods but is also a test to see what hypothetical loss generation process is most consistent with the data in the triangle.

The chain ladder has one parameter for each age, which is less than for the other emergence patterns listed in implication 2. This gives it an initial advantage, but if the other parameters improve the fit enough, they overcome this advantage. In testing the various patterns below, parameters will be fit by minimizing the sum of squared residuals. In some cases this will require an iterative procedure.

Alternative Emergence Pattern 1: Linear with Constant

The first alternative mentioned is just to add a constant term to the model. This is often significant in the age 0 to age 1 stage, especially for highly variable and slowly reporting lines, such as excess reinsurance. In fact, in the experience of myself and other actuaries who have reported informally, the constant term has often been found to be more statistically significant than the factor itself. If the constant is significant and the factor is not, a different development process is indicated. For instance in some triangles earning of additional exposure could influence the 0-to-1 development. It is important in such cases to normalize the triangle as much as possible, e.g., by adjusting for differences among accident years in exposure and cost levels (trend). With these adjustments a purely additive rather than a purely multiplicative method could be more appropriate.

Again, the emergence assumption underlying the linear with constant method is:

$$\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)c(w,d) + g(d).$$

If the constant is statistically significant, this emergence pattern is more strongly supported than that underlying the chain ladder.

Alternative Emergence Pattern 2: Factor Times Parameter

The chain ladder model expresses the next period's loss emergence as a factor times losses emerged so far. An important alternative, suggested by Bornhuetter and Ferguson (BF) in 1972, is to forecast the future emergence as a factor times estimated ultimate losses. While BF use some external measure of ultimate losses in this process, others have tried to use the data triangle itself to estimate the ultimate (e.g., see Verrall [13]). In this paper, models that estimate emerging losses as a percent of ultimate will be called parameterized BF models, even if they differ from the original BF method in how they estimate the ultimate losses.

The emergence pattern assumed by the parameterized BF model is:

$$\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)h(w).$$

That is, the next period expected emerged loss is a lag factor $f(d)$ times an accident year parameter $h(w)$. The latter could be interpreted as expected ultimate for the year, or at least proportional to that. This model thus has a parameter for each accident year as well as for each age (one less actually, as you can assume the $f(d)$'s sum to one—which makes $h(w)$ an estimate of ultimate losses; thus multiplying all the $f(d)$'s, $d > 0$, by a constant and dividing all the h's by the same constant will not change the forecasts). For reserving purposes there is even one fewer parameter, as the age 0 losses are already in the data triangle, so $f(0)$ is not needed. Thus, for a complete triangle with n accident years the BF has $2n-2$ parameters, or twice the number as the chain ladder. This will result in a penalty to goodness of fit, so the BF has to produce much lower fit errors than the chain ladder to give a better test statistic.

Testing the parameterized BF emergence pattern against that of the chain ladder cannot be done just by looking at the statistical significance of the parameters, as it could with the linear plus constant method, as one is not a special case of the other. This testing is the role of the test statistic, the sum of squared residuals divided by the square of the degrees of freedom. If this statistic is better for the BF model, that is evidence that the emergence pattern of the BF is more applicable to the triangle being studied. That would suggest that loss emergence for that book can be more accurately represented as fluctuating around a proportion of ultimate losses rather than a percentage of previously emerged losses.

Stanard [10] assumed a loss generation scheme that resulted in the expected loss emergence for each period being proportional to the ultimate losses for the period. This now can be seen to be the BF emergence pattern. Then by generating actual loss emergence stochastically, he tested some loss development methods. The chain ladder method gave substantially larger estimation errors for ultimate losses than his other methods, which were basically different versions of BF estimation. This illustrates how far off reserves can be when one reserving technique is applied to losses that have an emergence process different from the one underlying the technique.

A simulation in accord with the chain ladder emergence assumption would generate losses at age j by multiplying the simulated emerged losses at age $j-1$ by a factor and then adding a random component. In this manner the random components influence the expected emergence at all future ages. This may seem an unlikely way for losses to emerge, but it is for the triangles that follow this emergence pattern that the chain ladder will be optimal. The fact that Stanard used the simulation method consistent with the BF emergence pattern, and this was not challenged by the reviewer, John Robertson, suggests that actuaries may be more comfortable with the BF emergence assumptions than with those of the chain ladder. Or perhaps it just means that no one would be likely to think of simulating losses by the chain ladder method.

An important special case of the parameterized BF was developed by some Swiss and American reinsurance actuaries at a meeting in Cape Cod, and is sometimes called the Cape Cod method (CC). It is given by setting $h(w)$ to just a single h for all accident years. CC seems to have one more parameter than the chain ladder, namely h. However, any change in h can be offset by inverse changes in all the f's. CC thus has the same number of parameters as the chain ladder, and so its fit measure is not as heavily penalized as that of BF. However a single h requires a relatively stable level of loss exposure across accident years. Again it would be necessary to adjust for known exposure and price level differences among accident years, if using this method. The chain ladder and BF can handle changes in level from year to year as long as the development pattern remains consistent.

The BF model often has too many parameters. The last few accident years especially are left to find their own levels based on sparse information. Reducing the number of parameters, and thus using more of the information in the triangle, can often yield better predictions, especially in predicting the last few years. It could be that losses follow the BF emergence pattern, but this is disguised in the test statistic due to too many parameters. Thus, testing for the alternate emergence pattern should also include testing reduced parameter BF models.

The full BF not only assumes that losses emerge as a percentage of ultimate, but also that the accident years are all at different mean levels and that each age has a different percentage of ultimate losses. It could be, however, that several years in a row, or all of them, have the same mean level. If the mean changes, there could be a gradual transition from one level to another over a few years. This could be modeled as a linear progression of accident year parameters, rather than separate parameters for each year. A similar process could govern loss emergence. For instance, the 9th through 15th periods could all have the same expected percentage development. Finding these relationships and incorporating them in the fitting process will help determine what emergence process is generating the development.

The CC model can be considered a reduced parameter BF model. The CC has a single ultimate value for all accident years, while the BF has a separate value for each year. There are numerous other ways to reduce the number of parameters in BF models. Simply using a trend line through the BF ultimate loss parameters would use just two accident year parameters in total instead of one for each year. Another method might be to group years using apparent jumps in loss levels and fit an h parameter separately to each group. Within such groupings it is also possible to let each accident year's h parameter vary somewhat from the group average, e.g., via credibility, or to let it evolve over time, e.g., by exponential smoothing.

Alternative Emergence Patterns Example

Table 1 shows incremental incurred losses by age for some excess casualty reinsurance. As an initial test, the statistical sig-

TABLE 1

INCREMENTAL INCURRED LOSSES

	Age									
Year	0	1	2	3	4	5	6	7	8	9
0	5,012	3,257	2,638	898	1,734	2,642	1,828	599	54	172
1	106	4,179	1,111	5,270	3,116	1,817	−103	673	535	
2	3,410	5,582	4,881	2,268	2,594	3,479	649	603		
3	5,655	5,900	4,211	5,500	2,159	2,658	984			
4	1,092	8,473	6,271	6,333	3,786	225				
5	1,513	4,932	5,257	1,233	2,917					
6	557	3,463	6,926	1,368						
7	1,351	5,596	6,165							
8	3,133	2,262								
9	2,063									

TABLE 2

STATISTICAL SIGNIFICANCE OF FACTORS

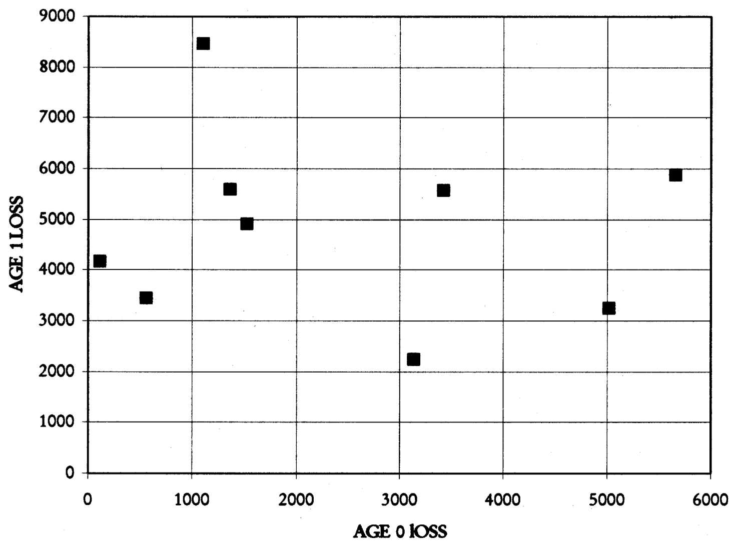
	0 to 1	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8
a	5,113	4,311	1,687	2,061	4,064	620	777	3,724
Std. Dev. a	1,066	2,440	3,543	1,165	2,242	2,301	145	0.000
b	−0.109	0.049	0.131	0.041	−0.100	0.011	−0.008	−0.197
Std. Dev. b	0.349	0.309	0.283	0.071	0.114	0.112	0.008	0.000

nificance of the factors was tested by regression of incremental losses against the previous cumulative losses. In the regression the constant is denoted by a and the factor by b. This provides a test of implication 1—significance of the factor, and also one test of implication 2—alternative emergence patterns. In this case the alternative emergence patterns tested are factor plus constant and constant with no factor. Here they are being tested by looking at whether or not the factors and the constants are significantly different from zero, rather than by any goodness-of-fit measure.

Table 2 shows the estimated parameters and their standard deviations. As can be seen, the constants are usually statistically

FIGURE 1

AGE 1 VS. AGE 0 LOSSES



significant (parameter nearly double its standard deviation, or more), but the factors never are. The chain ladder assumes the incremental losses are proportional to the previous cumulative, which implies that the factor is significant and the constant is not. The lack of significance of the factors and the significance of many of the constants both suggest that the losses to emerge at any age $d+1$ are not proportional to the cumulative losses through age d. The assumptions underlying the chain ladder model are thus not supported by this data. A constant amount emerging for each age usually appears to be a reasonable estimator, however.

Figure 1 illustrates this. A factor by itself would be a straight line through the origin with slope equal to the development factor, whereas a constant would give a horizontal line at the height of the constant. As an alternative, the parameterized BF model was fit to the triangle. As this is a non-linear model, fitting is a little more involved. A statistical package that includes non-linear regression could ease the estimation. A method of fitting the parameters without such a package will be discussed, followed by an analysis of the resulting fit.

To do the fitting, an iterative method can be used to minimize the sum of the squared residuals, where the (w,d) residual is $[q(w,d) - f(d)h(w)]$. Weighted least squares could also be used if the variances of the residuals are not constant over the triangle. For instance, the variances could be proportional to $f(d)^p h(w)^q$ for some values of p and q, usually 0, 1, or 2, in which case the regression weights would be $1/f(d)^p h(w)^q$.

A starting point for the f's or the h's is needed to begin the iteration. While almost any reasonable values could be used, such as all f's equal to $1/n$, convergence will be faster with values likely to be in the ballpark of the final factors. A natural starting point thus might be the implied $f(d)$'s from the chain ladder method. For ages greater than 0, these are the incremental age-to-age factors divided by the cumulative-to-ultimate factors. To get a starting value for age 0, subtract the sum of the other factors from unity. Starting with these values for $f(d)$, regressions were performed to find the $h(w)$'s that minimize the sum of squared residuals (one regression for each w). These give the best h's for that initial set of f's. The standard linear regression formula for these h's simplifies to:

$$h(w) = \sum_d f(d)q(w,d) \Big/ \sum_d f(d)^2.$$

Even though that gives the best h's for those f's, another regression is needed to find the best f's for those h's. For this step the usual regression formula gives:

$$f(d) = \sum_w h(w)q(w,d) \Big/ \sum_w h(w)^2.$$

TABLE 3

BF PARAMETERS

Age d	0	1	2	3	4	5	6	7	8	9
$f(d)$ 1st	0.106	0.231	0.209	0.155	0.117	0.083	0.038	0.032	0.018	0.011
$f(d)$ ult.	0.162	0.197	0.204	0.147	0.115	0.082	0.037	0.030	0.015	0.009
Year w	0	1	2	3	4	5	6	7	8	9
$h(w)$ 1st	17,401	15,729	23,942	26,365	30,390	19,813	18,592	24,154	14,639	12,733
$h(w)$ ult.	15,982	16,501	23,562	27,269	31,587	20,081	19,032	25,155	13,219	19,413

Now the h regression can be repeated with the new f's, etc. This process continues until convergence occurs, i.e., until the f's and h's no longer change with subsequent iterations. It may be possible that this procedure would converge to a local rather than the global minimum, which can be tested by using other starting values.

Ten iterations were used in this case, but substantial convergence occurred earlier. The first round of f's and h's and those at convergence are in Table 3. Note that the h's are not the final estimates of the ultimate losses, but are used with the estimated factors to estimate future emergence. In this case, in fact, $h(0)$ is less than the emerged to date. As the h's are unique only up to a constant of proportionality, which can be absorbed by the f's, it may improve presentations to set $h(0)$ to the estimated ultimate losses for year 0.

Standard regression assumes each observation q has the same variance, which is to say the variance is proportional to $f(d)^p h(w)^q$, with $p = q = 0$. If $p = q = 1$ the weighted regression formulas become:

$$h(w)^2 = \sum_d [q(w,d)^2/f(d)] \Big/ \sum_d f(d) \qquad \text{and}$$

$$f(d)^2 = \sum_w [q(w,d)^2/h(w)] \Big/ \sum_w h(w).$$

TABLE 4

DEVELOPMENT FACTORS

					Incremental				
Prior	0 to 1	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8	8 to 9
	1.22	0.57	0.26	0.16	0.10	0.04	0.03	0.02	0.01
					Ultimate				
	0 to 9	1 to 9	2 to 9	3 to 9	4 to 9	5 to 9	6 to 9	7 to 9	8 to 9
	6.17	2.78	1.77	1.41	1.21	1.10	1.06	1.03	1.01
					Incremental/Ultimate				
0.162	0.197	0.204	0.147	0.115	0.082	0.037	0.030	0.015	0.009

For comparison, the development factors from the chain ladder are shown in Table 4. The incremental factors are the ratios of incremental to previous cumulative. The ultimate ratios are cumulative to ultimate. Below them are the ratios of these ratios, which represent the portion of ultimate losses to emerge in each period. The zeroth period shown is unity less the sum of the other ratios. These factors were the initial iteration for the $f(d)$s shown above.

Having now estimated the BF parameters, how can they be used to test what the emergence pattern of the losses is?

A comparison of this fit to that from the chain ladder can be made by looking at how well each method predicts the incremental losses for each age after the initial one. The SSE adjusted for number of parameters will be used as the comparison measure, where the parameter adjustment will be made by dividing the SSE by the square of the difference between the number of observations and the number of parameters, as discussed earlier. Here there are 45 observations, as only the predicted points count as observations. The adjusted SSE was 81,169 for the BF, and 157,902 for the chain ladder. This shows that the emergence pattern for the BF (emergence proportional to ultimate) is much more consistent with this data than is the chain ladder emergence pattern (emergence proportional to previous emerged).

TABLE 5

FACTORS IN CC METHOD

Age d	0	1	2	3	4	5	6	7	8	9
$f(d)$	0.109	0.220	0.213	0.148	0.124	0.098	0.038	0.028	0.013	0.008

The CC method was also tried for this data. The iteration proceeded similarly to that for the BF, but only a single h parameter was fit for all accident years. Now:

$$h = \sum_{w,d} f(d)q(w,d) \Big/ \sum_{w,d} f(d)^2.$$

This formula for h is the same as the formula for $h(w)$ except the sum is taken over all w. The estimated h is 22,001, and the final factors f are shown in Table 5. The adjusted SSE for this fit is 75,409. Since the CC is a special case of the BF, the unadjusted SSE is necessarily worse than that of the BF method (in this case 59M vs. 98M), but with fewer parameters in the CC, the adjustment makes them similar. These are close enough that which is better depends on the adjustment chosen for extra parameters. The BIC also favors the CC, but the AIC is better for the BF. As is often the case, the statistics can inform decision-making but not determine the decision.

Intermediate special cases could be fit similarly. If, for instance, a single factor were sought to apply to just two accident years, the sum would be taken over those years to estimate that factor, etc.

This is a case where the BF has too many parameters for prediction purposes. More parameters fit the data better but use up information. The penalty in the fit measure adjusts for this problem, and the penalty used finds the CC to be a somewhat better model. Thus the data is consistent with random emergence around an expected value that is constant over the accident years.

TABLE 6

TERMS IN ADDITIVE CHAIN LADDER

Age d	1	2	3	4	5	6	7	8	9
$g(d)$	4,849.3	4,682.5	3,267.1	2,717.7	2,164.2	839.5	625.0	294.5	172.0

Again, the CC method would probably work even better for loss ratio triangles than for loss triangles, as then a single target ultimate value makes more sense. Adjusting loss ratios for trend and rate level could increase this homogeneity.

In addition, an additive development was tried, as suggested by the fact that the constant terms were significant in the original chain ladder, even though the factors were not. The development terms are shown in Table 6. These are just the average loss emerged at each age. The adjusted sum of squared residuals is 75,409. This is much better than the chain ladder, which might be expected, as the constant terms were significant in the original significance-test regressions while the factors were not. The additive factors in Table 6 differ from those in Table 2 because there is no multiplicative factor in Table 6.

Is it a coincidence that the additive chain ladder gives the same fit accuracy as the CC? Not really, in that they both estimate each age's loss levels with a single value. Let $g(d)$ denote the additive development amount for age d. As the notation suggests, this does not vary by accident year. The CC method fits an overall h and a factor $f(d)$ for each age such that the estimated emergence for age d is $f(d)h$. Here too the predicted development varies by age but is a constant for each accident year. If you have estimated the CC parameters you can just define $g(d) = f(d)h$. Alternatively, if the additive method has been fit, no matter what h is estimated, the f's can be defined as $f(d)h = g(d)$. As long as the parameters are fit by least-squares they have to come out the same: if one came out lower, you could have used the equations in the two previous sentences to get this same lower value for

TABLE 7

BF-CC PARAMETERS

Age d	0	1	2	3	4	5	6	7	8	9
$f(d)$	$*$	0.230	0.230	0.160	0.123	0.086	0.040	0.040	0.017	0.017
Year w	0	1	2	3	4	5	6	7	8	9
$h(w)$	14,829	14,829	20,962	25,895	30,828	20,000	20,000	20,000	20,000	20,000

the other. The two models have the same age and accident year relationships and so will always come out the same when fit by least-squares. They are defined differently, however, and so other methods of estimating the parameters may come up with separate estimates, as in Stanard [10]. In the remainder of this paper, the models will be used interchangeably.

Finally, an intermediate BF-CC pattern was fit as an example of the possible approaches of this type. In this case ages 1 and 2 are assumed to have the same factor, as are ages 6 and 7 and ages 8 and 9. This reduces the number of f parameters from 9 to 6. The number of accident year parameters was also reduced: years 0 and 1 have a single parameter, as do years 5 through 9. Year 2 has its own parameter, as does year 4, but year 3 is the average of those two. Thus there are 4 accident year parameters, and so 10 parameters in total. Any one of these can be set arbitrarily, with the remainder adjusted by a factor, so there are really just 9. The selections were based on consideration of which parameters were likely not to be significantly different from each other.

The estimated factors are shown in Table 7. The factor to be set arbitrarily was the accident year factor for the last 5 years, which was set to 20,000. The other factors were estimated by the same iterative regression procedure as for the BF, but the factor constraints change the simplified regression formula. The adjusted sum of squared residuals is 52,360, which makes it the best approach tried. This further supports the idea that claims emerge as a percent of ultimate for this data. It also indicates that the various accident years and ages are not all at different levels. The actual and fitted values from this, the chain ladder, and CC are in Exhibit 1. The fitted values in Exhibit 1 were calculated as follows. For the chain ladder, the factors from Table 4 were applied to the cumulative losses implied from Table 1. For the CC the fitted values are just the terms in Table 6. For the BF-CC they are the products of the appropriate f and h factors from Table 7. The parameters for all the models to this point are summarized in Exhibit 2.

Alternative Emergence Patterns-Summary

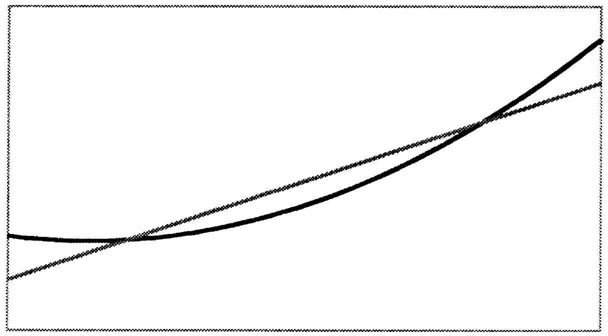
The chain ladder assumes that future emergence for an accident year will be proportional to losses emerged to date. The BF methods take expected emergence in each period to be a percentage of ultimate losses. This could be interpreted as regarding the emerged to date to have a random component that will not influence future development. If this is the actual emergence pattern, the chain ladder method will apply factors to the random component, and thus increase the estimation error.

The CC and additive chain ladder methods assume in effect that years showing low losses or high losses to date will have the same expected future dollar development. Thus a bad loss year may differ from a good one in just a couple of emergence periods, and have quite comparable loss emergence in all other periods. The chain ladder and the most general form of the BF, on the other hand, assume that a bad year will have higher emergence than a good year in most periods.

The BF and chain ladder emergence patterns are not the only ones that make sense. Some others will be reviewed when discussing diagonal effects below.

Which emergence pattern holds for a given triangle is an empirical issue. Fitting parameters to the various methods and looking at the significance of the parameters and the adjusted sum of squared residuals can test this.

FIGURE 2



RESIDUAL ANALYSIS—TESTING IMPLICATIONS 3 & 4

So far the first two of the six testable implications of the chain ladder assumptions have been addressed. Looking at the residuals from the fitting process can test the next two implications.

Implication 3: Test of Linearity—Residuals as Function of Previous

Figure 2 shows a straight line fit to a curve. The residuals can be seen to be first positive, then negative then all positive. This pattern of residuals is indicative of a non-linear process with a linear fit. The chain ladder model assumes the incremental losses at each age are a linear function of the previous cumulative losses.

A scatter plot of the incremental against the previous cumulative, as in Figure 3, can be used to check linearity; looking for this characteristic non-linear pattern (i.e., strings of positive and negative residuals) in the residuals plotted against the previous cumulative is equivalent. This can be tested for each age to see if a non-linear process may be indicated. Finding this would suggest that emergence is a non-linear function of losses to date. In

FIGURE 3

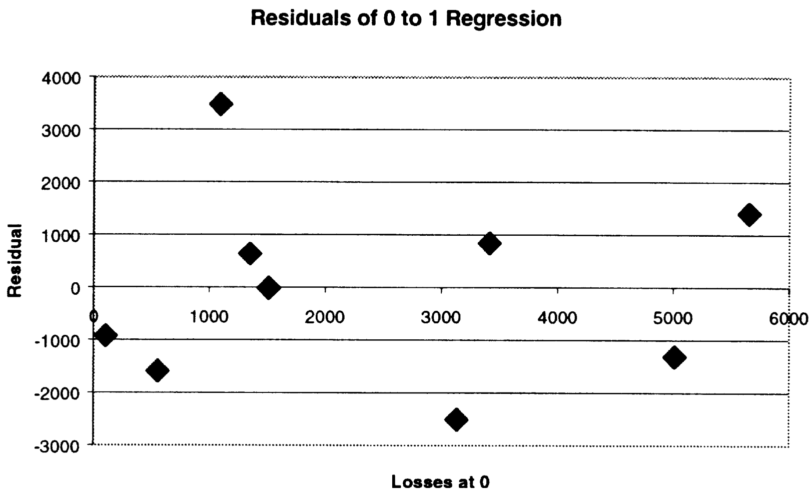


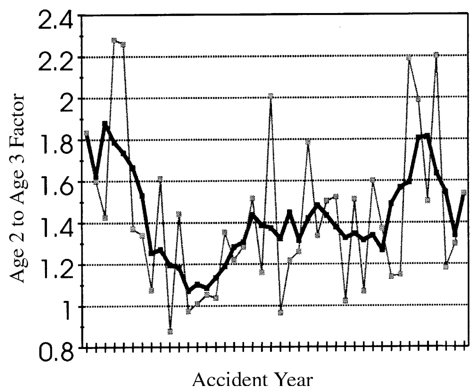
Figure 3 there are no apparent strings of consecutive positive or negative residuals, so non-linearity is not indicated.

Implication 4: Test of Stability—Residuals Over Time

If a similar pattern of sequences of high and low residuals is found when plotted against time, instability of the factors may be indicated. If the factors appear to be stable over time, all the accident years available should be used to calculate the development factors, in order to reduce the effects of random fluctuations. When the development process is unstable, the assumptions for optimality of the chain ladder are no longer satisfied. A response to unstable factors over time might be to use a weighted average of the available factors, with more weight going to the more recent years, e.g., just use the last 5 diagonals. A weighted average should be used when there is a good reason for it, e.g., when residual analysis shows that the factors are changing, but otherwise it will increase estimation errors by over-emphasizing some observations and under-emphasizing others.

FIGURE 4

2ND TO 3RD FIVE-TERM MOVING AVERAGE



Another approach to unstable development would be to adjust the triangle for measurable instability. For instance, Berquist and Sherman [1] suggest testing for instability by looking for changes in the settlement rate of claims. They measured this by looking at the changes in the percentage of claims closed by age. If instability is found, the triangle is adjusted to the latest pattern. The adjusted triangle, however, should still be tested for stability of development factors by residual analysis and as illustrated below.

Figure 4 shows the 2nd to 3rd factor by accident year from a large development triangle (data in Exhibit 3) along with its five-term moving average. The moving average is the more stable of the two lines, and is sometimes in practice called "the average of the last five diagonals." There is apparent movement of the factor over time as well as a good deal of random fluctuation. There is a period of time in which the moving average is as low as 1.1 and other times it is as high as 1.8. This is the kind of variability that would suggest using the average of recent diagonals instead of the entire triangle when estimating factors. This is not suggested due to the large fluctuations in factors, but rather because of the changes over time in the level around which the factors are fluctuating. A lot of variability around a fixed level would in fact suggest using all the data.

It is not clear from the data what is causing the moving average factors to drift over time. Faced with data like this, the average of all the data would not normally be used. Grouping accident years or taking weighted averages would be useful alternatives.

The state-space model in the Verall and Zehnwirth references provides a formal statistical treatment of the types of instability in a data triangle. This model can be used to help analyze whether to use all the data, or to adopt some form of weighted average that de-emphasizes older data. It is based on comparing the degree of instability of observations around the current mean to the degree of instability in the mean itself over time. While this is the main statistical model available to determine weights to apply to the various accident years of data, a detailed discussion is beyond the scope of this paper.

INDEPENDENCE—TESTING IMPLICATIONS 5 & 6

Implications 5 and 6 have to do with independence within the triangle. Mack's second assumption above is that, except for observations in the same accident year, the columns of incremental losses need to be independent. He developed a correlation test and a high-low diagonal test to check for dependencies. The data may have already been adjusted for known changes in the case reserving process. For instance, Berquist and Sherman recommend looking at the difference between paid and incurred case severity trends to determine if there has been a change in case reserve adequacy, and if there has, adjusting the data accordingly. Even after such adjustments, however, correlations may exist within the triangle.

TABLE 8

SAMPLE CORRELATION = $-1.35/(146.37 \times 0.20)^{1/2} = -.25$

Year	X = 0 to 1	Y = 1 to 2	$(X - \mathrm{E}[X])^2$	$(Y - \mathrm{E}[Y])^2$	$(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])$
1	0.65	0.32	54.27	0.14	2.78
2	39.42	0.26	986.46	0.19	−13.71
3	1.64	0.54	40.70	0.02	0.98
4	1.04	0.36	48.63	0.11	2.31
5	7.76	0.66	0.07	0.00	0.01
6	3.26	0.82	22.63	0.01	−0.57
7	6.22	1.72	3.24	1.05	−1.85
8	4.14	0.89	15.01	0.04	−0.74
Average	8.02	0.70	146.37	0.20	−1.35

Implication 5: Correlation of Development Factors

Mack developed a correlation test for adjacent columns of a development factor triangle. If a year of high emergence tends to follow one with low emergence, then the development method should take this into account. Another correlation test would be to calculate the sample correlation coefficients for all pairs of columns in the triangle, and then see how many of these are statistically significant, say at the 10% level. The sample correlation for two columns is just the sample covariance divided by the product of the sample standard deviations for the first n elements of both columns, where n is the length of the shorter column. The sample correlation calculation in Table 8 shows that for the triangle in Table 1 above, the correlation of the first two development factors is -25%.

Letting r denote the sample correlation coefficient, define $T = r[(n-2)/(1-r^2)]^{1/2}$. A significance test for the correlation coefficient can be made by considering T to be t-distributed with $n-2$ degrees of freedom. If T is greater than the t-statistic for 0.9 at $n-2$ degrees of freedom, for instance, then r can be considered significant at the 10% level. (See Miller and Wichern [7, p. 214].)

In this example, $T = -0.63$, which is not significant even at the 10% level. This level of significance means that 10% of the pairs of columns could show up as significant just by random happenstance. A single correlation at this level would thus not be a strong indicator of correlation within the triangle. If several columns are correlated at the 10% level, however, there may be a correlation problem.

To test this further, if m is the number of pairs of columns in the triangle, the number that display significant correlation could be considered a binomial variate in m and 0.1, which has standard deviation $0.3m^{1/2}$. Thus more than $0.1m + m^{1/2}$ significant correlations (mean plus 3.33 standard deviations) would strongly suggest there is actual correlation within the triangle. Here the 10% level and 3.33 standard deviations were chosen for illustration. A single correlation that is significant at the 0.1% level would also be indicative of a correlation problem, for example.

If there is such correlation, the product of development factors is not unbiased, but the relationship $\mathrm{E}[XY] = (\mathrm{E}[X])(\mathrm{E}[Y]) + \mathrm{Cov}(X,Y)$ could be used to correct the product, where here X and Y are development factors.

Implication 6: Significantly High or Low Diagonals

Mack's high-low diagonal test counts the number of high and low factors on each diagonal, and tests whether or not that is likely to be due to chance. Here another high-low test is proposed: use regression to see if any diagonal dummy variables are significant. This test also provides alternatives in case the pure chain ladder is rejected. An actuary will often have information about changes in company operations that may have created a diagonal effect. If so, this information could lead to choices of modeling methods—e.g., whether to assume the effect is permanent or temporary. The diagonal dummies can be used to measure the effect in any case, but knowledge of company operations will help determine how to use this effect. This is particularly so if the effect occurs in the last few diagonals.

A diagonal in the loss development triangle is defined by $w + d$ = constant. Suppose for some given data triangle, the diagonal $w + d = 7$ has been estimated to be 10% higher than normal. Then an adjusted BF estimate of a cell might be:

$$q(w,d) = 1.1f(d)h(w) \qquad \text{if} \quad w + d = 7, \qquad \text{and}$$

$$q(w,d) = f(d)h(w) \qquad \text{otherwise.}$$

This is an example of a multiplicative diagonal effect. Additive diagonal effects can also be estimated, using regression with diagonal dummies.

	Age			
Year	0	1	2	3
1	2	5	4	
3	8	9		
7	10			
7				

Incr. Ages 1–3	Cum. Age 0	Cum. Age 1	Cum. Age 2	Dummy 1	Dummy 2
2	1	0	0	0	0
8	3	0	0	1	0
10	7	0	0	0	1
5	0	3	0	1	0
9	0	11	0	0	1
4	0	0	8	0	1

The small sample triangle of incremental losses here will be used as an example of how to set up diagonal dummies in a chain ladder model. The goal is to get a matrix of data in the form needed to do a multiple regression. First the triangle (except the first column) is strung out into a column vector. This is the dependent variable, and forms the first column of the matrix above. Then columns for the independent variables are added. The second column is the cumulative losses at age 0 corresponding to the loss entries that are at age 1, and zero for the other loss entries. The regression coefficient for this column would be the 0 to 1 cumulative-to-incremental factor. The next two columns are cumulative losses at age 1 and age 2 corresponding to the age 2 and age 3 data in the first column. The last two columns are the diagonal dummies. They pick out the elements of the last two diagonals. The coefficients for these columns would be additive adjustments for those diagonals, if significant.

This method of testing for diagonal effects is applicable to many of the emergence models. In fact, if diagonal effects are found to be significant in chain ladder models, they probably are needed in the BF models of the same data. Thus tests of the chain ladder vs. BF should be done with the diagonal elements included. Some examples are given in the Appendix. Another popular modeling approach is to consider diagonal effects to be a measure of inflation (e.g., see Taylor [11]). In a payment triangle this would be a natural interpretation, but a similar phenomenon could occur in an incurred triangle. In this case the latest diagonal effects might be projected ahead as estimates of future inflation. An understanding of the aspects of company operations that drive the diagonal effects would help address these issues.

This approach incorporates diagonal effects right into the emergence model. For instance, an emergence model might be:

$$\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)g(w+d).$$

Here $g(w+d)$ is a diagonal effect, but every diagonal has such a factor. The usual interpretation is that g measures the cumulative claims inflation applicable to that diagonal since the first accident year. It would even be possible to add accident year effects $h(w)$ as well, e.g.,

$$\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)h(w)g(w+d).$$

There are clearly too many parameters here, but a lot of them might reasonably be set equal. For instance, the inflation might be the same for several years, or several accident years might be at the same level. Note that since g is cumulative inflation, a constant inflation level could be achieved by setting $g(w+d) = (1+j)^{w+d}$. Then j is the only inflation parameter to be estimated.

The age and accident year parameters might also be able to be written as trends rather than individual factors. If $f(d) = (1+i)^d$ and $h(w) = h \times (1+k)^w$, then the model reduces to four parameters h, i, j, and k. However it would be more usual to need more parameters than this, possibly written as changing trends. That is, i, j, and k might be constant for some periods, then change for others. Note that if they are constant for all periods, the estimator $h(1+i)^d(1+j)^{w+d}(1+k)^w$ is $h(1+i+j+ij)^d(1+k+j+jk)^w$, which eliminates the parameter j, as i becomes $i+j+ij$ and k becomes $k+j+jk$.

It might be better to start without the accident year trend and keep the calendar year trend, especially if the triangle has been normalized for accident year changes. The model for the (w,d) cell would then be $h(1+i)^d(i+j)^{w+d}$, which has just three parameters.

As with the BF model, the parameters of models with diagonal trends can be estimated iteratively. With reasonable starting values, fix two of the three sets of parameters, and fit the third by least squares, and rotate until convergence is reached. Alternatively, a non-linear search procedure could be utilized. As an example of the simplest of these approaches, modeling $\mathrm{E}[q(w,d+1) \mid \text{data to } w+d]$ as just $6{,}756(0.7785)^d$ gives an adjusted sum of squares of 57,527 for the reinsurance triangle above. This is not the best fitting model, but it is better than some and has only two parameters $h = 6{,}756$ and $i = -0.2215$.

Calendar year trend accounts for inflation in the time between loss occurrence and loss settlement, which many actuaries believe has an impact on ultimate losses. Whether it is influencing a given loss triangle can be investigated by testing for diagonal effects.

CONCLUSION

The first test that will quickly indicate the general type of emergence pattern faced is the test of significance of the cumulative-to-incremental factors at each age. This is equivalent to testing if the cumulative-to-cumulative factors are significantly different from unity. When this test fails, the future emergence is not proportional to past emergence. It may be a constant amount, or it may be proportional to ultimate losses, as in the BF pattern.

When this test is passed, the addition of an additive component may give an even better fit. Even when the test is failed, including an additive term may make the factor significant. In either case the BF emergence pattern may still produce a better fit. Reduced parameter BF models could also give better performance, as they will be less responsive to random variation. If an additive component is significant, then converting the triangle to on-level loss ratios may improve the forecasts.

Tests of stability and for diagonal effects may lead to further improvements in the model. However, if the emergence is stable, excluding data by using only the last n diagonals will lead to higher estimation errors on average.

An actuary might advise: "If the chain ladder doesn't work, try Bornhuetter-Ferguson." This is a reasonable conclusion, with the interpretation of "doesn't work" to mean "fails the assumptions of least-squares optimality," and "try" to mean "test the underlying assumptions of."

REFERENCES

[1] Berquist, James R., and Richard E. Sherman, "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," *PCAS* LXIV, 1977, pp. 123–184.

[2] Bornhuetter, Ronald L., and R. E. Ferguson, "The Actuary and IBNR," *PCAS* LIX, 1972, pp. 181–195.

[3] Gerber and Jones, "Credibility Formulas with Geometric Weights," *Society of Actuaries Transactions*, 1975.

[4] de Jong, P., and B. Zehnwirth, "Claims Reserving, State-space Models and the Kalman Filter," *Journal of the Institute of Actuaries*, 1983, pp. 157–181.

[5] Lütkepohl, *Introduction to Multiple Time Series Analysis*, Springer-Verlag, 1993.

[6] Mack, Thomas, "Measuring the Variability of Chain Ladder Reserve Estimates," Casualty Actuarial Society *Forum* 1, Spring 1994, pp. 101–182.

[7] Miller, Robert, and Dean Wichern, *Intermediate Business Statistics*, Holt, Rinehart and Winston, 1977.

[8] Murphy, Daniel M., "Unbiased Loss Development Factors," *PCAS* LXXXI, 1994, pp. 154–222.

[9] Popper, Karl R., *Conjectures and Refutations*, Routledge, 1969.

[10] Stanard, James N., "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques," *PCAS* LXXII, 1985, pp. 124–148.

[11] Taylor, Greg C., "Separation of Inflation and Other Effects from the Distribution of Non-Life Insurance Claim Delays," *ASTIN Bulletin* 9, 1977, pp. 219–230.

[12] Verrall, Richard, "A State Space Representation of the Chain Ladder Linear Model," *Journal of the Institute of Actuaries* 116, Part III, December 1989, pp. 589–610.

[13] Zehnwirth, Ben, "A Linear Filtering Theory Approach to Recursive Credibility Estimation," *ASTIN Bulletin* 15, April 1985, pp. 19–36.

[14] Zehnwirth, Ben, "Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals, and Risk Based Capital," Casualty Actuarial Society *Forum* 2, Spring 1994, pp. 447–606.

EXHIBIT 1

COMPARATIVE FITS

Chain Ladder

	1	2	3	4	5	6	7	8	9
Actual	3,257	2,638	898	1,734	2,642	1,828	599	54	172
Fit	6,101	4,705	2,846	1,912	1,350	656	580	296	172
% Error	87%	78%	217%	10%	−49%	−64%	−3%	448%	0%
Actual	4,179	1,111	5,270	3,116	1,817	−103	673	535	
Fit	129	2,438	1,408	1,728	1,374	632	499	257	
% Error	−97%	119%	−73%	−45%	−24%	−714%	−26%	−52%	
Actual	5,582	4,881	2,268	2,594	3,479	649	603		
Fit	4,151	5,116	3,619	2,614	1,868	900	736		
% Error	−26%	5%	60%	1%	−46%	39%	22%		
Actual	5,900	4,211	5,500	2,159	2,658	984			
Fit	6,883	6,574	4,113	3,444	2,336	1,057			
% Error	17%	56%	−25%	60%	−12%	7%			
Actual	8,473	6,271	6,333	3,786	225				
Fit	1,329	5,442	4,131	3,591	2,588				
% Error	−84%	−13%	−35%	−5%	1,050%				
Actual	4,932	5,257	1,233	2,917					
Fit	1,842	3,667	3,053	2,095					
% Error	−63%	−30%	148%	−28%					
Actual	3,463	6,926	1,368						
Fit	678	2,287	2,856						
% Error	−80%	−67%	109%						
Actual	5,596	6,165							
Fit	1,644	3,953							
% Error	−71%	−36%							
Actual	2,262								
Fit	3,814								
% Error	69%								

CC

	1	2	3	4	5	6	7	8	9
Actual	3,257	2,638	898	1,734	2,642	1,828	599	54	172
Fit	4,364	3,746	2,287	1,631	1,082	336	188	59	17
% Error	34%	42%	155%	−6%	−59%	−82%	−69%	9%	−90%
Actual	4,179	1,111	5,270	3,116	1,817	−103	673	535	
Fit	4,364	3,746	2,287	1,631	1,082	336	188	59	
% Error	4%	237%	−57%	−48%	−40%	−426%	−72%	−89%	
Actual	5,582	4,881	2,268	2,594	3,479	649	603		
Fit	4,364	3,746	2,287	1,631	1,082	336	188		
% Error	−22%	−23%	1%	−37%	−69%	−48%	−69%		
Actual	5,900	4,211	5,500	2,159	2,658	984			
Fit	4,364	3,746	2,287	1,631	1,082	336			
% Error	−26%	−11%	−58%	−24%	−59%	−66%			
Actual	8,473	6,271	6,333	3,786	225				

EXHIBIT 1

(CONTINUED)

Fit	4,364	3,746	2,287	1,631	1,082				
% Error	−48%	−40%	−64%	−57%	381%				
Actual	4,932	5,257	1,233	2,917					
Fit	4,364	3,746	2,287	1,631					
% Error	−12%	−29%	85%	−44%					
Actual	3,463	6,926	1,368						
Fit	4,364	3,746	2,287						
% Error	26%	−46%	67%						
Actual	5,596	6,165							
Fit	4,364	3,746							
% Error	−22%	−39%							
Actual	2,262								
Fit	4,364								
% Error	93%								
BF-CC									
	1	2	3	4	5	6	7	8	9
Actual	3,257	2,638	898	1,734	2,642	1,828	599	54	172
Fit	3,411	3,411	2,373	1,824	1,275	593	593	252	252
% Error	5%	29%	164%	5%	−52%	−68%	−1%	367%	47%
Actual	4,179	1,111	5,270	3,116	1,817	−103	673	535	
Fit	3,411	3,411	2,373	1,824	1,275	593	593	252	
% Error	−18%	207%	−55%	−41%	−30%	−676%	−12%	−53%	
Actual	5,582	4,881	2,268	2,594	3,479	649	603		
Fit	4,821	4,821	3,354	2,578	1,803	838	838		
% Error	−14%	−1%	48%	−1%	−48%	29%	39%		
Actual	5,900	4,211	5,500	2,159	2,658	984			
Fit	5,956	5,956	4,143	3,185	2,227	1,036			
% Error	1%	41%	−25%	48%	−16%	5%			
Actual	8,473	6,271	6,333	3,786	225				
Fit	7,090	7,090	4,932	3,792	2,651				
% Error	−16%	13%	−22%	0%	1,078%				
Actual	4,932	5,257	1,233	2,917					
Fit	4,600	4,600	3,200	2,460					
% Error	−7%	−12%	160%	−16%					
Actual	3,463	6,926	1,368						
Fit	4,600	4,600	3,200						
% Error	33%	−34%	134%						
Actual	5,596	6,165							
Fit	4,600	4,600							
% Error	−18%	−25%							
Actual	2,262								
Fit	4,600								
% Error	103%								

EXHIBIT 1

(CONTINUED)

Additive with Multiplicative Diagonals and Accident Years

	1	2	3	4	5	6	7	8	9
Actual	3,257	2,638	898	1,734	2,642	1,828	599	54	172
Fit	3,185	3,185	2,148	2,730	1,995	660	660	660	477
% Error	−2%	21%	139%	57%	−24%	−64%	10%	1,122%	177%
Actual	4,179	1,111	5,270	3,116	1,817	−103	673	535	
Fit	3,185	3,185	3,465	2,730	1,995	660	660	477	
% Error	−24%	187%	−34%	−12%	10%	−741%	−2%	−11%	
Actual	5,582	4,881	2,268	2,594	3,479	649	603		
Fit	4,036	6,508	4,390	3,460	2,529	836	604		
% Error	−28%	33%	94%	33%	−27%	29%	0%		
Actual	5,900	4,211	5,500	2,159	2,658	984			
Fit	6,508	6,508	4,390	3,460	2,529	604			
% Error	10%	55%	−20%	60%	−5%	−39%			
Actual	8,473	6,271	6,333	3,786	225				
Fit	5,136	5,136	3,465	2,730	1,442				
% Error	−39%	−18%	−45%	−28%	541%				
Actual	4,932	5,257	1,233	2,917					
Fit	5,136	5,136	3,465	1,972					
% Error	4%	−2%	181%	−32%					
Actual	3,463	6,926	1,368						
Fit	5,136	5,136	2,503						
% Error	48%	−26%	83%						
Actual	5,596	6,165							
Fit	5,136	3,710							
% Error	−8%	−40%							
Actual	2,262								
Fit	3,710								
% Error	64%								

EXHIBIT 2

SUMMARY OF PARAMETERS

	0	1	2	3	4	5	6	7	8	9
BF $f(d)$	0.162	0.197	0.204	0.147	0.115	0.082	0.037	0.030	0.015	0.009
BF $h(w)$	15,982	16,501	23,562	27,269	31,587	20,081	19,032	25,155	13,219	19,413
CC $f(d)$	0.109	0.220	0.213	0.148	0.124	0.098	0.038	0.028	0.013	0.008
Additive Chain	—	4,849.3	4,682.5	3,267.1	2,717.7	2,164.2	839.5	625.0	294.5	172.0
BF-CC $f(d)$	—	0.230	0.230	0.160	0.123	0.086	0.040	0.040	0.017	0.017
BF-CC $h(w)$	14,829	14,829	20,962	25,895	30,828	20,000	20,000	20,000	20,000	20,000

EXHIBIT 3

2ND TO 3RD FACTORS FROM LARGE TRIANGLE

2nd to 3rd→		1.81	1.60	1.41	2.29	2.25	1.38
	1.36	1.07	1.60	0.89	1.42	0.99	1.01
	1.03	1.02	1.35	1.21	1.28	1.51	1.17
	2.00	0.98	1.21	1.24	1.79	1.32	1.48
	1.51	1.01	1.51	1.06	1.60	1.10	1.11
	2.20	2.00	1.50	2.20	1.19	1.28	1.52

APPENDIX

DIAGONAL EFFECTS IN BF MODELS

As an example, a test for diagonal effects in the CC model was made in the reinsurance triangle as follows. The CC is the same as the additive chain ladder, so it can be expressed as a linear model. This can be estimated via a single multiple regression in which the dependent variable is the entire list of incremental losses for ages 1 to 9 and all accident years—45 items in all. That is, the triangle beyond age 0 is strung out into a single vector. Age and diagonal dummy independent variables can be established in a design matrix to pick out the right elements of the parameter vector of age and diagonal terms to estimate each incremental loss cell. For the additive chain ladder, the column dummy variables will be 1 or 0, as opposed to cumulative losses or 0 in the chain ladder example. Then the coefficient of that column will be the additive element for the given age.

The later columns of the design matrix would be diagonal dummies, as in the chain ladder example. By doing a multiple linear regression for the incremental loss column in terms of the age and diagonal dummies, additive terms by age and by diagonal will be estimated. The regression can tell which terms are statistically significant, and the others can be dropped from the specification.

With the reinsurance triangle tested above, the first three diagonals turned out to be lower than the others, as was the last diagonal. Also, the first two ages were not significantly different from each other, nor were the last four. This produced a model with five age parameters and two diagonal parameters—one for the first three diagonals combined, and one for the last diagonal. The parameters are shown in Table 9.

The sum of squared residuals for this model is 49,673.4 when adjusted for seven parameters used. This is considerably better

TABLE 9

TERMS IN ADDITIVE CHAIN LADDER WITH DIAGONAL EFFECTS

Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Diag 1–3	Diag 9
5,569.0	5,569.0	3,739.2	2,881.8	2,361.1	993.3	993.3	993.3	993.3	−2,319.9	−984.7

than the model without diagonal effects. The multiple regression found the diagonals to be statistically significant and adding them to the model improved the fit.

A problem with the diagonal analysis is how to use them in forecasting. One reason for diagonal effects is a change in company practice, particularly in the claims handling process. If the age effects are considered the dominant influence with occasional distortion by diagonal effects, then including diagonal dummy variables will give better estimates for the underlying age terms. Then these, but not the diagonal effects, would be used in forecasting.

Having identified the significant diagonal effects through linear regression, it may be more reasonable to convert them to multiplicative effects through non-linear regression. The model could be of the form:

$$q(w,d) = f(d)g(w+d),$$

where $f(d)$ is the additive age term for age d, and $g(w+d)$ is the factor for the $w+d$th diagonal. Again this can be estimated iteratively by fixing the f's to estimate the g's by linear regression, then fixing those g's to estimate the next iteration of f's, until convergence is reached. The previous model was refit with the diagonals as factors with the result in Table 10. This had a slightly better adjusted sum of squared residuals of 49,034.8.

Diagonal factors can be used in conjunction with accident year factors as in:

$$q(w,d) = f(d)g(w+d)h(w).$$

TABLE 10

ADDITIVE CHAIN LADDER WITH MULTIPLICATIVE DIAGONAL EFFECTS

Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Diag 1–3	Diag 9
5,692.3	5,692.3	3,823.0	2,816.1	2,416.7	672.1	672.1	672.1	672.1	.5598	.6684

TABLE 11

ADDITIVE CHAIN LADDER WITH MULTIPLICATIVE DIAGONAL & AY EFFECTS

Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Diag 1-3	Diag 9	AY 3-4
5,135.6	5,135.6	3,464.7	2,730.1	1,995.4	660.1	660.1	660.1	660.1	.6201	.7225	1.2672

As an example, a factor was added to the above model to represent accident years 3 and 4, and the 4th age term was forced to be the average of the 3rd and 5th. The result is in Table 11.

The adjusted sum of squared residuals came down to 44,700.9, which is considerably better than the previous best-fitting model, and almost twice as good as in the original BF model, which in turn was almost twice as good as the chain ladder. It appears that accident year effects and diagonal effects are significant in this data. The fit is shown as the last section of Exhibit 1. The numerous examples fit to this data were for the sake of illustration. Some models of the types discussed may still fit better than the particular ones shown here.

AGGREGATION OF CORRELATED RISK PORTFOLIOS: MODELS AND ALGORITHMS

SHAUN S. WANG, PH.D.

Abstract

This paper presents a set of tools for modeling and combining correlated risks. Various correlation structures are generated using copula, common mixture, component, and distortion models. These correlation structures are specified in terms of (i) the joint cumulative distribution function or (ii) the joint characteristic function and lend themselves to efficient methods of aggregation by using Monte Carlo simulation or fast Fourier transform.

ACKNOWLEDGEMENT

The author wishes to thank the CAS Committee on Theory of Risk for initiating and supporting this research project. A special thanks goes to Phil Heckman and Glenn Meyers for providing numerous stimulating comments while overseeing the progress of this project. Phil Heckman, Stephen Mildenhall, Don Mango, Louise Francis, and the CAS Editorial Committee, in their thorough review of this paper, have provided many valuable comments which led to significant improvement of an earlier draft report. The author also thanks Ole Hesselager, Julia Wirch, and Ken Seng Tan for comments.

1. INTRODUCTION

A good introduction for this research paper is the original Request For Proposal (RFP) drafted by the CAS Committee on Theory of Risk. In the following paragraph, the original RFP is restated with minor modification.

> Aggregate loss distributions are probability distributions of the total dollar amount of loss under one or

> a block of insurance policies. They combine the separate effects of the underlying frequency and severity distributions. In the actuarial literature, a number of methods have been developed for modeling and computing the aggregate loss distributions (see Heckman and Meyers [8], Panjer [21], and Robertson [23]). The main issue underlying this research project is how to combine aggregate loss distributions for separate but correlated classes of business.
>
> Assume a book of business is the union of disjoint classes of business each of which has an aggregate distribution. These distributions may be given in many different ways. Among other ways, they may be specified parametrically, e.g., lognormal or transformed beta with given parameters; they may be given by specifying separate frequency and severity distributions; e.g., negative binomial frequency and Pareto severity with given parameters. The classes of business are *not* independent. For this project, assume that we are given a correlation matrix (or some other easily obtainable measure of dependency) and that the correlation coefficients vary among different pairs of classes. The problem is how to calculate the aggregate loss distribution for the whole book.

In traditional actuarial theory, individual risks are usually assumed to be independent, mainly because the mathematics for correlated risks is less tractable. The CAS recognizes the importance of modeling and combining correlated risks and wishes to enhance the development of tools and models that improve the accuracy of the estimation of aggregate loss distributions for blocks of insurance risks. The modeling of dependent risks has special relevance to the current on-going project of Dynamic Financial Analysis (DFA).

In general, combining correlated loss variables requires knowledge of their joint (multivariate) probability distribution. However, the available data regarding the association between loss variables is often limited to some summary statistics (e.g., correlation matrix). In the special case of a multivariate normal distribution, the covariance matrix and the mean vector, as summary statistics, completely specify the joint distribution. For general loss frequency or severity distributions, specific dependency models have to be used in conjunction with summary statistics. Given fixed marginal distributions and a correlation matrix, one can construct infinitely many joint distributions. Ideally, models for dependency structure should be easy to implement and require relatively few input parameters. As well, the choice of the dependency model and its parameter values should reflect the underlying correlation-generating mechanism.

In developing dependency models, we are aiming at simple implementation by Monte Carlo simulation or by fast Fourier transform. To this end, we will take the following approaches to modeling and combining correlated risks:

Sections 2 to 5 serve as a background before major correlation models are discussed in later parts of this paper. Section 2 reviews some basic concepts for a discrete probability distribution, including probability generating function and fast Fourier transform (FFT). Section 3 reviews the aggregate loss model and the FFT method of calculating aggregate loss distributions. Section 4 introduces some basic concepts and tools for multivariate variables, including the joint cumulative distribution function and the joint probability generating function, which will form the basis of the whole paper. Section 5 reviews some basic measures of dependency, including (Pearson) correlation coefficients, Kendall's tau, and Spearman's rank correlation coefficient.

Sections 6, 7, and 8 investigate various correlation structures by using the concept of copulas (i.e., multivariate uniform distributions) as well as the associated simulation techniques. In particular, the Cook-Johnson copula and the normal copula lead to efficient simulation techniques.

Sections 9, 10, and 11, with due consideration to the underlying correlation-generating mechanism, present a variety of dependency models by using common mixtures and common shocks. These dependency models allow simple methods of aggregation by Monte Carlo simulation or by fast Fourier transform.

Section 12 presents a multivariate negative binomial model which lends itself to an efficient FFT method of combining the correlated risk portfolios. Section 13 gives an example of this method.

For the reader's convenience, an inventory of commonly used univariate distributions is given in Appendix A, including both discrete and continuous distributions. As a convention, X, Y, and Z represent any random variables (discrete, continuous, or mixed), while N and K represent only discrete variables defined on non-negative integers.

2. PROBABILITY GENERATING FUNCTION AND FFT

This section introduces some basic concepts for discrete probability distributions.

2.1. Discrete Probability Distributions

Let X be a discrete random variable defined on non-negative integers, $0, 1, 2, \ldots$. It may represent

- the *number of claims* arising from a specified block of insurance contracts within a pre-specified time period (such as one year); or
- the *claim amount* from a single claim count, with a pre-specified convenient monetary unit (such as $1,000).

The random variable X can be fully described by a probability vector

$$\mathbf{f}_X = [f_X(0), f_X(1), f_X(2), \ldots, f_X(R)],$$

or simply

$$\mathbf{f}_X = [f_0, f_1, f_2, \ldots, f_R],$$

with $f_X(i) = f_i = \Pr\{X = i\}$. In this representation, the maximal possible value of X cannot exceed R. When R is finite, X has infinitely many vector representations of the form

$$[f_0, f_1, f_2, \ldots, f_R, 0, 0, \ldots, 0],$$

where a number of zeros are added to the right.

For a discrete variable X with a probability vector $\mathbf{f}_X = [f_0, f_1, f_2, \ldots, f_R]$, the *probability generating function* (p.g.f.) is defined by a symbolic series:

$$P_X(t) = f_0 + f_1 t^1 + f_2 t^2 + f_3 t^3 + \cdots + f_R t^R,$$

which is also the expected value of t^X; i.e., $\mathrm{E}[t^X]$.

Example 2.1 If a discrete variable N has the following probabilities

$$\Pr\{N = 0\} = 0.5, \qquad \Pr\{N = 2\} = 0.4, \qquad \Pr\{N = 5\} = 0.1, \tag{2.1}$$

then it can be represented by a vector

$$\mathbf{f}_N = [0.5, 0, 0.4, 0, 0, 0.1, 0, \ldots, 0],$$

and it has a probability generating function

$$P_N(t) = 0.5 + 0.4t^2 + 0.1t^5.$$

Example 2.2 If a discrete variable K has the following probabilities

$$\Pr\{K = 1\} = 0.4, \qquad \Pr\{K = 2\} = 0.3, \qquad \Pr\{K = 3\} = 0.3, \tag{2.2}$$

then it can be represented by a vector

$$\mathbf{f}_K = [0, 0.4, 0.3, 0.3, 0, \ldots, 0],$$

and it has a probability generating function

$$P_K(t) = 0.4t + 0.3t^2 + 0.3t^3.$$

2.2. Fast Fourier Transforms

First we need to review some basics of complex numbers. Let $\mathrm{i} = \sqrt{-1}$ represent a symbol with the property that $\mathrm{i}^2 = -1$. The complex multiplication is defined as

$$(a + b\mathrm{i})(c + d\mathrm{i}) = (ac - bd) + (ad + bc)\mathrm{i}.$$

An important formula for complex numbers is the Euler formula

$$e^{\mathrm{i}z} = \cos(z) + \mathrm{i}\sin(z).$$

Now we are ready to define the fast Fourier transform. The following description of the FFT method draws on Klugman, Panjer, and Willmot [18] and Brigham [1].

The fast Fourier transform is a one-to-one mapping of n points into n points. For any n-point vector $(f_0, f_1, \ldots, f_{n-1})$, the *fast Fourier transform* is the mapping

$$\mathrm{FFT} : \mathbf{f} = [f_0, f_1, \ldots, f_{n-1}] \mapsto \tilde{\mathbf{f}} = [\tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_{n-1}]$$

defined by

$$\tilde{f}_k = \sum_{j=0}^{n-1} f_j \exp\left(\frac{2\pi\mathrm{i}}{n} jk\right), \qquad k = 0, 1, \ldots, n-1. \tag{2.3}$$

This one-to-one mapping has an inverse mapping:

$$f_j = \frac{1}{n}\sum_{k=0}^{n-1} \tilde{f}_k \exp\left(-\frac{2\pi\mathrm{i}}{n} kj\right), \qquad j = 0, 1, \ldots, n-1. \tag{2.4}$$

Note that the inverse fast Fourier transform (IFFT) is almost identical to the FFT except for a sign change and a division by n. In general, the FFT depends on the vector length n.

The fast Fourier transform in Equation 2.3 can also be viewed as a simple matrix multiplication:

$$\tilde{\mathbf{f}} = \mathbf{W}\mathbf{f} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \vdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \vdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{pmatrix} \mathbf{f},$$

where $\omega = \exp(2\pi \mathrm{i}/n)$.

The inverse FFT in Equation 2.4 is just $\mathbf{W}^{-1}\tilde{\mathbf{f}}$, where

$$\mathbf{W}^{-1} = \frac{1}{n} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \vdots & \omega^{-(n-1)} \\ 1 & \omega^{-2} & \omega^{-4} & \vdots & \omega^{-2(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \omega^{-(n-1)} & \omega^{-2(n-1)} & \cdots & \omega^{-(n-1)^2} \end{pmatrix}.$$

EXAMPLE 2.3 Reconsider the vector associated with the probability distribution in Equation 2.1. If we use the 5-point vector representation

$$\mathbf{f} = [0.5, 0, 0.4, 0, 0, 0.1],$$

the fast Fourier transform yields

$$\tilde{\mathbf{f}} = [1, 0.35 - 0.2598\mathrm{i}, 0.25 + 0.433\mathrm{i}, \\ 0.8, 0.25 - 0.433\mathrm{i}, 0.35 + 0.2598\mathrm{i}].$$

If we use the 6-point vector representation by (p)adding an additional zero

$$\mathbf{f} = [0.5, 0, 0.4, 0, 0, 0.1, 0],$$

the fast Fourier transform yields a different vector $\tilde{\mathbf{f}}$ as

$$[1, 0.3887 - 0.2925\mathrm{i}, 0.0495 + 0.1302\mathrm{i}, 0.8117 + 0.2345\mathrm{i},$$
$$0.8117 - 0.2345\mathrm{i}, 0.0495 - 0.1302\mathrm{i}, 0.3887 + 0.2925\mathrm{i}].$$

The fast Fourier transform is a "fast" computing algorithm because of the following properties: a fast Fourier transform of length $n = 2^r$ can be rewritten as the sum of two fast Fourier transforms, each of length $n/2 = 2^{r-1}$, the first consisting of the even numbered points and the second the odd numbered points.

$$\begin{aligned}\tilde{f}_k &= \sum_{j=0}^{n-1} f_j \exp\left(\frac{2\pi\mathrm{i}}{n} jk\right)\\ &= \sum_{j=0}^{n/2-1} f_{2j} \exp\left(\frac{2\pi\mathrm{i}}{n} 2jk\right) + \sum_{j=0}^{n/2-1} f_{2j+1} \exp\left(\frac{2\pi\mathrm{i}}{n}(2j+1)k\right)\\ &= \sum_{j=0}^{m-1} f_{2j} \exp\left(\frac{2\pi\mathrm{i}}{m} jk\right) + \exp\left(\frac{2\pi\mathrm{i}}{n} k\right) \sum_{j=0}^{m-1} f_{2j+1} \exp\left(\frac{2\pi\mathrm{i}}{m} jk\right),\end{aligned}$$

where $m = n/2 = 2^{r-1}$. Hence

$$\tilde{f}_k = \tilde{f}_k^a + \exp\left(\frac{2\pi\mathrm{i}}{n} k\right) \tilde{f}_k^b. \qquad (2.5)$$

Each of $\tilde{f}_k^a$ and $\tilde{f}_k^b$ can, in turn, be written as the sum of two transforms of length $m/2 = 2^{r-2}$. This can be continued successively.

The successive splitting of transforms into transforms of half the length will result, after r times, in transforms of length 1. Knowing the transform of length 1 will allow one to successively compose the transforms of length 2, 2^2, $2^3, \ldots, 2^r$ by using Equation 2.5.

Based on the above observations, the following comments are in order:

- To fully utilize the FFT speed, it is better to use a probability vector of length $n = 2^r$. This can be easily done by adding a number of zeros to the right.
- Thanks to the fact that many computer packages have already programmed FFT as a built-in function, we don't have to carry out the above steps by ourselves. The main purpose of the above paragraph is to illustrate why FFT is a fast algorithm.

It should be pointed out that many authors define the transform in Equation 2.3 as a *discrete* Fourier transform. The *fast* Fourier transform is simply a method for computing the *discrete* Fourier transform. On the other hand, in some applications such as Microsoft Excel, the term FFT is used to refer to the more general discrete Fourier transform. To simplify the terminology, this paper uses the term FFT for both the transform in Equation 2.3 and the special evaluation technique when $n = 2^r$.

As a theoretical note, the FFT should be viewed as a discretized version of the Fourier transform or characteristic function:

$$\phi(z) = \int_{-\infty}^{\infty} f(x) e^{izx} \, dx.$$

The characteristic function maps a continuous probability density function to a complex-valued continuous function, while the FFT maps a vector of n values to a vector of n values of complex numbers. This analog to characteristic functions is crucial to the understanding of various FFT algorithms presented in this paper.

2.3. Convolution

Suppose that N and K are independent discrete random variables defined on non-negative integers. Let $J = N + K$ represent the sum of N and K. The probability distribution of J represents the *convolution* of the probability distributions of N and K and is defined by

$$\Pr\{J=j\} = \sum_{n=0}^{j} \Pr\{N=n\}\Pr\{K=j-n\}, \qquad j=0,1,2,\ldots.$$

EXAMPLE 2.4 For the random variables defined in Equations 2.1 and 2.2, we have

$$\Pr\{J=5\} = \Pr\{N+K=5\} = \sum_{n=0}^{5} \Pr\{N=n\}\Pr\{K=5-n\}.$$

Since many of the terms are zero, we have

$$\Pr\{J=5\} = 0+0+\Pr\{N=2\}\Pr\{K=3\}+0+0+0 = 0.12.$$

Now let X represent a discrete claim severity distribution defined on non-negative integers. For a fixed number of k claims, the total claim amount has a distribution that can be evaluated through repeated convolutions

$$f_X^{*k}(x) = \sum_{y=0}^{x} f_X^{*(k-1)}(x-y) f_X(y), \qquad x=1,2,\ldots, \tag{2.6}$$

with the convention that

$$f^{*0}(0) = 1.$$

We call f^{*k} the kth fold convolution of f.

2.3.1. Convolution by probability generating function

Note that

$$P_{N+K}(t) = \mathrm{E}[t^{N+K}] = \mathrm{E}[t^N \cdot t^K] = \mathrm{E}[t^N]\mathrm{E}[t^K] = P_N(t)\cdot P_K(t)$$

due to the independence of N and K. In other words, the probability generating function of the sum $N+K$ is the product of $P_N(t)$ and $P_K(t)$.

EXAMPLE 2.5 For the random variables defined in Equations 2.1 and 2.2, in terms of probability generating function we have

$$P_J(t) = P_N(t) \cdot P_K(t) = (0.5 + 0.4t^2 + 0.1t^5)(0.4t + 0.3t^2 + 0.3t^3).$$

After expansion we get

$$\begin{aligned} P_J(t) &= 0.20t + 0.15t^2 + 0.31t^3 + 0.12t^4 \\ &\quad + 0.12t^5 + 0.04t^6 + 0.03t^7 + 0.02t^8. \end{aligned}$$

The coefficients of t^j give the probability that $J = j$; e.g., $\Pr\{J = 5\} = 0.12$.

2.3.2. Convolution by FFT

In terms of a characteristic function we have

$$\begin{aligned} \phi_{N+K}(t) &= \mathrm{E}[e^{\mathrm{i}t(N+K)}] = \mathrm{E}[e^{\mathrm{i}tN} \cdot t^{itK}] \\ &= \mathrm{E}[t^{\mathrm{i}tN}]\mathrm{E}[t^{\mathrm{i}tK}] = \phi_N(t) \cdot \phi_K(t) \end{aligned}$$

due to the independence of N and K. In other words, the characteristic function of the sum $N + K$ is the product of N and K.

Because of this relation in terms of characteristic function, FFT can also be used to perform convolutions. The FFT for the sum of two independent discrete random variables is the product of the FFTs of two individual variables, *provided that enough zeros are added (or padded) to each individual probability vector.* Note that FFT is a one-to-one mapping from n points to n points, which requires that input and output vectors be the same length. On the other hand, a longer vector is generally required for a discrete representation of the sum variable than for each component, since the sum variable will take on larger values with non-zero probability. If there is not enough room in the discrete vector, then the tail probabilities for the sum will wrap around and reappear at the beginning. Therefore, it is crucial to add enough zeros to the right of each individual probability vector.

2.3.3. *FFT Algorithm of convolution*

If $\mathbf{f} = [f_0, f_1, \ldots, f_{m-1}]$ and $\mathbf{g} = [g_0, g_1, \ldots, g_{k-1}]$ represent two probability vectors, then the following process can be used to evaluate their convolution:

- Pad zeros to the given vectors $\mathbf{f}$ and $\mathbf{g}$ such that each is of length $n \geq m + k$.
- Apply FFT to each of the vectors: $\tilde{\mathbf{f}} = \text{FFT}(\mathbf{f})$ and $\tilde{\mathbf{g}} = \text{FFT}(\mathbf{g})$.
- Take the product (complex number multiplication), element by element, of the two vectors: $\tilde{\mathbf{h}} = \tilde{\mathbf{f}} \cdot \tilde{\mathbf{g}}$.
- Apply IFFT to $\tilde{\mathbf{h}}$ to recover a probability vector, as the convolution of $\mathbf{f}$ and $\mathbf{g}$.

3. AGGREGATE LOSS MODELS AND THE FFT METHOD

In evaluating insurance losses for a book of business, the frequency/severity approach is the most flexible method, where the estimated mean frequency and mean severity are used to estimate the average aggregate loss. In order to facilitate a dynamic analysis of the underlying risk, the aggregate loss distribution is needed to quantify the inherent variability in the aggregate loss cost. In such situations, in addition to an estimate of the mean frequency and mean severity, probability distributions are needed to describe the possible variations in the number of claims and in the dollar amount of each individual claim. The *aggregate loss distribution* combines the effects of both the claim frequency and claim severity distributions.

This section introduces the basics of aggregate loss models and how FFT can be used to calculate the aggregate loss distribution.

3.1. Claim Frequency Distributions

In modeling the frequency of random processes in many fields of applications, the Poisson distribution is usually the starting point, although the parameter uncertainty regarding the Poisson mean often leads to a negative binomial frequency distribution (see Appendix A.3). Actuaries have found that in most cases the claim frequency can be modeled by the Poisson or negative binomial distributions.

- A Poisson distribution with mean $\lambda > 0$ is defined by a probability function:

$$\Pr\{N = n\} = e^{-\lambda}\frac{\lambda^n}{n!}, \qquad n = 0,1,2,\ldots.$$

  The Poisson (λ) distribution has a probability generating function

$$P_N(t) = \mathrm{E}[t^N] = e^{\lambda(t-1)},$$

  with mean and variance both equal to λ; i.e., $\mathrm{E}[N] = \mathrm{Var}[N] = \lambda$.

- A negative binomial distribution, with parameters $\alpha, \beta > 0$, has a probability function:

$$p_n = \Pr\{N = n\} = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)n!}\left(\frac{1}{1+\beta}\right)^{\alpha}\left(\frac{\beta}{1+\beta}\right)^{n}, \qquad n = 0,1,2,\ldots.$$

  It has a probability generating function

$$P_N(t) = [1-\beta(t-1)]^{-\alpha},$$

  with $\mathrm{E}[N] = \alpha\beta$ and $\mathrm{Var}[N] = \alpha\beta(1+\beta)$. In general, for a negative binomial distribution, the variance exceeds the mean. The variance to mean ratio is

$$\frac{\mathrm{Var}[N]}{\mathrm{E}[N]} = 1 + \beta.$$

Some actuaries consistently use the variance to mean ratio to specify a negative binomial distribution. Heckman and Meyers [8] used the contagion parameter, c, to specify a negative binomial distribution, where

$$\mathrm{Var}[N] = \mathrm{E}[N](1 + c \cdot \mathrm{E}[N]).$$

3.2. Claim Severity Distributions

Models for claim severity are very diverse. In many cases, a theoretical loss distribution is used. A list of the most commonly used theoretical distributions is given in Appendix A, including Pareto, gamma, Weibull and lognormal distributions. Among the commonly used two-parameter distributions, the ordering of heaviness (from most heavy to least heavy) of tails is as follows (see Wang [25]):

Distribution	Ranking
Pareto	1
lognormal	2
exponential inverse Gaussian	3
inverse Gaussian	4
Weibull	5
gamma	6

If a large data set is available, an empirical loss distribution can be used.

Once a severity distribution is selected, in order for fast computer implementation, it is necessary to construct a discrete severity distribution on multiples of a convenient monetary unit h, the *span*. If a theoretical continuous distribution is employed, the following methodology can be used to approximate it by a discrete distribution.

3.2.1. The rounding method

Suppose that we are given a continuous distribution with cumulative distribution $F_X(t) = \Pr\{X \leq t\}$. Choose a span h as appropriate (such that the number of points are sufficient but not excessive). Let f_j denote the probability placed at jh, $j = 0, 1, 2, \ldots$. Then set

$$\begin{aligned} f_0 &= F_X\left(\frac{h}{2}\right), \\ f_j &= F_X\left(jh + \frac{h}{2}\right) - F_X\left(jh - \frac{h}{2}\right), \qquad j = 1, 2, \ldots . \end{aligned} \tag{3.1}$$

This method splits the probability between $(j+1)h$ and jh and assigns it to $j+1$ and j. This, in effect, rounds all amounts to the nearest convenient monetary unit, h, the span of the distribution. For example, the span h can be chosen as every \$1000, \$5000, or \$10,000. As the monetary unit of measurement becomes small, the discrete distribution function needs to approach the true distribution function.

While the main advantage of this rounding method is its simplicity, it has a drawback of not preserving the mean severity of the continuous distribution.

3.2.2. The matching-mean method

To avoid the drawback of mismatch of mean severity, one can use a method that forces the matching of the mean.

For a severity distribution with cumulative distribution function F_X, we first evaluate the limited expected values at multiples of h:

$$\mathrm{E}[X; j \cdot h] = \int_0^{j \cdot h} [1 - F_X(u)]\, du, \qquad \text{for} \quad j = 1, 2, \ldots .$$

Then we calculate the probability vector by:

$$f_0 = \Pr\{X = 0 \cdot h\} = 1 - \mathrm{E}[X;h]/h, \tag{3.2}$$

$$\begin{aligned} f_j &= \Pr\{X = j \cdot h\} \\ &= (2\mathrm{E}[X;j \cdot h] - \mathrm{E}[X;(j-1)\cdot h] - \mathrm{E}[X;(j+1)\cdot h])/h, \\ &\qquad\qquad\qquad\qquad j = 1,2,\ldots, \end{aligned} \tag{3.3}$$

By doing so, the mean severity of the continuous distribution is preserved in the discrete distribution. One can verify this by taking the sum of f_i, $i = 0,1,2,\ldots$.

Recall that by taking the second-order derivative of the limited expected value function we get a probability density function. In the above method, we first obtain a discrete vector of limited expected values; by taking the second-order finite difference, we get a discrete probability function.

3.3. The Aggregation of Frequency and Severity

The aggregate losses are represented as a sum, Z, of a random number, N, of individual payment amounts $(X_1, X_2, \ldots, X_N)$.

The random sum

$$Z = X_1 + X_2 + \cdots + X_N \tag{3.4}$$

has a probability distribution

$$\begin{aligned} f_Z(x) &= \Pr(Z = x) \\ &= \sum_{n=0}^{\infty} \Pr(N = n)\Pr(Z = x \mid N = n) \\ &= \sum_{n=0}^{\infty} \Pr(N = n) f_X^{*n}(x), \end{aligned} \tag{3.5}$$

where $f_X(x) = \Pr(X = x)$ is the common probability distribution of the X_js.

A direct evaluation by Equation 3.5 of the aggregate loss distribution is usually very complicated and time consuming, even with today's fast-speed computers. The next subsection introduces the FFT technique for computing the aggregate loss distribution.

3.3.1. Computing aggregate loss distribution by FFT

In the aggregate loss model in Equation 3.4, we have in terms of characteristic function:

$$\begin{aligned}\phi_Z(t) &= \mathrm{E}[e^{\mathrm{i}t(Z)}] = \mathrm{E}_N[\mathrm{E}[e^{\mathrm{i}t(X_1+\cdots+X_N)} \mid N]]\\ &= \mathrm{E}_N[\phi_X(t)^N] = P_N(\phi_X(t)),\end{aligned}$$

where P_N is the probability generating function of N. This relation in terms of characteristic function suggests the following FFT algorithm for calculating the aggregate loss distribution:

1. Choose $n = 2^r$ for some integer r; n is the number of points desired in the distribution $f_Z(x)$ of aggregate losses. In other words, the aggregate loss distribution has negligible probability outside the range $[0,n]$. This range should be determined before one knows the exact aggregate loss distribution. Knowledge of the mean and standard deviation of the aggregate loss amount should be helpful.

2. Transform the severity probability distribution from a continuous one to a discrete one. The selection of the span h should depend upon the probable range of the severity distribution, as well as the intended application (central range or the extreme right tail). Let $(f_0, f_1, \ldots, f_{m-1})$ represent the discrete claim severity distribution.

   Add zeros to the given severity probability vector so that it is of length n. We denote the padded discrete severity distribution by

$$\mathbf{f}_X = [f_X(0), f_X(1), \ldots, f_X(n-1)].$$

3. Apply FFT to the severity probability vector: $\tilde{\mathbf{f}}_X = \mathrm{FFT}(\mathbf{f}_X)$.
4. Apply the probability generating function of the frequency, element by element, to the FFT of the severity vector: $\tilde{\mathbf{f}}_Z = P_N(\tilde{\mathbf{f}}_X)$.
5. Apply IFFT to recover the aggregate loss distribution: $\mathbf{f}_Z = \mathrm{IFFT}(\tilde{\mathbf{f}}_X)$.

As a simple example of the above algorithm, let severity be the degenerate distribution \$1 with certainty, and let frequency be negative binomial. Thus the aggregate distribution is the negative binomial. By choosing the number of points, n, the discrete severity distribution is an n-point vector $(0,1,0,\ldots,0)$. The FFT of the severity vector gives a vector of roots of unity. One can check that the FFT algorithm closely reproduces the negative binomial distribution if the number of points used is sufficiently large.

The FFT and IFFT algorithms are available in many computer software packages including Microsoft Excel. This makes the implementation of the FFT method widely accessible.

3.4. Techniques for Combining Multiple Lines of Business

3.4.1. Combining two lines of business by convolution

Suppose that we are combining two lines of business:

- Line 1 has a claim frequency N and a discrete claim severity X.
- Line 2 has a claim frequency K and a discrete claim severity Y.
- Assume that N, X, K, and Y are mutually independent.

- We are interested in the probability distribution of the aggregate losses for the combined portfolio:

$$Z = (X_1 + \cdots + X_N) + (Y_1 + \cdots + Y_K).$$

Under the above assumptions, we have

$$\phi_Z(t) = P_N(\phi_X(t)) \cdot P_K(\phi_Y(t)).$$

This relation in terms of characteristic function suggests the following FFT procedure:

Let $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{h}}$ represent the FFT of the aggregate loss distributions for Line 1 and Line 2, respectively:

$$\tilde{\mathbf{g}} = P_N(\tilde{\mathbf{f}}_X), \qquad \tilde{\mathbf{h}} = P_K(\tilde{\mathbf{f}}_Y).$$

Before applying IFFT to each of $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{h}}$, we take the complex product (element by element) of $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{h}}$. Then apply the IFFT to the product $\tilde{\mathbf{g}} \cdot \tilde{\mathbf{h}}$ to recover the aggregate loss distribution for Line 1 and Line 2 combined:

$$\mathbf{f}_Z = \text{IFFT}(\tilde{\mathbf{g}} \cdot \tilde{\mathbf{h}}).$$

Under this approach, the aggregate claim frequency is the convolution of individual claim frequency distributions. If each individual line has a negative binomial frequency distribution, the aggregate frequency distribution obtained by convolution may no longer be a negative binomial distribution.

3.4.2. The Poisson model

Here is a basic Poisson model for combining different lines of business:

Assume that we are combining k lines of business. For $j = 1, 2, \ldots, k$, assume that Line j has a Poisson frequency with mean λ_j and a severity distribution F_j. We assume that losses from different lines of business are *independent*.

In terms of characteristic function we have

$$\begin{aligned}\phi_Z(t) &= \prod_{j=1}^{k} P_{N_j}(\phi_{X_j}(t))\\ &= \prod_{j=1}^{k} e^{\lambda_j(\phi_{X_j}(t)-1)}\\ &= e^{\lambda(\phi_X(t)-1)},\end{aligned}$$

where $\lambda = \lambda_1 + \cdots + \lambda_k$, and

$$\phi_X(t) = \frac{\lambda_1}{\lambda}\phi_{X_1}(t) + \cdots + \frac{\lambda_k}{\lambda}\phi_{X_k}(t).$$

Therefore, the aggregate losses for the k lines of business combined have a Poisson frequency with mean

$$\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_k \tag{3.6}$$

and a severity distribution that is a weighted average of each individual severity distribution:

$$F(x) = \frac{\lambda_1}{\lambda}F_1(x) + \frac{\lambda_2}{\lambda}F_2(x) + \cdots + \frac{\lambda_k}{\lambda}F_k(x). \tag{3.7}$$

In summary, under the assumption of mutual independence between lines and a Poisson frequency model for each line, the aggregate loss distribution fork lines can be calculated as if you had a single line, provided that the frequency and severity are adjusted using Equations 3.6 and 3.7.

Next, we must consider the following complications: (i) the presence of parameter risk and (ii) possible correlation between lines. These two factors are often interrelated.

As an alternative to the Poisson model, the negative binomial distribution is commonly used to adjust for parameter uncertainty. Recall that a negative binomial distribution can be ob-

TABLE 1

FREQUENCY/SEVERITY DISTRIBUTIONS

Line	Mean Frequency	Frequency Var/Mean Ratio	Severity Distribution
1	$E(N_1)$	$1+\beta_1$	F_1
2	$E(N_2)$	$1+\beta_2$	F_2
⋮	⋮	⋮	⋮
k	$E(N_k)$	$1+\beta_k$	F_k

tained by assuming a gamma distribution for the unknown Poisson mean (see Appendix A.3).

With the presence of parameter uncertainty, we have to re-evaluate the independence assumption between lines. The common parameter uncertainty may have a similar effect (i.e., over- or under-estimate) on our estimates of individual line mean frequencies. In such cases, the individual claim frequencies may be correlated as a result of the common estimation error (due to the same underlying data quality, variations of the underwriting and claim handling practices of an insurer from the industry average, or bias in the trend and development factors used).

3.4.3. Negative binomial model

In general, consider k lines of business with the frequency/severity distributions shown in Table 1.

Regardless of which specific frequency model is used, the following general relationships hold:

- The mean of the aggregate frequency is the sum of each individual line mean frequency:

$$E[N_{agg}] = E[N_1] + E[N_2] + \cdots + E[N_k]. \qquad (3.8)$$

- The total variance of the aggregate frequency can be calculated by:

$$\mathrm{Var}[N_{agg}] = \mathrm{Var}\left[\sum_{i=1}^{k} N_i\right] = \sum_{i=1}^{k}\mathrm{Var}[N_i] + 2\sum_{i<j}\mathrm{Cov}[N_i,N_j]. \tag{3.9}$$

A simple and direct approach to incorporating claim count parameter uncertainty is to assume a negative binomial distribution for the aggregate frequency for all lines combined. With this aggregate approach, the negative binomial parameters can be readily estimated from $\mathrm{E}[N_{agg}]$ and $\mathrm{Var}[N_{agg}]$ in Equations 3.8 and 3.9. The severity distribution for all lines combined can be calculated as the weighted average of individual severity distributions:

$$F(x) = \frac{\mathrm{E}[N_1]}{\mathrm{E}[N_{agg}]}F_1(x) + \frac{\mathrm{E}[N_1]}{\mathrm{E}[N_{agg}]}F_2(x) + \cdots + \frac{\mathrm{E}[N_k]}{\mathrm{E}[N_{agg}]}F_k(x). \tag{3.10}$$

Here is the rationale for this approach: Suppose that after applying trend factors and development factors to losses by line of business, we blend all the trended ultimate losses (or in a reinsurance application, losses in an excess layer) from all lines combined. By considering these consolidated losses from all lines of business, the empirical aggregate frequency has a mean as given in Equation 3.8, and the empirical aggregate severity has a severity distribution as in Equation 3.10. The only difference between the aggregate and individual approaches is the following: The individual approach assumes that each line of business has a negative binomial frequency, while the aggregate approach assumes that the aggregate frequency for all lines combined has a negative binomial distribution.

One major advantage of this approach is its simplicity. By simply adjusting the variance to mean ratio in the aggregate negative binomial frequency, one can easily take account of the parameter uncertainty for each line, as well as correlations between lines. Suppose that we have the following correlation matrix between N_js:

$$\begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{k1} & \rho_{k2} & \cdots & \rho_{kk} \end{pmatrix},$$

then we can calculate the overall variance to mean ratio using Equation 3.9 and the relation

$$\mathrm{Cov}[N_i, N_j] = \rho_{ij}\sqrt{\mathrm{Var}[N_i]}\sqrt{\mathrm{Var}[N_j]}.$$

The above method may not be theoretically exact if each individual line (instead of the aggregate of all lines) has a negative binomial frequency, as assumed in the Request for Proposal by the CAS Committee on Theory of Risk. Sections 11 and 12 discuss some exact methods for combining individual lines, each having a negative binomial frequency.

3.5. Other Methods for Calculating the Aggregate Loss Distributions

Over the past two decades there have developed a number of methods for calculation of the aggregate loss distribution from given frequency and severity distributions.

1. Panjer's [21] recursive algorithm is easy to explain and implement. In Appendix C we give a brief introduction of this method.
2. The Heckman–Meyers [8] method utilizes direct inversion of characteristic functions.
3. Robertson [23] presented a FFT method using piecewise uniform severity distributions, instead of a discrete severity distribution.

4. The proposed FFT method in this paper uses a discrete severity distribution that has, after padding zeros, $n = 2^r$ points. This is to exploit the fast speed of the FFT algorithm and to facilitate spreadsheet calculations. Another advantage of the FFT method is that it allows a direct extension to multivariate variables, as we will see later in this paper.

5. The recursive method may encounter some numerical problems such as overflow/underflow with a large expected claim count. On the other hand, the Heckman–Meyers method performs well with large claim frequencies. Panjer and Willmot [22] discuss ways of dealing with large frequency problems for the recursive method. For the FFT method, the problem with a large claim count is setting the span small enough to capture features of the severity distribution, but large enough that n times span gives enough room for the aggregate distribution. For divisible frequency distributions like the Poisson and negative binomial, one can get around the problem by building the aggregate distribution in pieces (say a small number of claims at a time) and adding the resulting distributions by convolution.

4. SOME TOOLS FOR MULTIVARIATE DISTRIBUTIONS

4.1. Review of Univariate Case

Let X be a non-negative random variable of discrete, continuous, or mixed type. Let $f_X(x)$ be the probability (density) function of X; i.e.,

$$f_X(x) = \begin{cases} \Pr\{X = x\}, & \text{if} \quad X \text{ is discrete} \\ \dfrac{d}{dx}F_X(x), & \text{if} \quad X \text{ is continuous.} \end{cases}$$

- The *probability generating function* (p.g.f.) of X is defined by

$$P_X(t) = \mathrm{E}[t^X] = \begin{cases} \sum f_X(x)t^x & \text{if } X \text{ is discrete} \\ \int f_X(x)t^x\,dx & \text{if } X \text{ is continuous.} \end{cases}$$

- The *moment generating function* (m.g.f.) of X is defined by

$$M_X(t) = \mathrm{E}[e^{tX}] = P_X(e^t).$$

- The *characteristic function* (ch.f.), also called Fourier transform, is defined by

$$\phi_X(t) = \mathrm{E}[e^{\mathrm{i}tX}] = P_X(e^{\mathrm{i}t}) = M_X(\mathrm{i}t),$$

  where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit.

- It holds that $P_X(1) = M_X(0) = \phi_X(0) = 1$, and

$$\mathrm{E}[X] = \left[\frac{d}{dt}P_X(t)\right]_{t=1} = \left[\frac{d}{dt}M_X(t)\right]_{t=0} = -\mathrm{i}\left[\frac{d}{dt}\phi_X(t)\right]_{t=0}.$$

4.2. Multivariate Framework

For a set of random variables $(X_1,\ldots,X_k)$, let $f_{X_1,\ldots,X_k}$ be their *joint* probability (density) function; i.e.,

$$f_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \begin{cases} \Pr\{X_1 = x_1,\ldots,X_k = x_k\}, & \text{if the } X_j \text{ are discrete} \\ \dfrac{\partial^k}{\partial x_1\cdots\partial x_k}F_{X_1,\ldots,X_k}(x_1,\ldots,x_k), & \text{if the } X_j \text{ are continuous.} \end{cases}$$

For any subset of $\{X_1,X_2,\ldots,X_k\}$, their (joint) probability distribution is called a *marginal* probability distribution of $f_{X_1,X_2,\ldots,X_k}$. As special cases, f_{X_1} is a univariate marginal distribution of $f_{X_1,X_2,\ldots,X_k}$, and f_{X_1,X_2} is a bivariate marginal distribution of $f_{X_1,X_2,\ldots,X_k}$.

As standard tools for multivariate random variables $(X_1,\ldots,X_k)$, the *joint probability generating function*, *joint moment gen-*

erating function, and *joint characteristic function* are defined as follows (see Johnson et al., [16, pp. 2–12]):

$$P_{X_1,\dots,X_k}(t_1,\dots,t_k) = \mathrm{E}[t_1^{X_1}\cdots t_k^{X_k}];$$
$$M_{X_1,\dots,X_k}(t_1,\dots,t_k) = \mathrm{E}[e^{t_1X_1+\cdots+t_kX_k}] = P_{X_1,\dots,X_k}(e^{t_1},\dots,e^{t_k});$$
$$\phi_{X_1,\dots,X_k}(t_1,\dots,t_k) = \mathrm{E}[e^{\mathrm{i}(t_1X_1+\cdots+t_kX_k)}] = P_{X_1,\dots,X_k}(e^{\mathrm{i}t_1},\dots,e^{\mathrm{i}t_k}).$$

Note that in terms of the probability (density) function we have

$$P_{X_1,\dots,X_k}(t_1,\dots,t_k) = \begin{cases} \sum_{(x_1,\dots,x_k)} f_{X_1,\dots,X_k}(x_1,\dots,x_k)t_1^{x_1}\cdots t_k^{x_k}, & \text{discrete case} \\ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1,\dots,X_k}(u_1,\dots,u_k)t_1^{u_1}\cdots t_k^{u_k}\,du_1\cdots du_k, & \text{continuous case.} \end{cases}$$

The joint probability generating function $P_{X_1,\dots,X_k}$ or the joint characteristic function $\phi_{X_1,\dots,X_k}$ completely specifies a multivariate distribution. Equivalent results are obtained either in terms of probability generating function or in terms of characteristic function.

- The probability generating function or characteristic function for the univariate marginal distribution F_{X_j} can be obtained by

$$P_{X_j}(t_j) = P_{X_1,\dots,X_j,\dots,X_k}(1,\dots,1,t_j,1,\dots,1),$$
$$\phi_{X_j}(t_j) = \phi_{X_1,\dots,X_j,\dots,X_k}(0,\dots,0,t_j,0,\dots,0).$$

- If the variables $X_1,\dots,X_k$ are mutually independent, then

$$P_{X_1,\dots,X_k}(t_1,\dots,t_k) = \prod_{j=1}^{k} P_{X_j}(t_j).$$

- If two sets of variables $\{X_1,\dots,X_m\}$ and $\{Y_1,\dots,Y_n\}$ are independent, then

$$P_{X_1,\dots,X_m,Y_1,\dots,Y_n}(t_1,\dots,t_m,s_1,\dots,s_n) = P_{X_1,\dots,X_m}(t_1,\dots,t_m)P_{Y_1,\dots,Y_n}(s_1,\dots,s_n).$$

- The covariances can be evaluated by $\mathrm{Cov}[X_i, X_j] = \mathrm{E}[X_i X_j] - \mathrm{E}[X_i]\mathrm{E}[X_j]$ with

$$\mathrm{E}[X_i X_j] = \frac{\partial^2}{\partial t_i \partial t_j} P_{X_1,\ldots,X_m}(1,\ldots,1)$$
$$= -\frac{\partial^2}{\partial t_i \partial t_j} \phi_{X_1,\ldots,X_m}(0,\ldots,0).$$

  This can be seen from the expression

$$\frac{\partial^2}{\partial t_i \partial t_j} P_{X_1,\ldots,X_k}(t_1,\ldots,t_k)$$
$$= \sum x_i x_j f_{X_1,\ldots,X_k}(x_1,\ldots,x_k) t_1^{x_1} \cdots t_i^{x_i - 1} \cdots t_j^{x_j - 1} \cdots t_k^{x_k}.$$

- For a discrete multivariate distribution, the joint probability function is

$$f_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \frac{\partial^{x_1+\cdots+x_k}}{(\partial t_1)^{x_1} \cdots (\partial t_k)^{x_k}} P_{X_1,\ldots,X_k}(0,\ldots,0) \prod_{i=1}^{k} \frac{1}{x_i!}.$$

4.3. Aggregation of Correlated Variables

THEOREM 1 *For any k correlated variables $X_1,\ldots,X_k$ with joint probability generating function $P_{X_1,\ldots,X_k}$ and joint characteristic function $\phi_{X_1,\ldots,X_k}$, the sum $Z = X_1 + \cdots + X_k$ has a probability generating function and a characteristic function:*

$$P_Z(t) = P_{X_1,\ldots,X_k}(t,\ldots,t), \qquad \phi_Z(t) = \phi_{X_1,\ldots,X_k}(t,\ldots,t).$$

Proof $P_Z(t) = \mathrm{E}[t^{X_1+\cdots+X_k}] = \mathrm{E}[t^{X_1} \cdots t^{X_K}] = P_{X_1,\ldots,X_k}(t,\ldots,t).$

If we know the joint characteristic function of the k correlated variables $X_1,\ldots,X_k$, it is straightforward to get the characteristic function for their sum $\phi_Z(t) = \phi_{X_1,\ldots,X_k}(t,\ldots,t)$. Then the probability distribution of Z can be obtained by inverse

Fourier transform. In actual computer implementation, a discrete version (FFT method) can be used. The relation $\phi_{X_1+\cdots+X_k}(t) = \phi_{X_1,\ldots,X_k}(t,\ldots,t)$, along with its associated FFT algorithm, can be used to

- combine correlated risk portfolios if we let X_i represent the aggregate loss distributions for each individual risk portfolio,
- evaluate the total claim number distribution if we let X_i represent the claim frequency for each individual risk portfolio, or
- combine individual claims if we let X_i represent the claim size for each individual risk.

4.4. Aggregation of Risk Portfolios with Correlated Frequencies

Consider the aggregation of two correlated risk portfolios:

$$Z = (X_1 + \cdots + X_N) + (Y_1 + \cdots + Y_K),$$

where N and K are correlated, while the pair (N,K) is independent of the claim sizes X and Y, and the X_is and Y_js are mutually independent. We have

$$\begin{aligned} P_Z(t) &= \mathrm{E}[t^Z] = \mathrm{E}[t^{(X_1+\cdots+X_N)+(Y_1+\cdots+Y_K)}] \\ &= \mathrm{E}_{N,K}\mathrm{E}[t^{(X_1+\cdots+X_n)+(Y_1+\cdots+Y_m)} \mid N = n,\ K = m] \\ &= \mathrm{E}_{N,K}[P_X(t)^N P_Y(t)^K] \\ &= P_{N,K}(P_X(t), P_Y(t)). \end{aligned}$$

In terms of characteristic function we have

$$\phi_Z(t) = P_{N,K}(\phi_X(t), \phi_Y(t)). \tag{4.1}$$

5. MEASURES OF CORRELATION

5.1. Pearson Correlation Coefficients

For random variables X and Y, the *(Pearson) correlation coefficient*, defined by

$$\rho(X,Y) = \frac{\mathrm{Cov}[X,Y]}{\sigma[X]\sigma[Y]},$$

always lies in the range $[-1,1]$. The Pearson correlation coefficient is also called a linear correlation coefficient. Note that $\rho(X,Y) = 1$ if and only if $X = aY + b$ for some constants $a > 0$ and b. If there is no linear relationship between X and Y, the permissible range of $\rho(X,Y)$ is further restricted.

EXAMPLE 5.1 Consider the case that $\log X \sim N(\mu,1)$ and $\log Y \sim N(\mu\sigma,\sigma^2)$. The maximum correlation between X and Y is obtained when the deterministic relation $Y = X^\sigma$ holds. Thus, for random variables with these fixed marginal distributions we have [see Appendix A.4.2]

$$\max\{\rho(X,Y)\} = \frac{\exp(\sigma)-1}{\sqrt{\exp(\sigma^2)-1}\,\sqrt{e-1}}.$$

Observe that

- $\max\{\rho(X,Y)\} = 1$ when $\sigma = 1$ (i.e., $X = Y$),
- $\max\{\rho(X,Y)\}$ decreases to zero as σ increases to ∞, and
- $\max\{\rho(X,Y)\}$ decreases to $1/\sqrt{e-1}$ as σ decreases to 0.

For a set of k random variables $X_1,\ldots,X_k$, the correlation matrix

$$\begin{pmatrix} \rho(X_1,X_1) & \cdots & \rho(X_1,X_k) \\ \vdots & \vdots & \vdots \\ \rho(X_k,X_1) & \cdots & \rho(X_k,X_k) \end{pmatrix}, \qquad -1 \le \rho(X_i,X_j) \le 1,$$

is always positive definite, as it is symmetric and diagonally dominant.

5.2. Covariance Coefficients

For non-negative random variables X and Y, we define the *covariance coefficient* as

$$\omega(X,Y) = \frac{\mathrm{Cov}[X,Y]}{\mathrm{E}[X]\,\mathrm{E}[Y]} = \rho(X,Y)\frac{\sigma[X]}{\mathrm{E}[X]}\frac{\sigma[Y]}{\mathrm{E}[Y]}$$
$$= \rho(X,Y)\,\mathrm{CV}(X)\,\mathrm{CV}(Y),$$

where CV refers to the coefficient of variation. Note that the permissible range of $\omega(X,Y)$ depends on the shape of the marginal distributions.

EXAMPLE 5.2 Reconsider the variables X and Y in Example 5.1. It can be shown that

$$\max\{\omega(X,Y)\} = e^{\sigma} - 1.$$

Observe that

- $\max\{\omega(X,Y)\} = e - 1$ when $\sigma = 1$ (i.e., $X = Y$),
- $\max\{\omega(X,Y)\}$ increases to infinity as σ increases to infinity, and
- $\max\{\omega(X,Y)\}$ decreases to zero as σ decreases to zero.

For k non-negative random variables, $X_1,\ldots,X_K$, we define the matrix of covariance coefficients as

$$\begin{pmatrix} \omega(X_1,X_1) & \cdots & \omega(X_1,X_k) \\ \vdots & \vdots & \vdots \\ \omega(X_k,X_1) & \cdots & \omega(X_k,X_k) \end{pmatrix}.$$

One should exercise caution when choosing a parameter value for $\omega(X,Y)$, as its permissible range is sensitive to the marginal distributions. A practical method for obtaining the maximal positive and negative covariances between risks X and Y is given in Equations 5.1 and 5.2.

5.3. Frechet Bounds, Comonotonicity, and Maximal Correlation

Now consider the bivariate random variables (X,Y). Let

$$F_{X,Y}(x,y) = \Pr\{X \le x, Y \le y\}, \qquad S_{X,Y}(x,y) = \Pr\{X > x,\ Y > y\}$$

be the joint cumulative distribution function and the joint survivor function of (X,Y), respectively. Note that

$$\begin{aligned} &F_{X,Y}(x,\infty) = F_X(x),\\ &F_{X,Y}(\infty,y) = F_Y(y), \qquad \text{for} \quad -\infty < x,y < \infty\\ &S_{X,Y}(x,y) = 1 - F_X(x) - F_Y(y) + F_{X,Y}(x,y) \ne 1 - F_{X,Y}(x,y). \end{aligned}$$

If X and Y are independent, then $F_{X,Y}(x,y) = F_X(x)\cdot F_Y(y)$ and $S_{X,Y}(x,y) = S_X(x)\cdot S_Y(y)$. In general, the joint cumulative distribution function $F(x,y)$ is constrained from above and below.

LEMMA 1 *For any bivariate cumulative distribution function $F_{X,Y}$ with given marginal distributions F_X and F_Y, we have*

$$\max[F_X(x) + F_Y(y) - 1, 0] \le F_{X,Y}(x,y) \le \min[F_X(x), F_Y(y)].$$

Proof The first inequality results from the fact that $S(x,y) \ge 0$, and the second inequality can be proven using $P(A \cap B) \le \min[P(A), P(B)]$.

The upper bound

$$F_u(x,y) = \min[F_X(x), F_Y(y)]$$

and the lower bound

$$F_l(x,y) = \max[F_X(x) + F_Y(y) - 1, 0]$$

are called Frechet bounds.

Closely associated with Frechet bounds is the concept of comonotonicity. The upper Frechet bound is reached if X and Y are comonotonic. The lower Frechet bound is reached if X and $-Y$ are comonotonic.

DEFINITION 1 *Two random variables X and Y are comonotonic if there exists a random variable Z such that*

$$X = u(Z), \qquad Y = v(Z), \qquad \textit{with probability one},$$

where the functions u and v are non-decreasing.

Recall that X and Y are positively perfectly correlated if and only if $Y = aX + b$, $a > 0$. This linear condition is quite restrictive. Comonotonicity is an extension of the concept of perfect correlation to random variables with non-linear relations. Consider the following excess reinsurance arrangement of risk Z: the ceding company retains the first portion of any loss, and the reinsurer pays the excess portion. Putting it mathematically, the payments of the ceding company and the reinsurer will be

$$X = \begin{cases} Z, & Z \le d \\ d, & Z > d, \end{cases} \qquad Y = \begin{cases} 0, & Z \le d \\ Z - d, & Z > d \end{cases}$$

respectively. Note that X and Y are *not* perfectly correlated since one cannot be written as a function of the other. However, since X and Y are always non-decreasing functions of the original risk Z, they are *comonotonic*. They are bets on the same event, and neither of them is a hedge against the other.

5.4. Comonotonicity and Monte Carlo Simulation

The concept of comonotonicity can also be explained in terms of Monte Carlo simulation by inversion of random uniform numbers.

Assume that X has a cumulative distribution function F_X and a survivor function $S_X(x) = 1 - F_X(x)$. We define F_X^{-1} and S_X^{-1} as follows:

$$F_X^{-1}(q) = \min\{x : F_X(x) \geq q\}, \qquad 0 < q < 1$$
$$S_X^{-1}(q) = \min\{x : S_X(x) \leq q\}, \qquad 0 < q < 1.$$

Note that F_X^{-1} is non-decreasing, S_X^{-1} is non-increasing, and $S_X^{-1}(q) = F_X^{-1}(1-q)$.

The traditional Monte Carlo simulation method is based on the following result.

LEMMA 2 *For any random variable X and any random variable U which is uniformly distributed on* $(0,1)$*, X and* $F_X^{-1}(U)$ *have the same cumulative distribution function.*

Proof $\mathrm{P}\{F_X^{-1}(U) \leq x\} = \mathrm{P}\{U \leq F_X(x)\} = F_X(x)$.

A Monte Carlo simulation of a random variable X can be achieved by first drawing a random uniform number u from $U \sim \text{Uniform}(0,1)$ and then inverting u by $x = F_X^{-1}(u)$.

In order to simulate comonotonic risks X and Y, the same sample of random uniform numbers can be used in an inversion by F_X and F_Y, respectively. By contrast, if X and Y are independent, two independent samples of random uniform numbers have to be used in an inversion by F_X and F_Y, respectively.

For given marginal distributions F_X and F_Y, the maximal possible correlation exists when X and Y are comonotonic. Based on the Monte Carlo method of generating comonotonic risks, we can calculate the maximal possible covariance between two risks with given marginal probability distributions by:

$$\frac{1}{n}\sum_{j=1}^{n} F_X^{-1}\left(\frac{j}{n+1}\right) F_Y^{-1}\left(\frac{j}{n+1}\right) - \left(\frac{1}{n}\sum_{j=1}^{n} F_X^{-1}\left(\frac{j}{n+1}\right)\right)\left(\frac{1}{n}\sum_{j=1}^{n} F_Y^{-1}\left(\frac{j}{n+1}\right)\right) \quad (5.1)$$

for some large number n. The maximal negative correlation exists when X and $-Y$ are comonotonic, in which case an approximation of the covariance can be obtained from

$$\frac{1}{n}\sum_{j=1}^{n} F_X^{-1}\left(\frac{j}{n+1}\right) F_Y^{-1}\left(\frac{n+1-j}{n+1}\right) - \left(\frac{1}{n}\sum_{j=1}^{n} F_X^{-1}\left(\frac{j}{n+1}\right)\right)\left(\frac{1}{n}\sum_{j=1}^{n} F_Y^{-1}\left(\frac{n+1-j}{n+1}\right)\right) \tag{5.2}$$

for some large number n.

As we have seen, the permissible range for the Pearson correlation coefficient can be quite limited and subject to change under a transformation of the random variable. To overcome the shortcomings of the (linear) correlation coefficient, we can use distribution-free measures of correlation such as Kendall's tau and Spearman's rank correlation coefficient.

5.5. Kendall's Tau and Spearman's Rank Correlation Coefficient

Kendall's tau is a nonparametric correlation measure defined as

$$\begin{aligned}\tau &= \tau(X,Y)\\ &= \Pr\{(X_2 - X_1)(Y_2 - Y_1) \geq 0\} - \Pr\{(X_2 - X_1)(Y_2 - Y_1) < 0\},\end{aligned}$$

in which (X_1,Y_1) and (X_2,Y_2) are two independent realizations of a joint distribution.

Another nonparametric correlation measure is Spearman's rank correlation coefficient:

$$\text{RankCorr}(X,Y) = 12\text{E}\left[(F_X(X) - 0.5)(F_Y(Y) - 0.5)\right].$$

Both Kendall's tau and Spearman's rank correlation coefficient satisfy the following properties (see for example, Genest and Mackay [7]):

- $-1 \le \tau \le 1$; $-1 \le \text{RankCorr} \le 1$,
- if X and Y are comonotonic, then $\tau = 1$ and RankCorr = 1,
- if X and $-Y$ are comonotonic, then $\tau = -1$ and RankCorr = -1,
- if X and Y are independent, then $\tau = 0$ and RankCorr = 0,
- τ is invariant under strictly monotone transforms, that is, if f and g are strictly increasing (or decreasing) functions, then $\tau(f(X), g(Y)) = \tau(X, Y)$ and
$$\text{RankCorr}(f(X), g(Y)) = \text{RankCorr}(X, Y),$$
- if F_X and F_Y are the cumulative distribution functions of two continuous random variables, we have $\tau(F_X(X), F_Y(Y)) = \tau(X, Y)$ and $\text{RankCorr}(F_X(X), F_Y(Y)) = \text{RankCorr}(X, Y)$. Thus, Kendall's tau and rank correlation coefficient are often measured in terms of uniform random variables over $[0,1] \times [0,1]$.

Kendall's tau can be calculated, with due attention to singularity, as

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 F_{X,Y}(x, y) d^2 F_{X,Y}(x, y) - 1.$$

Assume that we have available a random sample of bivariate observations, (X_i, Y_i), $i = 1, \ldots, k$. A non-parametric estimate of Kendall's tau is

$$\hat{\tau}(X, Y) = \frac{2}{k(k-1)} \sum_{i<j} \text{sign}[(X_i - X_j)(Y_i - Y_j)],$$

where sign$[z]$ equals 1, 0, or -1 when z is positive, zero, or negative, respectively.

TABLE 2

A SAMPLE OF INCURRED LOSSES AND ALAE

Claimant #	Amount of Incurred Losses	Amount of ALAE
98001	50	5.0
98002	65	4.0
98003	28	0.0
98004	75	6.5
98005	38	4.5
Average	51.2	4
Std. Dev.	17.15	2.168

EXAMPLE 5.3 Suppose that we have a set of data for incurred losses and allocated loss adjusted expense as shown in Table 2.

The Pearson correlation coefficient can be estimated by:

$$\frac{(50-51.2)(5.0-4.0)+\cdots+(38-51.2)(4.5-4.0)}{5(17.15)(2.168)} = 0.78.$$

Kendall's tau can be estimated by

$$\hat{\tau}(X,Y) = \frac{2}{k(k-1)}\sum_{i<j}\text{sign}[(X_i - X_j)(Y_i - Y_j)] = 0.6.$$

To calculate the rank correlation coefficient, we first rank each claim by the ordering of losses and ALAE as shown in Table 3.

The Spearman rank correlation coefficient can be calculated as the ordinary Pearson correlation coefficient between the ranks of the losses and ALAE:

$$\frac{(3-3)(4-3)+\cdots+(2-3)(3-3)}{5\sqrt{2}\sqrt{2}} = 0.7.$$

The choice between Kendall's tau and the rank correlation coefficient depends on their relative simplicity for the intended application. Some commonly used random number generators

TABLE 3

RANK ORDERING OF LOSSES AND ALAE

Claimant #	Rank of Incurred Losses	Rank of ALAE
98001	3	4
98002	4	2
98003	1	1
98004	5	5
98005	2	3
Median	3	3
Average	3	3
Std. Dev.	$\sqrt{2}$	$\sqrt{2}$

(e.g., Palisade @Risk, which is a Microsoft Excel add-in) have implemented a method from Iman and Conver [12] to induce a given rank correlation structure.

6. THE CONCEPT OF COPULA

Recall that a Monte Carlo simulation of a random variable X can be achieved by first drawing a random uniform number u from $U \sim \text{Uniform}(0,1)$ and then inverting u by $x = F_X^{-1}(u)$. In a similar way, a Monte Carlo simulation of k variables, $(X_1,\ldots,X_k)$, usually starts with k uniform random variables, $(U_1,\ldots,U_k)$. If the variables $(X_1,\ldots,X_k)$ are independent (or correlated), then we need k independent (or correlated) uniform random variables $(U_1,\ldots,U_k)$. For a set of given marginal distributions, the correlation structure of the variables $(X_1,\ldots,X_k)$ is completely determined by the correlation structure of the uniform random variables, $(U_1,\ldots,U_k)$.

DEFINITION 2 *A copula is defined as the joint cumulative distribution function of k uniform random variables*

$$C(u_1,\ldots,u_k) = \Pr\{U_1 \le u_1,\ldots,U_k \le u_k\}.$$

For any set of arbitrary marginal distributions, the formula

$$F_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = C(F_{X_1}(x_1),\ldots,F_{X_k}(x_k)) \qquad (6.1)$$

defines a joint cumulative distribution function with marginal cumulative distributions $F_{X_1},\ldots,F_{X_k}$. The formula

$$S_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = C(S_{X_1}(x_1),\ldots,S_{X_k}(x_k)) \qquad (6.2)$$

defines a joint survivor function with marginal survivor function $S_{X_1},\ldots,S_{X_k}$.

The multivariate distributions given by Equations 6.1 and 6.2 are usually different, although they both have the same set of Kendall's tau and the same set of rank correlation coefficients.

7. THE COOK–JOHNSON FAMILY OF DISTRIBUTIONS

Let $(U_1,\ldots,U_k)$ be a k-dimensional uniform distribution with support on the hypercube $(0,1)^k$ and having the joint cumulative distribution function

$$F^{(\alpha)}_{U_1,\ldots,U_k}(u_1,\ldots,u_k) = \left\{\sum_{j=1}^{k} u_j^{-1/\alpha} - k + 1\right\}^{-\alpha}, \qquad (7.1)$$

where $u_j \in (0,1)$, $j = 1,\ldots,k$, and $\alpha > 0$. This multivariate uniform distribution has a Kendall's tau:

$$\tau(X_i,X_j) = \tau(U_i,U_j) = \frac{1}{1+2\alpha}.$$

On the other hand, for this family of multivariate distributions, there is no simple analytic form for the rank correlation coefficient.

Cook and Johnson [3] studied the family of multivariate uniform distributions given by Equation 7.1. They showed that

$$\lim_{\alpha\to 0} F^{(\alpha)}_{U_1,\ldots,U_k}(u_1,\ldots,u_k) = \min[u_1,\ldots,u_k],$$

and

$$\lim_{\alpha \to \infty} F^{(\alpha)}_{U_1,\ldots,U_k}(u_1,\ldots,u_k) = \prod_{j=1}^{k} u_j.$$

Thus, the correlation approaches its maximum (i.e., comonotonicity) when α decreases to zero, and the correlation approaches zero when α increases to infinity.

Cook and Johnson also gave the following simple simulation algorithm for the multivariate uniform distribution given by Equation 7.1:

STEP 1 Let $Y_1,\ldots,Y_k$ be independent and each have an exponential (1) distribution.

STEP 2 Let Z have a gamma$(\alpha,1)$ distribution.

STEP 3 Then the variables

$$U_j = [1 + Y_j/Z]^{-\alpha}, \qquad j = 1,\ldots,k, \tag{7.2}$$

have a joint cumulative distribution function given by Equation 7.1.

For a set of arbitrary marginal distributions, $F_{X_1},\ldots,F_{X_k}$, we can define a joint cumulative distribution function by

$$F_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \left\{ \sum_{j=1}^{k} F_{X_j}(x_j)^{-1/\alpha} - k + 1 \right\}^{-\alpha}. \tag{7.3}$$

Alternatively, we can define a joint survivor function by

$$S_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \left\{ \sum_{j=1}^{k} S_{X_j}(x_j)^{-1/\alpha} - k + 1 \right\}^{-\alpha}. \tag{7.4}$$

Note that Kendall's tau for this multivariate distribution is also $1/(1+2\alpha)$, which is determined by the underlying copula and is invariant under monotone transforms.

Consider the task of aggregating k risk portfolios, $(X_1,\ldots,X_k)$, where each X_j may represent the aggregate loss amount for the jth risk portfolio. If we assume that $(X_1,\ldots,X_k)$ have a multivariate distribution given by Equation 7.3, a simulation of $(X_1,\ldots,X_k)$ can be easily implemented by:

STEP 4 Invert the $(U_1,\ldots,U_k)$ in Equation 7.2 using $(F_{X_1}^{-1},\ldots, F_{X_k}^{-1})$.

Alternatively, if we assume that $(X_1,\ldots,X_k)$ have a multivariate distribution given by Equation 7.4, a simulation of $(X_1,\ldots,X_k)$ can be easily implemented by:

STEP 4* Invert the $(U_1,\ldots,U_k)$ in Equation 7.2 using $(S_{X_1}^{-1},\ldots, S_{X_k}^{-1})$.

In the multivariate uniform distribution given by Equation 7.1, all correlations are positive. Negative correlations can be accommodated by applying the transforms $U_i^* = 1 - U_i$ to some, but not all, uniform variables in Equation 7.2.

In this dependency model, no restriction is imposed on the marginal distributions, F_{X_j} or S_{X_j}, $j = 1,\ldots,k$. However, the correlation parameters are quite restricted in the sense that the Kendall's taus have to be the same for any pair of risks. To overcome this restriction in the correlation parameters, the *normal copula* permits arbitrary correlation parameters, $\tau_{ij} = \tau(X_i,X_j)$. It is explained in the next section.

8. THE NORMAL COPULA AND MONTE CARLO SIMULATION

In general, the modeling and combining of correlated risks are most straight-forward if the correlated risks have a multivariate normal distribution. In this section, we will use the multivariate normal distribution to construct the normal copula, and then use it to generate multivariate distributions with arbitrary marginal distributions. The normal copula enjoys much flexibility in the selection of correlation parameters. As well, it lends itself to simple Monte Carlo simulation techniques.

Assume that $(Z_1,\ldots,Z_k)$ have a multivariate normal distribution with standard normal marginal distribution $Z_j \sim N(0,1)$ and a positive definite correlation matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & & \vdots \\ \rho_{k1} & \rho_{k2} & \cdots & 1 \end{pmatrix},$$

where $\rho_{ij} = \rho_{ji}$ is the correlation coefficient between Z_i and Z_j. Then $(Z_1,\ldots,Z_k)$ have a joint probability density function:

$$f(z_1,\ldots,z_k) = \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}} \exp\left\{-\tfrac{1}{2}\mathbf{z}'\boldsymbol{\Sigma}^{-1}\mathbf{z}\right\}, \qquad \mathbf{z} = (z_1,\ldots,z_k). \tag{8.1}$$

From the correlation matrix $\boldsymbol{\Sigma}$ we can construct a lower triangular matrix

$$\mathbf{B} = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kk} \end{pmatrix},$$

such that $\boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}'$. In other words, the correlation matrix $\boldsymbol{\Sigma}$ equals the matrix product of $\mathbf{B}$ and its transpose $\mathbf{B}'$. The elements of the matrix $\mathbf{B}$ can be calculated from the following Choleski's algorithm (see Burden and Faires [2, Section 6.6]; Johnson [17, Section 4.1]):

$$b_{ij} = \frac{\rho_{ij} - \sum_{s=1}^{j-1} b_{is}b_{js}}{\sqrt{1 - \sum_{s=1}^{j-1} b_{js}^2}}, \qquad 1 \le j \le i \le n, \tag{8.2}$$

with the convention that $\sum_{s=1}^{0}(\cdot) = 0$. It is noted that:

- For $i > j$, the denominator of Equation 8.2 equals b_{jj}.
- The elements of **B** should be calculated from top to bottom and from left to right.

The following simulation algorithm can be used to generate multivariate normal variables with a joint probability density function given by Equation 8.1. See Herzog [9], and Fishman [5, pp. 223–224].

STEP 1 Construct the lower triangular matrix $\mathbf{B} = (b_{ij})$ by Equation 8.2.

STEP 2 Generate a column vector of independent standard normal variables $\mathbf{Y} = (Y_1, \ldots, Y_k)'$.

STEP 3 Take the matrix product $\mathbf{Z} = \mathbf{BY}$ of $\mathbf{B}$ and $\mathbf{Y}$. Then $\mathbf{Z} = (Z_1, \ldots, Z_k)'$ has the required joint probability density function given by Equation 8.1.

Let $\Phi(\cdot)$ represent the cumulative distribution function of the standard normal distribution:

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt.$$

Then $\Phi(Z_1), \ldots, \Phi(Z_k)$ have a multivariate uniform distribution with Kendall's tau (e.g., Frees and Valdez [6, pp. 25])

$$\tau(\Phi(Z_i), \Phi(Z_j)) = \tau(Z_i, Z_j) = \frac{2}{\pi} \arcsin(\rho_{ij}),$$

and (Spearman) rank correlation coefficient

$$\text{RankCorr}(\Phi(Z_i), \Phi(Z_j)) = \text{RankCorr}(Z_i, Z_j) = \frac{6}{\pi} \arcsin\left(\frac{\rho_{ij}}{2}\right),$$

where $\arcsin(x)$ is an inverse trigonometric function such that $\sin(\arcsin(x)) = x$.

Let's state this result more formally as a theorem due to its importance.

THEOREM 2 *Assume that* $(Z_1,\ldots,Z_k)$ *have a multivariate normal joint probability density function given by Equation* 8.1*, with correlation coefficient* $\rho_{ij} = \rho(Z_i,Z_j)$*. Let* $H(z_1,\ldots,z_k)$ *be their joint cumulative distribution function. Then*

$$C(u_1,\ldots,u_k) = H(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_k))$$

defines a multivariate uniform cumulative distribution function—called the normal copula.

For any set of given marginal cumulative distribution functions $F_1,\ldots,F_k$*, the variables*

$$X_1 = F_1^{-1}(\Phi(Z_1)),\ldots,X_k = F_k^{-1}(\Phi(Z_k))$$

have a joint cumulative distribution function

$$F_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = H(\Phi^{-1}(F_1(x_1)),\ldots,\Phi^{-1}(F_k(x_k)))$$

with marginal cumulative distribution functions $F_1,\ldots,F_k$*. The multivariate variables* $(X_1,\ldots,X_k)$ *have Kendall's tau*

$$\tau(X_i,X_j) = \tau(Z_i,Z_j) = \frac{2}{\pi}\arcsin(\rho_{ij})$$

and Spearman's rank correlation coefficients

$$\text{RankCorr}(X_i,X_j) = \text{RankCorr}(Z_i,Z_j) = \frac{6}{\pi}\arcsin\left(\frac{\rho_{ij}}{2}\right).$$

Although the normal copula does not have a simple analytical expression, it lends itself to a very simple Monte Carlo simulation algorithm.

Suppose that we are given a set of correlated risks $(X_1,\ldots,X_k)$ with marginal cumulative distribution functions $F_{X_1},\ldots,F_{X_k}$ and Kendall's tau $\tau_{ij} = \tau(X_i,X_j)$ or rank correlation coefficient

RankCorr(X_i, X_j). If we assume that $(X_1, \ldots, X_k)$ can be described by the normal copula in Theorem 2, then the following Monte Carlo simulation procedure can be used:

STEP 1 Convert the given Kendall's tau or rank correlation coefficient to our usual measure of correlation for multivariate normal variables:

$$\rho_{ij} = \sin\left(\frac{\pi}{2}\tau_{ij}\right) = 2\sin\left(\frac{\pi}{6}\text{RankCorr}(X_i, X_j)\right),$$

and construct the lower triangular matrix $\mathbf{B} = (b_{ij})$ by Equation 8.2.

STEP 2 Generate a column vector of independent standard normal variables $\mathbf{Y} = (Y_1, \ldots, Y_k)'$.

STEP 3 Take the matrix product of $\mathbf{B}$ and $\mathbf{Y}$: $\mathbf{Z} = (Z_1, \ldots, Z_k)' = \mathbf{BY}$.

STEP 4 Set $u_i = \Phi(Z_i)$ for $i = 1, \ldots, k$.

STEP 5 Set $X_i = F_{X_i}^{-1}(u_i)$ for $i = 1, \ldots, k$.

Theorem 2 and the associated simulation algorithm provide a powerful tool for generating correlated variables. The normal copula is very flexible as it allows any (symmetric, positive definite) matrix of rank correlation coefficients (or alternatively, Kendall's tau). The use of this algorithm implicitly assumes that the underlying variables can be described by a normal copula. Of course, there are many correlation structures that differ from a normal copula, for example, the Cook–Johnson distribution in Equation 7.1. In many practical situations, we only have some indication of the correlation parameters without knowing the exact underlying multivariate distribution. In such situations, a normal copula leads to a simple method of simulating the correlated variables.

Appendix B gives an overview of various other families of copulas and the associated Monte Carlo simulation techniques.

9. COMMON MIXTURE MODELS

In many situations, individual risks are correlated since they are subject to the same claim generating mechanism or are influenced by changes in the common underlying economic/legal environment. For instance, in property insurance, risk portfolios in the same geographic location are correlated, where individual claims are contingent on the occurrence and severity of a natural disaster (hurricane, tornado, earthquake, or severe weather condition). In liability insurance, new court rulings or social inflation may set new trends that affect the settlement of all liability claims for one line of business.

One way of modeling situations where the individual risks $\{X_1, X_2, \ldots, X_n\}$ are subject to the same external mechanism is to use a secondary mixing distribution. The uncertainty about the external mechanism is then described by a structure parameter, θ, which can be viewed as a realization of a random variable Θ. The aggregate losses of the risk portfolio can then be seen as a two-stage process: First the external parameter $\Theta = \theta$ is drawn from the distribution function, F_Θ, of Θ. Next, the claim frequency (or severity) of each individual risk X_i ($i = 1, 2, \ldots, n$) is obtained as a realization from the conditional distribution function, $F_{X_i|\Theta}(x_i \mid \theta)$, of $X_i \mid \Theta$.

9.1. *Common Poisson Mixtures*

Consider k discrete random variables $N_1, \ldots, N_k$. Assume that there exists a random parameter Θ such that

$$(N_j \mid \Theta = \theta) \sim \text{Poisson}(\theta\lambda_j), \qquad j = 1, \ldots, k,$$

where the variable Θ has a probability density function $\pi(\theta)$ and a moment generating function M_Θ. For any given $\Theta = \theta$, the variables $(N_j \mid \theta)$ are independent and Poisson $(\lambda_j\theta)$ distributed with a conditional joint probability generating function

$$\begin{aligned} P_{N_1,\ldots,N_k|\Theta}(t_1,\ldots,t_k \mid \theta) &= \mathrm{E}[t_1^{N_1}\cdots t_k^{N_k} \mid \Theta = \theta] \\ &= e^{\theta[\lambda_1(t_1-1)+\cdots\lambda_k(t_k-1)]}. \end{aligned}$$

However, unconditionally, $N_1,\ldots,N_k$ are correlated as they depend upon the same random parameter Θ. The unconditional joint probability generating function for $N_1,\ldots,N_k$ is

$$\begin{aligned} P_{N_1,\ldots,N_k}(t_1,\ldots,t_k) &= \mathrm{E}_\Theta[\mathrm{E}[t_1^{N_1}\cdots t_k^{N_k} \mid \Theta]] \\ &= \int_0^\infty e^{\theta[\lambda_1(t_1-1)+\cdots+\lambda_k(t_k-1)]}\pi(\theta)\,d\theta \\ &= M_\Theta(\lambda_1(t_1-1)+\cdots+\lambda_k(t_k-1)). \end{aligned}$$

It has marginal probability generating functions $P_{N_j}(t_j) = M_\Theta(\lambda_j(t_j-1))$ with $\mathrm{E}[N_j] = \lambda_j \mathrm{E}[\Theta]$.

Note that

$$\begin{aligned} \mathrm{Cov}[N_i,N_j] &= \mathrm{E}_\Theta \mathrm{Cov}[N_i \mid \Theta, N_j \mid \Theta] + \mathrm{Cov}[\mathrm{E}[N_i \mid \Theta],\mathrm{E}[N_j|\Theta]] \\ &= \mathrm{Cov}[\Theta\lambda_i,\Theta\lambda_j] = \lambda_i\lambda_j \mathrm{Var}[\Theta]. \end{aligned}$$

The covariance coefficient between N_i and N_j $(i \neq j)$ is

$$\omega(N_i,N_j) = \frac{\mathrm{Cov}[N_i,N_j]}{\mathrm{E}[N_i]\mathrm{E}[N_j]} = \frac{\mathrm{Var}[\Theta]}{\{\mathrm{E}[\Theta]\}^2},$$

where ω is the same for all i and j.

EXAMPLE 9.1 If Θ has a gamma$(\alpha,1)$ distribution with moment generating function $M_\Theta(z) = (1-z)^{-\alpha}$, then

$$P_{N_1,\ldots,N_k}(t_1,\ldots,t_k) = [1-\lambda_1(t_1-1)-\cdots-\lambda_k(t_k-1)]^{-\alpha} \tag{9.1}$$

defines a multivariate negative binomial with marginal distributions NB(α,λ_j) and covariance coefficients $\omega(N_i,N_j) = 1/\alpha$.

EXAMPLE 9.2 If Θ has an inverse Gaussian distribution, $IG(\beta,1)$, with a moment generating function $M_\Theta(z) = e^{1/\beta[1-\sqrt{1-2\beta z}]}$, then

$$P_{N_1,\dots,N_k}(t_1,\dots,t_k) = \exp\left\{\frac{1}{\beta} - \frac{1}{\beta}\sqrt{1-2\beta[\lambda_1(t_1-1)+\cdots+\lambda_k(t_k-1)]}\right\}$$

defines a multivariate Poisson inverse Gaussian with marginal distributions P-IG$(\beta\lambda_j,\lambda_j)$ and covariance coefficients $\omega(N_i,N_j) = \beta$.

Consider combining k risk portfolios. Assume that the frequencies N_j, $j = 1,\dots,k$, are correlated via a common Poisson-gamma mixture and have a joint probability generating function given by Equation 9.1. If the severities X_j, $j = 1,\dots,k$, are mutually independent and independent of the frequencies, there is a simple method of combining the aggregate loss distributions. Given $\lambda = \lambda_1 + \cdots + \lambda_k$ and

$$P_X(t) = \frac{\lambda_1}{\lambda}P_{X_1}(t) + \cdots \frac{\lambda_k}{\lambda}P_{X_k}(t),$$

then

$$P_{N_1,\dots,N_k}(P_{X_1}(t),\dots,P_{X_k}(t)) = [1-\lambda(P_X(t)-1)]^{-\alpha}.$$

In other words, the total loss amount for the combined risk portfolios has a compound negative binomial distribution with the severity distribution being a weighted average of individual severity distributions. In this case, dependency does not complicate the computation; in fact, it simplifies the calculation. It is simpler than combining independent compound negative binomial distributions.

In this multivariate Poisson-gamma mixture model, the k marginal distributions, negative binomial (α,λ_j), are required to have the same parameter α. This requirement limits its applicability in combining risk portfolios; in many practical cases the frequencies, negative binomial (α_j,λ_j), have different parameter values, α_j. Section 10 and Section 12 overcome this limitation by extending the Poisson-gamma mixture model to allow arbitrary negative binomial frequencies.

Similar arguments can be made about the Poisson inverse Gaussian distributions.

9.2. *Common Exponential Mixtures*

Consider k continuous random variables $X_1,\ldots,X_k$. Assume that there exists a random parameter Θ such that $(X_j \mid \Theta = \theta)$ is exponentially distributed with parameter $\lambda_j\theta$ and survivor function

$$S_{X_j|\Theta}(t_j \mid \theta) = \Pr\{X_j > t_j \mid \Theta = \theta\} = e^{-\theta\lambda_j t_j}, \qquad j = 1,\ldots,k,$$

where the variable Θ has a probability density function $\pi(\theta)$ and a moment generating function M_Θ.

For any given $\Theta = \theta$, the variables $(X_j \mid \theta)$, $j = 1,\ldots,k$, are conditionally independent and have a conditional joint survivor function

$$\begin{aligned} S_{X_1,\ldots,X_k|\Theta}(t_1,\ldots,t_k \mid \theta) &= \Pr\{X_1 > t_1,\ldots,X_k > t_k \mid \Theta = \theta\} \\ &= e^{-\theta[\lambda_1 t_1 + \cdots + \lambda_k t_k]}. \end{aligned}$$

However, unconditionally, $X_1,\ldots,X_k$ are correlated as they depend upon the same random parameter Θ. The unconditional joint survivor function for $X_1,\ldots,X_k$ is

$$\begin{aligned} S_{X_1,\ldots,X_k}(t_1,\ldots,t_k) &= \int_0^\infty e^{-\theta[\lambda_1 t_1 + \cdots + \lambda_k t_k]}\pi(\theta)\,d\theta \\ &= M_\Theta(-\lambda_1 t_1 - \cdots - \lambda_k t_k). \end{aligned}$$

EXAMPLE 9.3 If Θ has a gamma$(\alpha,1)$ distribution with moment generating function $M_\Theta(z) = (1-z)^{-\alpha}$, this defines a family of multivariate Pareto distributions

$$S_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = [1 + \lambda_1 t_1 + \cdots + \lambda_k t_k]^{-\alpha},$$

with marginal distributions being Pareto$(\alpha, 1/\lambda_j)$.

EXAMPLE 9.4 If Θ has an inverse Gaussian distribution with moment generating function $M_\Theta(z) = e^{1/\beta[1-\sqrt{1-2\beta z}]}$, this defines a family of multivariate exponential inverse Gaussian distributions

$$S_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = \exp\left[\frac{1}{\beta} - \frac{1}{\beta}\sqrt{1 + 2\beta(\lambda_1 t_1 + \cdots + \lambda_k t_k)}\right],$$

with marginal distributions being exponential inverse Gaussian, E-IG$(\beta\lambda_j, \lambda_j)$.

Now we consider the aggregation of k individual claim amounts. Suppose that the k individual claim amounts $X_1,\ldots,X_k$ are identically distributed with $X_i \sim$ Pareto(α,β). But they are correlated by a common exponential-gamma mixture with a joint survivor function

$$S_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = \left[1 + \frac{1}{\beta}(t_1 + \cdots + t_k)\right]^{-\alpha}.$$

Then the sum $X_1 + \cdots + X_k$ has a Pareto$(\alpha, n\beta)$ distribution. This is because, for any given $\Theta = \theta$, $(X_1 + \cdots + X_k \mid \theta) \sim$ exponential(θ/n).

Alternatively, this common exponential mixture model can be obtained by applying the Cook–Johnson copula to k identical marginal survivor functions, Pareto(α,β). In other words, the Cook–Johnson copula can be viewed as an extension of the common exponential mixture model.

10. EXTENDED COMMON POISSON MIXTURE MODELS

The common Poisson mixture model in the previous section has a simple correlation structure and is easy to use. However, it is quite restricted in the sense that it does not permit arbitrary parameter values in the marginal distributions. In this section we extend the common Poisson mixture model so that the marginal distributions may have arbitrary parameter values. This extended model permits simple implementation by Monte Carlo simulation.

Suppose that there exist random variables $(\Theta_1,\ldots,\Theta_k)$ such that for a given set of values $(\Theta_1 = \theta_1,\ldots,\Theta_k = \theta_k)$, the conditional variables $(N_1,\ldots,N_k)$ are independent Poisson(θ_j) variables with

$$P_{N_1,\ldots,N_k|(\Theta_1,\ldots,\Theta_k)}(t_1,\ldots,t_k \mid \theta_1,\ldots,\theta_k)$$
$$= \prod_{j=1}^{k} P_{N_j}(t_j \mid \theta_j) = \prod_{j=1}^{k} e^{\theta_j(t_j-1)},$$

where $M_{\Theta_1,\ldots,\Theta_k}(t_1,\ldots,t_k) = \mathrm{E}_{\Theta_1,\ldots,\Theta_k}[e^{t_1\Theta_1+\cdots+t_k\Theta_k}]$ is the joint moment generating function of $(\Theta_1,\ldots,\Theta_k)$.

The unconditional joint probability generating function is

$$P_{N_1,\ldots,N_k}(t_1,\ldots,t_k) = \mathrm{E}_{(\Theta_1,\ldots,\Theta_k)}P_{N_1,\ldots,N_k}(t_1,\ldots,t_k \mid \Theta_1,\ldots,\Theta_k)$$
$$= M_{\Theta_1,\ldots,\Theta_k}((t_1-1),\ldots,(t_k-1)).$$

By taking the first and second order partial derivatives of this joint probability generating function at $(1,\ldots,1)$, we obtain

$$\mathrm{E}[N_i] = \mathrm{E}[\Theta_i] \qquad \text{and} \qquad \mathrm{Cov}[N_i,N_j] = \mathrm{Cov}[\Theta_i,\Theta_j].$$

We observe a one-to-one correspondence between the correlation structures of the variables $(N_1,\ldots,N_k)$ and the mixing parameters $(\Theta_1,\ldots,\Theta_k)$.

Now consider the case that $\Theta_j \sim \mathrm{gamma}(\alpha_j,\beta_j)$ and thus $N_j \sim \mathrm{NB}(\alpha_j,\beta_j)$, with arbitrary parameter values, $\alpha_j,\beta_j > 0$. We further assume that the variables Θ_j, $j = 1,\ldots,k$, are comonotonic and thus can be simulated by using the same set of uniform random numbers. For $i \neq j$, the covariance $\mathrm{Cov}[\Theta_i,\Theta_j]$ can be numerically calculated by using Equation 5.1. For this dependency model, we have a simple Monte Carlo simulation algorithm:

STEP 1 Generate a uniform number, u, from $U \sim \mathrm{Uniform}(0,1)$.

STEP 2 Let $\theta_j = F_{\Theta_j}^{-1}(u)$, where $\Theta_j \sim \mathrm{gamma}(\alpha_j,\beta_j)$, $j = 1,\ldots,k$.

STEP 3 Simulate $(N_1,\ldots,N_k)$ from k independent Poisson(θ_j) variables, $j = 1,\ldots,k$.

If the α_js are the same, we get the common Poisson mixture model in Example 9.1.

11. COMPONENT MODELS

Consider the aggregation of different lines of business. For a multi-line insurer, the correlation between lines of business may differ from one region to another. Therefore, it may be more appropriate to divide each line into components and model the correlation separately for each component (e.g., by geographic region). There may exist higher correlations between lines in a high catastrophe risk region where the presence of the catastrophe risk may generate a common shock or a common mixture.

Note that many families of frequency and severity distributions are infinitely divisible. A family of distributions is infinitely divisible if any member can be obtained as an independent sum of other members in the same family. Let $X \oplus Y$ represent the sum of two independent random variables and $F_Y \oplus F_Y$ represent the convolution of two probability distributions. We have

- Poisson$(\lambda_1) \oplus$ Poisson(λ_2) = Poisson$(\lambda_1 + \lambda_2)$
- negative binomial: $\mathrm{NB}(\alpha_1,\beta) \oplus \mathrm{NB}(\alpha_2,\beta) = \mathrm{NB}(\alpha_1 + \alpha_2,\beta)$
- Poisson inverse Gaussian:

$$\text{P-IG}(\beta,\mu_1) \oplus \text{P-IG}(\beta,\mu_2) = \text{P-IG}(\beta,\mu_1 + \mu_2)$$

- gamma$(\alpha_1,\beta) \oplus$ gamma(α_2,β) = gamma$(\alpha_1 + \alpha_2,\beta)$
- inverse Gaussian: $\mathrm{IG}(\beta,\mu_1) \oplus \mathrm{IG}(\beta,\mu_2) = \mathrm{IG}(\beta,\mu_1 + \mu_2)$.

Infinitely divisible distributions are especially useful for dividing risks into independent components. Consider k infinitely divisible risks $X_j(\alpha_j)$, $j = 1,\ldots,k$, with α_j as the divisible parameter.

Consider a decomposition:

$$\begin{matrix} X_1(\alpha_1) = X_{11}(\alpha_{11}) \oplus \cdots \oplus X_{1n}(\alpha_{1n}) \\ \vdots \qquad \vdots \qquad \vdots \\ X_k(\alpha_k) = X_{k1}(\alpha_{k1}) \oplus \cdots \oplus X_{kn}(\alpha_{kn}) \end{matrix}, \qquad \alpha_{js} \geq 0. \quad (11.1)$$

Then we can generate correlation structures component by component:

$$P_{X_1,\ldots,X_k} = \prod_{s=1}^{n} Q_{X_{1s},\ldots,X_{ks}},$$

where the joint probability generating function $Q_{X_{1s},\ldots,X_{ks}}$ for the sth components can be modeled by using a common mixture, a common shock (described below), or by assuming independence, as appropriate. It can be verified that for the component model in Equation 11.1 we have

$$\mathrm{Cov}[X_i, X_j] = \sum_{s=1}^{n} \mathrm{Cov}[X_{is}, X_{js}].$$

11.1. Common Shock Models

Let $X_j = X_{ja} \oplus X_{jb}$, $j = 1,\ldots,k$, be a decomposition into two *independent* components

$$P_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = \mathrm{E}[t_1^{X_{1a}} \cdots t_k^{X_{ka}}]\mathrm{E}[t_1^{X_{1b}} \cdots t_k^{X_{kb}}].$$

If $X_{1a} = \cdots = X_{ka} = X_0$, we obtain

$$P_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = \mathrm{E}[(t_1 \cdots t_k)^{X_0}]\mathrm{E}[t_1^{X_{1b}} \cdots t_k^{X_{kb}}].$$

In particular, if the X_{ib}s are independent, we have $\mathrm{Cov}[X_i, X_j] = \mathrm{Var}[X_0]$. The only source of correlation comes from the common shock variable X_0.

EXAMPLE 11.1 Consider the aggregation of two correlated compound Poisson distributions:

- Portfolio 1. The claim frequency N_1 has a Poisson(λ_1) distribution, and the claim severity X has a probability function $f_1(x)$.
- Portfolio 2. The claim frequency N_2 has a Poisson(λ_2) distribution, and the claim severity Y has a probability function $f_2(y)$.
- Assume that X, Y are independent and both are independent of (N_1,N_2). However, N_1 and N_2 are correlated via a common shock model

$$N_1 = N_0 \oplus N_{1b}, \qquad N_2 = N_0 \oplus N_{2b},$$

where $N_0 \sim \text{Poisson}(\lambda_0)$, $N_{1b} \sim \text{Poisson}(\lambda_1 - \lambda_0)$, and $N_{2b} \sim \text{Poisson}(\lambda_2 - \lambda_0)$.

In this common shock model (N_1,N_2) have a joint probability generating function:

$$\begin{aligned} P_{N_1,N_2}(t_1,t_2) &= \mathrm{E}[t_1^{N_1} t_2^{N_2}] \\ &= \exp[\lambda_1(t_1-1) + \lambda_2(t_2-1) + \lambda_0(t_1-1)(t_2-1)], \end{aligned}$$

with $\mathrm{Cov}[N_1,N_2] = \mathrm{Var}[X_0] = \lambda_0$. It can be shown that the aggregate losses for the combined risk portfolio,

$$S = (X_1 + \cdots + X_{N_1}) + (Y_1 + \cdots + Y_{N_2}),$$

have a compound Poisson($\lambda_1 + \lambda_2 - \lambda_0$) distribution with a severity probability function

$$\begin{aligned} f(z) = {} & \frac{\lambda_1 - \lambda_0}{\lambda_1 + \lambda_2 - \lambda_0} f_1(z) + \frac{\lambda_2 - \lambda_0}{\lambda_1 + \lambda_2 - \lambda_0} f_2(z) \\ & + \frac{\lambda_0}{\lambda_1 + \lambda_2 - \lambda_0} f_{1*2}(z), \end{aligned}$$

where f_{1*2} represents the convolution of f_1 and f_2. Thus, existing methods can be applied.

This common shock model can be easily extended to any higher dimension ($k > 2$). For illustrative purposes, we now give an example involving three frequency variables.

EXAMPLE 11.2 The joint probability generating function

$$P_{N_1,N_2,N_3}(t_1,t_2,t_3) = \exp\left\{\sum_{i=1}^{3}\lambda_{ii}(t_i-1)+\sum_{i<j}\lambda_{ij}(t_it_j-1)+\lambda_{123}(t_1t_2t_3-1)\right\} \tag{11.2}$$

defines a multivariate Poisson distribution with marginal distributions

$$N_j \sim \text{Poisson}\left(\lambda_{123}+\sum_{i=1}^{3}\lambda_{ij}\right), \qquad j = 1,2,3,$$

and for $i \neq j$, $\text{Cov}[N_i,N_j] = \lambda_{ij} + \lambda_{123}$.

We let

- $K_{ii} \sim \text{Poisson}(\lambda_{ii})$, for $i = 1,2,3$,
- $K_{ij} \sim \text{Poisson}(\lambda_{ij})$, for $1 \leq i < j \leq 3$,
- $K_{ij} = K_{ji}$, for $1 \leq i,j \leq 3$,
- $K_{123} \sim \text{Poisson}(\lambda_{123})$,
- $N_j = K_{1j} \oplus K_{2j} \oplus K_{3j} \oplus K_{123}$, for $j = 1,2,3$.

Then the resulting (N_1,N_2,N_3) have a joint probability generating function given by Equation 11.2. In this model, K_{123} represents the common shock among all three variables (N_1,N_2,N_3). In addition, for $i \neq j$, $K_{ij} = K_{ji}$ represents the extra common shock between N_i and N_j.

Note that we can easily simulate the correlated frequencies, (N_1, N_2, N_3), component by component.

Subject to scale transforms, the common shock multivariate Poisson model can be extended to gamma variables.

EXAMPLE 11.3 Consider two variables $X_1 \sim \text{gamma}(\alpha_1, \beta_1)$ and $X_2 \sim \text{gamma}(\alpha_2, \beta_2)$. Suppose there is a decomposition

$$X_1 = \beta_1(X_0 \oplus X_{1b}), \qquad X_2 = \beta_2(X_0 \oplus X_{2b}),$$

where $X_0 \sim \text{gamma}(\alpha_0, 1)$ with $\alpha_0 \leq \min\{\alpha_1, \alpha_2\}$, $X_{1b} \sim \text{gamma}(\alpha_1 - \alpha_0, 1)$ and $X_{2b} \sim \text{gamma}(\alpha_2 - \alpha_0, 1)$. Then $\text{Cov}[X_1, X_2] = \beta_1\beta_2\text{Var}[X_0] = \alpha_0\beta_1\beta_2$, and

$$X_1 + X_2 = (\beta_1 + \beta_2)X_0 \oplus \beta_1 X_{1b} \oplus \beta_2 X_{2b}.$$

11.2. Peeling Method

Recall that the common Poisson-gamma mixture requires that the marginal distributions $N_j \sim \text{NB}(\alpha, \lambda_j)$ must have the same parameter value α. Now we shall illustrate that, by using the component method, we can construct correlated multivariate negative binomials with arbitrary parameters (α_j, λ_j).

Suppose that we are given k marginal negative binomial distributions:

$$N_1 \sim \text{NB}(\alpha_1, \lambda_1), \ldots, N_k \sim \text{NB}(\alpha_k, \lambda_k).$$

Model 1. Let $\alpha_0 \leq \min\{\alpha_1, \ldots, \alpha_k\}$, and let each N_j $(j = 1, \ldots, k)$ have a decomposition:

$$N_j = N_{ja} \oplus N_{jb}, \qquad N_{ja} \sim \text{NB}(\alpha_0, \lambda_j), \qquad N_{jb} \sim \text{NB}(\alpha_j - \alpha_0, \lambda_j).$$

Note that the N_{ja}s have the same parameter α_0, and thus can be modeled by a common Poisson-gamma mixture

$$P_{N_{1a},\ldots,N_{ka}}(t_1, \ldots, t_k) = \{1 - \lambda_1(t_1 - 1) - \cdots - \lambda_k(t_k - 1)\}^{-\alpha_0}.$$

If we assume that the N_{jb}s are independent, then $(N_1,\ldots,N_k)$ have a joint probability generating function

$$P_{N_1,\ldots,N_k}(t_1,\ldots,t_k) = \{1-\lambda_1(t_1-1)-\cdots-\lambda_k(t_k-1)\}^{-\alpha_0} \times \prod_{j=1}^{k}\{1-\lambda_j(t_j-1)\}^{\alpha_0-\alpha_j}.$$

Note that

$$\mathrm{Cov}[N_i,N_j] = \alpha_0\lambda_i\lambda_j = \frac{\alpha_0}{\alpha_i\alpha_j}\mathrm{E}[N_i]\mathrm{E}[N_j].$$

Simple methods exist for combining the individual aggregate loss distributions, provided that the severities are mutually independent and independent of $(N_1,\ldots,N_k)$.

Model 2. Assume that the α_j are in an ascending order, $\alpha_1 \leq \cdots \leq \alpha_k$. The decomposition

$$\mathrm{NB}(\alpha_j,\lambda_j) = \mathrm{NB}(\alpha_1,\lambda_j)\oplus \mathrm{NB}(\alpha_2-\alpha_1,\lambda_j) \oplus\cdots\oplus \mathrm{NB}(\alpha_j-\alpha_{j-1},\lambda_j)$$

can be used in conjunction with common mixture models to generate the following joint probability generating function:

$$\begin{aligned}P_{N_1,\ldots,N_k}(t_1,\ldots,t_k) &= \{1-\lambda_1(t_1-1)-\cdots-\lambda_k(t_k-1)\}^{-\alpha_1}\\ &\quad\times\{1-\lambda_2(t_2-1)-\cdots-\lambda_k(t_k-1)\}^{\alpha_1-\alpha_2}\\ &\quad\times\cdots\times\{1-\lambda_k(t_k-1)\}^{\alpha_{k-1}-\alpha_k}.\end{aligned}$$

It can be verified that the marginal univariate probability generating function is $P_{N_j}(t_j) = [1-\lambda_j(t_j-1)]^{-\alpha_j}$ and the marginal bivariate probability generating function is

$$P_{N_i,N_j}(t_i,t_j) = \{1-\lambda_i(t_i-1)-\lambda_j(t_j-1)\}^{-\alpha_i} \times\{1-\lambda_j(t_j-1)\}^{\alpha_i-\alpha_j}, \qquad i<j,$$

with

$$\mathrm{Cov}[N_i,N_j] = \alpha_i\lambda_i\lambda_j = \frac{1}{\alpha_j}\mathrm{E}[N_i]\mathrm{E}[N_j].$$

11.3. Mixed Correlation Models

Assume that the joint probability generating functions $P_{X_1,\ldots,X_k}$ and $Q_{X_1,\ldots,X_k}$ have the same set of marginal probability generating functions $P_{X_1},\ldots,P_{X_k}$. Then the mixed joint probability generating function

$$qP_{X_1,\ldots,X_k}(t_1,\ldots,t_k) + (1-q)Q_{X_1,\ldots,X_k}(t_1,\ldots,t_k), \qquad (0 < q < 1),$$

also has marginal probability generating functions $P_{X_1},\ldots,P_{X_k}$. For this mixed joint probability generating function, we have

$$\mathrm{Cov}[X_i,X_j] = (1-q)\mathrm{Cov}^P[X_i,X_j] + q\,\mathrm{Cov}^Q[X_i,X_j],$$

where Cov^P and Cov^Q represent the covariances implied by the joint probability generating functions P and Q, respectively.

A mixture of joint probability generating functions can be used to represent a set of possible scenarios. For instance, we can let P represent the joint probability generating function under the scenario of major catastrophe occurrence, Q correspond to zero catastrophe occurrence, and q represent the probability of the catastrophe occurrence.

12. THE DISTORTION METHOD

Let $X_1,\ldots,X_k$ be k random variables (discrete, continuous, or multivariate variables) with probability generating functions $P_{X_1}(t_1),\ldots,P_{X_k}(t_k)$, respectively. If the X_js are mutually independent, we have

$$P_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = \prod_{j=1}^{k} P_{X_j}(t_j).$$

Let g be a strictly increasing function over $[0,1]$ with $g(1) = 1$ and whose inverse function is g^{-1}. In a quite loose sense, we assume that $g \circ P_{X_1,\ldots,X_k}$ specifies a joint probability generating function with marginal probability generating functions $g \circ P_{X_j}$, $(j = 1,\ldots,k)$. By assuming that the distorted joint probability generating function $g \circ P_{X_1,\ldots,X_k}$ has non-correlated marginal probability generating functions, namely,

$$g \circ P_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = \prod_{j=1}^{k} g \circ P_{X_j}(t_j),$$

a correlation structure is introduced to the original joint probability generating function:

$$P_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = g^{-1}\left\{\prod_{j=1}^{k} g \circ P_{X_j}(t_j)\right\}.$$

For mathematical convenience we introduce $h(x) = \ln g(x)$ which is a strictly increasing function over $[0,1]$ with $h(1) = 0$. In terms of h, the above equation can be expressed as

$$P_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = h^{-1}\left\{\sum_{j=1}^{k} h \circ P_{X_j}(t_j)\right\}. \qquad (12.1)$$

Note that Equation 12.1 may not define a proper multivariate distribution, as the only constraint on the joint probability (density) function is that it sums to one. It defines a proper multivariate distribution if and only if the joint probability (density) function, $f_{X_1,\ldots,X_k}$, is non-negative everywhere.

Recall that for a discrete distribution,

$$f_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \frac{\partial^{x_1+\cdots+x_k}}{(\partial t_1)^{x_1}\cdots(\partial t_k)^{x_k}} P_{X_1,\ldots,X_k}(0,\ldots,0)\prod_{i=1}^{k}\frac{1}{x_i!},$$

which can also be derived using multivariate Taylor series expansion. Thus, $P_{X_1,\ldots,X_k}$ defines a proper joint probability distribution if and only if its partial derivatives at $t_1 = \cdots = t_k = 0$ are all non-negative.

THEOREM 3 *Suppose that Equation* 12.1 *defines a joint probability generating function; we have*

$$\mathrm{Cov}[X_i, X_j] = -\left\{\frac{h''(1)}{h'(1)} + 1\right\}\mathrm{E}[X_i]\mathrm{E}[X_j].$$

Proof We take the second order partial derivative, $\partial^2/\partial t_i \partial t_j$, $(i \neq j)$, on both sides of the equation

$$h \circ P_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = \sum_{j=1}^{k} h \circ P_{X_j}(t_j).$$

We obtain zero by taking the second order partial derivative, $\partial^2/\partial t_i \partial t_j$, $(i \neq j)$, on the right-hand side. Thus we should also get zero for the second order partial derivative on the left-hand side:

$$0 = \frac{\partial^2}{\partial t_i \partial t_j}\left\{h \circ P_{X_1,\ldots,X_k}\right\} = \frac{\partial}{\partial t_i}\left\{h'(P_{X_1,\ldots,X_k})\frac{\partial P_{X_1,\ldots,X_k}}{\partial t_j}\right\},$$

which further yields that

$$h''(P_{X_1,\ldots,X_k})\frac{\partial P_{X_1,\ldots,X_k}}{\partial t_i}\frac{\partial P_{X_1,\ldots,X_k}}{\partial t_j} + h'(P_{X_1,\ldots,X_k})\frac{\partial^2 P_{X_1,\ldots,X_k}}{\partial t_i \partial t_j} = 0.$$

Setting the values $t_s = 1$ for $s = 1,\ldots,k$, we get

$$h''(1)\mathrm{E}[X_i]\mathrm{E}[X_j] + h'(1)\mathrm{E}[X_i X_j] = 0.$$

This family of multivariate distributions has a symmetric structure in the sense that ω_{ij} is the same for all $i \neq j$. It would be suitable for combining risks in the same class, where any two individual risks share the same covariance coefficient.

Questions remain as to which distortion function to use and whether the distortion method in Equation 12.1 defines a proper multivariate distribution. In general, the feasibility of the distortion method depends on the marginal distributions.

The next section shows how the distortion method is inherently connected to the common Poisson-mixture models.

12.1. Links with the Common Poisson Mixtures

Reconsider the common Poisson mixture model in Section 9: for any given θ, $(N_j \mid \Theta = \theta)$, $j = 1, \ldots, k$, are conditionally independent Poisson variables with mean $\lambda_j\theta$. If the random parameter Θ has a moment generating function M_Θ, then $(N_1, \ldots, N_k)$ has an unconditional joint probability generating function

$$P_{N_1,\ldots,N_k}(t_1, \ldots, t_k) = M_\Theta(\lambda_1(t_1 - 1) + \cdots + \lambda_k(t_k - 1)),$$

with marginal probability generating function

$$P_{N_j}(t_j) = M_\Theta(\lambda_j(t_j - 1)).$$

LEMMA 3 *For a non-negative random variable Θ, the inverse of the moment generating function, M_Θ^{-1}, is well defined over the range* $[0,1]$ *with* $(d/du)M_\Theta^{-1}(u) > 0$, $M_\Theta^{-1}(0) = -\infty$, *and* $M_\Theta^{-1}(1) = 0$.

If we define $h(y) = M_\Theta^{-1}(y)$, then the joint probability generating function for the common Poisson mixture model satisfies

$$P_{N_1,\ldots,N_k}(t_1, \ldots, t_k) = h^{-1}\left\{\sum_{j=1}^{k} h \circ P_{N_j}(t_j)\right\}.$$

EXAMPLE 12.1 If Θ has a gamma$(1/\omega, 1)$ distribution with moment generating function $M_\Theta(z) = (1 - z)^{-1/\omega}$, then $h(y) = 1 - y^{-\omega}$, and we get the following joint probability generating function:

$$P^{(\omega)}_{N_1,\ldots,N_k}(t_1, \ldots, t_k) = \left\{P_{N_1}(t_1)^{-\omega} + \cdots + P_{N_k}(t_k)^{-\omega} - k + 1\right\}^{-1/\omega}, \quad \omega \neq 0,$$

with

$$\mathrm{Cov}[N_i, N_j] = \omega\, \mathrm{E}[N_i]\mathrm{E}[N_j] \qquad \text{and}$$

$$\lim_{\omega \to 0} P^{(\omega)}_{N_1,\ldots,N_k} = P_{N_1}(t_1) \cdots P_{N_k}(t_k).$$

EXAMPLE 12.2 If Θ has an inverse Gaussian distribution, $\mathrm{IG}(\omega, 1)$, with a moment generating function

$$M_\Theta(z) = \exp\left\{\frac{1}{\omega}[1 - \sqrt{1 - 2\omega z}]\right\},$$

then $h(y) = \ln y - \omega/2(\ln y)^2$, and we get the following **joint** probability generating function:

$$P^{(\omega)}_{N_1,\ldots,N_k}(t_1,\ldots,t_k)$$
$$= \exp\left\{\frac{1}{\omega} - \sqrt{\frac{1}{\omega^2} - \sum_{j=1}^{k}\left[\frac{2}{\omega}\ln P_{N_j}(t_j) - (\ln P_{N_j}(t_j))^2\right]}\right\},$$

with

$$\mathrm{Cov}[N_i, N_j] = \omega\,\mathrm{E}[N_i]\mathrm{E}[N_j] \qquad \text{and}$$

$$\lim_{\omega\to 0} P^{(\omega)}_{N_1,\ldots,N_k} = P_{N_1}(t_1)\cdots P_{N_k}(t_k).$$

12.2. A Family of Multivariate Negative Binomial Distributions

As an example of the distortion method, we now discuss a family of multivariate distributions with arbitrary negative binomial marginal distributions, $\mathrm{NB}(\alpha_j, \beta_j)$, $j = 1,\ldots,k$.

THEOREM 4 *The joint probability generating function*

$$P_{N_1,\ldots,N_k}(t_1,\ldots,t_k) = \left\{\sum_{j=1}^{k}[1 - \beta_j(t_j - 1)]^{\alpha_j\omega} - k + 1\right\}^{-1/\omega},$$
$$\omega \neq 0, \qquad (12.2)$$

defines a multivariate negative binomial distribution with marginal distributions $\mathrm{NB}(\alpha_j, \beta_j)$ *when either of the following conditions holds:*

- $0 < \omega < \min\{1/\alpha_j,\ j = 1,\ldots,k\}$,

- $\omega < 0$ such that $P_{N_1,\ldots,N_k}(0,\ldots,0) > 0$ and $1/\omega$ is a negative integer.

Proof Equation 12.2 can be rewritten as

$$P_{N_1,\ldots,N_k}(t_1,\ldots,t_k) = Q(t_1,\ldots,t_k)^{-1/\omega},$$

where

$$Q(t_1,\ldots,t_k) = \sum_{j=1}^{k}[1+\beta_j-\beta_j t_j]^{\alpha_j\omega} - k + 1.$$

(i) For $0 < \omega < \min\{1/\alpha_j,\ j=1,\ldots,k\}$, we have $\alpha_j\omega \le 1$; and the partial derivatives $(\partial^{x_1+\cdots+x_k}/(\partial t_1)^{x_1}\cdots(\partial t_k)^{x_k})P_{N_1,\ldots,N_k}$ are the sum of terms of the following form:

$$aQ(t_1,\ldots,t_k)^{-b}\prod_{j=1}^{k}[1+\beta_j-\beta_j t_j]^{-c_j}, \qquad a,b,c_j \ge 0.$$

Thus, the joint probability function

$$f_{N_1,\ldots,N_k}(x_1,\ldots,x_k) = \frac{\partial^{x_1+\cdots+x_k}}{(\partial t_1)^{x_1}\cdots(\partial t_k)^{x_k}}P_{N_1,\ldots,N_k}(0,\ldots,0)\prod_{i=1}^{k}\frac{1}{x_i!}$$

is always non-negative. Therefore Equation 12.2 does define a proper joint distribution.

(ii) When $\omega < 0$ such that $P_{N_1,\ldots,N_k}(0,\ldots,0) > 0$ and $1/\omega$ is a negative integer, we have

$$P(t_1,\ldots,t_k) = Q(t_1,\ldots,t_k)^n,$$
$$\text{where} \quad n = -1/\omega \text{ is a positive integer,}$$

which can be viewed as the n-fold convolutions of $Q(t_1,\ldots,t_k)$. Note that $[1+\beta_j-\beta_j t_j]^{\alpha_j\omega}$ represents the probability generating function of $\text{NB}(-\alpha_j\omega,\beta_j)$. Thus, $Q(t_1,\ldots,t_k)$ defines a proper multivariate distribution as long as $P_{N_1,\ldots,N_k}(0,\ldots,0) > 0$.

Note that the joint probability generating function in Equation 12.2 requires that ω_{ij} be the same for all i and j, but it allows arbitrary marginal negative binomial distributions, NB(α_j, λ_j). In the special case that all α_j are the same, $\alpha_j = \alpha$, the family of joint distributions in Equation 12.2 returns to the common Poisson-Gamma mixture model with $\omega = 1/\alpha$. This special case corresponds to the usual definition of multivariate negative binomial distributions in Johnson, Kotz and Balakrishnan [16, p. 93]. Thus, Equation 12.2 extends the usual class of multivariate negative binomial distributions.

Remark Consider k individual risk portfolios that are specified by their frequencies and severities: (N_j, X_j), $j = 1, \ldots, k$. Assume that $(N_1, \ldots, N_k)$ has a joint probability generating function as in Equation 12.2, and the only correlation exists between the frequencies. Based on Equation 4.1, the aggregate loss, Z, for the combined risk portfolios has a characteristic function

$$\phi_Z(t) = \left\{ \sum_{j=1}^{k} [1 - \beta_j(\phi_{X_j}(t) - 1)]^{\alpha_j \omega} - k + 1 \right\}^{-1/\omega}, \qquad \omega \neq 0.$$

Thus FFT can be used to evaluate the aggregate loss distribution.

13. AN EXAMPLE OF CORRELATED FREQUENCIES

Consider two correlated risk portfolios with frequency/severity distributions specified as follows:

- Portfolio 1 has a negative binomial frequency with mean = 10 and variance = 20. It has a probability generating function: $P_{N_1}(t) = [1 - (t-1)]^{-10}$. Portfolio 1 has a Pareto($\alpha = 2$, $\beta =$ 50,000) severity subject to a policy limit of \$200,000. Its average severity is \$39,960.
- Portfolio 2 has a negative binomial frequency with mean = 6 and variance = 15. It has a probability generating function: $P_{N_2}(t) = [1 - 2(t-1)]^{-4}$. Portfolio 2 has a Pareto($\alpha = 1.5$, $\beta =$

40,000) severity subject to a policy limit of \$300,000. Its average severity is \$52,560.

- The two claim frequencies are correlated with a covariance coefficient $\omega_{12} = 0.2$; i.e., $\text{Cov}[N_1,N_2] = 0.2 \cdot \text{E}[N_1] \cdot \text{E}[N_2]$.
- The claim severities X_1 and X_2 for the two risk portfolios are mutually independent; they are also independent of the frequencies (N_1,N_2).

Method I. We approximate the combined frequency N by a negative binomial distribution with

$$\text{E}[N] = \text{E}[N_1] + \text{E}[N_2] = 16,$$

and

$$\text{Var}[N] = \text{Var}[N_1] + \text{Var}[N_2] + 2\text{Cov}[N_1,N_2] = 59.$$

This negative binomial distribution has a probability generating function:

$$P_N(t) = \left[1 - \frac{256}{43}(t-1)\right]^{-43/16}.$$

The combined severity distribution can be calculated by

$$f_X(x) = \frac{\text{E}[N_1]}{\text{E}[N]} f_{X_1}(x) + \frac{\text{E}[N_2]}{\text{E}[N]} f_{X_2}(x),$$

where f_{X_1} and f_{X_2} are the severity distributions for Portfolios 1 and 2, respectively.

Method II. Assume that N_1 and N_2 have a bivariate negative binomial distribution with a joint probability generating function (see Equation 12.2):

$$P_{N_1,N_2}(t_1,t_2) = \{[1-(t_1-1)]^2 + [1-2(t_2-1)]^{0.8} - 1\}^{-5}.$$

Based on the earlier result in Equation 4.1, the aggregate loss, Z, for the combined risk portfolios has a characteristic function

$$\phi_Z(t) = \{[1-(\phi_{X_1}(t)-1)]^2 + [1-2(\phi_{X_2}(t)-1)]^{0.8} - 1\}^{-5}.$$

Thus FFT can be used to evaluate the aggregate loss distribution.

Some details of the calculation steps are as follows:

1. First we approximate the severity distribution by a discrete probability distribution. We choose the number of points for the FFT computation at $4096 = 2^{12}$. This is the maximum number of points for the Microsoft Excel FFT routine. In some other computer languages such as MATLAB, a higher number of points is allowed. We choose a span of $h = \$1{,}000$ and use the "matching-mean" method to approximate each individual severity distribution by a discrete one. For a severity distribution with cumulative distribution function F_X, we first evaluate the limited expected values at multiples of h:

$$\mathrm{E}[X; j\cdot h] = \int_0^{j\cdot h} [1 - F_X(u)]\,du, \qquad \text{for} \quad j = 1,2,\ldots.$$

   Then we apply the following method:

$$\begin{aligned}\Pr\{X = 0\cdot h\} &= 1 - \mathrm{E}[X;h]/h,\\ \Pr\{X = j\cdot h\} &= (2\mathrm{E}[X;j\cdot h] - \mathrm{E}[X;(j-1)\cdot h]\\ &\qquad - \mathrm{E}[X;(j+1)\cdot h])/h, \qquad j = 1,2,\ldots.\end{aligned}$$

   Note that the severity distribution for the two risk portfolios are subject to policy limits of \$200,000 and \$300,000, respectively. Given that the span was chosen at \$1,000, the maximum severity points with non-zero probabilities are 200 and 300, respectively. *It is critical to pad (i.e., add) enough zeros to the discrete severity vectors so that each severity vector has the same length, 4096 in this case, as the target aggregate loss distribution.* Let $\mathbf{f}_{X_1}$ and $\mathbf{f}_{X_2}$ represent the discrete severity vectors for the two risk portfolios, each of which is of length 4,096.

One should exercise caution in the selection of the span, h, for the discrete severity distributions. Too large a span would affect the accuracy of the discrete distribution. Too small a span may produce some "wrapping" (non-zero probabilities at the high points near 4,096) in the calculated aggregate loss distributions.

2. **Method I:** Let

$$\mathbf{f}_X(j) = \tfrac{10}{16}\mathbf{f}_{X_1}(j) + \tfrac{6}{16}\mathbf{f}_{X_2}(j), \qquad j = 0,1,\ldots,4095.$$

Perform FFT on the severity vector $\mathbf{f}_X$. Let $\tilde{\mathbf{f}}_X = \text{FFT}(\mathbf{f}_X)$ represent the resulting vector (of length 4,096). Apply the frequency probability generating function, element by element, to the vector $\tilde{\mathbf{f}}_X$:

$$\tilde{\mathbf{f}}_Z(j) = [1 - \tfrac{256}{43}(\tilde{\mathbf{f}}_X(j) - 1)]^{-43/16}.$$

Finally, perform an inverse FFT on $\tilde{\mathbf{f}}_Z$, and let $\mathbf{f}_Z = \text{IFFT}(\tilde{\mathbf{f}}_Z)$. Note that $\mathbf{f}_Z$ is a probability vector with a span of \$1,000, which approximates the aggregate loss distribution for the combined risk portfolios.

3. **Method II.** Perform FFT on each of the severity vectors, $\mathbf{f}_{X_1}$ and $\mathbf{f}_{X_2}$. Let $\tilde{\mathbf{f}}_{X_1} = \text{FFT}(\mathbf{f}_{X_1})$ and $\tilde{\mathbf{f}}_{X_2} = \text{FFT}(\mathbf{f}_{X_2})$ represent the resulting vectors (each of length 4,096). Apply the bivariate frequency probability generating function:

$$\tilde{\mathbf{f}}_Z(j) = \{[1-(\tilde{\mathbf{f}}_{X_1}(j)-1)]^2 + [1-2(\tilde{\mathbf{f}}_{X_2}(j)-1)]^{0.8} - 1\}^{-5}, \qquad j = 0,1,\ldots,4095. \qquad (13.1)$$

Finally, perform an inverse FFT on $\tilde{\mathbf{f}}_Z$, and let $\mathbf{f}_Z = \text{IFFT}(\tilde{\mathbf{f}}_Z)$. Note that $\mathbf{f}_Z$ is a probability vector with a span of \$1,000, which approximates the aggregate loss distribution for the combined risk portfolios.

4. **Independence Case:** For comparison purposes, we can also calculate the aggregate loss distribution under the assumption of independence between the frequencies. In this case, we repeat Method II except that Equation 13.1 is replaced by the following formula:

$$\tilde{\mathbf{f}}_Z(j) = [1 - (\tilde{\mathbf{f}}_{X_1}(j) - 1)]^{-10} \cdot [1 - 2(\tilde{\mathbf{f}}_{X_2}(j) - 1)]^{-4}, \quad j = 0, 1, \ldots, 4095.$$

Table 4 lists some values of the calculated aggregate loss distributions.

We can draw two conclusions regarding this specific example:

1. Methods I and II result in two very close aggregate loss distributions.
2. In both methods, correlation has a significant impact on the tail probabilities (quantiles).

14. CONCLUSIONS

This paper has presented a set of tools for modeling and combining correlated risks. A number of correlation structures have been generated using copula, common mixture, component, and distortion models. A good understanding of the claim generating process should be helpful in choosing a model, as well as in selecting correlation parameters. These correlation models are often specified by (i) the joint cumulative distribution function (i.e., a copula) or (ii) the joint characteristic function. The copula construction leads to efficient simulation techniques which can be implemented readily on a spreadsheet. The characteristic function specification leads to simple methods of aggregation by using fast Fourier transforms.

In the high-dimension world of multivariate variables, one may encounter very diverse correlation structures. Regardless of the complexity of the situation, Monte Carlo simulation can al-

TABLE 4

COMPARISON OF VARIOUS METHODS

Loss Amount in Dollars	Method I Single NB	Method II Bivariate NB	Independence Case
x	$\Pr\{Z \leq x\}$	$\Pr\{Z \leq x\}$	$\Pr\{Z \leq x\}$
0	0.00046	0.00032	0.00003
250,000	0.11014	0.11129	0.06888
500,000	0.34756	0.35292	0.30621
750,000	0.59539	0.59897	0.59178
1,000,000	0.77954	0.77937	0.80217
1,250,000	0.89125	0.88894	0.91753
1,500,000	0.95038	0.94777	0.96941
1,750,000	0.97872	0.97672	0.98964
2,000,000	0.99132	0.99006	0.99674
2,250,000	0.99661	0.99590	0.99903
2,500,000	0.99872	0.99836	0.99972
2,750,000	0.99953	0.99936	0.99993
3,000,000	0.99983	0.99976	0.99998
3,250,000	0.99994	0.99991	0.99999
3,500,000	0.99998	0.99997	1.00000
3,750,000	0.99999	0.99999	1.00000
4,000,000	1.00000	1.00000	1.00000
Aggregate Moments	**Method I Single NB**	**Method II Bivariate NB**	**Independence Case**
$E[Z]$	715,355	715,349	715,361
$CV[Z]$	0.584	0.593	0.503
$E[(Z - E[Z])^3]$	6.948×10^{12}	7.731×10^{12}	3.837×10^{12}

ways be employed in an analysis of the correlation risk. For instance, in some situations, the frequency and severity variables are correlated. With the assistance of Monte Carlo simulation, the common mixture model in Section 9 can be adapted to describe the association between the frequency and severity random variables, if both depend on the same external parameter. This external parameter may be chosen to represent the Richter scale of an earthquake, the velocity of wind speed, or several scenarios of legal climate, etc., depending on the underlying claim environment.

Dependency has always been a fascinating research subject, as well as part of reality. A good understanding of the impact of correlation on the aggregate loss distribution is essential for the dynamic financial analysis of an insurance company. It is hoped that the set of tools developed in this paper will be useful to actuaries in quantifying the aggregate risks of a financial entity. It is also hoped that this research will stimulate more scientific investigations on this subject in the future.

REFERENCES

[1] Brigham, E. O., *The Fast Fourier Transform and Its Applications*, Englewood, NJ, Prentice Hall, 1988.

[2] Burden, R. L., and J. D. Faires, *Numerical Analysis*, 4th edition, Boston, PWS-KENT Publishing Company, 1989.

[3] Cook, R. D., and M. E. Johnson, "A Family of Distributions for Modelling Non-elliptically Symmetric Multivariate Data," *Journal of the Royal Statistical Society*, B43, 1981, pp. 210–218.

[4] Crow, E. L., and K. Shimizu, *Lognormal Distributions: Theory and Applications*, New York, Marcel Dekker, Inc., 1988.

[5] Fishman, G. S., *Monte Carlo: Concepts, Algorithms, and Applications*, New York, Springer-Verlag, 1996.

[6] Frees, E. W., and E. A. Valdez, "Understanding Relationships Using Copulas," *North American Actuarial Journal* 2, 1, 1998, pp. 1–25.

[7] Genest, C., and J. Mackay, "The joy of copulas: bivariate distributions with uniform marginals," *The American Statistician* 40, 4, 1986, pp. 280–283.

[8] Heckman, P. E., and G. G. Meyers, "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *PCAS* LXX, 1983, pp. 22–61.

[9] Herzog, T. N., *An Introduction to Stochastic Simulation*, CAS 4B Study Note, 1986.

[10] Hesselager, O., S. S. Wang, and G. E. Willmot, "Exponential and Scale Mixtures and Equilibrium Distribution," *Scandinavian Actuarial Journal* 2, 1997, pp. 125–142.

[11] Hogg, R., and S. Klugman, *Loss Distributions*, New York, John Wiley, 1984.

[12] Iman, R. L., and W. J. Conver, "A Distribution Free Approach to Inducing Rank Correlation Among Input Variables," *Communications in Statistics: Simulation Computation* 11, 3, 1982, pp. 311–334.

[13] Johnson, N., and S. Kotz, *Distributions in Statistics: Continuous Multivariate Distributions*, New York, John Wiley & Sons, Inc., 1972.

[14] Johnson, N., and S. Kotz, "On Some Generalized Farlie-Gumbel-Morgenstern Distributions," *Communications in Statistics* 1977, pp. 415–427.

[15] Johnson, N., and S. Kotz, "On some Generalized Farlie-Gumbel-Morgenstern Distributions–II Regression, Correlation and Further Generalizations," *Communications in Statistics* 1977, pp. 485–496.

[16] Johnson, N., S. Kotz, and N. Balakrishnan, *Discrete Multivariate Distribution*, New York, John Wiley & Sons, Inc., 1997.

[17] Johnson, M. E., *Multivariate Statistical Simulation*, New York, Wiley, 1987.

[18] Klugman, S., H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions*, New York, John Wiley & Sons, Inc., 1998.

[19] Marshall, A. W., and I. Olkin, "Families of Multivariate Distributions," *Journal of the American Statistical Association* 83, 1988, pp. 834–841.

[20] Oakes, D., "Bivariate Survival Models Induced by Frailties," *Journal of the American Statistical Association* 84, 1989, pp. 487–493.

[21] Panjer, H. H., "Recursive Evaluation of a Family of Compound Distributions," *ASTIN Bulletin* 12, 1981, pp. 22–26.

[22] Panjer, H. H., and G. E. Willmot, "Computational Aspects of Recursive Evaluation of Compound Distributions," *Insurance: Mathematics and Economics* 5, 1986, pp. 113–116.

[23] Robertson, J, "The Computation of Aggregate Loss Distributions," *PCAS* LXXIX, 1992, pp. 57–133.

[24] Vaupel, J. W., K. G. Manton, and E. Stallard, "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality," *Demography* 16, 3, 1979, pp. 439–454.

[25] Wang, S. S., "An Actuarial Index of the Right-Tail Risk," *North American Actuarial Journal* 2, 2, 1998, pp. 88–101.

[26] Willmot, G. E., "The Poisson-Inverse Gaussian as an Alternative to the Negative Binomial," *Scandinavian Actuarial Journal* 1987, pp. 113–127.

APPENDIX A

AN INVENTORY OF UNIVARIATE DISTRIBUTIONS

A.1. Counting Distributions

- The Poisson distribution, Poisson(λ), $\lambda > 0$, is defined by a probability function:

$$p_n = \Pr\{N = n\} = e^{-\lambda}\frac{\lambda^n}{n!}, \qquad n = 0,1,2,\ldots .$$

  It has a probability generating function

$$P_N(t) = \mathrm{E}[t^N] = e^{\lambda(t-1)},$$

  and $\mathrm{E}[N] = \mathrm{Var}[N] = \lambda$.

- The negative binomial distribution, NB(α,β), $\alpha,\beta > 0$, has a probability function:

$$p_n = \Pr\{N = n\} = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)n!}\left(\frac{1}{1+\beta}\right)^{\alpha}\left(\frac{\beta}{1+\beta}\right)^{n}, \qquad n = 0,1,2,\ldots .$$

  It has a probability generating function

$$P_N(t) = [1-\beta(t-1)]^{-\alpha},$$

  with $\mathrm{E}[N] = \alpha\beta$ and $\mathrm{Var}[N] = \alpha\beta(1+\beta)$.

  When $\alpha = 1$, the negative binomial distribution NB($1,\beta$) is called the geometric distribution.

- The Poisson inverse Gaussian distribution, P-IG(β,μ), has a probability generating function

$$P_N(t) = \mathrm{E}[t^N] = \exp\left\{-\frac{\mu}{\beta}[\sqrt{1+2\beta(1-t)}-1]\right\}.$$

It can be verified that $\mathrm{E}[N] = \mu$ and $\mathrm{Var}[N] = \mu(1+\beta)$. The probabilities can be calculated via a simple recursion (Willmot, [26]):

$$p_n = \frac{2\beta}{1+2\beta}\left(1 - \frac{3}{2n}\right)p_{n-1} + \frac{\mu^2}{n(n-1)(1+2\beta)}p_{n-2}, \qquad n = 2,3,\ldots,$$

with starting values

$$p_0 = e^{-\mu/\beta[\sqrt{1+2\beta}-1]}, \qquad p_1 = \frac{\mu}{\sqrt{1+2\beta}}p_0.$$

A.2. Continuous Distributions

- The exponential distribution, exponential(λ), is defined by

$$S(x) = 1 - F(x) = e^{-\lambda x}, \qquad x > 0,$$

  with $\mathrm{E}[X] = 1/\lambda$ and $\mathrm{Var}[X] = 1/\lambda^2$.

- The gamma distribution, gamma(α,β), $\alpha,\beta > 0$, has a probability density function

$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, \qquad x > 0.$$

  It has a moment generating function

$$M_X(t) = \mathrm{E}[e^{tX}] = (1-\beta t)^{-\alpha},$$

  and $\mathrm{E}[X] = \alpha\beta$, and $\mathrm{Var}[X] = \alpha\beta^2$.

- The Pareto distribution, Pareto(α,β), $\alpha,\beta > 0$, has a survivor function

$$S(x) = 1 - F(x) = \left(\frac{\beta}{x+\beta}\right)^{\alpha} = (1 + x/\beta)^{-\alpha}.$$

  The mean $\mathrm{E}[X] = \beta/\alpha - 1$ exists only if $\alpha > 1$.

- The Weibull distribution, Weibull(β,τ), $\beta,\tau > 0$, has a survivor function

$$S(x) = 1 - F(x) = e^{-(x/\beta)^\tau},$$

with $E[X] = \beta\Gamma(1 + \tau^{-1})$ and $E[X^2] = \beta^2[\Gamma(1 + 2\tau^{-1})]$.

- The inverse Gaussian distribution, IG(β,μ), has a probability density function

$$f(x) = \mu(2\pi\beta x^3)^{-1/2} \exp\left\{-\frac{(x-\mu)^2}{2\beta x}\right\}, \qquad x > 0.$$

It has a moment generating function

$$M(t) = e^{\mu/\beta[1-\sqrt{1-2\beta t}]},$$

and $E[X] = \mu$, and $\mathrm{Var}[X] = \mu\beta$.

- The exponential inverse Gaussian distribution, E-IG(β,μ), has a survivor function:

$$S(x) = 1 - F(x) = e^{\mu/\beta\{1-(1+2\beta x)^{1/2}\}}, \qquad x > 0,$$

with moments (Hesselager/Wang/Willmot, [10]):

$$E[X] = \frac{\beta + \mu}{\mu^2}, \qquad \mathrm{Var}[X] = \frac{5\beta^2 + 4\beta\mu + \mu^2}{\mu^4}.$$

- The lognormal distribution, LN(μ,σ^2), has a probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{1}{2}\left[\frac{\log(x) - \mu}{\sigma}\right]^2\right), \qquad x > 0,$$

with

$$E[X] = \exp[\mu + \sigma^2/2] \qquad \text{and}$$
$$\mathrm{Var}[X] = \exp[2\mu + \sigma^2][\exp(\sigma^2) - 1].$$

A.3. Parameter Uncertainty and Mixture Models

In modeling insurance losses, actuaries are called upon to pick a frequency distribution and a severity distribution based on past claim data and their own judgement. Actuaries are fully aware of the presence of parameter uncertainty in the assumed models. As a way of incorporating parameter uncertainty, mixture models are often employed.

- The most popular frequency distributions are the negative binomial family of distributions. In modeling claim frequency, the negative binomial NB(α,β) can be interpreted as a mixed Poisson model, where the Poisson parameter λ has a gamma(α,β) distribution. This can be seen from the probability generating function

$$P_N(t) = \mathrm{E}[t^N] = \mathrm{E}_\lambda[\mathrm{E}(t^N \mid \lambda)] = \mathrm{E}_\lambda[e^{\lambda(t-1)}]$$
$$= M_\lambda(t-1) = \{1-\beta(t-1)\}^{-\alpha}.$$

- A popular claim severity distribution is the Pareto distribution which has a thick right tail representing large claims. The Pareto(α,β) distribution can be interpreted as a mixed exponential distribution, where the exponential parameter λ has a gamma$(\alpha,1/\beta)$ distribution. This can be seen from the survivor function

$$S(x) = \mathrm{E}_\lambda[e^{-\lambda x}] = M_\lambda(-x) = (1+x/\beta)^{-\alpha} = \left(\frac{\beta}{\beta+x}\right)^{\alpha}.$$

- A more flexible family of claim severity distributions are the Burr distributions (including Pareto as a special case). The Burr(α,β,τ) distributions can be expressed as a Weibull-gamma mixture. This can be seen from the survivor function

$$S(x) = \mathrm{E}_\lambda[e^{-\lambda x^\tau}] = M_\lambda(-x^\tau) = (1+x^\tau/\beta)^{-\alpha} = \left(\frac{\beta}{\beta+x^\tau}\right)^{\alpha}.$$

  The Burr(α,β,τ) family includes the Pareto(α,β) as a special member when $\tau = 1$.

For $\tau > 1$, the Burr(α,β,τ) distribution has a lighter tail than its Pareto(α,β) counterpart.
For $\tau < 1$, the Burr(α,β,τ) distribution has a thicker tail than its Pareto(α,β) counterpart.

A.4. Lognormal Distributions

A.4.1. Univariate lognormal distributions

The normal distribution, $N(\mu,\sigma^2)$, has a probability density function

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right), \qquad -\infty < x < \infty.$$

It has a moment generating function

$$M_X(t) = \mathrm{E}[e^{tX}] = \exp[\mu t + \tfrac{1}{2}\sigma^2 t^2].$$

If $X \sim N(\mu,\sigma^2)$, then $Y = e^X$ has a lognormal distribution with a probability density function

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma y}\exp\left(-\frac{1}{2}\left[\frac{\log(y)-\mu}{\sigma}\right]^2\right), \qquad y > 0.$$

The moments of Y are

$$\begin{aligned}\mathrm{E}[Y^n] &= \mathrm{E}[\exp(nX)] = M_X(n)\\ &= \exp\left[n\mu + \frac{n^2\sigma^2}{2}\right], \qquad n = 1,2\ldots.\end{aligned}$$

Specifically,

$$\begin{aligned}\mathrm{E}[Y] &= \exp\left[\mu + \frac{\sigma^2}{2}\right],\\ \mathrm{Var}[Y] &= \exp[2\mu+\sigma^2][\exp(\sigma^2)-1],\\ \mathrm{E}[Y-\mathrm{E}[Y]]^3 &= \exp\left[3\mu + \frac{3\sigma^2}{2}\right]\left[\exp(3\sigma^2) - 3\exp(\sigma^2) + 2\right].\end{aligned}$$

A.4.2. Bivariate lognormal distributions

Let X_1 and X_2 have a bivariate normal distribution with joint probability density function

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp\left\{-\frac{1}{2(1-\rho)^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\}.$$

Then X_1 and X_2 have marginal distributions $N(\mu_1,\sigma_1^2)$ and $N(\mu_2,\sigma_2^2)$, respectively. (X_1,X_2) has a covariance matrix

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

where ρ is the correlation coefficient between X_1 and X_2. Note that $\rho = 1$ if and only if $\Pr\{X_1 = aX_2 + b\} = 1$ with $a > 0$.

Now consider the variables $Y_1 = \exp(X_1)$ and $Y_2 = \exp(X_2)$. Note that $\log(Y_1Y_2)$ has a $N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)$ distribution. We have

$$\begin{aligned}\mathrm{Cov}[Y_1,Y_2] &= \mathrm{E}[Y_1Y_2] - \mathrm{E}[Y_1]\mathrm{E}[Y_2] \\ &= \exp\{(\mu_1+\mu_2) + \tfrac{1}{2}[\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2]\} \\ &\quad - \exp[\mu_1+\sigma_1^2+\mu_2+\sigma_2^2] \\ &= \exp[\mu_1+\mu_2+\tfrac{1}{2}(\sigma_1^2+\sigma_2^2)]\{\exp(\rho\sigma_1\sigma_2)-1\}.\end{aligned}$$

Therefore, the correlation coefficient of Y_1 and Y_2 is

$$\rho_{Y_1,Y_2} = \frac{\exp(\rho\sigma_1\sigma_2)-1}{\sqrt{\exp(\sigma_1^2)-1}\sqrt{\exp(\sigma_2^2)-1}},$$

where $\rho = \rho_{X_1,X_2}$.

A.4.3. Multivariate lognormal distributions

Consider a vector $(X_1,\ldots,X_n)'$ of positive random variables. Assume that $(\log X_1,\ldots,\log X_n)'$ has an n-dimensional normal distribution with mean vector and variance-covariance matrix

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \vdots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix},$$

respectively.

The distribution of $(X_1,\ldots,X_n)'$ is said to be an n-dimensional lognormal distribution with parameters $(\boldsymbol{\mu},\boldsymbol{\Sigma})$ and denoted by $\Lambda_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$. The probability density function of $\mathbf{X} = (X_1,\ldots,X_n)'$ having $\Lambda_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$ is (see Crow and Shimizu, [4, Chapter 1]):

$$f(x_1,\ldots,x_n) = \begin{cases} \dfrac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|\prod_{i=1}^n x_i}} \exp\{-\frac{1}{2}(\log\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\log\mathbf{x}-\boldsymbol{\mu})\}, & \mathbf{x}\in(0,\infty)^n \\ 0, & \text{otherwise.} \end{cases}$$

From the moment generating function for the multivariate normal distribution we have

$$\mathrm{E}[X_1^{s_1}\cdots X_n^{s_n}] = \exp(\mathbf{s}'\boldsymbol{\mu} + \tfrac{1}{2}\mathbf{s}'\boldsymbol{\Sigma}\mathbf{s}),$$

where $\mathbf{s} = (s_1,\ldots,s_n)'$. Specifically, we have for any $i = 1,2,\ldots,n$,

$$\mathrm{E}[X_i^r] = \exp(r\mu_i + \tfrac{1}{2}r^2\sigma_{ii}^2)$$

and for any $i,j = 1,2,\ldots,n$,

$$\mathrm{Cov}[X_i,X_j] = \exp\{\mu_i + \mu_j + \tfrac{1}{2}(\sigma_{ii}^2 + \sigma_{jj}^2)\}\{\exp(\sigma_{ij}) - 1\}.$$

A simulation of this multivariate lognormal distribution can be easily achieved by first generating a sample from a multivariate normal distribution and then taking the logarithms.

APPENDIX B

MORE ON COPULAS AND SIMULATION METHODS

This appendix presents greater detail on the construction of copulas and the associated simulation techniques. For simplicity, we confine the discussion to the bivariate case. The discussion here can be readily extended to higher dimensions ($k > 2$).

B.1. Bivariate Copulas

A bivariate *copula* refers to a joint cumulative distribution function $C(u,v) = \Pr\{U \leq u; V \leq v\}$ with uniform marginals $U,V \sim \text{Uniform}[0,1]$. It links the marginal distributions to their multivariate joint distribution. Let $S_{X,Y}(x,y)$ be a joint survivor function with marginals S_X and S_Y. Then there is a copula C such that

$$S_{X,Y}(x,y) = C(S_X(x), S_Y(y)), \qquad \text{for all} \quad x,y \in (-\infty, \infty).$$

Conversely, given any marginals S_X, S_Y, and a copula C, $S_{X,Y}(x,y) = C(S_X(x), S_Y(y))$ defines a joint survivor function with marginals S_X and S_Y. Furthermore, if S_X and S_Y are continuous, then C is unique.

Note that $S_X(X)$ and $S_Y(Y)$ are uniformly distributed random variables. The association between X and Y can be described by the association between uniform variables $U = S_X(X)$ and $V = S_Y(Y)$. In theory, if one can first generate a sample pair (u_i, v_i) from the bivariate uniform distribution of (U,V), one can simulate a sample pair of (X,Y) by the inverse transforms: $x_i = S_X^{-1}(u_i)$ and $y_i = S_Y^{-1}(v_i)$. Unfortunately, there is no simple way of generating a set of bivariate uniform numbers that works for all copulas. In reality, each type of copula needs a different simulation technique.

Note that

$$F_{X,Y}(x,y) = C(F_X(x), F_Y(y)) \qquad \text{and}$$
$$S_{X,Y}(x,y) = C(S_X(x), S_Y(y))$$

may imply different bivariate distributions. Here we assume that a copula is applied to the survivor functions unless otherwise mentioned.

B.2. Distortion of the Joint Survivor Function

Let $g : [0,1] \to [0,1]$ be an increasing function with $g(0) = 0$ and $g(1) = 1$. If $S_{X,Y}(x,y)$ is a joint survivor function with marginals S_X and S_Y, then $g[S_{X,Y}(x,y)]$ defines another joint survivor function with marginals $g \circ S_X$ and $g \circ S_Y$. If we assume that, after applying a distortion g, $g[S_{X,Y}(x,y)]$ has non-correlated marginals:

$$g[S_{X,Y}(x,y)] = g[S_X(x)]\, g[S_Y(y)],$$

then we have

$$S_{X,Y}(x,y) = g^{-1}(g[S_X(x)] \cdot g[S_Y(y)]), \tag{B.1}$$

which corresponds to the copula

$$C(u,v) = g^{-1}[g(u)g(v)]. \tag{B.2}$$

If we let $h(t) = -\log g(t)$, then Equation B.1 gives the following relation: $S_{X,Y}(x,y) = h^{-1}(h[S_X(x)] + h[S_Y(y)])$, which gives the Archimedean family of copulas (see Genest and Mackay, [7]).

For a bivariate copula C, Kendall's tau is

$$\tau = 4 \int_0^1 \int_0^1 C(u,v)\, dC(u,v) - 1.$$

If a copula $C = g^{-1}(g(u)g(v))$ is defined by a distortion g, then

$$\tau = 1 + 4 \int_0^1 \frac{g(t) \log g(t)}{g'(t)}\, dt.$$

EXAMPLE B.1 The distortion function $g(t) = \exp\{1 - t^{-\alpha}\}$, $\alpha > 0$, corresponds to the Clayton family of copulas with

$$\begin{aligned} C_\alpha(u,v) &= \{u^{-\alpha} + v^{-\alpha} - 1\}^{-1/\alpha}, \\ C_\infty(u,v) &= \lim_{\alpha\to\infty} C_\alpha(u,v) = \min[u,v], \\ C_0(u,v) &= \lim_{\alpha\to 0+} C_\alpha(u,v) = uv. \end{aligned} \tag{B.3}$$

Thus, C_∞ and C_0 are the copulas for the Frechet upper bounds and the independent case, respectively.

EXAMPLE B.2 The distortion function $g(t) = \exp\{-(-\log t)^\alpha\}$, $\alpha \geq 1$, corresponds to the Hougaard family of copulas with

$$\begin{aligned} C_\alpha(u,v) &= \exp\{-[(-\log u)^\alpha + (-\log v)^\alpha]^{1/\alpha}\}, \\ C_\infty(u,v) &= \lim_{\alpha\to\infty} C_\alpha(u,v) = \min[u,v], \\ C_1(u,v) &= uv. \end{aligned} \tag{B.4}$$

Thus, C_∞ and C_1 are the copulas for the Frechet upper bounds and the independent case, respectively.

EXAMPLE B.3 The distortion function $g(t) = \alpha^t - 1/\alpha - 1$, $\alpha > 0$, corresponds to the Frank family of copulas with

$$\begin{aligned} C_\alpha(u,v) &= [\log\alpha]^{-1}\log\left\{1 + \frac{(\alpha^u - 1)(\alpha^v - 1)}{\alpha - 1}\right\}, \qquad 0 < \alpha < \infty, \\ C_\infty(u,v) &= \lim_{\alpha\to\infty} C_\alpha(u,v) = \max[u + v - 1, 0], \\ C_0(u,v) &= \lim_{\alpha\to 0+} C_\alpha(u,v) = \min[u,v], \\ C_1(u,v) &= \lim_{\alpha\to 1} C_\alpha(u,v) = uv. \end{aligned} \tag{B.5}$$

Thus, C_∞, C_0, and C_1 are the copulas for the Frechet lower and upper bounds and the independent case, respectively.

B.3. Common Frailty Models

Vaupel, Manton, and Stallard [24] introduced the concept of *frailty* in their discussion of a heterogeneous population. Each individual in the population is associated with a frailty, r. The frailty varies across individuals and thus is modeled as a random variable R with cumulative distribution function $F_R(r)$. The conditional survival function of lifetime T, given r, is

$$\Pr\{T > t \mid R = r\} = B(t)^r,$$

in which $B(t)$ is the base line survivor function (for a standard individual with $r = 1$). The unconditional survivor function for a heterogeneous population is

$$\Pr\{T > t\} = \int_0^\infty B(t)^r \, dF_R(r) = M_R(\log B(t)),$$

where M_R is the moment generating function of R.

Oakes [20] uses a bivariate frailty model to describe associations between two random variables as follows. Assume that X and Y both can be modeled by frailty models

$$S_X(x) = \int_0^\infty B_1(x)^r \, dF_R(r) = M_R(\log B_1(x)) \qquad \text{and}$$

$$S_Y(y) = \int_0^\infty B_2(y)^r \, dF_R(r) = M_R(\log B_2(y)),$$

respectively, in which B_1 and B_2 are the base line survivor functions. Assume that X and Y are conditionally independent, given frailty $R = r$. However, X and Y are associated as they depend on the common frailty variable R. The bivariate frailty model yields the following joint survivor function

$$\begin{aligned} S_{X,Y}(x,y) = \Pr\{X > x, Y > y\} &= \int_0^\infty [B_1(x) \cdot B_2(y)]^r \, dF_R(r) \\ &= M_R(\log[B_1(x) \cdot B_2(y)]). \end{aligned}$$

For $g(u) = \exp[M_R^{-1}(u)]$, we have

$$g[S_{X,Y}(x,y)] = B_1(x) \cdot B_2(y) = g[S_X(x)] \cdot g[S_Y(y)].$$

In other words, a bivariate frailty model yields a joint distribution that can also be obtained by using the distortion function g.

EXAMPLE B.4 Assume that the frailty R has a gamma distribution with $M_R(z) = (1/1-z)^{1/\alpha}$, $\alpha > 0$. Then $M_R^{-1}(u) = 1 - u^{-\alpha}$, and $g(u) = \exp[1 - u^{-\alpha}]$. Therefore, the common gamma frailty model yields the Clayton family of copulas given by Equation B.3:

$$S_{X,Y}(x,y) = \left\{\frac{1}{S_X(x)^\alpha} + \frac{1}{S_Y(y)^\alpha} - 1\right\}^{-1/\alpha}, \qquad 0 < \alpha < \infty.$$

This family is particularly useful for constructing bivariate Burr (including Pareto) distributions (see Johnson and Kotz, [13, p. 289]).

EXAMPLE B.5 If the frailty R has a stable distribution with $M_R(z) = \exp\{-(-z)^{1/\alpha}\}$, $\alpha \geq 1$, the corresponding joint survivor function is given by Equation B.4:

$$S_{X,Y}(x,y) = \exp\{-[(-\log S_X(x))^\alpha + (-\log S_Y(y))^\alpha]^{1/\alpha}\}.$$

This family of copulas is particularly useful for constructing bivariate Weibull (including exponential) distributions.

EXAMPLE B.6 If the frailty R has a logarithmic (discrete) distribution on positive integers with $M_R(z) = [\log\alpha]^{-1}\log[1 + e^z(\alpha - 1)]$, then we get the Frank family of copulas given by Equation B.5.

Marshall and Olkin [19] proposed a simulation algorithm for copulas with a frailty construction. This algorithm is applicable to copulas with arbitrary dimensions ($k \geq 2$):

STEP 1 Generate a value r from the random variable R having moment generating function M_R.

STEP 2 Generate independent uniform $(0,1)$ numbers $U_1, \ldots, U_k$.

STEP 3 For $j = 1,\ldots,k$, set $U_j^* = M_R(r^{-1}\log U_j)$, and calculate $X_j = S_j^{-1}(U_j^*)$.

B.4. The Morgenstern Copula

The Morgenstern copula is defined by

$$C(u,v) = uv[1 + \alpha(1-u)(1-v)], \qquad -1 \le \alpha \le 1.$$

This copula cannot be generated by distortion or frailty models.

The following simulation algorithm for the Morgenstern copula can be found in Johnson [17, p. 185]:

STEP 1 Generate independent uniform variables V_1, V_2, and set $U_1 = V_1$.

STEP 2 Calculate $A = \alpha(2U_1 - 1) - 1$ and $B = [1 - \alpha(2U_1 - 1)]^2 + 4\alpha V_2(2U_1 - 1)$.

STEP 3 Set $U_2 = 2V_2/(\sqrt{B} - A)$.

For the Morgenstern copula, Kendall's tau is

$$\tau(\alpha) = \tfrac{2}{9}\alpha, \qquad -1 \le \alpha \le 1,$$

which is limited to the range $(-\frac{2}{9}, \frac{2}{9})$. Thus, the Morgenstern copula can be used only in situations with weak dependence.

An extension of the Morgenstern copula to arbitrary dimensions has been given by Johnson and Kotz [14, 15].

B.5. Summary and Comments

Table B.1 lists the most commonly used bivariate copulas. Except for the reverse monotone copula, they can readily be

TABLE B.1

A SUMMARY OF POPULAR COPULAS

Associated Names	Function Form $C(u,v)$	Kendall's Tau
Independence	uv	0
Common monotone	$\min[u,v]$	1
Reverse monotone	$\max[u+v-1,0]$	-1
Cook–Johnson, Clayton	$[u^{-1/\alpha}+v^{-1/\alpha}-1]^{-\alpha} \quad (\alpha > 0)$ $[u^{-\alpha}+v^{-\alpha}-1]^{-1/\alpha} \quad (\alpha > 0)$	$\frac{1}{1+2\alpha}$ $\frac{\alpha}{\alpha+2}$
Frank	$\log_\alpha\left\{1+\frac{(\alpha^u-1)(\alpha^v-1)}{\alpha-1}\right\} \quad (0<\alpha<\infty)$	*
Farlie, Gumbel, Morgenstern	$uv[1+\alpha(1-u)(1-v)] \quad (-1 \le \alpha \le 1)$	$\frac{2}{9}\alpha$
Gumbel–Hougaard	$\exp\{-[(-\ln u)^\alpha+(-\ln v)^\alpha]^{1/\alpha}\} \quad (\alpha \ge 1)$	$1-\alpha^{-1}$
normal	$H(\Phi^{-1}(u),\Phi^{-1}(v))^{**} \quad (-1 \le \rho \le 1)$	$\frac{2}{\pi}\arcsin(\rho)$

* For the Frank copula,

$$\tau(\alpha) = 1 + \frac{4}{-\log(\alpha)}\left[\frac{1}{-\log(\alpha)}\int_0^{-\log(\alpha)} \frac{t}{e^t-1}dt - 1\right].$$

** H is the joint cumulative distribution function for a bivariate standard normal distribution with a correlation coefficient ρ.

Note that the Cook–Johnson copula with parameter α is the same as the Clayton copula with parameter $1/\alpha$.

extended to multivariate cases ($k > 2$). In higher dimensions, the Cook–Johnson copula requires that all taus are the same, while the normal copula allows complete freedom in selecting Kendall's tau.

Some comments on higher dimension extensions of the listed copulas are in order.

1. The independence copula and the common monotone copula both have a unique extension to higher dimensions, while the reverse monotone has multiple possible extensions.

2. Recall that the Cook–Johnson copula, the Clayton copula, the Frank copula, and the Gumbel–Hougaard copula can be generated by the distortion method. They can be readily generated to higher dimensions by the relation

$$g[S_{X_1,X_2,\ldots,X_k}(t_1,t_2,\ldots,t_k)] \\ = g[S_{X_1}(t_1)]\cdot g[S_{X_2}(t_2)]\cdots g[S_{X_k}(t_k)].$$

   However, the correlation structure is restricted in a sense that Kendall's tau (or rank correlation coefficient) is the same for any pair of variables.

3. The Morgenstern copula can be generated to higher dimensions, but the parameter values are further restricted.

4. The normal copula stands out among others for its extremely flexible correlation structure at higher dimensions. It allows complete freedom in selecting Kendall's taus or rank order coefficients, as we have seen in Section 8.

Frees and Valdez [6] have written a good survey paper on the use of copulas, including the associated simulation techniques. In general, the use of copulas permits simple implementation by Monte Carlo simulation, thus aggregating correlated risks.

B.6. The Use of Mixed Copulas

Assume that joint cumulative distribution functions $F_{X_1,\ldots,X_k}$ and $G_{X_1,\ldots,X_k}$ have the same marginals $F_{X_1},\ldots,F_{X_k}$. Then the mixed joint cumulative distribution function

$$(1-q)F_{X_1,\ldots,X_k}(t_1,\ldots,t_k) + qG_{X_1,\ldots,X_k}(t_1,\ldots,t_k), \qquad 0 < q < 1,$$

also has a marginal cumulative distribution function $F_{X_1}, \ldots, F_{X_k}$. For this mixed joint distribution, we have

$$\tau[X_i, X_j] = (1-q)\tau^F[X_i, X_j] + q\tau^G[X_i, X_j],$$

where τ^F and τ^G represent Kendall's taus implied by the joint cumulative distribution functions F and G, respectively. In particular, if we let $F_{X_1,\ldots,X_k}(t_1, \ldots, t_k) = \prod_{j=1}^{k} t_j$ represent the independent copula and $G_{X_1,\ldots,X_k}(t_1, \ldots, t_k) = \min[t_1, \ldots, t_k]$ represent the comonotonic copula, then $\tau[X_i, X_j] = q$.

The mixture of joint cumulative distribution functions can be used to adjust, up or down, Kendall's tau. For example, if we feel that a common mixture joint cumulative distribution function F would give too strong a correlation, then we can mix it with an independent joint cumulative distribution function G. If we feel that a common mixture joint cumulative distribution function F would give too weak a correlation, then we can mix it with a comonotonic joint cumulative distribution function G'.

APPENDIX C

PANJER'S RECURSIVE METHOD

As an alternative to the FFT method, we introduce Panjer's recursive method for evaluation of aggregate loss distributions.

Suppose that the severity distribution $f_X(x)$ is defined on $0,1,2,\ldots$, representing multiples of some convenient monetary unit.

Suppose that the frequency distribution is a member of the (a,b) class satisfying

$$\Pr\{N = k\} = \left(a + \frac{b}{k}\right)\Pr\{N = k-1\}, \qquad k = 1,2,3,\ldots . \tag{C.1}$$

Note that the Poisson and negative binomial distributions are included in this family.

For the Poisson distribution, we have $a = 0$ and $b = \lambda$.

For the negative binomial (α,β) distribution, we have

$$a = \frac{\beta}{1+\beta} \qquad \text{and} \qquad b = \frac{(\alpha-1)\beta}{1+\beta}.$$

Panjer [21] has shown that the aggregate loss distribution $f_S(x)$ can be evaluated recursively

$$f_S(x) = \left[\sum_{y=1}^{x}\left(a + \frac{by}{x}\right) f_X(y) f_S(x-y)\right](1 - a f_X(0))^{-1}. \tag{C.2}$$

The starting value of the recursive algorithm is $f_S(0) = P_N(f_X(0))$.

In the case of the Poisson distribution, it reduces to

$$f_S(x) = \frac{\lambda}{x}\sum_{y=1}^{x} y f_X(y) f_S(x-y), \qquad x = 1,2,\ldots, \tag{C.3}$$

with starting value

$$f_S(0) = e^{-\lambda(1-f_X(0))}. \tag{C.4}$$

The recursive method is fairly easy to program.

For example, suppose that the claim frequency N has a Poisson distribution with mean $\lambda = 3$ and the claim severity X has a probability distribution

$$\Pr\{X = 1\} = 0.5, \qquad \Pr\{X = 2\} = 0.3, \qquad \Pr\{X = 3\} = 0.2, \tag{C.5}$$

then the probability distribution of the aggregate loss S can be calculated recursively

$$f_S(0) = e^{-\lambda} = 0.04979,$$

$$f_S(1) = \frac{\lambda}{1}[f_X(1)f_S(0)] = 0.07468,$$

$$f_S(2) = \frac{\lambda}{2}[f_X(1)f_S(1) + f_X(2)f_S(0)] = 0.07842,$$

$$f_S(3) = \frac{\lambda}{3}[f_X(1)f_S(2) + f_X(2)f_S(1) + f_X(3)f_S(0)] = 0.07157.$$

APPENDIX D

SOME FREQUENTLY ASKED QUESTIONS

Q1. In this paper a lot of discussion has been given to correlated frequency models. What about models of correlation between claim severities?

A1. We have a general correlation model—normal copula—which can be used to model correlated claim severities. In fact, @Risk (which is an Excel add-in application) can be readily used to carry out such simulations. But one should keep in mind that the correlation parameters used in @Risk are rank correlation coefficients.

Correlation in claim severities often comes from the uncertainty in the future claim inflation and loss development. A simple method of quantifying this correlation is to use a common multiplier:

$$X_1 = B \cdot Y_1, \ldots, X_j = B \cdot Y_j, \ldots X_N = B \cdot Y_N,$$

where

- the Y_js are independent,
- the number of claims N may be fixed or random but independent from the severity Y_js and the multiplier B, and
- the common multiplier B may be assigned a probability distribution (discrete or continuous).

The sum of the k losses is

$$Z = X_1 + X_2 + \cdots + X_N = B \cdot (Y_1 + Y_2 + \cdots + Y_N).$$

Thus, one can first evaluate the independent sum of the Y_js, and then combine it with the multiplier B. In this model, we have $\mathrm{Cov}[X_i, X_j] = \mathrm{Var}(B) \cdot \mathrm{E}[Y_i] \cdot \mathrm{E}[Y_j]$.

Q2. In some situations, the claim frequency and severity are correlated. How would one construct such a model?

A2. In some catastrophe modeling it might be plausible to consider the dependency between claim frequency and claim severity. For instance, the Richter scale value of an earthquake may affect the claim frequency and severity simultaneously, and for hurricane losses, the wind speed would affect both the claim frequency and severity in the same direction. For a modeling of such catastrophe losses, a good understanding of the underlying loss generating mechanism is essential, which requires sound knowledge in meteorology, construction engineering, population density, insurance coverage, etc. Some have observed that demand surge after a major catastrophe may also generate correlation between claim frequency and severity. Nevertheless, mathematics can serve as a tool to quantify our understanding of the underlying loss generating mechanism.

For reinsurance excess-of-loss modeling, the frequency and severity of a given layer may be positively correlated in a high inflation environment. This is due to the leverage effect of inflation. This correlation can be quantified by using a random trending factor for ground-up losses and then quantifying the frequency/severity for losses for a reinsurance layer.

IMPLEMENTATION OF PROPORTIONAL HAZARDS TRANSFORMS IN RATEMAKING

SHAUN WANG

Abstract

This article introduces a relatively new method for calculating risk load in insurance ratemaking: the use of proportional hazards (PH) transforms. This method is easy to understand, simple to use, and supported by theoretical properties as well as economic justification. After an introduction of the PH-transform method, examples show how it can be used in pricing ambiguous risks, excess-of-loss coverages, increased limits, risk portfolios, and reinsurance treaties.

ACKNOWLEDGEMENT

I would like to thank Gary Venter for his inspiration and numerous helpful comments and suggestions. I have also benefited from comments by James Higgins, Israel Krakowski, John Kulik, and the Committee on Review of Papers.

1. INTRODUCTION

Recently, there has been considerable interest in and extensive discussion of risk loads within the Casualty Actuarial Society. These discussions have focused on measures of risk and methods to arrive at a 'reasonable' risk load. Although there are diverse opinions on the appropriate measurement of risk, there is general agreement on the distinction between process risk and parameter risk, and on the importance of parameter risk in ratemaking. (See Finger [5], Miccolis [18], McClenahan [15], Feldblum [4], Philbrick [21], Meyers [16], Robbin [25], and Bault [1].)

Consistent with previous papers, this paper will consider only pure risk-adjusted premiums (the expected loss plus risk load, excluding all expenses and commissions). These pure risk-adjusted premiums are sometimes referred to as premiums in the paper.

Following Venter's [28] advocacy of adjusted distribution methods, Wang [30] proposes using proportional hazards (PH) transforms in the calculation of the risk-adjusted premium. This paper focuses on the practical aspects of implementation of PH-transforms in ratemaking. More specifically, the paper shows how PH-transforms can be used to quantify process risk, parameter risk, and dependency risk. It also discusses economic justification after introducing implementation issues.

To utilize the PH-transform in ratemaking, a probability distribution for claims is needed. A probability distribution can often be estimated from industry claim data or by computer simulations. Even though a probability distribution can be obtained from past claim data, sound and knowledgeable judgements are always required to ensure that the estimated loss distribution is valid for ratemaking.

It is safe to say that no theoretical risk-load formula can claim to be the *right* one, since subjective elements always exist in any practical exercise of ratemaking. However, a good theoretical risk-load formula can assist actuaries and help maintain logical consistency in the ratemaking process. In this respect, it is hoped that the PH-transform method becomes a useful tool for practicing actuaries in insurance ratemaking.

The remainder of this paper is divided into four sections. Section 2 introduces the PH-transform method and applies it to the pricing of a single risk (including excess cover and increased limits ratemaking). Section 3 discusses the use of PH-transforms in pricing risk portfolios and reinsurance treaties. Section 4 discusses two simple mixtures of PH-transforms. The first mixture can yield a minimal rate-on-line, and the second mixture suggests a new measure for the right tail risk. Section 5 briefly reviews the leading economic theories of risk and uncertainty, and their relationship to insurance ratemaking.

2. PROPORTIONAL HAZARDS TRANSFORM

In the pricing of insurance risks, it is common for the actuary to first obtain a best-estimate loss distribution based on all possible information (e.g., empirical data) and/or judgement. The best estimate loss distribution, serving as an anchor, is then transformed into a heavier-tailed distribution, and the mean from the latter is used to price the business, thereby producing a risk load. Venter [28] advocated the adjusted distribution principle and gave a theoretical justification by using a no-arbitrage pricing argument. He observed that the only methods of premium calculation that preserve layer additivity are those that can be generated from transformed distributions, where the premium for any layer is the expected loss for that layer under the transformed distribution. Inspired by Venter's insightful observation, Wang [30] proposed the proportional hazards transform method which is also the topic of this paper.

An insurance risk refers to a non-negative loss random variable X, which can be described by the decumulative distribution function (ddf): $S_X(u) = \Pr\{X > u\}$. An advantage of using the ddf is the unifying treatment of discrete, continuous, and mixed-type distributions. In general, for a risk X, the expected loss can be evaluated directly from its ddf:

$$\mathrm{E}[X] = \int_0^\infty S_X(u)\,du.$$

(A proof of this statement is given in Appendix A.) In practice, the actuary does not know the true underlying loss distribution, but instead may have a best-estimate loss distribution based on available information. The PH-transform is a method for adjusting the best-estimate distribution according to the levels of uncertainty, market competition, and portfolio diversification.

DEFINITION 1 *Given a best-estimate loss distribution* $S_X(u) = \Pr\{X > u\}$*, for some exogenous* ***index*** r *(*$0 < r \leq 1$*), the* ***proportional hazards (PH) transform*** *refers to a mapping* $S_Y(u) = [S_X(u)]^r$*, and the* ***PH-mean*** *refers to the expected value under the transformed distribution:*

$$H_r[X] = \int_0^{\infty} [S_X(u)]^r \, du, \qquad (0 < r \leq 1).$$

The PH-mean was introduced by Wang to represent the risk-adjusted premium (the expected loss plus risk load). As we shall see, the PH-mean is quite sensitive to the choice of the index r. It could be infinite for some unlimited loss distributions and choices of r.

EXAMPLE 1 The following three loss distributions,

$$S_U(u) = 1 - u/(2b), \qquad 0 \leq u \leq 2b \qquad \text{(uniform)},$$

$$S_V(u) = e^{-u/b} \qquad \text{(exponential), and}$$

$$S_W(u) = b^2/(b+u)^2 \qquad \text{(Pareto)},$$

have the same expected loss, b. One can easily verify that

$$H_r[U] = \frac{2b}{1+r},$$

$$H_r[V] = \frac{b}{r},$$

$$H_r[W] = \begin{cases} \dfrac{b}{2r-1}, & r > 0.5; \\ \infty, & r \leq 0.5. \end{cases}$$

The PH-mean, interpreted as the risk-adjusted premium, preserves the usually accepted ordering of riskiness based on heaviness of tail (see Table 1). Here it is assumed that the distributions are known to be of the type shown, whereas uncertainty about the type of distribution could contribute further risk.

TABLE 1

SOME VALUES OF THE PH-MEAN $H_r[\cdot]$

	U	V	W
$r_1 = \frac{5}{6}$	$1.09b$	$1.2b$	$1.5b$
$r_2 = \frac{2}{3}$	$1.2b$	$1.5b$	$3.0b$

EXAMPLE 2 When X has a Pareto distribution with parameters (α, λ),

$$S_X(u) = \left(\frac{\lambda}{\lambda + u}\right)^{\alpha}, \qquad \text{and}$$

the PH-transform $S_Y(u)$ also has a Pareto distribution with parameters $(r\alpha, \lambda)$.

When X has a Burr distribution with parameters (α, λ, τ),

$$S_X(u) = \left(\frac{\lambda}{\lambda + u^{\tau}}\right)^{\alpha}, \qquad \text{and}$$

the PH-transform $S_Y(u)$ also has a Burr distribution with parameters $(r\alpha, \lambda, \tau)$.

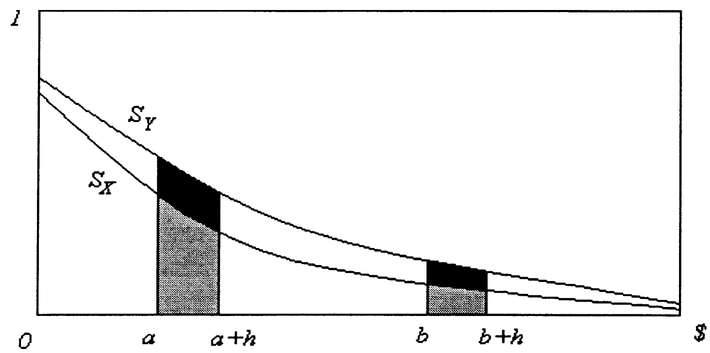
When X has a gamma (or lognormal) distribution, the PH-transform $S_Y(u)$ is no longer a gamma (or lognormal). In such cases, numerical integration may be required to evaluate the PH-mean.

2.1. *Pricing of Ambiguous Risks*

In practice, the underlying loss distribution is seldom known with precision. There are always uncertainties regarding the best-estimate loss distribution. Insufficient data or poor quality data often result in sampling errors. Even if a large amount of high-quality data is available, due to changes in the claim generating mechanisms, past data may not fully predict the future claim distribution. The PH-transform can be adjusted to give a higher risk load when this parameter uncertainty is greater.

FIGURE 1

MARGINS FOR PARAMETER UNCERTAINTY BY PH-TRANSFORMS



As illustrated in Figure 1, the PH-transform, $S_Y(u) = [S_X(u)]^r$, can be viewed as an upper confidence limit for the best-estimate loss distribution $S_X(u)$. A smaller index r yields a wider range between the curves S_Y and S_X. This upper confidence limit interpretation has support in statistical estimation theory (see Appendix B). The index r can be assigned accordingly with respect to the level of confidence in the estimated loss distribution. The more ambiguous the situation is, the lower the value of r that should be used.

EXAMPLE 3 Consider the following experiment conducted by Hogarth and Kunreuther [6]. An actuary is asked to price warranties on the performance of a new line of microcomputers. Suppose that the cost of repair is \$100 per unit, and there can be at most one breakdown per period. Also, suppose that the risks of breakdown associated with any two units are independent. The best-estimate of the probability of breakdown has three scenarios:

$$\theta = 0.001, \qquad \theta = 0.01, \qquad \theta = 0.1.$$

The level of confidence regarding the best estimate has two scenarios:

> Non-ambiguous: There is little ambiguity regarding the best-estimate loss distribution. Experts all agree with confidence on the chances of a breakdown.
>
> Ambiguous: There is considerable ambiguity regarding the best-estimate loss distribution. Experts disagree and have little confidence in the estimate of the probabilities of a breakdown.

Note that the loss associated with a computer component can only assume two possible values, either zero or \$100. For any fixed $u < 100$, the probability that the loss exceeds u is the same as the probability of being exactly \$100, namely θ. For a fixed $u \geq 100$, it is impossible that the loss exceeds u. Thus, the best-estimate ddf of the insurance loss cost is

$$S_X(u) = \begin{cases} \theta, & 0 < u < 100; \\ 0, & 100 \leq u. \end{cases}$$

The PH-transform with index r yields a risk-adjusted premium of $100\theta^r$.

In both cases a risk load is needed because there is frequency uncertainty, but more load is needed in the ambiguous case. If we choose $r = 0.97$ for the non-ambiguous case, and $r = 0.87$ for the ambiguous case, we get the premium structure shown in Table 2.

For comparison purposes, Table 2 also shows the premium structure using the standard deviation method[1] set to agree with the PH-mean at the 0.01 frequency. Note that for the Bernoulli type of risks in this example, the standard deviation loads vary more by frequency. However, as we shall see in Section 3.2, this

[1]The traditional standard deviation method calculates a risk-adjusted premium by the formula $E[X] + \beta\sigma[X]$, where $\beta \geq 0$ is an exogenous constant.

TABLE 2

THE RATIO OF THE RISK-ADJUSTED PREMIUM TO THE EXPECTED LOSS

PH-Transform Method	$\theta = 0.001$	$\theta = 0.01$	$\theta = 0.1$
Non-ambiguous ($r = 0.97$)	1.23	1.15	1.07
Ambiguous ($r = 0.87$)	2.45	1.82	1.35
Standard Deviation Method	$\theta = 0.001$	$\theta = 0.01$	$\theta = 0.1$
Non-ambiguous ($\beta = 0.01508$)	1.48	1.15	1.05
Ambiguous ($\beta = 0.0824$)	3.60	1.82	1.25

pattern no longer holds for continuous-type risks. The main problem with standard deviation is in its lack of additivity when a risk is divided into sub-layers.

In summary, the PH-transform can be used as a means of provision for estimation errors. The actuary can subsequently set up a table for the index r according to different levels of ambiguity, such as the following:

Ambiguity Level	Index r
Slightly Ambiguous	0.960 – 1.000
Moderately Ambiguous	0.900 – 0.959
Highly Ambiguous	0.800 – 0.899
Extremely Ambiguous	0.500 – 0.799

Note that the premium developed is particularly sensitive to the choice of r, especially for small r, so care should be exercised in its selection.

2.2. *Pricing Excess Layers of a Single Risk*

Since most insurance contracts contain clauses such as a deductible and a maximum limit, it is convenient to use the general language of excess-of-loss layers. A layer $(a, a+h]$ of a risk X is defined by the loss function:

$$X_{(a,a+h]} = \begin{cases} 0, & 0 \le X < a; \\ (X-a), & a \le X < a+h; \\ h, & a+h \le X, \end{cases}$$

where a is the *attachment point (retention)*, and h is the *limit*.

In this subsection, we restrict our discussion to a single risk X (individual or aggregate). For instance, X may represent an underlying risk for facultative reinsurance, or the aggregate loss amount for a risk portfolio being priced. Under this restriction, there will be either no or one claim to a given layer. In other words, the claim frequency to a given layer is Bernoulli. In Section 3 we will discuss the pricing of excess layers of reinsurance treaties where there can be multiple claims to a given layer.

One can verify that the loss variable $X_{(a,a+h]}$ has a ddf of

$$S_{X_{(a,a+h]}}(u) = \begin{cases} S_X(a+u), & 0 \le u < h \\ 0, & h \le u, \end{cases}$$

and that the average loss cost for the layer $(a, a+h]$ is

$$\mathrm{E}[X_{(a,a+h]}] = \int_0^h S_X(a+u)\,du = \int_a^{a+h} S_X(u)\,du.$$

Under the PH-transform $S_Y(u) = [S_X(u)]^r$, the PH-mean for the layer $(a, a+h]$ is

$$\begin{aligned} H_r[X_{(a,a+h]}] &= \int_0^\infty [S_{X_{(a,a+h]}}(u)]^r\,du \\ &= \int_0^h [S_X(a+u)]^r\,du = \int_a^{a+h} [S_X(u)]^r\,du. \end{aligned}$$

In other words, the expected loss and the risk-adjusted premium for the layer $(a, a+h]$ are represented by the areas over the interval $(a, a+h]$ under the curves $S_X(u)$ and $S_Y(u)$, respectively (see Figure 1).

In Wang [30], it is shown that, for $0 < r < 1$, the PH-mean has the following properties:

- Positive loading: $H_r[X_{(a,a+h]}] \geq \mathrm{E}[X_{(a,a+h]}]$.

- Decreasing risk-adjusted premiums:

$$\text{For} \quad a < b, \qquad H_r[X_{(a,a+h]}] \geq H_r[X_{(b,b+h]}].$$

- Increasing relative loading:

$$\text{For} \quad a < b, \qquad \frac{H_r[X_{(a,a+h]}]}{\mathrm{E}[X_{(a,a+h]}]} \leq \frac{H_r[X_{(b,b+h]}]}{\mathrm{E}[X_{(b,b+h]}]}.$$

These properties are consistent with market premium structures (Patrick, [20]; Venter, [28]).

EXAMPLE 4 A single (ground-up) risk has a 10% chance of incurring a claim, and, if a claim occurs, the claim size has a Pareto distribution ($\lambda = 2{,}000$, $\alpha = 1.2$). Putting the Bernoulli frequency and the Pareto severity together, we have a ground-up loss distribution

$$\begin{aligned} S_X(u) &= \Pr\{X > u\} \\ &= \text{Probability of occurrence} \times \Pr\{\text{Loss Size } > u\} \\ &= 0.1 \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2}. \end{aligned}$$

The actuary is asked to price various layers of the (ground-up) risk. Suppose that the actuary infers an index, say $r = 0.92$, from individual risk analysis and market conditions. The actuary may need to consider the risk loads for other contracts with similar characteristics in the insurance and/or financial markets.

TABLE 3

LAYER PREMIUMS USING PH-TRANSFORMS

Layer $X_{(a,b]}$ 1K = $1,000	Expected Loss	$H_{0.92}[X_{(a,b]}]$ (r = 0.92)	Percentage Loading	$H_{0.90}[X_{(a,b]}]$ (r = 0.90)	Percentage Loading
(0, 1K]	77.89	95.47	22.6%	100.45	29.0%
(5K, 6K]	20.51	27.99	36.5%	30.25	47.5%
(10K, 11K]	11.098	15.91	43.3%	17.41	56.8%
(50K, 51K]	1.982	3.26	64.5%	3.69	86.3%
(100K, 101K]	0.888	1.56	75.4%	1.79	102%
(500K, 501K]	0.132	0.269	104%	0.322	144%
(1,000K, 1,001K]	0.058	0.126	118%	0.152	165%

The PH-transform with $r = 0.92$ yields a ddf of

$$S_Y(u) = 0.1^{0.92} \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2\times 0.92}.$$

For any excess layer $[a, a+h)$, the expected loss to the layer is

$$\mathrm{E}[X_{(a,a+h]}] = \int_a^{a+h} 0.1 \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2} du,$$

and the risk-adjusted premium by using a PH-transform (r = 0.92) is

$$H_r[X_{(a,a+h]}] = \int_a^{a+h} 0.1^{0.92} \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2\times 0.92} du.$$

Risk-adjusted premiums for various layers are shown in Table 3. In Table 3 we also list the prices by using a slightly different $r = 0.90$. Note that the developed prices are sensitive to the index r.

2.3. Increased Limits Ratemaking

In commercial liability insurance, a policy generally covers a loss (it may include allocated loss adjustment expense) up to a specified maximum dollar amount that will be paid on any individual loss. In the U.S., it is general practice to publish rates for some standard limit called the basic limit (historically $25,000, but now $100,000). Increased limit rates are calculated by applying increased limit factors (ILFs). Without risk load, the increased limit factor is the expected loss at the increased limit divided by the expected loss at the basic limit. The increased limit factor with risk load is the sum of the expected loss and the risk load at the increased limit divided by the sum of the expected loss and the risk load at the basic limit:

$$ILF(\omega) = \frac{\mathrm{E}[X;\omega] + RL_{(0,\omega]}}{\mathrm{E}[X;100{,}000] + RL_{(0,100{,}000]}}.$$

It is widely felt that ILFs should satisfy the following conditions (see Rosenberg [26], Meyers [16], and Robbin [25]). They implicitly assume that insureds who buy different limits are nevertheless subject to the same loss distributions.

1. The relative loading with respect to the expected loss is higher for higher limits.
2. ILFs should produce the same price under any arbitrary division of layers.
3. The ILFs should exhibit a pattern of declining marginal increases as the limit of coverage is raised. In other words, when $x < y$,

$$ILF(x+h) - ILF(x) \geq ILF(y+h) - ILF(y).$$

In the U.S., many companies use the ILFs published by the Insurance Services Office (ISO). Traditionally, only the severity distribution is used when producing ILFs. Until the mid-1980s, ISO used the variance of the loss distribution to calculate risk loads, a method proposed by Robert Miccolis [18]. From the mid-1980s to the early 1990s, ISO used the standard deviation of the loss distribution to calculate risk loads (e.g., Feldblum [4]).

Meyers [16] presents a competitive market equilibrium approach, which yields a variance-based risk load method; however, some authors have questioned the appropriateness of the variance-based risk load method for the calculation of ILFs (e.g., Robbin [25]).

The following is an illustrative example to show how the PH-transform method can be used in increased limits ratemaking.

EXAMPLE 5 Assume that the claim severity distribution has a Pareto distribution with ddf

$$S_X(u) = \left(\frac{\lambda}{\lambda + u}\right)^{\alpha},$$

with $\lambda = 5{,}000$ and $\alpha = 1.1$. This is the same distribution used by Meyers, although he also considered parameter uncertainty.

Assume that, based on the current market premium structure, the actuary feels that (for illustration only) an index $r = 0.9$ provides an appropriate provision for parameter uncertainty. When using a Pareto severity distribution, there is a simple analytical formula for the ILFs:

$$ILF(\omega) = \frac{1 - \left(\dfrac{\lambda}{\lambda + \omega}\right)^{r\alpha - 1}}{1 - \left(\dfrac{\lambda}{\lambda + 100{,}000}\right)^{r\alpha - 1}}.$$

One can then easily calculate the increased limit factors at any limit (see Table 4).

2.4. *Some Properties of the PH-Mean*

For a single risk X and for $0 \le r \le 1$, the PH-mean has the following properties (see Wang [30]):

- $\mathrm{E}[X] \le H_r[X] \le \max[X]$. When r declines from one to zero, $H_r[X]$ increases from the expected loss, $\mathrm{E}[X]$, to the maximum possible loss, $\max[X]$.

TABLE 4

INCREASED LIMIT FACTORS USING PH-TRANSFORM

Policy Limit ω	Expected Loss $E[X;\omega]$	ILF Without RL	Risk Load	ILF With RL
100,000	13,124	1.00	2,333	1.00
250,000	16,255	1.24	3,796	1.30
500,000	18,484	1.41	5,132	1.53
750,000	19,726	1.50	6,000	1.66
1,000,000	20,579	1.57	6,653	1.76
2,000,000	22,543	1.71	8,343	2.00

- Scale and translation invariant: $H_r[aX+b] = aH_r[X]+b$, for $a,b \geq 0$.

- Sub-additivity: $H_r[X+Y] \leq H_r[X]+H_r[Y]$.

- Layer additivity: when a single risk X is split into a number of layers
$$\{(x_0,x_1],(x_1,x_2],\ldots\},$$
the layer premiums are additive (the whole is the sum of the parts):
$$H_r[X] = H_r[X_{(x_0,x_1]}] + H_r[X_{(x_1,x_2]}] + \cdots.$$

Pricing often assumes that a certain degree of diversification will be reached through market efforts. In real life examples, risk-pooling is a common phenomenon. It is assumed that, in a competitive market, the benefit of risk-pooling is transferred back to the policyholders (in the form of premium reduction). In the PH-model, the layer-additivity and the scale-invariance have already taken into account the effect of risk-pooling. To illustrate, consider a single risk with a maximum possible loss of \$100 million. Suppose one insurer is asked to quote premium rates for each of the following as stand-alone coverages: sub-layers $(0,10], (10,20], \ldots, (90,100]$ and the whole risk $(0,100]$. The quoted premium for the entire risk $(0,100]$ may exceed the sum of individual premiums for each sub-layer. This is because the limit of $100 million may be a lot for a single insurer to carry without a substantial profit margin. However, the market mechanism would facilitate risk-sharing schemes among several insurers (say, ten insurers each take a sub-layer). Thus, when this risk pooling effect is transferred back to the policyholder, the premiums should be additive for different layers. Likewise, if one insurer is asked to quote premium rates for a 10% quota-share of this risk as opposed to the whole risk, the quoted premiums may exhibit non-linearity. However, the market risk-sharing scheme would force the premiums to be scale invariant—i.e., a 10% quota share demands 10% of the total premium.

Theoretically, in an efficient market (no transaction expenses in risk-sharing schemes) with complete information, the optimal cooperation among insurers is to form a market insurance portfolio (like the Dow Jones index), and each insurer takes a layer or quota-share of the market insurance portfolio.

In real life, however, the insurance market is *not* efficient. This is mainly because of incomplete information (ambiguity) and extra expenses associated with the risk-sharing transactions. There exist distinctly different local market climates in different geographic areas and in different lines of insurance. Catastrophe risk varies from region to region. In some geographic regions, due to high concentration and lack of information (ambiguity), existing risk-sharing schemes are not sufficient to diversify the risk to the extent one would wish. As a result the market would demand a higher risk load (a smaller value of r in the PH-model).

In summary, the index r may vary with respect to the local market climate which is characterized by *the levels of ambiguity*, *risk concentration*, and *competition*.

3. PRICING RISK PORTFOLIOS AND REINSURANCE TREATIES

When pricing a (re)insurance contract that covers a group of risks, the actuary often estimates claim frequency and claim severity separately, due to the type of information available. One straight-forward approach is to apply PH-transforms to the frequency and severity distributions separately, and then take the product of the loaded frequency and the loaded severity. An alternative is to first calculate the aggregate loss distribution from the estimated frequency and severity distributions, and then apply the PH-transform to the aggregate loss distribution. This section will discuss and compare both approaches.

3.1. Frequency/Severity Approach to Pricing Group Insurance

Let N denote the claim frequency with probability function $p_k = \Pr\{N = k\}$ and ddf $S_N(k) = p_{k+1} + p_{k+2} + \cdots$, $(k = 0,1,2,\ldots)$. The PH-mean for the frequency can be calculated as the sum

$$H_r[N] = S_N(0)^r + S_N(1)^r + S_N(2)^r + \cdots,$$

where convergence is required if N is unlimited (e.g., a Poisson frequency is unlimited).

Depending on the available information, the actuary may have different levels of confidence in the estimates for the frequency and severity distributions. According to the level of confidence in the estimated frequency and severity distributions, the actuary can choose an index r_1 for the frequency and an index r_2 for the severity. As a result, the actuary can calculate the risk-adjusted premium for the risk portfolio as

$$H_{r_1}[N] \times H_{r_2}[X].$$

EXAMPLE 6 Consider a group coverage of liability insurance. The actuary has estimated the following loss distributions: (i) the claim frequency has a Poisson distribution with $\lambda = 2.0$, and (ii) the claim severity is modeled by a lognormal distribution with a mean of \$50,000 and coefficient of variation of 3 (which was used by Finger [5] for a liability claim severity distribution). Here we also assume a coverage limit of one million dollars per claim. Suppose that the actuary has low confidence in the estimate of claim frequency, but higher confidence in the estimate of the claim severity distribution, and thus chooses $r_1 = 0.85$ for the claim frequency and $r_2 = 0.9$ for the claim severity. The premium can be calculated using numerical integrations:

$$H_{0.85}[N] = 2.227, \qquad \text{and} \qquad H_{0.9}[X] = 58{,}080.$$

Thus, the required total premium is

$$H_{0.85}[N] \times H_{0.9}[X] = 129{,}344.$$

Kunreuther et al. [13] discussed the ambiguities associated with the estimates for claim frequencies and severities. They mentioned that for some risks such as playground accidents, there are considerable data on the chances of occurrence but much uncertainty about the potential size of the loss due to arbitrary court awards. On the other hand, for some risks such as satellite losses or new product defects, the chance of a loss occurring is highly ambiguous due to limited past claim data. However, the magnitude of such a loss is reasonably predictable.

3.2. Frequency/Severity Approach to Pricing Per Risk Excess-of-Loss Reinsurance Treaties

A reinsurance excess-of-loss treaty normally covers a block of underlying policies where the attachment point and the policy limit apply on a per risk basis. For such reinsurance treaties, the claim frequency usually has a non-Bernoulli type distribution—that is, the number of claims may exceed one. For some low limit working layers where a substantial number of claims is expected, the major uncertainty might be in the claim frequency rather than in the severity.

In the market, reinsurance brokers often structure the coverage in a number of layers. It is important to have consistent pricing on all layers. Here we give an example.

EXAMPLE 7 Consider a reinsurance excess-of-loss treaty. The projected ceding company subject earned premium (SEP) for the treaty is $10,000,000. The actuary is asked to price the following excess layers which are all on a per risk basis:

1. $400K xs $100K,
2. $500K xs $500K, and
3. The combined layer $900K xs $100K.

Suppose that, based on past loss data of the ceding company, after appropriate trending and development, the actuary has come up with the following best-estimates:

- The number of claims which cut into the first layer has a Poisson distribution with mean $\lambda = 6$.
- The size of losses greater than $100K can be modeled by a single parameter Pareto with ddf

$$S_X(u) = \left(\frac{100}{u}\right)^{1.647}, \qquad u > 100.$$

Under this ground-up severity distribution, the loss to the first layer has a Pareto (100, 1.647) distribution truncated at 400 with a ddf of

$$S_1(u) = \begin{cases} \left(\dfrac{100}{u+100}\right)^{1.647}, & 0 < u < 400; \\ 0, & 400 \leq u, \end{cases}$$

which has a mean severity of $100,001.

The loss to the second layer has a Pareto (500, 1.647) distribution truncated at 500 with a ddf of

$$S_2(u) = \begin{cases} \left(\dfrac{500}{u+500}\right)^{1.647}, & 0 < u \leq 500; \\ 0, & 500 \leq u, \end{cases}$$

which has a mean severity of \$279,284.

In general, the frequency and severity distributions both change with the attachment point. To ensure consistency, it is important to work with the frequency and severity distributions for losses above the *minimum* attachment point. For convenience we refer to them as "ground-up" distributions, although they are not "real" ground-up distributions. In practice, reinsurers are usually supplied with data of large losses only, the "real" ground-up loss distribution below the attachment point is seldom known to the reinsurer.

By transforming the "ground up" frequency and severity distributions separately, we can load for the different frequency/severity risks accordingly.

For numerical illustration, we use the same PH-index $r = 0.95$ for both frequency and severity.

The PH-mean for a Poisson(6) distribution is 6.119, which represents a 1.98% frequency loading. First, we apply the PH-transform ($r = 0.95$) to the "ground-up" severity distribution and allocate the loaded costs to each layer (see Table 5). Under the Pareto (100, 1.647) "ground up" severity curve, the average severity in the layer 500 xs 500 is \$279,284, and the probability of cutting into the second layer given that a loss has cut into the first layer is 0.706. Therefore, the average loss to the second layer, *among all claims that have cut into the first layer*, can be calculated as the product $0.706 \times \$279,284 = \$19,717$.

TABLE 5

TRANSFORMING THE "GROUND UP" SEVERITY DISTRIBUTION

	Layer	Average Loss to the Layer Before Transform	Average Loss to the Layer After Transform $r = 0.95$	Relative Loading Ratio
(1)	400 xs 100	\$100,001	\$105,726	1.057
(2)	500 xs 500	\$ 19,717	\$ 23,117	1.172
(3)	900 xs 100	\$119,718	\$128,843	1.076
(1) + (2)	900 xs 100	\$119,718	\$128,843	1.076

TABLE 6

COMBINING LOADED "GROUND UP" FREQUENCY AND SEVERITY (PH-INDEX $r = 0.95$)

	Layer	Burning Cost (expected loss) As % of SEP	Loaded Rate As % of SEP $H_r[N] \times H_r[X]$	Relative Loading Ratio
(1)	400 xs 100	6.000%	6.469%	1.078
(2)	500 xs 500	1.183%	1.414%	1.196
(3)	900 xs 100	7.183%	7.883%	1.098
(1) + (2)	900 xs 100	7.183%	7.883%	1.098

Finally, we multiply the loaded "ground up" frequency and loaded severity in each layer to get the premium rate for each reinsurance layer (see Table 6). As a convention in reinsurance, the burning costs and premium rates are expressed as a percentage of the subject earned premium (SEP), \$10,000,000 in this example.

Note that, with this approach, we get premiums that are layer-additive. In other words, the total premium would not change regardless of how we divide the coverage into layers.

3.3. Aggregate Approach to Pricing Per-Risk Excess Treaty

As an alternative approach, the actuary can calculate/simulate the aggregate loss distribution from the best-estimate frequency and severity distributions, and subsequently apply the PH-transform to the aggregate loss distribution.

For given frequency N and severity X, let

$$Z = X_1 + X_2 + \cdots X_N$$

represent the aggregate loss amount for the risk portfolio. Various numerical and simulation techniques are available for calculating the aggregate loss distribution (e.g., Heckman and Meyers [7], and Panjer [19]).

In general, we get different results by transforming the frequency and severity distributions separately *versus* transforming the aggregate loss distribution. For the collective risk model, where claim severities are assumed mutually independent and independent of the frequency, we have the following inequality:

$$H_r[Z] \leq H_r[N]H_r[X], \qquad 0 < r < 1.$$

This is because, conditional on $N = n$, we always have

$$H_r[Z \mid N = n] = H_r[X_1 + \cdots + X_n] \leq nH_r[X].$$

In other words, the PH-transform of the aggregate loss distribution takes account of the fact that the variability regarding the aggregate loss is reduced in the pooling of N independent losses. However, one should carefully examine the validity of the independence assumption, especially with the presence of ambiguity (parameter uncertainties) in the best-estimate loss distributions. Parameter uncertainty can generate some correlation effect, although the claim processes may be independent provided that the true underlying distributions are known.

EXAMPLE 7 revisited: Now we re-consider the reinsurance treaty example using an aggregate approach. For ease of computation,

TABLE 7

APPLY PH-TRANSFORM TO THE AGGREGATE LOSS DISTRIBUTIONS OF EACH PER RISK EXCESS LAYER ($r = 0.9025$)

	Layer in 000's	Burning Cost As % of SEP	Indicated Rate As % of SEP	Relative Loading Ratio
(1)	400 xs 100	6.000%	6.384%	1.064
(2)	500 xs 500	1.183%	1.408%	1.190
(3)	900 xs 100	7.183%	7.742%	1.078
(1) + (2)	900 xs 100	7.183%	7.792%	1.085

here we assume independence among the individual claims in the calculation of the aggregate loss distribution for each layer. For a numerical comparison with the separate adjustment of frequency and severity, we apply a PH-transform with an index $r = r_1 \times r_2 = 0.95 \times 0.95 = 0.9025$ to the aggregate loss distribution of each layer. The indicated rate for each layer is given in Table 7.

We give some modeling details regarding this specific example. The claim frequency for the upper layer 500 xs 500 has a Poisson distribution with mean 0.424. This can be derived from the Poisson frequency for the lower layer 400 xs 100 and the probability of cutting into the second layer given that a loss has already cut into the first layer. Recall that the claim severity distribution for the layer 500 xs 500 has a Pareto (500, 1.647) distribution truncated at the policy limit 500. This can be verified using a conditional probability argument.

In this aggregate approach, we used a more severe PH-index $r = 0.9025$ as compared to $r = 0.95$. The aggregate approach produces a premium structure similar to that obtained by transforming frequency and severity separately (see Table 7 and Table 6). The use of a more severe index offsets the risk reduction as a result of pooling independent loss sizes.

Another important observation can be made from Table 7. With the aggregate approach, the premium rates are *not* additive for layers. The premium rate for the first layer (6.384%) plus that for the second layer (1.408%) is 7.792%, which is greater than the rate for the combined layer (7.742%). This lack of layer additivity may be a drawback of the aggregate approach in pricing per risk excess reinsurance treaties.

3.4. Aggregate Approach to Pricing Aggregate Contracts

Some reinsurance contracts are written in aggregate terms where the coverage triggers when the aggregate loss (or loss ratio) for the whole book exceeds some specified amount. Usually these contracts specify the attachment point and coverage limit in aggregate terms. In pricing such aggregate treaties, a natural approach would be to use the aggregate loss distribution, simply because the coverage trigger is the aggregate loss amount. In other words, the actuary needs to calculate/simulate a probability distribution for the aggregate loss $Z = X_1 + \cdots + X_N$. Based on the claim generating mechanism as well as the level of ambiguity, the actuary may assume some correlation between individual risks. The PH-transform of the aggregate loss distribution will automatically take into account the effect of correlation. The higher the correlation between individual risks, the greater the PH-mean for the aggregate loss distribution.

For some CAT events it might be plausible to consider the dependency between claim frequency and claim severity. For instance, the Richter scale value of an earthquake may affect both the frequency and severity simultaneously, and for hurricane losses, the wind velocity would affect both the frequency and severity simultaneously. Regardless of the dependency structure, computer simulation methods can always be used to model the aggregate losses based on a given geographic concentration. The PH-transform of the aggregate loss distribution can capture the correlation risk in the developed prices.

4. MIXTURE OF PH-TRANSFORMS

While a single index PH-transform has one parameter r to control the relative premium structure, one can obtain more flexible premium structures by using a mixture of PH-transforms:

$$p_1 H_{r_1} + p_2 H_{r_2} + \cdots + p_n H_{r_n},$$

$$\sum_{j=1}^{n} p_j = 1, \qquad 0 \leq r_j \leq 1 \quad (j = 1, \ldots, n).$$

The PH-index mixture can be interpreted as a collective decision-making process. Each member of the decision-making 'committee' chooses a value of r, and the index mixture represents different rs chosen by different members. It also has interpretations as (i) an index mixture chosen by a rating agency according to the indices for all insurance companies in the market; (ii) an index mixture which combines an individual company's index with the rating agency's index mixture.

A mixture of PH-transforms has the same properties as that for a single index PH-transform (see Section 2.4). For ratemaking purposes, a mixture of PH-transforms enjoys more flexibility than a single index PH-transform. Now we shall discuss two special mixtures of the form

$$(1-\alpha) H_{r_1}[X] + \alpha H_{r_2}[X], \qquad 0 \leq \alpha \leq 1, \quad r_1, r_2 \leq 1.$$

4.1. Minimum Rate-on-Line

In most practical circumstances, very limited information is available for claims at extremely high layers. In such highly ambiguous circumstances, most reinsurers adopt a survival rule of minimum rate-on-line. The rate-on-line is the premium divided by the coverage limit, and most reinsurers establish a minimum they will accept for this ratio (see Venter [28]).

TABLE 8

LAYER PREMIUMS UNDER AN INDEX MIXTURE

Layer $X_{(a,b]}$ 1K=\$1,000	Expected Loss	Risk-adjusted Premium	Percentage Loading
(0, 1K]	77.892	131.56	45.8%
(5K, 6K]	20.512	47.43	131%
(10K, 11K]	11.098	35.59	220%
(50K, 51K]	1.982	23.20	1,070%
(100K, 101K]	0.888	21.53	2,324%
(500K, 501K]	0.132	20.26	15,276%
(1,000K, 1,001K]	0.058	20.12	34,875%

By using a two-point mixture of PH-transforms with $r_1 \leq 1$ and $r_2 = 0$, the premium functional

$$(1-\alpha)H_{r_1}[X] + \alpha H_0[X] = (1-\alpha)H_{r_1}[X] + \alpha \max[X]$$

can yield a minimum rate-on-line at α.

EXAMPLE 8 Reconsider Example 4. The best-estimate loss distribution (ddf) is

$$S_X(u) = 0.1 \times \left(\frac{2{,}000}{2{,}000+u}\right)^{1.2}.$$

By choosing a two-point mixture with $r_1 = 0.92$, $r_2 = 0$, and $\alpha = 0.02$, we get an adjusted distribution:

$$S_Y(u) = 0.98 \times 0.1^{0.92} \times \left(\frac{2{,}000}{2{,}000+u}\right)^{1.2\times 0.92} + 0.02.$$

Note that S_Y being a proper loss distribution requires a finite upper layer limit.

As shown in Table 8, this two-point mixture guarantees a minimum rate-on-line at 0.02 (1 full payment out of 50 years). Note that the average index $\bar{r} = (1-\alpha)r_1 + \alpha r_2 = 0.9016 \approx 0.90$.

We can see that this method yields distinctly different premiums from those in Table 3 where the single indices $r = 0.92$ and $r = 0.90$ are used.

4.2. *The Right-Tail Deviation*

Consider a two-point mixture of PH-transforms with $r_1 = 1$ and $r_2 = \frac{1}{2}$:

$$(1-\alpha)H_1[X] + \alpha H_{1/2}[X] = \mathrm{E}[X] + \alpha(H_{1/2}[X] - \mathrm{E}[X]), \qquad 0 < \alpha < 1,$$

which is similar in form to the standard deviation method of $\mathrm{E}[X] + \alpha\sigma[X]$.

Now we introduce a new risk-measure analogous to the standard deviation.

DEFINITION 2 *The right-tail deviation is defined as*

$$D[X] = H_{1/2}[X] - \mathrm{E}[X] = \int_0^\infty \sqrt{S_X(u)}\,du - \int_0^\infty S_X(u)\,du.$$

Analogous to the standard deviation, the right-tail deviation $D[X]$ satisfies the following properties:

- If $\Pr\{X = b\} = 1$, then $D[X] = 0$.
- Scale-invariant: $D[cX] = cD[X]$ for $c > 0$.
- Shift-invariant: $D[X + b] = D[X]$ for any constant b.
- Sub-additivity: $D[X + Y] \le D[X] + D[Y]$.
- If X and Y are perfectly correlated, then $D[X + Y] = D[X] + D[Y]$.

It is shown in Appendix A that, for a small layer $(a, a+h]$, $D[X_{(a,a+h]}] \le \sigma[X_{(a,a+h]}]$, and $D[X_{(a,a+h]}]$ converges to $\sigma[X_{(a,a+h]}]$

TABLE 9

THE RIGHT-TAIL DEVIATION VERSUS THE STANDARD DEVIATION

Layer 1K = $1,000	Expected loss	Std-deviation of the loss	Right-tail deviation	Percentage difference
L	$\mathrm{E}[L]$	$\sigma[L]$	$D[L]$	$\frac{\sigma[L]}{D[L]} - 1$
(0, 1K]	77.89	256.0	200.5	27.7%
(1K, 2K]	51.56	214.3	175.2	22.3%
(10K, 11K]	11.10	103.9	94.24	10.3%
(100K, 101K]	.8879	29.76	28.91	2.93%
(1,000K, 1,001K]	.05754	7.584	7.528	0.75%
(10,000K, 10,001K]	.003640	1.908	1.904	0.19%

at upper layers (i.e., the relative error goes to zero when a becomes large). As a result, the right-tail deviation $D[X]$ is finite if and only if the standard deviation $\sigma[X]$ is finite.

EXAMPLE 9 Re-consider the loss distribution in Example 4 with a ddf of

$$S_X(u) = 0.1 \times \left(\frac{2{,}000}{2{,}000 + u}\right)^{1.2}.$$

For different layers with fixed limit at 1000, Table 9 compares the standard deviation with the right-tail deviation.

Having stated a number of similarities, here we point out two crucial differences between the right-tail deviation $D[X]$ and the standard deviation $\sigma[X]$ (see Wang [31]):

- $D[X]$ is layer-additive, but $\sigma[X]$ is *not* additive.
- $D[X]$ preserves some natural ordering of risks such as first stochastic dominance,[2] but $\sigma[X]$ does not.

[2]Risk X is smaller than risk Y in first stochastic dominance if $S_X(u) \le S_Y(u)$ for all $u \ge 0$; or equivalently, Y has the same distribution as $X + Z$ where Z is another non-negative random variable.

TABLE 10

LAYER ADDITIVITY: A COMPARISON

Layer	Expected Losses	Standard Deviation	Right-Tail Deviation
(0, 10K]	301	1378	1355
(10K, 20K]	80	834	809
(0, 20K]	381	2035	2164
Result	Additive	Sub-Additive	Additive

These two crucial differences give the right-tail deviation an advantage over the standard deviation in pricing insurance risks.

Although for a small layer $(a, a+h]$ we have $D[X_{(a,a+h]}] \leq \sigma[X_{(a,a+h]}]$, for the entire risk X the right-tail deviation often exceeds the standard deviation, since the right-tail deviation is layer-additive while the standard deviation is not. For example, consider two sub-layers $(0, 10\text{K}]$ and $(10\text{K}, 20\text{K}]$, and a combined layer $(0, 20\text{K}]$. The right-tail deviation exceeds the standard deviation for the combined layer (0, 20K], although the reverse relation holds for each sub-layer (see Table 10).

Remark For a layer $(a, a+h]$, the loss-to-limit ratio is defined as the ratio of incurred loss to the limit of the layer. When the layers are refined (h becomes small), the loss-to-limit ratio approaches the ddf at that layer, which is also the frequency of hitting the layer. This can be seen from the relation

$$\lim_{h \to 0} \frac{\int_a^{a+h} S_X(u)\, du}{h} = S_X(a).$$

If a ground-up risk is divided into small adjacent layers, the empirical loss-to-limit ratios at various layers yield an approximation to the underlying ddf. As a pragmatic method for computing risk loads, it has been a longstanding practice of some reinsurers to adjust the empirical loss-to-limit ratio by adding a multiple of the square root of the empirical loss-to-limit ratio. As the layers are refined (h becomes small), this pragmatic method approaches the following:

$$\mathrm{E}[X_{(a,a+h]}] + \alpha(D[X_{(a,a+h]}] + \mathrm{E}[X_{(a,a+h]}]).$$

5. ECONOMIC THEORIES OF RISK LOAD

In this section we review some economic theories and show how the PH-transform fits in.

5.1. Expected Utility Theory

Traditionally, expected utility (EU) theory has played a dominant role in modeling decisions under risk and uncertainty. To a large extent, the popularity of EU was attributed to the axioms of von Neumann and Morgenstern [29].

Let V represent a random economic prospect and let $S_V(u) = \Pr\{V > u\}$ (i.e., the probability that the random economic prospect V exceeds value u). Let the symbols $\succ$ and $\sim$ represent strict preference and indifference, respectively. Von Neumann and Morgenstern proposed five axioms of decision under risk:

- EU.1 If V_1 and V_2 have the same probability distribution, then $V_1 \sim V_2$.
- EU.2 Weak order: $\succeq$ is reflective, transitive, and connected.
- EU.3 $\succeq$ is continuous in the topology of weak convergence.
- EU.4 If $V_1 \geq V_2$ with probability one, then $V_1 \succeq V_2$.
- EU.5 If $S_{V_1} \succeq S_{V_2}$, and for any $p \in [0,1]$, the probabilistic mixture satisfies

$$(1-p)S_{V_1} + pS_{V_3} \succeq (1-p)S_{V_2} + pS_{V_3}.$$

Von Neumann and Morgenstern showed that any decision-making which is consistent with these five axioms can be modeled by a utility function u such that '$V_1 \succ V_2$ if and only if $\mathrm{E}[u(V_1)] \geq \mathrm{E}[u(V_2)]$.'

When EU is applied to produce an insurance premium for a risk X, the minimum premium P that an insurance company will accept for full insurance satisfies the EU-equation

$$u(w) = \mathrm{E}[u(w + P - X)]$$

in which u and w refer to the insurer's utility function and wealth (see Bowers et al. [2]). The premium P from the EU-equation does not satisfy layer-additivity. Thus, the PH-transform does not fit in the expected utility framework.

5.2. The Dual Theory of Yaari

Modern economic theory questions the assumption that a firm can have a utility function, even when it accepts that individuals do. Yaari [32] proposed an alternative theory of decision under risk and uncertainty.

While the first four EU axioms are apparently reasonable, many people challenged the fifth axiom in the expected utility theory. While keeping the first four EU axioms unchanged, Yaari proposed an alternative to the fifth EU axiom:

DU.5* If V_1, V_2, and V_3 are co-monotone and $V_1 \succeq V_2$, for any $p \in [0,1]$, the outcome mixture satisfies

$$(1-p)V_1 + pV_3 \succeq (1-p)V_2 + pV_3.$$

Two risks X and Y are *co-monotone* if there exists a random variable Z and non-decreasing real functions u and v such that $X = u(Z)$ and $Y = v(Z)$ with probability one. Co-monotonicity is a generalization of the concept of perfect correlation to random variables without linear relationships. Note that perfectly correlated risks are co-monotone, but the converse does not hold. Consider two layers $(a, a+h]$ and $(b, b+h]$ for a continuous variate X. The layer payments $X_{(a,a+h]}$ and $X_{(b,b+h]}$ are co-monotone since both are non-decreasing functions of the original risk X. They are bets on the same event and neither of them is a *hedge* against the other. On the other hand, for $a \neq b$, $X_{(a,a+h]}$ and $X_{(b,b+h]}$ are *not* perfectly correlated since neither can be written as a linear function of the other.

Under axioms EU.1–4 & EU.5*, Yaari showed that there exists a distortion function $g : [0,1] \rightarrow [0,1]$ such that a certainty equivalent to a random economic prospect V on interval $[0,1]$ is

$$\int_0^1 g[S_V(y)]\,dy.$$

In other words, the certainty equivalent to a random economic prospect, $0 \leq V \leq 1$, is just the expected value under the distorted probability distribution, $g[S_V(y)]$, $0 \leq y \leq 1$.

Regarding the concept of risk-aversion, Yaari made the following observations:

> At the level of fundamental principles, risk-aversion and diminishing marginal utility of wealth, which are synonymous under expected utility theory, are horses of different colors. The former expresses an attitude towards risk (increased uncertainty hurts) while the latter expresses an attitude towards wealth (the loss of a sheep hurts more when the agent is poor than when the agent is rich). [32, p. 95]

The PH-transform fits in Yaari's economic theory with $g(x) = x^r$.

5.3. Schmeidler's Ambiguity-Aversion

As early as 1921, John Keynes identified a distinction between the *implication* of evidence (the implied likelihood) and the *weight* of evidence (confidence in the implied likelihood). Frank Knight [10] also drew a distinction between *risk* (with known probabilities) and *uncertainty* (ambiguity about the probabilities). A famous example on ambiguity-aversion is Ellsberg's [3] paradox which can be briefly described as follows: There are two urns each containing 100 balls. One is a non-ambiguous urn which has 50 red and 50 black balls; the other is an ambiguous urn which also contains red and black balls but with unknown proportions. When subjects are offered \$100 for betting on a red draw, most subjects choose the non-ambiguous urn (and the same for the black draw). Such a pattern of preference *cannot* be explained by EU (Quiggin, [24, p. 42]).

Ellsberg's work has spurred much interest in dealing with ambiguity factors in risk analysis. Schmeidler [27] brought to economists *non-additive probabilities* in his axiomization of preferences under uncertainty. For instance, in Ellsberg's experiment, the non-ambiguous urn with 50 red and 50 black balls is preferred to the ambiguous urn with unknown proportions of red or black balls. This phenomenon can be explained if we assume that one assigns a subjective probability $\frac{3}{7}$ to the chance of getting a red draw (or black draw). Since $\frac{3}{7} + \frac{3}{7} = \frac{6}{7}$ which is less than one, the difference $1 - \frac{6}{7} = \frac{1}{7}$ may represent the magnitude of ambiguity aversion.

In his axiomization of acts or risk preferences, Schmeidler obtained essentially the same mathematical formulation (axioms and theorems) as that of Yaari. A certainty equivalent to a random economic prospect V $(0 \leq V \leq 1)$ can be evaluated as

$$H[V] = \int_0^1 g[S_V(u)]\,du,$$

where $g : [0,1] \mapsto [0,1]$ is a distortion (increasing, non-negative) function, and $g[S_X(u)]$ represents the subjective probabilities.

The major difference between the Schmeidler model and the Yaari model lies in the interpretation (Quiggin, [24]). Yaari assumes that the objective distributions (e.g., S_X) are known and one applies a distortion (i.e., g) to the objective distribution. Schmeidler argues that it is illogical to assume an objective distribution; instead, he interprets $g \circ S_X$ as non-additive subjective probabilities which can be inferred from acts.

However, economic interpretations are important. For instance, if the underlying distribution is assumed to be known, then the process risks can be diversified away in a risk portfolio. Ambiguity is uncertainty regarding the best-estimate probability distribution, and thus may not be diversifiable in a risk portfolio.

The PH-transform fits in Schmeidler's economic theory with an interpretation of aversion to ambiguity (parameter risk).

5.4. No-Arbitrage Theory of Pricing

No-arbitrage is a fundamental principle in financial economic theory, which requires linearity of prices. The theories of Yaari and Schmeidler can be viewed as a more relaxed (or more general) version of the no-arbitrage theory, i.e., they only require no arbitrage (linearity) on co-monotone risks (e.g., different layers of the same risk). Using a market argument, Venter [28] discussed the no-arbitrage implications of reinsurance pricing. He observed that in order to ensure additivity when layering a risk, it is necessary to adjust the loss distribution so that layer premiums are expected losses under the adjusted loss distribution. Venter's observation is in agreement with the theories of Yaari and Schmeidler. In fact, the PH-transform is a specific transform which conforms to Venter's adjusted distribution method.

6. SUMMARY

In this paper we have introduced the basic methodologies of the PH-transform method and have shown by example how it can be used in insurance ratemaking. We did not discuss how to decide the overall level of contingency margin, which depends greatly on market conditions. An important avenue for future research is to link the overall level of risk load with the required surplus for supporting the written contract. Some pioneering work in this direction can be found in Kreps [11], [12] and Philbrick [22].

The use of adjusted/conservative life tables has long been practiced by life actuaries (see Venter, [28]). To casualty actuaries, the PH-transform method contributes a theoretically sound and practically plausible way to adjust the loss distributions.

REFERENCES

[1] Bault, Todd, Discussion of Feldblum: "Risk Loads for Insurers," *PCAS* LXXXII, pp. 78–96.

[2] Bowers, Newton L., Hans U. Gerber, James C. Hickman, Donald A. Jones, and Cecil J. Nesbitt, *Actuarial Mathematics*, Itasca, IL: Society of Actuaries, 1986.

[3] Ellsberg, Daniel, "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 1961, pp. 643–669.

[4] Feldblum, Sholom, "Risk Loads for Insurers," *PCAS* LXXVII, 1990, pp. 160–195.

[5] Finger, Robert J., "Estimating Pure Premiums by Layer—An Approach," *PCAS* LXIII, 1976, pp. 34–52.

[6] Hogarth, Robin M. and Howard Kunreuther, "Price Insurance and Warranties: Ambiguity and Correlated Risks," *The Geneva Papers on Risk and Insurance Theory*, 17, 1, 1992, pp. 35–60.

[7] Heckman, Philip and Glenn Meyers, "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *PCAS* LXX, 1983, pp. 22–61.

[8] Hogg, Robert and Stuart Klugman, *Loss Distributions*, John Wiley, New York, 1984.

[9] Keynes, John M., *A Treatise on Probability*, London: Macmillian, 1921.

[10] Knight, Frank H., *Risk, Uncertainty and Profit*, Boston: Houghton Mifflin, 1921.

[11] Kreps, Rodney, "Reinsurer Risk Loads from Marginal Surplus Requirements," *PCAS* LXXVII, 1990, pp. 196–203.

[12] Kreps, Rodney, "Investment-Equivalent Reinsurance Pricing," *PCAS* LXXXV, 1998, pp. 101–131.

[13] Kunreuther, Howard, Robin Hogarth, and Jacqueline Meszaros, "Insurer Ambiguity and Market Failure," *Journal of Risk and Uncertainty*, 7, 1993, pp. 71–87.

[14] Lawless, Jerry F., *Statistical Models and Methods For Lifetime Data*, John Wiley & Sons, New York, 1982.

[15] McClenahan, Charles L., "Ratemaking," in *Foundations of Casualty Actuarial Science*, Chapter 2, Casualty Actuarial Society, New York, 1990.

[16] Meyers, Glenn G., "The Competitive Market Equilibrium Risk Load Formula for Increased Limits Ratemaking," *PCAS* LXXVIII, 1991, pp. 163–200.

[17] Meyers, Glenn G., "Managing the Catastrophe Risk," *Incorporating Risk Factors in Dynamic Financial Analysis*, Casualty Actuarial Society Discussion Paper Program, May 1995, pp. 111–150.

[18] Miccolis, Robert S., "On the Theory of Increased Limits and Excess-of-Loss Pricing," *PCAS* LXIV, 1977, pp. 27–59.

[19] Panjer, Harry H., "Recursive Evaluation of a Family of Compound Distributions," *ASTIN Bulletin*, 12, 1981, pp. 22–26.

[20] Patrick, S. Gary, "Reinsurance," in *Foundations of Casualty Actuarial Science*, Chapter 6, Casualty Actuarial Society, New York, 1990.

[21] Philbrick, Stephen, Discussion of Feldblum: "Risk Loads for Insurers," *PCAS* LXXVIII, 1991, pp. 56–63.

[22] Philbrick, Stephen W., "Accounting For Risk Margins," Casualty Actuarial Society *Forum*, 1, Spring 1994, pp. 4–87.

[23] Quiggin, John, "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 4, 1982, pp. 323–343.

[24] Quiggin, John, *Generalized Expected Utility Theory: The Rank-Dependent Model*, Kluwer Academic Publishers, Boston, 1993.

[25] Robbin, Ira, Discussion of Meyers: "The Competitive Market Equilibrium Risk Load Formula for Increased Limits Ratemaking," *PCAS* LXXIX, 1992, pp. 367–384.

[26] Rosenberg, Sheldon, Discussion of Miccolis: "On the Theory of Increased Limits and Excess-of-Loss Pricing," *PCAS* LXIV, 1977, pp. 60–73.

[27] Schmeidler, David, "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57, 1989, pp. 571–587.

[28] Venter, Gary G., "Premium Calculation Implications of Reinsurance Without Arbitrage," *ASTIN Bulletin*, 21, 1991, pp. 223–230.

[29] von Neumann, John and Oskar Morgenstern, *Theory of Games and Economic Behavior*, (2nd edition), Princeton, NJ: Princeton University Press, 1947.

[30] Wang, Shaun, "Insurance Pricing and Increased Limits Ratemaking by Proportional Hazards Transforms," *Insurance Mathematics and Economics*, 17, 1995, pp. 43–54.

[31] Wang, Shaun, "Premium Calculation by Transforming the Layer Premium Density," *ASTIN Bulletin*, 26, 1, 1996, pp. 71–92.

[32] Yaari, Menahem E., "The Dual Theory of Choice Under Risk," *Econometrica*, 55, 1987, pp. 95–115.

APPENDIX A: PROOFS OF SOME STATED RESULTS

THEOREM 1 *For any non-negative random variable X (discrete, continuous, or mixed), we have*

$$E[X] = \int_0^\infty S_X(u)\,du.$$

Proof For $x \geq 0$ it is true that

$$x = \int_0^\infty I(x > u)\,du,$$

where I is the indicator function (assuming values of 0 and 1). For a non-negative random variable it holds that

$$X = \int_0^\infty I(X > u)\,du.$$

By taking expectation on both sides of the equation one gets

$$E[X] = \int_0^\infty E[I(X > u)]\,du = \int_0^\infty S_X(u)\,du. \qquad \square$$

THEOREM 2 *For a small layer* $[a, a+h)$ *with* h *being a small positive number, we have*

- $D[X_{(a,a+h]}] \leq \sigma[X_{(a,a+h]}]$, *and*
- $\lim_{a\to\infty} D[X_{(a,a+h]}]/\sigma[X_{(a,a+h]}] = 1$.

Proof Let $p = S_X(u)$ be the probability of hitting the layer $[a, a+h)$. Note that $p \to 0$ as $u \to \infty$. The payment by the small layer $[a, a+h)$ has approximately a Bernoulli type distribution:

$$\Pr\{X_{(a,a+h]} = 0\} = 1 - p, \qquad \Pr\{X_{(a,a+h]} = h\} = p.$$

Thus, $\sigma(X_{(a,a+h]}) = \sqrt{p - p^2}h$ and $D[X_{(a,a+h]}] = (\sqrt{p} - p)h$. The two results come from the fact that $\sqrt{p} - p \leq \sqrt{p - p^2}$ for $0 \leq p \leq 1$ and

$$\lim_{p\to 0} \frac{\sqrt{p} - p}{\sqrt{p - p^2}} = 1. \qquad \square$$

APPENDIX B: AMBIGUITY AND PARAMETER RISK

Most insurance risks are characterized by the uncertainty about the estimate of the tail probabilities. This is often due to data sparsity for rare events (small tail probabilities), which in turn causes the estimates for tail probabilities to be unreliable.

To illustrate, assume that we have a finite sample of n observations from a class of identical insurance policies. The empirical estimate for the loss distribution is

$$\hat{S}(u) = \frac{\#\text{ of observations} > u}{n}, \qquad u \geq 0.$$

Let $S(u)$ represent the *true* underlying loss distribution, which is generally unknown and different from the empirical estimation $\hat{S}(u)$. From statistical estimation theory (e.g., Lawless [14, p. 402], Hogg and Klugman [8]), for some specified value of u, we can treat the quantity

$$\frac{\hat{S}(u) - S(u)}{\sigma[\hat{S}(u)]},$$

as having a standard normal distribution for large values of n, where

$$\sigma[\hat{S}(u)] \approx \frac{\sqrt{\hat{S}(u)[1-\hat{S}(u)]}}{\sqrt{n}}.$$

The η% *upper confidence limit (UCL)* for the true underlying distribution $S(u)$ can be approximated by

$$UCL(u) = \hat{S}(u) + \frac{q_\eta}{\sqrt{n}}\sqrt{\hat{S}(u)[1-\hat{S}(u)]},$$

where q_η is a quantile of the standard normal distribution: $\Pr\{N(0,1) \leq q_\eta\} = \eta$. Keeping n fixed and letting $t \to \infty$, the ratio of the UCL to the best-estimate $\hat{S}(u)$ is

$$\frac{UCL(u)}{\hat{S}(u)} = 1 + \frac{q_\eta}{\sqrt{n}}\sqrt{\frac{1-\hat{S}(u)}{\hat{S}(u)}} \to \infty,$$

which grows without bound as u increases.

As a means of dealing with ambiguity regarding the best-estimate, the PH-transform

$$\hat{S}_Y(u) = [\hat{S}_X(u)]^r, \qquad r \leq 1,$$

can be viewed as an upper confidence limit (UCL) for the best-estimate $\hat{S}_X(u)$. It automatically gives higher relative safety margins for the tail probabilities, and the ratio

$$\frac{[\hat{S}_X(u)]^r}{\hat{S}_X(u)} = [\hat{S}_X(u)]^{r-1} \to \infty, \qquad \text{as} \quad u \to \infty,$$

increases without bound to infinity.

DISCUSSION BY GARY G. VENTER

1. INTRODUCTION

Dr. Wang has provided a good case for the use of the mean of the PH-transform of a loss distribution as the risk-loaded premium. I would like to comment on several issues: 1) the need for consistency of the adjustment among contracts; 2) alternative transforms; 3) calibration; 4) the need for arbitrage free methods; 5) links to other premium loading methods; 6) minimum rates on line; and 7) connecting to Yaari and Schmeidler.

2. CONSISTENCY AMONG CONTRACTS

In Venter [5] I argued that the following aspects of a risk load formula are all equivalent:

1. It is arbitrage free.
2. It is strictly additive for both independent and correlated risks.
3. It can be calculated as the expected value of an adjusted probability distribution.

Note however that this does not guarantee a positive load over expected values. That depends on the adjusted distribution chosen and the cover being priced.

The PH-transform would be one choice for generating such an adjusted distribution. When considering various contracts and portfolios, however, the use of adjusted distributions produces additive and arbitrage-free pricing only if a single adjustment of probability is made for each event and this event probability is kept fixed when looking at the various contracts. That is, to avoid arbitrage, once a transformed probability has been selected for an event, that probability has to be used for the entire calculation.

This rule is violated, for example, in Wang's Table 7. He applies PH-transformed probabilities to the aggregate losses in the occurrence layers 400 xs 100 and 500 xs 500, and calculates the mean loss in each layer under those probabilities (6.384% and 1.408% respectively). Given these probabilities, the loss to the layer 900 xs 100 then has to be the sum, or 7.792%. This is regardless of independence—the means are always additive. But as Wang shows, the PH-transform of the 900 xs 100 layer using the same parameter gives a different mean, and so must imply different probabilities for the layer losses.

This reasoning may lead to inconsistencies any time the PH-transform is applied to aggregate losses, especially if any hypothetical contract has to be priced. This would suggest applying transformations separately to frequency and severity. In any case, for use in reinsurance pricing, it would be useful to have separate transforms of the frequency and severity distributions, so that the transformed distributions can be used in excess, proportional, and aggregate treaty pricing.

For this reason, and for ease of calculation, the PH-transform is probably best applied to the severity distribution only. Frequency can be adjusted more simply, perhaps by changing the parameters of the frequency distribution. The additivity that results from adjusting severity with a single transformation applied to the entire severity distribution is illustrated in Wang's Table 5.

The need for consistent adjustment of probabilities in order to produce additivity can also be illustrated using an example from Delbaen and Haezendonck [3]. One of the transforms they consider is the adjusted distribution

$$f^*(x) = f(x)[1 + h(x - \mathrm{E}(X))].$$

This gives the adjusted mean

$$\mathrm{E}^*(X) = \mathrm{E}(X) + h\,\mathrm{Var}(X).$$

Thus, they seem to have shown that a variance load is a form of adjusted probability. However, for this to be a true variance load, inconsistent probability adjustments are sometimes required. For instance, if Y is adjusted by the same method, the price for Y would be the adjusted mean

$$\mathrm{E}^*(Y) = \mathrm{E}(Y) + h\,\mathrm{Var}(Y).$$

These probability adjustments would then determine the adjusted mean for $X+Y$ to be

$$\mathrm{E}^*(X+Y) = \mathrm{E}(X+Y) + h\,\mathrm{Var}(X) + h\,\mathrm{Var}(Y),$$

as means are additive even for correlated variates. If X and Y are in fact correlated, this is not the same as the variance load

$$\mathrm{E}(X+Y) + h\,\mathrm{Var}(X+Y).$$

This load could be achieved as an adjusted mean for $X+Y$, but different probabilities for X and Y would be needed. Thus for an adjusted mean to always give the variance load, the transformed probabilities must change during the calculation. If the variable X is kept fixed when transforming other variables, this transformation becomes a covariance load that produces the variance load only for X itself.

Even variables that depend only on X do not always receive a variance loading under this transformation. For instance, a 2% quota share of X would have mean under f^* of

$$\mathrm{E}^*(0.02X) = 0.02[\mathrm{E}(X) + h\,\mathrm{Var}(X)],$$

but under a variance load it would have a price of

$$0.02\mathrm{E}(X) + h(0.02)^2\,\mathrm{Var}(X).$$

This is only 1/50th as much loading, so with a variance load a risk could be 100% ceded to 50 reinsurers and 98% of the profit retained by the cedant. With the adjusted probability method, the whole profit would go to the reinsurers.

3. AN ALTERNATIVE TRANSFORM

Buhlmann [2] suggested using the Esscher transform to calculate risk load. This transformation is

$$f^*(x) = e^{hx} f(x)/\mathrm{E}(e^{hX}).$$

More recently, Gerber and Shiu [4] suggest using the Esscher transform on $\ln X$. They show how lognormally distributed security prices are transformed into another lognormal distribution with just an adjustment to the μ parameter. The transformed μ parameter is $\mu + h\sigma^2$. If you then calibrate h to produce the current security price as the discounted transformed mean of the original security, they show that you recover the Black–Scholes option pricing formula as the adjusted expected present value of the option. This makes the Esscher transform on logs interesting, as in this case it ties in with known financial theory. It can sometimes be calculated more simply as the equivalent power transform of the original distribution $f^*(x) = x^r f(x)/\mathrm{E}(X^r)$.

However, for Poisson variables, they show that the Esscher transform of the original variable (not the log) agrees with option prices for jump processes. Here the transform is also Poisson. A reasonable approach for compound frequency/long-tailed severity processes may be to apply the Esscher transform to frequency and to the log of severity.

4. CALIBRATION

Both the PH and Esscher transforms have a free constant that determines the level of the risk load. The value of this constant will depend on market conditions. However, if the price for basic primary coverage is known, it can be used to determine the value of the constant, which then can be used to price different reinsurance or excess layers for the same risk. This is basically the approach Black and Scholes use to price derivative contracts based on the pricing of the underlying asset.

As an example, suppose a portfolio of commercial property risks has a disappearing deductible of 1 (in appropriate units), and severity distribution $F(x) = 1 - x^{-a}$; i.e., the simple Pareto, for $x > 1$. The distribution for $\ln X$ is the exponential distribution $G(x) = 1 - e^{-ax}$. For this, the moment generating function is

$$M(h) = \mathrm{E}(e^{hx}) = a/(a-h).$$

Then the Esscher transform on G is

$$g^*(x) = e^{hx} g(x)(a-h)/a.$$

Since $g(x) = ae^{-ax}$, this shows that

$$g^*(x) = (a-h)e^{-(a-h)x},$$

which is another exponential distribution. Thus $F^*(x) = 1 - x^{-(a-h)}$ is the transformed severity distribution. But this distribution can also be reached as a PH-transform of the original Pareto. In this case, then, the log-Esscher and PH-transform can be calibrated to give the same distribution.

To determine the constant h, consider the loading for the portfolio. The mean for the simple Pareto is $a/(a-1)$. Suppose $a = 2$, and a loading of 10% of premium is built into primary pricing. Assume no loading is made for frequency. Then $\mathrm{E}(X) = 2$, and h is needed so that $(a-h)/(a-h-1) = 2.2$, or $(2-h)/(1-h) = 2.2$. This gives $h = 1/6$. Thus the Pareto with $a = 1.833$ gives the primary price as its mean, and so can be used for consistent pricing of reinsurance layers. The corresponding r for the PH-transform would be $r = 11/12 = 0.9166$, as $(x^{-2})^{0.9166} = x^{-1.833}$. Note that the transformed distribution is zero below $x = 1$, as is the original. This is an important detail omitted from the above discussion: in order to avoid arbitrage the transformed distribution must give zero probability to the same events as does the original distribution.

A problem with this calculation is that a loading should be made for frequency, as the frequency risk will not be carried for free. Reinsurance contracts for aggregate losses or for aggregate coverage on small per-occurrence layers would carry substantial frequency risk, which would have to be priced. The above calculation of the severity r or h can be done for any assumed frequency load. Rather than no load, another assumption might be that the frequency and severity loads are the same, or are in some pre-selected proportion. Thus, there is still some judgment involved in trying to find the adjusted probabilities that support the underlying pricing.

The PH and log-Esscher transforms are not always the same. For instance, for the inverse Weibull distribution

$$F(x) = \exp(-(\theta/x)^{\tau}),$$

the PH-transform is

$$1 - [1 - \exp(-(\theta/x)^{\tau})^{r},$$

whereas the log-Esscher transform is inverse transformed gamma. These will tend to be similar in typical cases, however, depending on the calibration. This suggests that the choice between PH and log-Esscher transforms for severity may be largely a matter of ease of calculation. For example, the log-Esscher transform of the lognormal is lognormal, whereas the PH-transform is more complicated.

5. SHOULD ACTUARIES ALWAYS USE ARBITRAGE FREE METHODS?

Insurers may want to try to build arbitrage possibilities into prices; e.g., by using variance loadings. If these loadings succeed in the market, this might give the insurers arbitrage opportunities. Exploiting arbitrage opportunities is usually regarded as producing an improvement in market functioning, as it tends to compete away those opportunities. However, this policy would need to be monitored carefully, as the products with arbitrage might eventually lose out to other competitors, resulting in a loss of market share.

Some apparent arbitrage in the reinsurance market may not actually be such. For instance, a strongly capitalized insurer may cede a portion of its risk to a group of weaker reinsurers for less than it received in premium. But the insurer is bearing the insolvency risk of those reinsurers, while its customers are not.

6. LINKS TO OTHER RISK LOAD METHODS

I was interested in Dr. Wang's comment that while the PH-transform aims to give the market price, different insurers may want different prices, depending for instance on their current risk portfolios. While he does not discuss how an insurer's desired price would be calculated, a logical approach might be to try to quantify the degree of risk assumed, and look to market norms or carrier goals to determine what the return should be for carrying that risk. This in general terms is how some other risk load methodologies proceed; e.g., the papers of Kreps and Meyers that Dr. Wang's paper cites. Thus, those approaches could be considered to be aiming at an insurer's desired risk load, while the transformed distributions are looking for the market price.

7. MINIMUM RATES ON LINE

Wang shows that a mixture of PH-transforms can produce minimum rates on line as market prices for low-expected-value risks. For individual reinsurers, minimum rates on line may make sense due to capacity constraints. Minimums are problematic from a market theory viewpoint, however. If a group of risks with very low expected losses were each written at a market minimum rate on line, they could conceivably be packaged and ceded as a group at that same rate on line, generating an arbitrage profit. It could be that market minimums exist as barriers but not as actual prices, as they may serve only to stop purchases when the expected loss is sufficiently less than the rate. If so, the above problem could not arise, as two or more minimum rate risks would cost more than the minimum when combined as a group.

Even without the mixture, a single PH or Esscher transform can produce layer rates that decline very slowly as the retention increases, and the risk load as a percentage of expected losses can increase without limit. Although not giving a true minimum rate on line, this could approximate market behavior fairly reasonably.

8. CONNECTING TO YAARI AND SCHMEIDLER

Both of these approaches advocate pricing by the mean of a distortion of the ddf, denoted by Wang as $g[S_V(u)]$. This makes it look like the distorted probability should act directly on the ddf. However, any distortion of probability can be re-expressed to state its effect on the ddf. For instance, consider a scale transform of a Pareto with original ddf of $(1+u/b)^{-a}$ and transformed ddf of $(1+u/c)^{-a}$. If $g(x) = [1 + b(x^{-1/a} - 1)/c]^{-a}$, then it will produce the transformed ddf from the original ddf.

Perhaps surprisingly, an increased scale parameter does not always produce a positive loading. An example where it does not is attributed to Thomas Mack in Albrecht [1]. It turns out that if you have a disappearing deductible, and want a separate cover to buy that back to full coverage, the buyback of the deductible can be cheaper for a higher scale parameter. For instance, suppose severity is Pareto with $S(x) = (1+x/b)^{-a}$, and you want to buy a cover that pays the full loss X up to $X = c$, and nothing above c. The expected loss for this cover is the expected loss limited to c, $\mathrm{E}(X \mid c)$, less $cS(c)$. From

$$\mathrm{E}(X \mid c) = [1 - (1+c/b)^{1-a}]b/(a-1),$$

and taking $a = 2$, this simplifies to

$$bc^2/(b+c)^2.$$

This can be seen to be a decreasing function of b for $c < b$, so in that case an increased scale parameter would give an adjusted mean less than the actual mean. The scale transform is pushing probability up to higher loss levels, so there is less below the deductible.

Unfortunately, this can happen with the PH and log-Esscher transforms also. The PH-transform of a Pareto lowers the a parameter. For $b = 1$ the deductible expected value is

$$[1 - (1 + da)(1 + d)^{-a}]/(a - 1),$$

which is an increasing function of a in some ranges. Thus lowering a will lower the expected value, making the adjusted mean less than the true mean. The log-Esscher transform in h of the Pareto is the generalized Pareto distribution with parameters $a - h, b, h + 1$. The deductible mean under this is also sometimes less than the non-transformed mean.

Even Delbaen and Haezendonck's loading of a portion of the covariance of the deductible with full coverage would have the same problem, as when c is low enough, the deductible is negatively correlated with total losses due to it being zero for larger loss amounts. For all of these approaches the full coverage contract would have less risk load than coverage excess of a small disappearing deductible.

This may be appropriate when the losses under the deductible are negatively correlated with total losses. When the variance of total losses is less than that of the losses excess of the disappearing deductible, even a traditional standard deviation or variance loading will price a buyback of the deductible for an excess policyholder at less than the expected losses of the additional coverage. For instance, consider the above Pareto with $a = 2.5$ and $b = c = 1$. The variance of a full coverage loss is 2.22, while a loss excess of the deductible has variance 2.36. If losses are loaded by $h \times$ variance, an excess policyholder can buy back to full coverage for the expected value of the deductible less $0.14h$. In any case, the loading method chosen will always have to be checked for its practical application to the problem at hand.

9. CONCLUSION

The PH-transform appears to be a useful way to build in risk load, especially for severity. The power transform (i.e., the log of the Esscher transform) should also be considered, especially when it is easier to calculate. For frequency, the Esscher transform can be applied, which in practice will often just result in a change in the frequency parameters.

REFERENCES

[1] Albrecht, P., "A Note on a Contribution by G. Venter," *ASTIN Bulletin* 22, 1992, pp. 247–254.

[2] Buhlmann, H., "An Economic Premium Principle," *ASTIN Bulletin* 11, 1980, pp. 52–60.

[3] Delbaen, F., and J. Haezendonck, "A Martingale Approach to Premium Calculation Principles in an Arbitrage Free Market," *Insurance Mathematics and Economics* 8.4, December 1989, pp. 269–277.

[4] Gerber, H., and E. Shiu, "Martingale Approach to Pricing Perpetual American Options," *ASTIN Bulletin* 24.2, 1994, pp. 195–220.

[5] Venter, G., "Premium Calculation Implications of Reinsurance without Arbitrage," *ASTIN Bulletin* 21.2, 1991, pp. 223–230.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXXIII

THE COMPLEMENT OF CREDIBILITY

JOSEPH A. BOOR

DISCUSSION BY SHOLOM FELDBLUM

> It is the concept of credibility that has been the casualty actuaries' most important and most enduring contribution to casualty actuarial science.
>
> —Matthew Rodermund

1. INTRODUCTION

Credibility theory is the crown jewel of casualty actuarial science. The statistician measures the significance of empirical findings, and the businessman uses judgment to select among diverse recommendations. Credibility theory meshes these two traditions, enabling us to combine varied indications based upon the relative predictive power of each of them.

In the quotation above, Rodermund speaks of credibility theory itself. Joseph Boor reminds us that the determination of credibility is only half of the pricing actuary's task. The other half, of no less importance, is to choose the figure that receives the complement of credibility.

Boor focuses on the practical track. There are dozens of fine actuarial papers on the theory of credibility, not all of which are easily implemented by the practicing actuary. Boor's paper, in contrast, was written first for the actuarial student, as a study note for the CAS ratemaking examination. It is equally valuable for the experienced actuary, keeping our gaze focused on the practicalities of insurance pricing.

2. BAYESIAN VS. CLASSICAL CREDIBILITY

It is instructive to contrast Bayesian and classical credibility procedures in the light of Boor's paper. Classical credibility assigns a credibility value to the experience data based upon its predictive power: that is, based upon the probability that the indication derived from the historical experience will be relatively close to the true expectations (see Longley-Cook [5]). The term "indication" here refers to the claim frequency, the pure premium, the loss ratio, or any similar ratemaking figure. Following Boor's illustrations, this discussion also uses examples of pure premiums.

In the classical tradition, the credibility assigned to the experience data is independent of the figure that is accorded the complement of credibility. Indeed the qualities of the figure that is assigned the complement of credibility are not relevant for determining the credibility that it receives, and they are therefore not relevant for determining the final credibility-weighted indication.

For instance, traditional automobile liability ratemaking procedures may assign a credibility value to the experience loss ratio or the experience pure premium based upon the number of claims in the experience period, using a full credibility standard of 1,024 claims.[1] For experience with fewer than 1,024 claims, the credibility assigned to the experience data equals $\sqrt{N/1024}$, where N is the number of claims in the experience.[2] If the experience contains 164 claims, then the credibility formula requires

[1]The rationale for this approach stems from a normal approximation to a Poisson claim distribution, whereby 1,024 claims in the experience period with no subsequent changes in the parameters of the Poisson distribution provides a 95% confidence interval that the true claim frequency is within ±5% of the indicated claim frequency (Stern [8]; Longley-Cook [5]).

[2]The rationale for this partial credibility formula is that the variance of the experience indications varies with the square root of the volume of the experience. It is difficult to construct a partial credibility rule from this rationale, since classical credibility does not consider the predictive accuracy of the information which receives the complement of credibility.

that we assign 60% weight to some other information, such as the current pure premium, or the pure premium from some other classes, some other territories, or some other time periods. Boor asks: "How should we choose this other information? In particular, what characteristics should this other information have?"

We will come back to this question in a moment, to see whether it is indeed well formulated. Let us first consider the workings of classical credibility. Most fundamentally, the attributes of the information that receives the complement of credibility do not affect the amount of credibility to be assigned either to this data or to the experience data. Whether we use the current pure premium, or the pure premium from some other classes, some other territories, or some other time periods makes no difference. The experience data still receives 40% credibility, and the other information receives the remaining 60% credibility.

In contrast, Bayesian credibility procedures do not speak about the predictive power of the experience data. In Bayesian credibility, the experience data in one ratemaking scenario may be sparse and volatile, but they will be assigned high credibility simply because we have nothing else that is more accurate. In another ratemaking scenario, the experience data may be voluminous and steady, but they may be assigned a lower credibility because we have other, equally good information.

If Boor's paper is important for classical credibility, then it is doubly important for Bayesian credibility. In classical credibility, the determination of credibility comes first. Before we have chosen the data to be assigned the complement of credibility, we know the amount of credibility that will be assigned to the experience data. A proper choice of the data to be assigned the complement of credibility improves the ratemaking indication only because it improves the part of the indication stemming from the complement of credibility.

In Bayesian credibility, we do not know the amount of credibility to be assigned the experience data until we know the type of information to which the complement of credibility will be assigned. The attributes of this information, such as its predictive power and its freedom from bias, affect the weight that we assign both to this data and to the experience data.

Many actuaries conceive of classical credibility and Bayesian credibility as two points along a continuum. Actuarial ratemaking seeks to produce the most accurate indications possible. Classical credibility was an early attempt to achieve this, and it remains the most practical technique in most ratemaking environments. Bayesian credibility is a more refined method of achieving the same end. To summarize: Bayesian credibility is a statistically justified procedure for optimizing the accuracy of the rate indications. Classical credibility is an early, less sophisticated attempt to do the same.

3. LEAST FLUCTUATION CREDIBILITY AND GREATEST ACCURACY CREDIBILITY

Gary Venter [9] has put forth an alternative perspective. The aim of classical credibility is not solely the achievement of accurate rate indications. Rather, the aim of classical credibility is to limit the fluctuation of rate levels from year to year, unless there is good statistical justification for a change. As Venter says [9, pp. 383, 384]:

> The basic philosophical difference between these methods is as follows. The limited fluctuation approach aims to limit the effect that random fluctuations in the data can have on the estimate; the greatest accuracy approach attempts to make the estimation errors as small as possible. The most well developed approach to greatest accuracy credibility is *least squares credibility*, which seeks to minimize the expected value of the square of the estimation error. The term "classical credibility" has sometimes been used in North America to denote limited fluctuation credibility …

Let us clarify this distinction with an illustration. Suppose that we are making rates for a new coverage, and we have no historical experience and no prior expectations upon which to base the rate. It is a stated amount coverage for $1,000, with a maximum of one claim per policy period, so we know that the pure premium must be between $0 and $1,000 (depending on the claim frequency).[3]

Since there is no statistical information or prior expectations for this coverage, let us suppose that the regulator chooses a random number between $0 and $1,000 as the pure premium for the first five years. The random number is $670. We add risk load, expenses, and profit margins to set the rate.

After the five years go by, we have some historical experience, which indicates the pure premium should be $245. However, the historical experience is sparse, and the true expected pure premium may be much different. We want to use this experience and credibility theory to set the pure premium for the next five years.

Assume that we have no prior expectation and no external information. The only information we have is the $245 historical pure premium. Under Bayesian credibility theory, the historical experience gets full credibility, and our best estimate for the expected pure premium is $245.[4]

Classical credibility is designed to avoid undue fluctuations in the rates. Both the customer and the regulator are accustomed to a pure premium of $670. Because the data are sparse, the rate level will change significantly if we rely on Bayesian credibility. Classical credibility will change the rate only to the extent that we have "credible" historical experience for doing so.

[3]For example, consider term life insurance coverage on an insured whose mortality expectation is entirely unknown.

[4]We may want to set a higher risk load, since there is a greater chance that we will lose money with lower rates. However, the risk load is distinct from the expected pure premium.

Suppose that we determine that the classical credibility is 40%, based on the statistical parameters which we choose, such as the size of the confidence interval. The new pure premium is $40\% \times \$245 + 60\% \times \$670 = \$500$.

With Bayesian credibility, the indicated pure premium moves as rapidly as possible towards the true expected pure premium, though the sequence of pure premium changes may have wide swings. With classical credibility, the indication moves less rapidly towards the true expected pure premium, but the sequence of pure premium changes has fewer and narrower swings. Bayesian credibility emphasizes accuracy; classical credibility emphasizes stability.

This example is admittedly extreme. In general, we have *a priori* expectations for the pure premium, and the previously charged rate is rarely so different from the experience data. However, this distinction between classical and Bayesian credibility is true for any ratemaking scenario, though the differences between the two methods will rarely be as great.

4. THE COMPLEMENT OF CREDIBILITY: CLASSICAL APPROACH

The implications for Boor's thesis are important, assuming that we interpret the distinction between classical and Bayesian credibility in the manner proposed by Venter. For classical credibility, the primary concern is limiting the fluctuations in rates. If so, the choice of the information which should be accorded the complement of credibility is clear. It is the current rate, adjusted (if necessary) for factors unrelated to potentially random fluctuations in experience.

Since this is the essence of classical credibility, it warrants further explanation. Suppose that

- the underlying pure premium is \$100 per exposure,
- there is no monetary inflation affecting claims costs (i.e., no loss cost trend),

- there is no expected change in claim frequency, and
- there are no changes in the compensation system that might affect loss costs.

Random loss occurrences, however, affect the experience pure premiums. Sometimes the experience data indicate a $90 pure premium, and sometimes they indicate a $110 pure premium. We don't know whether the true expected pure premium per exposure is $90, $100, $110, or some other amount. Unless there is good actuarial justification for doing so, the insurance company will not change the underlying pure premium. To the extent that the experience is "credible," the company will indeed change the rate to bring it more in line with the historical experience.

Without reliable experience indications, the company and the regulator are reluctant to change rates because (i) the public has come to expect a $100 pure premium and (ii) there is no good actuarial justification for changing the rates. There may be some external factor affecting the expected pure premium. In that case, the pricing actuary aims to reflect that factor in the price change. For instance, if there is 10% monetary inflation affecting loss costs (i.e., the loss cost trend is +10%), then the company, the regulator, and the public expect a 10% increase in premium rates if we have no other information. This is the rationale for credibility weighting the experience pure premium (or experience loss ratio) not with the underlying pure premium (or the expected loss ratio) but with the *trended* underlying pure premium (or the trended expected loss ratio).

The same is true for any other external change affecting the expected loss costs, such as changes in the expected claim frequency, or changes in the insurance compensation system. In practice, these factors affect both the experience data and the underlying pure premium (or the expected loss ratio). For instance, if there is a non-zero loss cost trend, the trend factor is applied both to the experience data and to the underlying pure premium.

5. THE COMPLEMENT OF CREDIBILITY: IN PRACTICE

The central question of Boor's paper is "What are the desirable characteristics of the information to be assigned the complement of credibility?" For classical credibility, we have answered this question. If the goal is limited fluctuation, then the information to be assigned the complement of credibility should be the current rate (or the current pure premium, or the expected loss ratio), adjusted for all factors other than the uncertainty inherent in the insurance process.

This is indeed what is done in most primary lines of business. In automobile liability, for instance, the experience loss ratio is credibility weighted with the expected loss ratio, adjusted (if necessary) for loss cost trends. Similarly, the indicated territorial or classification rate relativity is credibility weighted with the current territorial or classification rate relativity.

Boor's illustration of Harwayne's method of determining workers compensation pure premiums is particularly instructive, since it demonstrates numerous aspects of good ratemaking technique. In Harwayne's method, there are three components given credibility weights in calculating the pure premium (see Harwayne [4]):

1. the indicated pure premium,
2. the national pure premium, and
3. the underlying pure premium.

The second component of the formula, the national pure premium, reflects greatest accuracy credibility. The third component of the formula, the underlying pure premium, reflects limited fluctuation credibility. We discuss in Section 9 the rationale for this rate making procedure, as well as the adjustments made to the national pure premium for the greatest accuracy component. For now, let it suffice to say that the underlying pure premium is adjusted for all external influences, as described above.

General liability has a more complex procedure for combining the experience pure premium (or loss ratio) with the underlying pure premium (or expected loss ratio). This procedure, termed the "C-factor" by Graves and Castillo [3], was introduced by the Insurance Services Office in the 1980s, and it is illustrated in Boor's paper.[5]

The procedure looks like limited fluctuation credibility, but its rationale is different. The desirable characteristics of the information receiving the complement of credibility follow directly from the rationale of this credibility procedure.

Boor's central concern is to determine the desirable characteristics of the information receiving the complement of credibility. Boor lists these characteristics at the beginning of his paper. He then discusses several commonly used credibility procedures, and he discusses how well each one measures up to these characteristics.

The primary purpose of this discussion is to show how the desirable characteristics of the information receiving the complement of credibility follow from the rationale of the credibility procedure. We do this for each separate use of credibility: limited fluctuation, proxy for past experience, greatest accuracy, and marketplace pricing. The results sometimes agree with the conclusion in Boor's paper, and sometimes they expand on them. Keep in mind this theme: if we wish to determine the desirable characteristics of the information receiving the complement of credibility, we must know why we are using credibility in the first place.

Let us illustrate the C-factor procedure, so that its workings are clear. We then differentiate it from limited fluctuation cred-

[5]Graves and Castillo, who use a loss ratio ratemaking procedure, credibility weight the experience loss ratio with the trended and adjusted expected loss ratio. Boor uses the same procedure, though he credibility weights the pure premium with a trended and adjusted underlying pure premium.

ibility, and we re-examine Boor's central question: "What are the desirable characteristics of the information that receives the complement of credibility?"

For the sake of clarity, let us simplify the illustration by ignoring the time lags that are needed for data collection and rate filings. Suppose that we are making rates for a policy to be effective on January 1, 1999, using experience from accident year 1998. (For simplicity, we are making rates for a single policy, not for a policy year. Were we making rates for a policy year, we would have an additional half year of trend in the mathematics below.) The current pure premium is $100 per exposure unit, which was filed and became effective on January 1, 1998. Using accident year 1998 experience, the developed pure premium, trended to the average effective date under the anticipated rates, is $135 per exposure unit. The loss cost trend is +10% per annum. The credibility to be assigned to the experience pure premium is 60%, based upon classical credibility procedures. What is the credibility weighted pure premium for the rate filing?

Limited fluctuation credibility says the following: the public and the regulator have seen a pure premium of $100 per exposure unit in 1998. Loss costs are increasing by +10% per annum, so they expect a pure premium of $110 per exposure unit in 1999. We are willing to change the pure premium to conform with our experience only to the extent that this experience is "credible" (regardless of the quality of other information). The classical credibility is 60%, so the credibility weighted pure premium is

$$60\% \times \$135 + (1 - 60\%) \times \$110 = \$125.$$

6. EXPONENTIAL SMOOTHING AND ACTUARIAL SHORT-CUTS

The formula above is correct, if our goal is limited fluctuation in rate levels. But limited fluctuation is not the only rationale for classical credibility. Let us change the interpretation of the credibility procedure; this in turn changes the appropriate formula.

The new pricing actuary asks: "How many years of data should one use for ratemaking?" The general answer is straightforward, though the specific parameters vary from case to case:

1. One should use as much experience as available, as long as it relates to the type of coverage presently being offered.
2. One should assign higher weight to the more recent experience, since it is likely to be a better predictor of future experience.[6]
3. The additional benefit of maintaining, trending, and adjusting older years of data declines rapidly, and this benefit is soon outweighed by the cost of this work. Actuarial short-cuts can improve the efficiency of the ratemaking process.

This actuarial short-cut is another use of credibility. Let us resume with the previous illustration. We are making rates for a policy to be issued on January 1, 1999, using data from accident year 1998, and assigning 60% credibility to the experience. For simplicity, let us assume that we have always assigned 60% credibility to the experience when making rates for this coverage.

Let PP_t be the pure premium charged in year t, and let EX_t be the pure premium indicated by the experience in year t. The pure premium charged in 1999 is

$$PP_{99} = 60\% \times EX_{98} + 40\% \times PP_{98}.$$

Assuming that the same 60% credibility value was used in the past, we substitute for PP_{98} to give

$$PP_{98} = 60\% \times EX_{97} + 40\% \times PP_{97}.$$

[6]This statement is more applicable for rapidly developing experience. Mahler [5] uses an illustration from baseball "win-loss" statistics, where there is no development. When significant and especially volatile development is expected, as in casualty excess-of-loss reinsurance, some actuaries are inclined to rely more heavily on older, more mature years of experience; compare Cook [2].

We combine the two equations to get

$$PP_{99} = 60\% \times EX_{98} + 40\% \times [60\% \times EX_{97} + 40\% \times PP_{97}].$$

We continue this substitution process to express the indicated pure premium for 1999 as a function of the experience pure premiums in all previous years for which we have data. If the credibility each year is Z, then the indicated pure premium for year k equals

$$PP_k = Z \times \sum_{t=1}^{\infty} \{(1-Z)^{t-1} \times EX_{k-t}\} \qquad (5.1)$$

where the summation runs over all preceding years for which data are available ($t = 1,2,3$, etc.).[7]

This equation says that the indicated pure premium for year k is a weighted average of the experience pure premiums in each preceding year, where the weights form a decreasing exponential series. Intuitively, this makes sense. All experience provides some information useful for determining the new rates, but the older the experience is, the less useful it is.[8]

There are three problems with using the general equation for PP_k given above:

1. To use the general equation, we must retain all past experience, and we must re-analyze it each year. This can be a cumbersome task, and the costs might outweigh the benefits.

---

[7]See Mahler [6], pp. 255–256. If there are w years of data available, then the sum of the coefficients in Equation (5.1) equals $1-(1-Z)^w$. Thus, in theory, all the coefficients should be multiplied by $1/[1-(1-Z)^w]$. In practice, this is about the same as using a slightly higher Z value. Since values of Z within a fairly broad range work about equally well, an attempt to optimize the value of Z by the correction noted here would not be cost-efficient. See Mahler [6], pp. 256–257, on the relative efficiency of different Z values.

[8]For a more complete exposition of this rationale for the ratemaking credibility procedure, see Mahler [6]. Different rationales for the credibility procedure lead to different credibility formulas. Mahler extends the analysis by showing how the covariance structure for the risk parameters affects the optimal credibility to be assigned to the experience.

2. The general equation used here assumes that the value of Z remains the same from year to year. In fact, the value of Z may change from year to year, particularly if the volume of business is changing from year to year. If it does, the mathematics become much more complex.

3. The ratemaking process cannot be reduced to a rote formula. Every rate review requires the careful judgment of the pricing actuary to discern anomalies in the data, shifts in the external environment or in the company's operations that might affect the anticipated loss costs, and changes in compensation systems or consumer behavior that might affect the company's ultimate claim payments. Each year the pricing actuary may subjectively adjust the experience indication up or down based upon an analysis of the data and of the insurance environment. The general equation would require us to somehow retain all these adjustment factors.

The last problem listed above is critical. We do want to use all the past experience, but we also want to use the actuarial judgment of the ratemakers who analyzed this past experience.

The Credibility "Short-Cut"

The solution to all three problems is the same, as is clear from the derivation of the formula above. The underlying pure premium serves as a proxy for the experience of all prior years. The historical experience itself need not be retained by the company or re-analyzed each year by the pricing actuary. Credibility weights may have varied from year to year, and at each filing the pricing actuary may have adjusted the indications. The effects of all these factors are retained in the underlying pure premium.

Let us examine the rationale of this credibility formula in order to address Boor's fundamental question. We are credibility weighting the experience pure premium with the current pure premium, so we are tempted to think of limiting the fluctuation in rates. But the current pure premium is used here as a proxy for the experience pure premiums from past years. We use all this past experience in order to produce the most accurate indication. Our goal is greatest accuracy, not limited fluctuation.

Proxy Problems

The information properly assigned the complement of credibility is the underlying pure premium—as long as the underlying pure premium is indeed an accurate proxy for the indicated pure premiums from past years. If this is not true—that is, if it is not an accurate proxy—then a different complement is required.

There are two ways in which the current (underlying) pure premium may not be an accurate proxy for the indicated pure premiums from past years:

1. The pricing actuary, when reviewing the experience from past years, judgmentally adjusted the data and erred in doing so. In this illustration, we have implicitly assumed that there were no errors; we trust the previous pricing actuary's judgment. The current pure premium is the best proxy for the indicated pure premium from past experience years, after adjustment for data outliers, system changes, operational changes, and so forth.

2. The pricing actuary, after reviewing the experience from past years, filed one pure premium, but the state insurance department approved only part of the rate request. If we trust the judgment of the pricing actuary, we would use the filed pure premium, not the approved one. We are assuming that the state insurance department's actions were motivated by non-actuarial concerns, such as a political desire not to raise rates more than a certain amount.

Let us consider a numerical example for the second proxy problem. Suppose that the 1998 pure premium is $100 per exposure, and there is a +10% annual loss cost trend. Based upon the 1998 experience, the indicated pure premium for the 1999 policy is $135. The credibility to be assigned to the experience is 60%. As demonstrated above, if our goal is limited fluctuation in rate levels, then the credibility weighted indicated pure premium is

$$60\% \times \$135 + (1 - 60\%) \times \$110 = \$125.$$

If, instead, we are using the current pure premium as a proxy for the indicated pure premium based on past experience years, then we must know the filed and approved pure premiums for the 1998 policy. Suppose that the pricing actuary had filed for a +50% rate increase, but the insurance department had granted only a +25% rate increase.

These figures tell us that the pure premium for the 1997 policy was $\$100 \div 1.25$, or $80. The indicated pure premium for the 1998 policy based on the 1997 experience was $\$80 \times 1.50$, or $120. If we want to use the current pure premium as a proxy for the indicated pure premium based on past experience, we must assign the complement of the credibility to the $120 adjusted for trend and calculate the credibility weighted indicated pure premium as

$$60\% \times \$135 + (1 - 60\%) \times \$120 \times 1.10 = \$133.80.$$

7. BAYESIAN CREDIBILITY

For applications of classical credibility, whether as limited fluctuation credibility or as actuarial short-hand for older experience, the information that receives the complement of credibility is determined by the purpose of the credibility procedure. So where do Boor's six characteristics come into play?

Venter describes Bayesian credibility as greatest accuracy credibility. If the rationale of the credibility procedure is to improve the accuracy of our indications, then characteristics such as predictive power, independence, and freedom from bias seem natural.

Nevertheless, a careful analysis leads to less firm conclusions. So let us tread gingerly over this terrain, beginning with the rationale for Bayesian credibility.[9]

There are some immediate problems with the question of desirable characteristics for the complement of credibility. In the Bayesian view, there is no qualitative distinction between the information that receives the credibility and the information that receives the complement of credibility. Pricing actuaries tend to think of the experience data as being assigned the credibility and some other data as being assigned the complement of credibility. Those of us who are steeped in classical credibility tend to think of credibility as a function of the reliability or the predictive power of the experience data.[10] This may indeed reflect the thinking of most pricing actuaries, but it is not a Bayesian view.

---

[9]When Boor speaks of Bayesian credibility, he uses an illustration of territorial relativities. This may confuse some readers, since there are two independent dimensions:

- Classical credibility versus Bayesian credibility, and
- Credibility for statewide indications versus credibility for territorial relativities.

When making rates for territorial relativities, most actuaries use classical credibility techniques, not Bayesian credibility. Since the aim is often to limit fluctuation in territorial relativities, the figure that receives the complement of credibility is the current territorial relativity. For territorial ratemaking, this objective is sometimes explicitly stated. See, for instance, Conger [1] on the objectives of personal automobile territorial ratemaking in Massachusetts.

[10]The CAS Statements of Principles and the American Academy of Actuaries Standards of Practice show how deeply ingrained this perspective is in the actuarial community. Theoretical actuaries may have discarded classical credibility in favor of its Bayesian counterpart, and the CAS examination syllabus extols the elegance of the Bayesian-Bühlmann procedures. Yet the CAS Statements of Principles and the ASB Standards of Practice show no trace of the Bayesian influence. The "Credibility" paragraphs in the Statements of Principles begin "Credibility is a measure of the predictive value that the actuary attaches to a particular body of data" (Statement of Principles Regarding Property and Casualty Insurance Ratemaking, lines 88-89; compare also Statement of Principles Regarding Property and Casualty Loss and Loss Adjustment Expense Reserves, line 193). The Actuarial Standard of Practice #25, "Credibility Procedures Applicable to Accident and Health, Group Term Life, and Property/Casualty Coverages," even defines the "full credibility" standard as "the level at which the subject experience is assigned full predictive value based on a selected confidence interval." This definition, albeit incorrectly worded (see the correct wording given earlier in this discussion), is based entirely on classical credibility theory; there is no concept of a "full credibility standard" in Bayesian credibility theory.

In the Bayesian view, we have two or more sets of data, each of which tells us something about the number that we seek to estimate. We would like to use all of these data to develop our estimate. That is, we seek a weighted average of the various estimators. The Bayesian credibility procedure give us relative weights to assign to each set of data.

We cannot speak of the desirable qualities of the information that receives the complement of credibility as if this information were somehow different from our other ratemaking data. No set receives the credibility with some other set receiving the complement of credibility. It is only by convention that we speak of the experience data as receiving the credibility and of some other data receiving the complement of credibility. This convention come from classical credibility, not from Bayesian credibility.

8. BAYESIAN RATEMAKING

Perhaps we can rephrase Boor's question as: "What are the desirable characteristics of the data that receives some portion of the credibility?" This is, indeed, a proper question for the pricing actuary to ask, and Boor's six characteristics are a valid set of characteristics. But this question has nothing to do with credibility. It is a question about data quality: "What are the desirable characteristics of ratemaking data?"

Actually, five of Boor's characteristics can apply to data quality. One of the characteristics deals more specifically with credibility. Boor's five characteristics of good ratemaking data are:

1. accuracy as a predictor of next year's mean loss costs,
2. absence of bias as a predictor of next year's mean subject expected losses,
3. availability of data,
4. ease of computation, and
5. clarity of relationship to the subject loss costs.

If one wishes to use two or more sets of ratemaking data, and to combine them by means of a Bayesian credibility procedure, then the ratemaking procedure is enhanced to the degree that the two or more sets of data are relatively *independent* of each other.

But we start with the data sets. For optimal ratemaking, we should use all the data available. Suppose we have three sets of data, A, B, and C. Set A is the historical experience. It is the best data, and it is the most acceptable data for the state insurance department, so we surely want to use set A. Sets A and B are relatively independent of each other. Sets A and C are relatively dependent. Under Boor's thesis we should assign the credibility to set A, and assign the complement of credibility to set B, not to set C. This will optimize the ratemaking procedure.

At first glance, Bayesian credibility doesn't say this at all. Rather, Bayesian credibility says that we should use all three sets of data and assign the proper weights to each of them. There is no constraint limiting us to only two data sets. In fact, it is common practice to use three or more data sets in many ratemaking applications. Property ratemaking uses five years of data (mandated by statute in many jurisdictions), with different weights applied to each year. Most common is a 10%–15%–20%–25%–30% weighting, with the higher weights applied to the more recent years. In theory, the optimal weights may be determined from a Bayesian analysis, though the accuracy of the final indication may not depend that strongly on the weights chosen, as long as they are within a reasonable range (see Mahler [6]). The experience loss ratio is then credibility weighted with a permissible loss ratio. Indicated territorial relativities are credibility weighted with the current relativities. In sum, the new rates for a particular territory are a weighted average of the indications from five separate years of experience, the current statewide rates, the current territorial relativity, and the indicated territorial relativity. These weights are not chosen by a Bayesian analysis. Rather, classical credibility procedures along with *ad hoc* weighting schemes are used. But for classical credibility, as noted above, Boor's paper is irrelevant. If Bayesian analysis were used for all the weights, then Boor's thesis does become relevant. However, we do not choose the two data sets that are most independent. We choose weights to optimize the accuracy of the indication, given all the data that we have available.

Workers compensation ratemaking provides another good example. The formula pure premiums are derived from six sets of data:

A. Financial Data (all classifications)
   - (A.1) Calendar year experience
   - (A.2) Accident year experience
   - (A.3) Policy year experience

B. Unit Statistical Plan experience (by classification)
   - (B.1) Indicated partial pure premiums
   - (B.2) Underlying partial pure premiums
   - (B.3) National partial pure premiums.

In fact, the procedure is even more complex, since between the unit statistical plan classification experience and the financial data statewide experience there is class group experience (manufacturing, contracting, and all other). In workers compensation, just as with property, the credibility weights stem from the early days of actuarial ratemaking, before Bayesian analysis caught the fancy of pure actuaries. In theory, though, Bayesian analysis could be used here as well.

9. HARWAYNE'S PROCEDURE

Harwayne's procedure for a three-way credibility weighting of workers compensation partial pure premiums is one of the most illuminating of Boor's examples. Indeed, Harwayne's procedure is a wonderful example of actuarial practice. It was a critical advance in workers compensation ratemaking, and it has since been applied to other lines of business as well.

How does it relate to Boor's thesis? It is a wonderful example, but what exactly does it show?

There are various elements in Harwayne's procedure. Some relate to ratemaking in general, some relate to credibility considerations, and some relate to the characteristics of the data that receives the complement of credibility. We must separate these strands, so that we can focus on the last of these issues.

To appreciate its elegance, Harwayne's procedure must be viewed in the history of workers compensation ratemaking.[11] The procedure uses three sets of information:

- the statewide classification experience, giving an *indicated* pure premium,
- the current classification pure premium (the *underlying* pure premium), and
- the classification experience from other states, giving the *national* pure premium.

Class Plan Refinement

The first two sets of information are routinely used in actuarial ratemaking. For instance, when making personal automobile insurance rates for the state of New York, the pricing actuary uses the New York experience and the current New York rates (perhaps adjusted for trend and similar influences). The pricing actuary would not consider Massachusetts personal auto experience or Illinois personal auto experience or national personal auto experience. So why is workers compensation different? Why does it combine the statewide pure premium with the national pure premium?

In personal auto, the classification scheme is a well structured, multi-dimensional system. For any classification, there is

[11] Harwayne's procedure is summarized in Boor's paper. It is presented in detail in Harwayne [4], along with the justification for its use.

generally sufficient experience in New York to set credible rates. Moreover, the Massachusetts and Illinois automobile compensation systems are so different from the New York system, that the Massachusetts and Illinois data won't help much. New York has a no-fault compensation system with a strong verbal tort threshold. Massachusetts has a no-fault compensation system with a much abused monetary tort threshold, and Illinois has a tort compensation system.

The states also have different classification systems, and they have different statutory constraints on underwriting, such as those relating to gender-based differentiation. Finally, they have vastly different rate filing systems. New York has a prior approval system, Illinois has open competition, and Massachusetts has a state rating bureau using a mandated financial pricing model.

In sum, the states are incomparable: Massachusetts experience and Illinois experience are nearly impossible to convert to "New York type" experience.

Workers compensation is almost the exact opposite of the personal auto situation. Workers compensation has a simple, one-dimensional classification system. Each state has about six or seven hundred classes, many of which have relatively little experience in any one state. Moreover, the states generally use similar class definitions.[12] Finally, the workers compensation systems seemed to be of the same type (at least to the founding members of the CAS), though there were differences in the parameters by state. Medical benefits are unlimited, and indemnity benefits are generally paid as some percentage of the pre-injury wage.

State workers compensation benefits were introduced rather suddenly in the early years of the twentieth century. Before the introduction of workers compensation laws, workplace accidents were handled through the tort liability system, with injured em-

---

[12]This is particularly true in the NCCI states, and it is even true in states which have their own rating bureaus.

ployees suing employers for negligence. The applicable insurance was employers' liability coverage, not workers compensation. When first setting workers compensation rates for a state, pricing actuaries had no prior experience from that state. For the smaller and medium sized classes—which comprised the majority of the workers compensation classes—a dozen years might elapse before there would be credible experience from the state under consideration. So how might one begin a workers compensation pricing structure?

Reduction Factors

Massachusetts was one of the first states to initiate a workers compensation system. As other states began their own systems, pricing actuaries took the Massachusetts experience and converted it to the benefit levels of the other states. For instance, suppose that the Massachusetts statute provided benefits equal to two-thirds of the pre-injury wage, and that the statute of another state provided benefits equal to 60% of the pre-injury wage. To set initial rates for the other state, begin with the Massachusetts rates and multiply them by 90% (= 60% ÷ 66.7%).

This procedure is straightforward and logical, enabling the efficient development of a complete workers compensation pricing structure. The founding members of the CAS meticulously calculated all the required "reduction" factors to convert rates from one state system to a second state system, considering not just different compensation rates but also various maximum and minimum benefit limitations, durations of compensation for various types of disability, and variations in state statutes regarding dependency awards. The result, as embodied in some of the first *Proceedings* papers, was a truly elegant actuarial procedure.

Unfortunately, it did not work. The founding fathers of the CAS spent months of painstaking work determining reduction factors to convert workers compensation loss costs from one state to another, only to have their results empirically invalidated by the emerging experience. The rigorous analysis, for instance, may have said that State X's loss costs should be 25% greater than State Y's, but the emerging experience showed that they were 15% lower. These results were surprising, but they were not wrong. There are many differences between compensation systems that are difficult to quantify. The administrative procedures in one compensation system, for instance, may encourage attorney involvement in workers compensation claims, while those in another state may discourage attorney involvement. The effects on loss costs can be dramatic, but these effects are rarely amenable to actuarial quantification.

Actuaries live by numbers. If one could not quantify the appropriate reduction factors, how could one use the experience of other states in setting rates? The first generation of actuaries rushed to develop reduction factors and to use the experience of other states in setting rates. The next generation of actuaries, discouraged by the empirical discrepancies, were ready to abandon these techniques and to use the experience of each state in isolation.

The flaw with the original procedure was the attempt to quantify *a priori* the reduction factor from one state to another. To the early actuaries, this had seemed essential: how could one use Massachusetts experience for a certain classification to help set the New York classification rate unless one knew how the Massachusetts classification loss costs would appear under the New York compensation system?

Harwayne saw a solution to this problem. Indeed, there are no reduction factors at all in Harwayne's procedure, because there are too many powerful but invisible factors that affect loss costs. Rather, Harwayne's procedure assumes that these invisible factors affect all classifications equally. Massachusetts loss costs may be unusually high because of greater attorney involvement in workers compensation claims, greater claims consciousness among the populace, or any such unquantifiable factor. But if we can empirically quantify the *overall* effect, then we can use the Massachusetts experience to help set other states' *classification* rates.

To highlight the advance made by Harwayne, let us consider a simplified example. Suppose that we have classification rates for State A, which is a large state with credible experience in most classes. We need to set classification rates for State B, which is a small state, with sparse experience in many classes.

If we look at the benefit structures in these two states, we might say that State B loss costs will be 25% higher than those in State A. This conclusion is not really helpful, since there are so many factors than affect the relative loss costs in the two states. Rather, we look at the overall empirical loss costs per exposure in the two states. We might find that the State B loss costs are, on average, 15% lower than those in State A. Using this figure as the implicit reduction factor, we multiply each classification rate from State A by 85% to get the indicated State B classification rates.

Unfortunately, this doesn't work either. We need an "overall loss costs per exposure" for each state. But there is no such thing as an overall loss cost per exposure. Some classes are high-risk, and they have high loss costs per dollar of payroll. Other classes are low-risk, and they have low loss costs per dollar of payroll. Perhaps State A has more high-risk classes and State B has more low-risk classes.

Think of the problem in the following fashion. We would like to derive the average loss cost per dollar of payroll. But the exposure base is not dollars of payroll. The exposure differs for each class: it is dollars of blacksmith payroll in the blacksmith class, dollars of carpentry payroll in the carpenters class, and so forth. One can not add blacksmith payroll to carpentry payroll.

Harwayne's solution was to translate every state to the same classification mix. Suppose that State A has two blacksmiths for each carpenter, and State B has two carpenters for each blacksmith. Harwayne's procedure calculates the overall loss costs per exposure in each state by taking 2/3 of that state's carpentry pure premium and 1/3 of that state's blacksmith pure premium. This puts the experience of both states on the same classification mix basis.

Harwayne's procedure solves the workers compensation problem, but this problem is unrelated to credibility considerations. Harwayne wants to use State A experience to set rates in State B by classification. He is not concerned with credibility.

As mentioned above, Harwayne's procedure deals with three issues:

1. The procedure adjusts the national pure premium to the benefit level of the state under consideration. This is the crux of the procedure. It relates to the general ratemaking issue of ensuring that the ratemaking data is not biased. It does not relate to issues of credibility.
2. Harwayne's procedure uses a complex three-way credibility weighting formula.
    A. The indicated partial pure premium has a full credibility standard based upon the expected losses in that classification, with the full credibility standard differing for serious indemnity, non-serious indemnity, and medical benefits.
    B. Partial credibility is set by a "three-halves" rule.[13] The three-halves rule says the following: If \$$X$ of

[13]The term "three-halves rule" stems from the obverse of this formula. If one needs \$$X$ of expected losses for full credibility, then for $Z\%$ credibility, one needs \$$X \times Z^{3/2}$.

One is tempted to delve into statistics textbooks to find a rationalization for the three-halves rule. In fact, the three-halves rule is used because it looks actuarial and it works. This justification of the three-halves rule has served admirably for over half a century now, and it should not be dismissed lightly. Any formula with a two-thirds power and used by actuaries all over the country must be mathematically unassailable; no one would simply make it up. And it works, in the sense that regulators and underwriters consistently defer to the pricing actuary's expertise in using this formula. They can't possibly argue with the formula, since they can't possibly understand it.

In a fascinating addendum to this school of thought, Howard Mahler has shown that the formula actually works. In fact, he shows that almost any formula works, as long as the credibility weights are within a reasonable range. Furthermore, they all work about equally well. Given the advantages to three-halves formula noted above, the formula is unimpeachable.

expected losses suffices for full credibility, then the credibility for $\$Y$ of expected losses is $(Y \div X)^{2/3}$.

C. The national partial pure premium has a full credibility standard based upon the national claim count in that classification. Once again, the full credibility standard differs for serious indemnity, non-serious indemnity, and medical benefits.

D. Partial credibility for the national partial pure premiums is set by a three-halves rule, similar to the rule for indicated partial pure premiums, except that claim counts are used instead of expected losses.

E. The credibility for the national partial pure premium may not exceed one-half of the complement of the credibility for the indicated partial pure premium. For instance, suppose that the indicated partial pure premium receives 40% credibility, and the three-halves rule would give a credibility of 50% for the national partial pure premium. The limit on the credibility for the national partial pure premium is $(1 - 40\%)/2 = 30\%$, so this is the credibility assigned to the national partial pure premium.

F. The remaining credibility is assigned to the underlying partial pure premium. In the example in the preceding paragraph, this remaining credibility is $(1 - 40\% - 30\%) = 30\%$ If this were a very small class, and the credibilities for the indicated and national partial pure premiums were 10% and 20%, respectively, then the underlying pure premium would receive $(1 - 10\% - 20\%) = 70\%$ credibility.

In Venter's terms [9], this procedure combines limited fluctuation credibility with greatest accuracy credibility. Since it does not purport to justify any of the parameters statistically, it would not be reasonable for us to rationalize the parameters after the fact.

Harwayne's Procedure and Boor's Thesis

The final issue in Harwayne's formula pertains to Boor's thesis. Why do we go to all the trouble of adjusting the national experience to the benefit level of the state under review? Why isn't it sufficient to credibility weight with the underlying pure premium, as is done in other lines of business?

Boor's paper provides the answer. The underlying pure premium is not independent of the indicated pure premium. This is particularly true for small classifications in workers compensation, more so than for most other blocks of business.

To see why this is so, let us consider a simple example. Suppose that we are setting rates for a new insured in State A in classification W. The classification is small; in fact, suppose that there are only five other insureds in classification W in State A. The historical experience is not fully credible. In other contexts, when we say that historical experience is not fully credible, we mean that random loss fluctuations may cause a significant disparity between the observed pure premium and the expected pure premium. In this case, the lack of full credibility has a more expansive meaning. Specifically, these five insureds may be better or worse than average, so we do not want to rely totally upon their experience to set rates for other insureds.

In actuarial terms, it is not simply that the historical experience is too volatile. Rather, we are afraid that the historical experience may be biased, though we do not know the magnitude of the potential bias or even the direction of the bias. To reduce the effects of the potential bias, we want to credibility weight the historical experience with additional information.

What other information should we use? The standard ratemaking answer is to use the underlying pure premium. In fact, many novice actuaries will indeed credibility weight with the underlying pure premium (or with the expected loss ratio). But this does not do the trick at all. The five risks in this classification have been insured for many years, and the underlying pure premium is based upon their experience in past years. We are concerned that they are not representative of the average risk. The underlying pure premium is just as problematic as the experience pure premium.

The national pure premium, however, is independent of the experience pure premium. It is based on the experience of other risks. The five risks in this classification in this state may be better or worse than average, but the hundred and fifty risks in this classification in the rest of the country are more likely to reflect the true average.

Contrast this workers compensation example with a corresponding personal automobile example. Suppose that we are making personal auto rates for a small classification W in state A. The classification is not fully credible, because there are only 500 drivers in this classification.

Here we are concerned with random loss fluctuations, not with bias. We are not worried that these 500 drivers may be better or worse than the average classification W driver that the company will insure. Rather, we are concerned with volatility. Perhaps the true expected claim frequency is 10%, so we should expect 50 claims. Actual experience may have been 40 claims or 60 claims, so the indicated rates may be 20% too low or too high.

The loss volatility affects each accident year separately. The most recent experience may be too high or too low, so we credibility weight with the underlying pure premium (or with the expected loss ratio), which reflects the experience of prior years, along with the business judgment of the past pricing actuaries. The underlying pure premium is not that interdependent with the historical pure premium, so there is less need to turn to external information.[14]

[14]The remarks made earlier about compensation system differences apply here as well. In workers compensation, if classification W has twice the average loss costs per dollar of payroll in State A compared to the statewide average, than it probably also has twice the average loss costs per dollar of payroll in State B compared to the statewide average.

We can now state Boor's thesis in rigorous terms:

> If the ratemaking data may be biased (though neither the magnitude of the bias nor even the direction of the bias are known), it is useful to credibility weight the experience indication with information that is relatively independent of the ratemaking data set.

Two characteristics of this revised thesis are of particular import:

1. All data sets used in ratemaking should have the five desirable characteristics drawn from Boor's paper. These characteristics are equally relevant for the information that receives the complement of credibility as they are for the basic ratemaking experience. The only difference is a practical one: often the data that receives the complement of credibility must be carefully adjusted in accordance with these characteristics, as is true in Harwayne's method.
2. Independence is particularly important when one believes that the historical experience may be biased, and especially when one does not know the magnitude or the direction of the bias. If the historical data is simply sparse, and random loss fluctuations may distort the indications, then independence is not of great concern. A larger volume of data is all that is required. It is the bias problem that demands a solution of independence.

In personal auto, the classification differentials are heavily dependent upon the compensation system and the underwriting structure. For instance, young unmarried male drivers may have an expected pure premium five times the statewide average in a tort liability state but only three times the statewide average in a no-fault state with a strong verbal tort threshold. Similarly, driver experience, or "years since first licensed," may be a more powerful classification variable in a state that does not permit underwriting by age of the driver than in a state which does allow such underwriting. Finally, the major classification dimension in personal auto is territory, which serves as a proxy for a host of hard to quantify loss cost drivers, such as attorney involvement in insurance claims and medical treatment of automobile injuries. Territorial relativities are peculiar to each state. One cannot credibility weight an indicated territorial relativity with information from other states.

10. RATEMAKING VERSUS PRICING

> … let me tell you how I use credibility. When I need a higher rate, I choose a credibility factor that gives me a higher rate. When I need a lower rate, I choose a credibility factor that gives me a lower rate.
>
> (A prominent American pricing actuary, 1988)

The previous discussions of credibility apply primarily to small and medium sized insurers whose experience is intermittently rocked by random loss occurrences. In personal automobile, most of the coverage in the United States is written by large carriers with thousands of claims in many states, such as State Farm, Allstate, USAA, GEICO, Farmers, and Liberty Mutual. The traditional formulas generally assign full credibility to their historical experience. Do they have any need for considering the complement of credibility?[15]

If the characteristics of the complement of credibility are important for the small insurer, they are crucial for the large insurer—though they are entirely different. The actuarial apprentice begins with traditional ratemaking, advances through financial pricing models and multi-year ratemaking procedures, and finally graduates to the tasks of the master actuary: marketplace competition, underwriting cycle movements, elasticity of supply and of demand, and the relationship of risk quality to price.

We want to examine the relationship of Boor's thesis to actual insurance pricing, not simply to traditional actuarial rate reviews. To understand the determinants of insurance pricing, we must first understand the economics of risk.

---

[15]In a similar vein, Richard Woll [10] points out that there is insignificant "process risk" in the claim costs of these large insurance companies, though "parameter risk" remains for them, just as it affects other insurers. Classical credibility theory—at least in the traditional treatment by Longley-Cook [5]—pertains to process risk, not to parameter risk. Indeed, these companies generally accord credibility of 100% to their historical experience in their formal rate reviews. However, the rate-setting practices of these insurers are far more market-oriented than are the corresponding rate-setting practices of the more traditional independent-agency companies.

Insurance Risk

Novice actuaries are often told that insurance operations are particularly risky, since the costs of coverage are not known until after the policy has expired. The nature of insurance risk has important implications for policy pricing and for Boor's thesis, so the dictum in the previous sentence warrants careful analysis.

Compare the auto manufacturer to the auto insurer. The auto manufacturer—so the argument goes—knows the costs of its inventory, its work force, its equipment, and its supplies before it sets a price for the final product. This price can be set as a fixed mark-up over the costs, ensuring a steady return for the manufacturer.

The automobile insurer, in contrast, needs actuaries to peer into the future—to convert raw historical records into prophecies of future costs. These prophecies are uncertain, so auto insurers need an extra margin of profit to compensate them for this risk.

This argument would be laughable if it were not so frequently repeated, in one guise or another, in actuarial circles. Yes, there are some risks that are indeed peculiar to insurers. Asbestos and pollution risks have hurt many large commercial lines carriers, and natural catastrophes have hurt many personal lines companies.[16] But these are the extraordinary events that have ruined the rare insurer: sometimes the overly aggressive insurer, sometimes simply the unlucky insurer. Insurers writing mostly the "bread-and-butter" lines with carefully considered reinsurance programs have largely avoided these risks.

Let us consider the true business risks to the manufacturer and to the insurer. Consider first the auto manufacturer. Most auto makers must design new model cars at least 36 months

[16]In truth, this argument sheds more light on the myopic view of many casualty actuaries and other businesspeople than on the attributes of the insurance industry. Asbestos has bankrupted its manufacturers, and pollution liabilities have devastated many chemical concerns. Asbestos and pollution have siphoned billions of dollars from the insurance industry, but most carriers will weather the storm.

before they are brought to the market.[17] The investment is enormous: retooling plants and equipment, sometimes building whole new factories, setting up production lines, producing hundreds of parts that will be needed with the new chassis, developing extensive advertising and promotional activities, educating an entire sales force of independent dealers with the characteristics of the new model.[18] Sometimes the new model will sell well, and the auto manufacturer will earn hundreds of millions of dollars. Sometimes the new model will flop, and the auto manufacturer will have lost hundred of millions of dollars.[19]

This is risk. It has nothing to do with Poisson distributions or inverse power curves.

Insurance does not have these risks. To produce insurance policies, the insurer must purchase a word processor and an office copier, hire an underwriter, and contract with an agent. It does not need a plant or a factory or a laboratory. The insurer does not spend tens of millions of dollars designing a product, buying parts, producing the final goods, advertising them in expensive campaigns. The insurer hangs out a shingle and sells the policy. Well … maybe it's not *that* simple. But the underlying principle is correct: the insurer does not face the large up-front capital commitment that represents manufacturing risk.[20]

---

[17]This time lag was about 60 months through the mid-1980's, until the intensified global competition from Japanese firms forced U.S. auto manufacturers to streamline their production schedules.

[18]As an example of the size of the investment, the decision to produce the Saturn automobile required General Motors to set up a new branch —the size of a major U.S. firm—many years before a single car would be sold.

[19]Other industries have equally great investments. Pharmaceutical companies, for instance, routinely spend tens of millions of dollars in research and development a dozen years or more before they expect to bring a new prescription drug to market.

[20]The formal economic expression of this is that manufacturers, utilities, pharmaceutical companies, and similar enterprises have high operating leverage, so their returns are sensitive to changes in market demand. Insurers have low operating leverage, since almost all their costs are variable. Even most of the expenses that casualty actuaries call "fixed expenses" are considered variable expenses by economists: they do not vary in direct proportion to premium, but they do vary with overall business volume. As a result, insurance profits are far less sensitive to changes in market demand.

In many industries, brand name differentiation adds to the business risks. It is not just the expense of manufacturing a new car that represents risk. To successfully bring a car to market, the auto manufacturer must convince dealers and consumers that the new model is superior to dozens of existing models. Insurance policies, in contrast, look more or less like one another across the industry. Product differentiation is hard to achieve in insurance.

Is insurance then riskless, or at least of low risk? Not at all, but the risk is of a different sort.

The ease of entry into the insurance market—or at least the apparent ease of entry into the insurance market—highlights the actual risk of insurance operations. Many insurance products are like commodities, with standard terms and multiple suppliers. Customer loyalty is high; that is, repeat sales are not as sensitive to price as new production is. As a result, many insurers are sometimes misled. They do not see high price elasticity in the majority of their business (that is, in the renewal customers), so they presume that customer service is more important than price.[21]

In fact, the opposite is true. Price is the dominant variable for new business production in most lines of business, and (because of high retention rates) new business production is of primary importance for overall volume and ultimately for the viability of the insurance enterprise.

Pricing: Cost-Based and Market-Based

The preceding paragraphs lay the groundwork; let us now return to Boor's thesis. The pricing actuary is in a quandary. The question is not what price best reflects the costs of the product. Actuarial ratemaking skills are so well-honed, and the law of large numbers so effectively eliminates much of the loss volatility, that actuarial techniques are accurate predictors of future

[21] It seems that every American insurer (by its own admission) provides exceptional service—or, at least, this is true for every failing American insurer.

costs. But the dilemma of the pricing actuary is different. If the price is too high, the insurer will lose market share: imperceptibly in the short run, but significantly in the long run. If the price is too low, the insurer will lose money on the policies that it sells.

The novice actuary retorts: "If our techniques work so well, the price will never be too high or too low." This actuary has confused ratemaking and pricing. Whether the rate indication is too high or too low depends on the technical skills of the actuary. Whether the price is too high or too low depends on market conditions (such as supply and demand) and the prices charged by competitors, which fluctuate with the underwriting cycle, not just with random loss occurrences.

Actuaries seem to espouse cost-based pricing to the exclusion of market-based pricing. This seems strange, since Western economists are virtually unanimous that market-based pricing—that is, pricing based on supply and demand considerations—is the linchpin of free-market capitalist systems. Cost-based pricing, in contrast, is not a rational pricing system for free markets. It has been used in regulated markets, such as in utility markets before the 1990's, but it would be useless in the competitive markets for property/casualty insurance that now prevail in most states.

In truth, the apparent predilection for cost-based pricing is an artifact of actuarial theory, not of actuarial practice. Actuarial theory emphasizes rigorous mathematical procedures. Cost-based pricing can be made as rigorous as desired, regardless of how relevant it is for the real world, so the actuarial literature is replete with formulas for cost-based pricing. Market-based pricing may be the crux of actual practice, but there are no theorems and few formulas, so the actuarial literature is devoid of papers on market-based pricing.

Boor's thesis is fundamental to the issues raised above. What is the ideal data that should receive the complement of credibility? The data from one's own company is inherently suspect, for two reasons. First, if the data relates to the coverage at issue, it is rarely independent of the historical experience. Second, such data gives us more information for cost-based pricing. A large insurer has all the information it needs for cost-based pricing. It needs instead information relevant to market-based pricing.

The rates charged by peer companies are the ideal data set for the complement of credibility, as long as they can be converted to the underwriting basis of one's own company. This conversion is critical for real-world pricing. Suppose that you are setting personal automobile insurance rates in a certain state. After working out the rate indications based on your own company's experience, you examine the rates of a major competitor. You find that your competitor's rates are about 40% higher than your own rate indications.

The first question should be: "Is the coverage the same? That is, are the underwriting criteria the same for the two companies?" Your company may be selling policies to standard or preferred risks, whereas your competitor may be selling to substandard risks. If your substandard rates are about 40% higher than your standard rates, then the disparity between your rates and your competitor's rates may be ascribed to underwriting differences, not to pricing differences.[22]

Competitors' rates tell us two things:

1. They tell us about the expected costs of the coverage, based on independent historical data, probably some differences in the ratemaking method, and differences in actuarial judgment.

[22]This is analogous to Harwayne's procedure. Harwayne adjusts for differing benefit levels and cost levels by jurisdiction. Here you are adjusting for differing underwriting practices by company.

2. They tell us a great deal about market place forces, competitive pressures, underwriting cycle movements, and supply and demand considerations.

Some actuaries are loath to incorporate market-based considerations into their rate recommendations. They say: "The actuary determines the proper rates—rates that are equitable for both insurers and consumers. Market-based pricing is irrational, based on seemingly bizarre underwriting cycle movements. Actuaries, as the champions of rigorous theory, should not be abetting irrational behavior."

This is a wonderful argument, but it is irrelevant. Real world insurers prefer market-based recommendations over mathematical elegance. Actuarial rigor is firmly established in traditional ratemaking departments. Actuaries who wish to be heard must seek the light of the marketplace.

Consider again the quotation at the beginning of this section. Yes, the language is a bit facetious: even actuaries should be allowed a sense of humor. But the underlying intent is serious. The actuary who made the remark—the chief actuary of one of the country's largest and most successful insurers—was particularly skilled at anticipating the rate movements of competitors, to know when it was safe to raise rates, and when other pricing or underwriting actions would have to substitute. He made this remark in response to a theoretical presentation on the credibility that should be assigned to the experience loss ratio. The elegant expositions so often heard at actuarial seminars and conferences are often irrelevant to real world pricing.

Let us rephrase the quotation in accordance with Boor's thesis. The pricing actuary ponders:

> My actuarial student, upon examining the company's experience, has obtained a rate indication of +6%. The marketing department says that our major competitors are about 8% to 10% above our rates. My guess is that

> our competitors will take rate increases of around 5% this year.
>
> This means that we could take a rate increase of as much as 14% or 15% this year without exceeding the market rate. Perhaps the rate indication of +6% is understated: maybe the trend estimate is too low, or maybe we had some particularly lucky experience this past year. Even if the +6% is accurate, we have all this leeway between +6% and +15%. Should we take something closer to +15% and reap the profits? Or should we take something closer to +6% and try to gain market share?

This is the essence of the complement of credibility thesis. Sophisticated pricing means weighting together independent indications to determine the rate request that is actually filed. If two indications stem from the same set of data, then these indications are probably interdependent, and they may contain little more information than a single data set would provide. If the two indications stem from different sources, and particularly if the rationale for the indications are different (e.g., one is cost-based and one is market-based), then the indications are probably independent, and the two indications provide more information than either one alone contains.

Actuaries well-versed in traditional rate-making techniques will object, saying: "How can one determine the proper credibility to assign to the historical data versus to the rates of peer companies? This is too subjective; there is no rigor in this." So these actuaries give up on real-world pricing, and they return to actuarial theory.

Quite the contrary is true. The traditional (classical) credibility figures are plucked out of the air. We say things like: "The full credibility standard is 1,024 claims, which gives a 95% probability that the historical claim frequency is within ±5% of the true claim frequency. If there are fewer than 1,024 claims, then the credibility assigned to the historical experience is determined by the square root rule, and the complement of credibility is assigned to the trended expected loss ratio."

There is no doubt that this impresses the layman. But what does "a 95% probability … " have to do with a firm that is trying to maximize profits? What relation does it have to pricing in a competitive market? The actuary is using cost-based pricing when the actual prices will be set by marketplace forces. No credibility formula will be correct, since the actuary has not asked the right questions.

The actuary should be asking: "If the indicated rate from my own experience is $1,000 per car for a certain classification and territory, and the corresponding average rate of my peer companies is $1,100 per car, what rate should I use?" This is the proper question, and this is a statistical question. The answer depends on (i) the price elasticity of demand, (ii) the persistency rate of insureds at different cost differentials, and (iii) the discount rate for future profits. At one extreme, with (i) a high price elasticity of demand, (ii) a low persistency rate of insureds at high cost differentials, and (iii) a low discount rate for future profits, it is wise to price below the competition (as long as one can do so profitably), pick up market share (both in new business production and in transfers from peer companies), and accrue the long term profits from the expanded book of business. At the other extreme, with (i) a low price elasticity of demand, (ii) a high persistency rate of insureds even at high cost differentials, and (iii) a high discount rate for future profits, it is better to move towards the market rate and to take the current profits from the redundant price.

This is a credibility question. At any given price elasticity of demand, persistency rate, discount rate, and differential between one's own indications and the rates of peer companies, there is a theoretically optimal credibility to assign to one's own experience. Of course, price elasticities are difficult to measure, and some companies do not keep track of persistency rates, but at least the pricing actuary is asking the right questions. Once casualty actuaries are turned in the right direction—that is, once they have formulated the questions correctly—they will make rapid progress on the solutions.

We have come full circle. Readers who skim lightly over Boor's paper receive the impression that the estimate of credibility is the crucial question, and the secondary consideration is to know what will receive the complement of credibility. On the contrary: until we define the purpose of the credibility procedure, we cannot know what should receive the complement of credibility. And until we know what will receive the complement of credibility, we cannot know the amount of credibility to assign to the experience.

11. CONCLUSION

Boor's paper leads in many directions, continually circling back to his thesis.

There are four rationales of credibility procedures: (i) limited fluctuation, (ii) proxy for past experience, (iii) greatest accuracy, and (iv) marketplace pricing tool. Each of these rationales implies a different formula for calculating the credibility, and each of these rationales implies a different set of data that should receive the complement of credibility.

Limited Fluctuation

Credibility may be used to limit the fluctuation in rate levels from year to year. This is particularly important in a regulated industry with great public concern about price increases and about alleged rate redundancies in some lines.[23] This rationale leads to the classical credibility procedures. The parameters of the full

[23]The author of this discussion, like most casualty actuaries, would dispute these allegations. Nevertheless, they continually recur, and they have great influence on many state legislators and regulatory officials.

credibility standard—that is, the size of the confidence interval and the probability constant—depend on how strictly one wishes to limit the fluctuation in rate levels.

The complement of credibility should be assigned to the "current rates:" i.e., to the underlying pure premium or to the expected loss ratio. The figure receiving the complement of credibility should first be adjusted for all factors other than random loss fluctuations, such as loss cost trends and changes in the insurance compensation system.

We may term this the Venter view of classical credibility. Some pure actuaries look with disdain upon this procedure, as a relic from unsophisticated actuarial practice. Nevertheless, it remains the prevailing standard in most lines of business.

Proxy for Past Experience

The credibility weighting procedure may be used as a proxy for the historical experience of older years.

The credibility assigned to the historical experience depends on the rapidity of shift of risk parameters over time. In more formal actuarial terms, it depends on the covariance structure of these risk parameters along a time dimension.[24]

(B) The complement of credibility should be assigned to the "current rates:" i.e., to the underlying pure premium or to the expected loss ratio, after adjustment for any part of the most recently filed rate revision that was not approved by the state insurance department. In addition, the figure receiving the complement of credibility should be adjusted for all factors other than

[24] See Mahler [6], pp. 261–263, for a full explanation. Based on Mahler's analysis, which examined baseball win-loss statistics, not insurance losses, a wide range of credibility figures may give equally good results. Mahler's sports results are probably valid for insurance experience as well, since they stem from the stochastic characteristics of random variables, not from any peculiarities of baseball. However, it is difficult to prove this assertion.

random loss fluctuations, such as loss cost trends and changes in the insurance compensation system.

This use of credibility is discussed by Mahler, and the adjustment to the expected loss ratio is documented by Graves and Castillo. The procedure is used by ISO for general liability ratemaking.

Greatest Accuracy

Credibility may improve the predictive accuracy of cost-based pricing.

This rationale is the underpinning for Bayesian or Bühlmann credibility methods. The credibility equals $M/(M+K)$, where M is a measure of business volume and K is proportional to the "within variance" divided by the "between variance."' This procedure is not concerned with deviations from the current rate.

In the Bayesian perspective, there is no conceptual difference between the credibility amount and the complement of credibility amount. There are as many estimators that may receive some credibility as there are ratemaking data sets. Ratemaking data sets are more useful to the extent that (i) they are accurate predictors of future experience, (ii) they are practical, and (iii) they are unbiased. Independence of these data sets avoids the costs of extra analysis that may have little benefit.

The Bühlmann credibility formula is commonly used in experience rating plans, though the K values are not always chosen by rigorous statistical analysis.[25] Bayesian credibility procedures have often been explored for territorial ratemaking, and K values have sometimes been estimated for these applications. Bayesian credibility analysis is not commonly used in practice for standard statewide ratemaking (class ratemaking), though many casualty

[25]In the past decade, there has been a trend toward a more actuarial selection of the K constants, particularly for rating bureau pricing procedures, such as those at ISO and at the NCCI.

actuaries and insurance companies have explored this topic and are using some of these procedures on a limited basis.

Pricing

Credibility may be used to combine cost-based and market-based pricing indications.

The goal of pricing is not to estimate the costs of the product but to optimize the long-term profits of the firm, or to meet other objectives of the firm. The credibility to be accorded to the company's historical experience depends on the price elasticity of demand, the persistency of insureds at different cost differentials, and the discount rate for future profits.

The estimate that should receive the complement of credibility is the marketplace price, for which the rates of major competitors (or peer companies) is often substituted. Adjustments must be made for underwriting differences among the peer companies.

The actuarial literature, which is replete with papers on ratemaking, is almost devoid of material on policy pricing. In practice, senior company actuaries provide both ratemaking and pricing recommendations for their employers.

Policy pricing is generally learned on the job, not from books and papers. Policy pricing is learned *from* experience; the price is not found *in* the experience.

The extension of actuarial expertise to real-world pricing problems in competitive markets is one of the most alluring tasks for the future casualty actuary. One of the primary questions is how much weight should be accorded to one's own indications. Boor's paper awakens us to the other, equally important question: to what information should we give the remaining weight?

Solutions to these two questions will help move the actuary's backroom desk to the forefront of insurance company operations.

REFERENCES

[1] Conger, Robert F., "The Construction of Automobile Rating Territories in Massachusetts," *PCAS* LXXIV, 1987, pp. 1–74.

[2] Cook, Charles F., "Trend and Loss Development Factors," *PCAS* LVII, 1970, pp. 1–26.

[3] Graves, Nancy C. and Richard Castillo, "Commercial General Liability Ratemaking for Premises and Operations," *Pricing*, Casualty Actuarial Society Discussion Paper Program, 1990, Volume II, pp. 631–696.

[4] Harwayne, Frank, "Use of National Experience Indications in Workers' Compensation Classification Ratemaking," *PCAS* LXIV, 1977, pp. 74–84.

[5] Longley-Cook, Lawrence H., "An Introduction to Credibility Theory," *PCAS* XXXIX, 1962, pp. 194–221.

[6] Mahler, Howard C., "An Example of Credibility and Shifting Risk Parameters," *PCAS* LXXVII, 1990, pp. 225–308.

[7] Rodermund, Matthew, *Foundations of Casualty Actuarial Science*, Second Edition, New York: Casualty Actuarial Society, 1992, pp. 1–24.

[8] Stern, Philipp K., "Ratemaking Procedures for Automobile Liability Insurance," *PCAS* LII, 1965, pp. 139–202.

[9] Venter, Gary G., "Credibility," *Foundations of Casualty Actuarial Science*, Second Edition, New York: Casualty Actuarial Society, 1992, pp. 375–483.

[10] Woll, Richard G., Review of Cozzolino and Kleinman: "A Capacity Management Model Based on Utility Theory," *Pricing, Underwriting and Managing the Large Risk*, Casualty Actuarial Society Discussion Paper Program, 1982.

ADDRESS TO NEW MEMBERS—NOVEMBER 9, 1998

MICHAEL A. WALTERS

First of all, let me congratulate you all again on this immense accomplishment of Fellowship in the Casualty Actuarial Society. Most of you no doubt spent an intense seven to nine years of sacrifice and dedication to achieve this milestone.

While you bask in your glory today, let me highlight some of the challenges ahead, adding the perspective that the training you have had thus far stands you in pretty good stead to meet the future.

I should really say training and selection. Of all the talented people who started down the path of actuarial exams, you made it to the finish line. The determination and judgment you have already demonstrated will come in handy in meeting the challenges ahead. Let's face it, the syllabus for our exams is still quite a formidable barrier to master completely. You had to exercise considerable judgment in finding the essence of those readings and deducing the likely exam questions.

One area that has not been tested in the actuarial exams is supervisory management skill—or the art of getting things done through others. Actuaries do not come by this skill naturally, because their first instinct is to solve problems themselves, instead of letting others do it.

Some time ago, when the CAS and SOA were discussing the possibility of expanding the core of common exams, the SOA was contemplating adding management courses to the syllabus, because new FSAs were struggling initially as managers. It seems that large life companies were placing their new FSAs immediately into management positions, without any supervisory experience. The CAS response was, don't put management on the actuarial exams. Instead, exhort companies to include supervisory management in the actuaries' training before promoting them.

Actually, in the future, supervisory management may be less critical, although still important. In the new information age, corporations don't need vast spans of control and complex organizational structures. These are being replaced by knowledge workers with vastly improved communication access. This lessens the need for supervisory apprenticeship, with more emphasis on conceptual and communication skills to go along with technical skills which actuaries have in abundance.

The casualty actuarial profession has enjoyed phenomenal growth over the past two decades meeting demands for their skills. But will this growth continue in the future? Will the casualty actuaries experience some of the problems our life brethren now face, wherein the demand for traditional service from 10,000 life actuaries may have peaked? Hence, life actuaries are reaching for new applications for their skills.

First, we won't reach 10,000 in number for a decade or two. Second, the demand for casualty skills doesn't seem to be slowing down. About ten years and 1,500 actuaries ago, we interviewed selected CEOs of major insurers to gauge future casualty actuarial demand. We focused on companies where the CEOs either were actuaries or hired a lot of actuaries—Bill Bailey of Aetna, Ed Budd of Travelers, Jack Byrne of Fireman's Fund, and Warren Buffett of Berkshire Hathaway.

Warren Buffett was actually a surprise interviewee, as Jack Byrne on his own had lined him up for a teleconference with us in the middle of his scheduled interview. In fact, we were told we could get him only for a short time, and so were limited to a few questions.

With our questions rationed, the first one we crafted was "What are the keys to success in the casualty insurance business." Buffett hesitated at first, asking what we meant. We improvised by observing that in some industries, the keys to success were "brains and guts." He thought for a moment and replied: "The keys to success in casualty insurance are brains and no guts."

He also offered a suggestion regarding the newly emerging requirements for every insurer to have an actuarial opinion on loss reserves. Taking a page out of Jack Byrne's perennial message that insurers need a disciplined balance sheet, he exhorted the appointed actuaries to have more courage in standing up to their companies' reserve committee in needed IBNR.

In fact, he recommended a unique reserve runoff test, which has surprisingly not yet been adopted by our profession. He said: "Compare the five-year run-off results of all the appointed actuaries. Then take the actuary with the worst record of underreserving and shoot him." He added, "You don't have to actually pull the trigger. You could whisk him off to a South Sea island, just so no one finds out you haven't actually eradicated him." This message of the need for more accountability for loss reserve opinions was passed along to the appropriate committees, including the Discipline Committee.

The other CEOs interviewed all concluded that casualty actuarial skill would continue to be in great demand, and even a fourfold increase would not be enough. A much greater supply would allow some of those trained as actuaries to become underwriters—perhaps an even more difficult job than an actuary.

Of course, this group could not have known then that two of their companies would later merge, and that consolidation of many insurers could threaten to curtail the number of actuarial positions available.

In fact, industry contraction has been a major, and not surprisingly parochial, concern of actuaries recently. A few years ago, the CAS began an annual survey of actuarial leaders to identify the top ten actuarial stories of the year. For two years running, the lead story was industry consolidation—first by primary insurers, and then by reinsurers. For years there weren't enough casualty actuaries to put even one in each company. Lately, the ratio is up to three per company. Not a cause for alarm just yet, but if the number of insurers drops below 50, there may be a need to speed up the nontraditional applications of actuarial science.

Appropriately, the top story of the past year was the growth of risk securitization. Talk about re-energizing actuarial need: Pushed to an extreme, every separate book of business by subline within a traditional insurer could potentially require an actuarial opinion on its expected value for a transaction from a risk originator to a risk bearer. The latter may not even be an insurance company.

A third top ten story was the emergence of enterprise risk management—another nontraditional actuarial venture. Why not take the tried and true precepts of risk management—risk identification, risk assessment, risk control, and risk financing, and translate them into the rest of corporate risk management? Even if some of those business risks are classically uninsurable, the expertise of actuaries can surely be applied to some of those ventures.

For example, the year 2000 problem lends itself to enterprise risk management. In fact, there is even a way to buy insurance for that computer risk. It may have liberal doses of risk control in it, but some risk transfer is still possible.

But the greatest potential expansion of actuarial demand—dynamic financial analysis (DFA)—did not even appear as a top ten story probably because of the way that list was compiled. Only external news stories were used, which had some actuarial implications. No internal CAS research or committee work was considered. This was to keep an external focus for potential long-range planning purposes and to avoid commenting on internal works in progress.

For almost eight years, the CAS has been actively working to nurture the DFA concept, which has the potential of becoming a major practice area for actuaries in the future, on a scale with ratemaking or loss reserving. It builds on ratemaking and reserving skills by adding a new dimension on the asset side of the ledger and its interaction with liabilities. It also adds a fourth dimension by measuring the risk or variability of results and its effect on capital needs.

DFA requires new skills not now in the repertory of most actuaries but which can be acquired. The CAS Syllabus 2000 will be adding these new financial topics. And there is a mushrooming set of continuing education guidelines to give some of our existing members the skills and techniques needed to practice in this area. What might also be needed is some marketing prowess, as this concept still has some hurdles to overcome before it is readily accepted by CEOs and CFOs.

Lastly, if these new opportunities don't generate enough jobs for actuaries, there are even more nontraditional possibilities. For example, last year Mavis and I were returning from a golf outing, and stopped at Romanelli's pizzeria to pick up the order my wife had phoned in. Much to my surprise, there were two pizzas in my name—a small pepperoni and a medium cheese. This did not look like a very good deal, because Romanelli's new jumbo pizza, with half pepperoni and half cheese, had a much lower price and appeared to provide about the same amount of pizza.

So I started quizzing the cashier. How big is the jumbo pizza? Answer: 18 inches. Is that radius or diameter? Ahh…around. That can't be the circumference; it must be the diameter. Then I started some quick calculations, out loud, to assess the relative areas. Let's see. A small pie is 12 inches diameter and eight dollars. At πr^2, that's 36π. A medium is 14 inches at ten dollars; πr^2 makes it 49π; and a jumbo is 18 inches at twelve dollars; πr^2 makes it 81π—or almost the same size as the two smaller pies but at a third less cost. So why didn't you suggest a jumbo pie when my wife called in? To which the English-as-second-language cashier replied: "Pies are not square; pies are round."

Meanwhile, I was attracting a crowd of interested parties, including the owner who asked if I was a mathematician? No, an actuary. He then asked if I was available to help him price his pizzas, since his view of the situation now was that he might be underpricing the new jumbo pies. My response was that we had to consider the fixed cost of pie preparation—not just the size. Also, from the hungry customer's viewpoint, there may be some adverse selection against the seller. We left it that I would have to get back to him when I was no longer hungry.

Never went back. The lesson learned? Every pricing problem is an actuarial problem; we just have to understand the environment a little better. And when banks and investment firms get into the insurance risk transfer business, we'll just have to get into their business at the same time.

Thank you and best of luck to the CAS Class of 1998.

PRESIDENTIAL ADDRESS—NOVEMBER 9, 1998

THERE IS MORE TO DO

MAVIS A. WALTERS

When Bob Anker handed me the gavel at this time last year, he said the next twelve months would be interesting. I expected they would also be challenging and demanding. I was not disappointed. This really has been an exciting time to be in a leadership role in the CAS.

Looking back over this past year, we've done a lot of things well, but some things we could have, and should have, done better. And I'm confident that my able successor, Steve Lehmann, will find a way to improve on what has been accomplished so far. After all, the CAS is a strong organization. Its members are dedicated and resourceful. Each generation of leaders, including committee volunteers, chairpersons, officers, and board members, has without exception led us to higher levels of achievement.

I have no doubt that with Steve's leadership and the strong support of Alice Gannon, the Executive Council, and the Board of Directors, the CAS will continue in that tradition. So, where we might not have done so well this past year—not to worry—our future leaders will remedy that in the years ahead.

There is more to do.

Before sharing with you some thoughts on CAS accomplishments over this past year, as well as the challenges ahead, let me acknowledge some very special people who provided me with enormous support and encouragement during the past year.

The first is Bob Anker, my predecessor, who taught me what being presidential was all about. Bob was always gracious, kind and respectful, even to those he disagreed with. He was and is a wonderful role model. Thanks, boss!

And to Bob's predecessor, Al Beer, special thanks for his advice and assistance in reviewing the CAS discipline procedures. Al's good judgment and encouragement were particularly valuable as we struggled with some fundamental issues.

And no CAS president could get through even the first month without recognizing the enormous support and guidance provided by Tim Tinsley, our Executive Director. Tim is absolutely the critical ingredient who holds the CAS operations together, who sees to it that everything functions smoothly, and he does this with style and demeanor that are truly first class. We are most fortunate to have him as a part of our team.

Tim, you are a treasure. Many, many thanks.

I would also like to acknowledge and thank my counterpart at the Society of Actuaries during this past year, Anna Rappaport. Anna, my friend, thanks much for all your hard work and efforts to help bring our two organizations into closer harmony.

Another past president of the CAS has served as a role model for me, someone whose fairness and objectivity I have always admired but never could quite measure up to—my brother Mike. Mike is completely unflappable. I think he got all the patience genes when we were born.

Mike, thanks for always being so supportive of me and for encouraging me in all kinds of ways. I have always been proud that you are my brother, and I'm particularly happy that we both have been able to serve the CAS as its president.

I also want to thank Mike's wife, Mary Anne, my matron-of-honor. Her enthusiasm and willingness to do whatever she could to assist me were unbounded. She's been available to take on any assignment to help things run smoothly for me, both in May in Marco Island and here in Toronto. Thank you, Mary Anne.

And finally, I want to thank my husband, Tom—for all his patience and understanding in general, but over the past two years, in particular.

We both do a fair amount of traveling but have usually been able to catch up with each other on weekends.

But with the Council of Presidents, CAS Board, International Presidents Group, and other travel over the past 24 months, I've spent a lot of weekends someplace other than home. Tom's career demands made it virtually impossible for him to travel with me, so we've spent a lot of time apart. I am happy that he has been able to come to both CAS meetings this year and share in some of the fun things that go with the CAS presidency. Thanks, Tom, for your good humor, your patience, and for every other way you have been so good to me.

Now I'd like to turn to the significant accomplishments of the CAS over the last year—and to what more there is to do.

During 1998 our membership surpassed 3,000, and at this meeting we are welcoming our largest Fellowship class ever—126 new Fellows. Our current membership now stands at 3,064. Our growth as a professional society has been quite impressive. From 1974 to 1984 our membership grew 83 percent to 1,112. From 1984 to 1994 membership doubled, and we have almost doubled again since 1989. Obviously something about a career as a casualty actuary is impressive to college students. And we have continued to attract students who are intelligent, well motivated and not intimidated by the prospect of sitting for 10 or 14 exams to achieve their credentials.

But we should not let our past success delude us into complacency.

There is more to do.

We must be prepared to take whatever steps are necessary to not only protect our franchise as casualty actuaries, but also to expand our expertise so we remain relevant and necessary in the changing marketplace. Our core competency today is well established and highly regarded. But the world around us is changing rapidly, and we have to keep pace.

During 1998 our Board of Directors approved several proposals to amend our constitution and bylaws. One proposal provides an affiliate class of membership that recognizes the training, experience, and interests of actuaries practicing in casualty work who have received the highest actuarial designation in another recognized organization.

But this new class of membership falls short of mutual recognition of Fellowship. At the last meeting of the International Presidents Group, such a proposal was advocated by the Institute of Actuaries of Australia. And the Faculty, and the Institute of Actuaries in the U.K. strongly endorsed it. Under the proposal an actuary who is a Fellow of any of the sponsoring, exam-giving organizations would, upon petition to any other similar sponsoring organization, be granted Fellowship in that organization, assuming the actuary practices in the area covered by the petitioned organization.

It is possible that the Society of Actuaries and the Canadian Institute of Actuaries may eventually find this proposal acceptable, perhaps with additional requirements. Certainly it will be pursued at the International Presidents Group meetings in the coming year.

So there is more to do.

I suspect many CAS members will strongly oppose mutual recognition, but I believe we should at least explore the concept and give it careful consideration. A way to make mutual recognition work to every nation's and to every organization's benefit may be possible.

Also this year, the Board of Directors approved changes to the constitution and bylaws addressing the discipline of members.

A complete package containing the proposed changes and the revised Rules of Procedure for Disciplinary Action went to the membership in July, and I discussed the new procedures in the August *Actuarial Review*.

To my surprise, only eleven CAS members actually commented on the proposals. Why only eleven? Were CAS members persuaded that the changes were improvements, and so there was no need to comment? Are you satisfied that the new process is fair and balanced? Do most members think they will never be subjected to disciplinary action, so they believe the subject is irrelevant to them? I think the answer to all these questions is yes. But that should not diminish the importance of the accomplishment.

Still, there is more to do.

To maintain that we truly are a profession, it is essential that we accept responsibility for disciplining those who do not abide by the highest standards of professionalism.

It is difficult and painful to have to pass judgment on one of our own members, but enforcing our code of professional conduct is an obligation we cannot avoid. The new procedures will serve us well as we fulfill that duty.

On the education front, this year we agreed to jointly sponsor the first four examinations with the Society of Actuaries under the new exam structure effective in 2000. This was explored about two years ago, but for a variety of reasons, including a series of misunderstandings that led to a lack of trust between the CAS and the SOA, those efforts fell short.

Thanks to the dedication and hard work of a lot of actuaries in both organizations, we were able to restart discussions that eventually led to both boards adopting the joint sponsorship that was announced to students in early July.

I believe joint sponsorship of the first four exams is a very positive step, not only for the CAS, but also for the actuarial profession. Now students will have an opportunity to learn something about all areas of actuarial practice before having to decide which series of exams to pursue. And in most cases students will have some workplace experience before they make that decision.

Of course there is still more to do to put the new syllabus and exam structure in place. Our committees are hard at work. I am confident they will reach or exceed their goal.

Another significant achievement this year: the excellent relationship that now exists between the leadership and the boards of the CAS and the Society of Actuaries.

Many of you are aware of the difficulties, indeed, hostility between our organizations early last year. But repairing the damage began almost immediately when Bob Anker and I were invited to a Society of Actuaries Board of Governors meeting to explain the CAS perspective on the events that led to the breakdown. And in turn, we invited the SOA's President and President-elect to CAS board meetings.

Since then, SOA President Anna Rappaport and I have communicated frequently and candidly about matters of mutual interest or concern. We have worked together on the American Academy Board of Directors, the Council of Presidents, and the International Presidents Group.

We've had dinner together on numerous occasions. And, just this past September, our two societies held a joint meeting of our boards. That provided a wonderful opportunity for the leadership of both organizations to explore topics of mutual interest and share views on the challenges facing *our* profession. I am confident that all who attended that meeting came away with increased respect and appreciation for each other professionally.

There is still more to do.

For the good of the profession we must make sure that mutual respect and good working relationships continue beyond the terms of any particular presidents. To that end, Anna and I have recommended that joint board meetings be held at least every other year, if not annually. We believe closer ties between us strengthen both societies. I am confident that Steve Lehmann and Howard Bolnick will continue to work well together, as will Alice Gannon and Norm Crowder the following year.

Another area where there is more to do is in the international arena.

During this past year Steve Lehmann and I met with the International Presidents Group in London and then went on to the International Congress of Actuaries meeting in Birmingham, England. Although about 950 actuaries from all over the world attended that conference, not a single CAS member presented a paper or participated in the formal program. And just last month a combined ASTIN and general insurance conference was held in Glasgow, where there was only one CAS presenter.

It is difficult to maintain the posture that the Casualty Actuarial Society represents the finest, best educated and most knowledgeable actuaries in the general insurance area when we don't participate in important international forums such as these. Fortunately, while in Glasgow, Steve Lehmann was able to make progress on a seminar to be jointly sponsored by the CAS, the Faculty, and the Institute next June in London. We will also have an excellent opportunity to make an impression at the next International Congress in Cancun in 2002.

Lastly, we have intensified our focus on nontraditional areas of practice: those wider fields where our members toil outside the familiar ratemaking and reserving areas.

In response to a survey by our External Communications Committee, our members most frequently identified marketing, underwriting, and risk management as the nontraditional actuarial duties they perform. In addition, 19 percent mentioned valuation; almost 17 percent, dynamic financial analysis; and 8 percent, investments.

And 22 percent of the respondents said they had "other" nontraditional duties, which included capital analysis, mergers and acquisitions, and strategic planning, to name a few.

The real growth in opportunities for casualty actuaries will be in nontraditional areas, and many of our members are already leading the way.

Many of us are involved in catastrophe modeling and exposure management, and this is leading some into exploring various risk-securitization techniques. The contribution that casualty actuaries can make to help find solutions to complex financial problems is unlimited.

There is more to do.

We must find ways to support our members as they face those new challenges and make sure our members have the tools to deal with those challenges appropriately.

We have a new Task Force on Nontraditional Practice Areas, chaired by Mike Miller, and we'll look forward to their report.

In conclusion, I'd like to share with you just a few personal observations.

It is a high honor to serve as your president. I've worked closely with the leadership of our organization. I've seen first-hand the enormous contribution so many members make, so I leave office this November with great confidence and pride in the Casualty Actuarial Society. We are blessed with a dedicated, talented, motivated membership that can meet any challenge, no matter how difficult.

The last 12 months have been a wonderful experience for me. I am proud to have served as your President. It was great fun, and I enjoyed almost every minute of it.

Thank You.

MINUTES OF THE 1998 CAS ANNUAL MEETING

November 8–11, 1998

SHERATON CENTRE TORONTO HOTEL

TORONTO, ONTARIO, CANADA

Sunday, November 8, 1998

The Board of Directors held their regular quarterly meeting from noon to 5:00 p.m.

Registration was held from 4:00 p.m. to 6:00 p.m.

From 5:30 p.m. to 6:30 p.m., there was a special presentation to new Associates and their guests. All 1998 CAS Executive Council members briefly discussed their roles in the Society with the new members. In addition, Michael L. Toothman, who is a past president of the CAS, gave a short talk on the American Academy of Actuaries (AAA) Casualty Practice Council.

A welcome reception for all members and guests was held from 6:30 p.m. to 7:30 p.m.

Monday, November 9, 1998

Registration continued from 7:00 a.m. to 8:00 a.m.

CAS President Mavis A. Walters opened the business session at 8:00 a.m. and introduced members of the Executive Council and the CAS Board of Directors. Ms. Walters also recognized past presidents of the CAS who were in attendance at the meeting, including: Robert A. Anker (1997), Irene K. Bass (1993), Albert J. Beer (1995), Phillip N. Ben-Zvi (1985), Ronald L. Bornhuetter (1975), Michael Fusco (1989), Frederick W. Kilbourne (1982), Michael L. Toothman (1991), and Michael A. Walters (1986).

Ms. Walters also recognized special guests in the audience: A. Norman Crowder, president-elect of the Society of Actuaries;

Peter F. Morse, president of the Canadian Institute of Actuaries; Anna M. Rappaport, immediate past president of the Society of Actuaries; Harris Schlesinger, immediate past president of the American Risk and Insurance Association; Michael L. Toothman, president-elect of the Conference of Consulting Actuaries; and Stuart F. Wason, president-elect of the Canadian Institute of Actuaries.

Ms. Walters then announced the results of the CAS elections. The next president will be Steven G. Lehmann, and the president-elect will be Alice H. Gannon. Members of the Executive Council for 1998—1999 will be: Curtis Gary Dean, Vice President–Administration; Kevin B. Thompson, Vice President–Admissions; Abbe Sohne Bensimon, Vice President–Continuing Education; David R. Chernick, Vice President–Programs and Communication; and Robert S. Miccolis, Vice President–Research and Development. New members of the CAS Board of Directors are Charles A. Bryan, John J. Kollar, Gail M. Ross, and Michael L. Toothman.

Curtis Gary Dean and Kevin B. Thompson announced the new Associates and Steven G. Lehmann announced the new Fellows. The names of these individuals follow.

NEW FELLOWS

John Porter Alltop
Lewis Victor Augustine
Barry Luke Bablin
Michael James Bednarick
Michael James Belfatti
Wayne F. Berner
Barry E. Blodgett
Kimberly Ann Bowen
Douglas J. Bradac
Ron Brusky
Christopher John Burkhalter
Tania Janice Cassell
Cindy C. M. Chu
Brian Arthur Clancy
Kay A. Cleary
Christopher G. Cunniff
Kenneth Scott Dailey
Smitesh Davé
Karen L. Davies
Jeffrey Wayne Davis
John David Deacon
Camley Ann Delach
Margaret E. Doyle
David L. Drury
Tammy Lynn Dye
Kristine M. Esposito
Sylvain Fauchon
Kendra Margaret Felisky-Watson
Stephen A. Finch
Walter H. Fransen
Kay L. Frerk
John Edward Gaines

David Evan Gansberg
Kathy Helene Garrigan
Margaret Wendy Germani
Moshe David Goldberg
John E. Green
Steven A. Green
Daniel Cyrus Greer
Daniel Eli Greer
Greg M. Haft
Ellen M. Hardy
Robert L. Harnatkiewicz
William Nesthus Herr Jr.
Daniel Leo Hogan Jr.
Jeffrey R. Hughes
Paul Ivanovskis
Christopher Donald Jacks
Joseph William Janzen
Jeremy M. Jump
Hsien-Ming Keh
Brandon Daniel Keller
Steven A. Kelner
Thomas Paul Kenia
Joseph P. Kilroy
Bradley James Kiscaden
Brian Scott Krick
Mary Downey Kroggel
Alexander Krutov
Kenneth Allen Kurtzman
Timothy John Landick
Robert John Larson
Guy Lecours
Thomas C. Lee
Thomas L. Lee
Scott Jay Lefkowitz
Steven Joel Lesser
Robert Glenn Lowery
Gary P. Maile
Anthony L. Manzitto
Richard Joseph Marcks
Peter Robert Martin
Dee Dee Mays
Stephen J. McGee
Jeffrey A. Mehalic
Brian James Melas
Anne Hoban Moore
Matthew Stanley Mrozek
Raymond D. Muller
Timothy O. Muzzey
Mindy Yu Nguyen
Mark A. O'Brien
Mary Beth O'Keefe
David J. Otto
Joseph Martin Palmer
Dmitry E. Papush
Thomas Passante
Abha B. Patel
Harry Todd Pearce
Lynne M. Peterson
Anne Marlene Petrides
Jennifer K. Price
David Scott Pugel
Kara Lee Raiguel
Kiran Rasaretnam
Natalie J. Rekittke
Dennis Louis Rivenburgh Jr.
Daniel Gregg Roth
Chet James Rublewski
Kevin L. Russell
Thomas A. Ryan
Elizabeth A. Sander
Manalur Sundaram Sandilya
Michael Bruce Schenk
Matt John Schmitt
Arthur J. Schwartz
Craig James Scukas
Gerson Smith
Mary Kathryn Smith
Alan M. Speert
Catherine Elaine Staats
Ilene Gail Stone
Scott Jay Swanay
Christopher Tait
Sebastian Yuan Yew Tan
Georgia A. Theocharides
Alice Underwood
Timothy John Ungashick
Jeffrey Alan Van Kley
Kimberley A. Ward
Wyndel S. White
William Robert Wilkins
Michael J. Williams
Kirby W. Wisian
Yoke Wai Wong
Charles John Yesker

NEW ASSOCIATES

Stephen Allan Alexander
Jennifer Ann Andrzejewski
Michele Segreti Arndt
Robert Daniel Bachler
Lee Matthews Bowron
John Carol Burkett
Matthew R. Carrier
Andrew K. Chu
Louise Chung-Chum-Lam
Mary Katherine Thérèse Dardis
Robert Earl Davis
Nathalie Dufresne
Carolyn F. Edmunds
Jane Eichmann
Jonathan Palmer Evans
Michelle Lynn Freitag
Donald Michael Gambardella
Charles Edward Gegax
Bradley Gordon Gipson
Natasha Cecilia Gonzalez
Francis Xavier Gribbo
Gary Michael Harvey
Kevin Blaine Held
Melissa Katherine Houck
Weidong Wayne Jiang
Susan Kay Johnston
Daniel R. Kamen
Mary Jo Kannon
Jeffrey Dale Kimble
Andrew M. Koren
Scott C. Kurban
Douglas H. Lacoss
James Peter Leise
Diane Lesage-Cantin
Xiaoying Liang
Jason Kirk Machtinger
James William Mann
Stephen Paul Marsden
John Vincent Mulhall
Moshe C. Pascher
Judith Diane Perr
Dylan Pamphilon Place
Ricardo Anthony Ramotar
Brian Paul Rucci
Asif Sardar
Kelvin Bryce Sederburg
Kelli Denise Shepard-El
John Haldane Soutar
Andrew K. Tran
Michael Charles Tranfaglia
Joel Andrew Vaag
Steven John Vercellini
Linda M. Waite
Robert Joseph Wallace
William Boyd Westrate
Matthew Michael White
L. Alicia Williams
Robin Davis Williams

Ms. Walters then introduced Michael A. Walters, a past president of the Society, who presented the Address to New Members.

A short awards program followed the address. Ms. Walters presented the 1998 CAS Matthew S. Rodermund Service Award to Richard H. Snader. Mr. Snader was chosen for his contributions to the actuarial profession. Ms. Walters also announced James A. Tilley, Ph.D. as the recipient of the 1998 CAS Charles A. Hachemeister Prize for his paper, “The Securitization of Catastrophic Property Risks.” Ms. Walters noted that Dr. Tilley was unable to attend this meeting but may present his paper at the 1999 Spring Meeting in Orlando, Florida.

Curtis Gary Dean announced Gene C. Lai, Robert C. Witt, Hung-Gay Fung, and Richard D. MacMinn as the winners of the CAS Online Services Prize. The group won for their contribution, "On Liability Insurance Crises," a paper in a portable document format with links to a Web page with animated graphics. The CAS Online Services Prize was established as the result of the 1998 Call for Contributions to the CAS Web Site. The call's purpose is to make CAS members aware of, and actively involved in the CAS Web Site and to establish the Web site as a primary forum for sharing news and ideas.

Gary R. Josephson, Chairperson of the CAS Committee on Review of Papers, presented the 1998 Woodward-Fondiller Prize to Donald F. Mango for his paper, "An Application of Game Theory: Property Catastrophe Risk Load." Mr. Josephson then presented the 1998 CAS Dorweiler Prize to Rodney E. Kreps for his paper, "Investment-Equivalent Reinsurance Pricing." Both papers are published in this edition of the *Proceedings*.

Mr. Josephson announced that nine *Proceedings* papers and one discussion of a November 1998 *Proceedings* paper would be presented at this meeting. In addition, one paper would be published in the 1998 *Proceedings* but would not be presented.

Patrick J. Grannan, CAS Vice President–Programs and Communications, announced Joan Lamm-Tennant and Mary A. Weiss as winners of the American Risk and Insurance Association Prize (ARIA) for their paper, "International Insurance Cycles: Rational Expectations/Institutional Intervention." Mr. Grannan then introduced Harris Schlesinger, Immediate Past President of ARIA, who spoke briefly about ARIA.

Ms. Walters then requested a moment of silence in honor of those CAS members who passed away since November 1997. They are: James M. Cahill, Clarence C. Coates, Charles Desjardins, K. Arne Eide, Gilbert W. Fitzhugh, Dave R. Holmes, Robert C. Perry, Paul E. Singer, Emil J. Strug, Paul A. Verhage, and Herbert E. Wittick.

In a final item of business, Ms. Walters acknowledged a donation of $10,000 from D. W. Simpson & Company to the CAS Trust (CAST). The donation was made October 7, 1998.

Ms. Walters then concluded the business session of the Annual Meeting and introduced the featured speaker, Alan J. Parisse. Mr. Parisse is a managing partner of his own firm, a former senior executive for Oppenheimer and other national investment firms, a guest lecturer at Stanford and Wharton, and the author of numerous articles and several technical books.

After a refreshment break, the first General Session was held from 10:45 a.m. to 12:15 p.m.

"Globalization"

Moderator/ Panelist:	Albert J. Beer President Munich-American RiskPartners
Panelists:	Wayne H. Fisher Head of Global Specialties Zurich Insurance Group
	James N. Stanard Chairman, President and Chief Executive Officer Renaissance Reinsurance, Ltd.

Following the general session, CAS President Mavis A. Walters gave her Presidential Address at the luncheon. At the luncheon's end, she officially passed on the CAS presidential gavel to the new CAS President, Steven G. Lehmann.

After the luncheon, the afternoon was devoted to presentations of concurrent sessions, which included presentations of the ARIA Prize paper and *Proceedings* papers. The panel presentations from

1:30 p.m. to 3:00 p.m. covered the following topics:

1. Y2K—Impact on D&O Insurance

Moderator/ Panelist:	Hilary Rowan Partner Thelen, Reid, & Priest
Panelist:	Michael T. Hasse Consultant Eon 2000 Division Double-E Computers

2. Workers Compensation Managed Care—Has Its Impact Been Felt?

Moderator/ Panelist:	Charles W. McConnell Consulting Actuary
Panelists:	N. Mike Helvacian Director of Research and Chief Economist National Council on Compensation Insurance
	Layne M. Onufer Principal Ernst & Young LLP

3. Canadian Catastrophes

Moderator:	Joseph A. Herbers Consulting Actuary Miller, Rapp, Herbers & Terry, Inc.
Panelists:	Paul Kovacs Vice President and Chief Economist Insurance Bureau of Canada
	Isabelle Perigny Consulting Actuary Tillinghast-Towers Perrin

Hugh G. White
Senior Vice President, Corporate Underwriting
Zurich Canada

4. Using Complex Models: A Proposed Actuarial Standard of Practice

Moderator: Karen F. Terry
Actuary
State Farm Fire & Casualty Company

Panelists: Paul E. Kinson
Consulting Actuary
Liscord, Ward & Roy, Inc.

Ronald Kozlowski
Consulting Actuary
Tillinghast-Towers Perrin

5. Introduction to the CAS Examination Committee

Moderator: David L. Menning
Chairperson, CAS Examination Committee
Senior Associate Actuary
State Farm Mutual Automobile Insurance Company

Panelists: Pierre Dionne
Property and Casualty Actuary
CIBC Insurance

J. Thomas Downey
Manager, Admissions
Casualty Actuarial Society

Thomas G. Myers
Vice President
Prudential Property & Casualty Insurance Co.

The following 1998 ARIA Prize Paper was presented:

"International Insurance Cycles: Rational Expectations/Institutional Intervention"

Authors: Joan Lamm-Tennant
Vice President
General Re New England Asset Management
Professor of Finance
Villanova University

Mary A. Weiss
Chair of Risk Management and Insurance
Temple University

The following 1998 *Proceedings* Papers were presented:

1. "Personal Automobile: Cost Drivers, Pricing, and Public Policy"

   Authors: John B. Conners
   Executive Vice President
   Liberty Mutual Group

   Sholom Feldblum
   Assistant Vice President and Senior Associate Actuary
   Liberty Mutual Group

2. "Studying Policy Retention Using Markov Chains"

   Author: Joseph O. Marker
   Vice President and Chief Actuary
   Citizens Insurance Company of America

After a refreshment break from 3:00 p.m. to 3:30 p.m., presentations of concurrent sessions and *Proceedings* papers continued. Certain call papers and concurrent sessions presented earlier were repeated. Additional concurrent sessions presented from 3:30 p.m.

to 5:00 p.m. were:

1. Mortgage Guaranty Insurance

   Moderator: John F. Gibson
   Principal
   PricewaterhouseCoopers LLP

   Panelists: Douglas Rivenburgh
   Assistant Vice President Structural Products
   United Guaranty Corporation

   Michael C. Schmitz
   Associate Actuary
   Milliman & Robertson, Inc.

2. Actuarial Supply and Demand

   Moderator: Frederick O. Kist
   Senior Vice President and Corporate Actuary
   CNA

   Panelists: Francis J. Lattanzio
   Partner
   Kelly and Lattanzio

   Isaac Mashitz
   Senior Vice President and Chief Actuary
   Zurich Reinsurance North America, Inc.

   Arlie J. Proctor
   Senior Consulting Actuary
   Scruggs Consulting

3. Questions and Answers With the CAS Board of Directors

   Moderator: Steven G. Lehmann
   CAS President-Elect
   Consulting Actuary
   Miller, Rapp, Herbers & Terry, Inc.

   Panelists: Regina M. Berens
   Consulting Actuary
   MBA, Inc.

David N. Hafling
Senior Vice President and Actuary
American States Insurance Companies

David L. Miller
Senior Vice President and Chief Actuary
Commercial Union Insurance Companies

Proceedings papers presented during this time were:

1. "Testing the Assumptions of Age-To-Age Factors"

   Author: Gary G. Venter
   Executive Vice President
   Sedgwick Re

2. "The Mechanics of a Stochastic Corporate Financial Model"

   Authors: Gerald S. Kirschner
   Associate Actuary
   Liberty Mutual Group

   William C. Scheel
   President
   DFA Technologies, LLC

3. "A Graphical Illustration of Experience Rating Credibilities"

   Author: Howard C. Mahler
   Vice President and Actuary
   Workers Compensation Rating and
   Inspection Bureau of Massachusetts

4. "The Myers–Cohn Profit Model, A Practical Application"

   Author: Howard C. Mahler
   Vice President and Actuary
   Workers Compensation Rating and
   Inspection Bureau of Massachusetts

An Officers' Reception for New Fellows and accompanying persons was held from 5:30 p.m. to 6:30 p.m. A general reception for all attendees followed from 6:30 p.m. to 7:30 p.m.

Tuesday, November 10, 1998

Registration continued from 7:00 a.m. to 8:00 a.m.

The following General Sessions were held from 8:00 a.m. to 9:30 a.m.:

"Integrating Financial and Insurance Products"

Moderator: Jeffrey H. Mayer
Principal
Head of Client Service Group
Swiss Re New Markets

Panelists: E. Randall Clauser
Executive Vice President and Head of
Zurich Corporate Solutions
Zurich American Insurance

Steven R. Fallon
Senior Vice President
Centre Solutions

Daniel Isaacs
Vice President and Investment Actuary
Falcon Asset Management

"Fraud-The Unknown Factor in Claim Payments"

Moderator/Panelist: Richard A. Derrig
Vice President–Research
Insurance Fraud Bureau of Massachusetts

Panelists: Keith J. Crocker
Professor of Economics
University of Michigan

Elizabeth A. Sprinkel
Senior Vice President and Chief
Research Officer
Insurance Research Council

Following a break from 9:30 a.m. to 10:00 a.m., certain concurrent sessions that had been presented earlier during the meeting were repeated from 10:00 a.m. to 11:30 a.m. Additional concurrent sessions were:

1. Securitization 102—Securitizing Noncatastrophe Risks

Moderator: John S. Bunt
Senior Vice President
American Re Financial Products

Panelists: Peter Bouyoucos
Principal
Morgan Stanley & Company, Inc.

John Kelly
Managing Director
Citicorp Securities, Inc.

Ronald G. Keenan
Managing Director
American Re Securities Corp.

2. Predicting the Auto Underwriting Roller-Coaster

Moderator/Panelist: Patrick B. Woods
Assistant Vice President
Insurance Services Office, Inc.

Panelist: Claudette Cantin
Consulting Actuary
Tillinghast-Towers Perrin

3. Mergers and Acquisitions—Beyond the Numbers

Moderator: Michael C. Dubin
Consulting Actuary
Milliman & Robertson, Inc.

Panelists: John C. Burville
Chief Actuary
ACE, Limited

Spencer M. Gluck
Head of Actuarial Services
Swiss Re New Markets

Gail M. Ross
Vice President
AM-RE Consultants

4. The Appointed Actuary in Canada

Moderator: James K. Christie
Partner
Ernst & Young LLP

Panelists: Barbara J. Addie
President and Chief Executive Officer
Canadian Surety Company

Andrew R. Cartmell
Vice President Personal Lines
The Co-operators

Michael Hale
Director, Actuarial Division
Office of the Superintendent of Financial Institutions of Canada

Cynthia M. Potts
Partner
Eckler Partners Ltd.

The following *Proceedings* papers were presented:

1. "Implementation of PH-Transforms in Ratemaking"

Author: Shaun Wang
Associate Actuary
SCOR Reinsurance Company

2. Discussion of "Implementation of PH-Transforms in Ratemaking"

   Shaun Wang (November 1998)

   Discussion by: Gary Venter
   Executive Vice President
   Sedgwick Re

Various committee meetings were held from 1:00 p.m. to 5:00 p.m. Certain concurrent sessions that had been presented earlier during the meeting were also repeated from 1:00 p.m. to 2:30 p.m. Additional concurrent sessions presented at this time were:

1. Internet Interaction

   Moderator: Israel Krakowski
   Senior Actuary
   Allstate Insurance Company

   Panelists: J. Michael Boa
   Communications and Research
   Coordinator Casualty Actuarial Society

   LeRoy A. Boison Jr.
   Senior Vice President
   Insurance Services Office, Inc.

   Robin A. Harbage
   General Manager
   Progressive Corporation

2. General Principles of Actuarial Science

   Moderator/Panelist: Stephen W. Philbrick
   Consulting Actuary

   Panelists: Linda L. Bell
   Senior Vice President and Chief Actuary
   The Hartford

   Michael A. Walters
   Consulting Actuary
   Tillinghast-Towers Perrin

A reception and buffet dinner were held from 6:30 p.m. to 10:00 p.m. at the Royal Ontario Museum.

Wednesday, November 11, 1998

Certain concurrent sessions were repeated from 8:00 a.m. to 9:30 a.m., and the following *Proceedings* Papers were presented:

1. "Credibility With Shifting Risk Parameters, Risk Heterogeneity and Parameter Uncertainty"

   Author: Howard C. Mahler
   Vice President and Actuary
   Workers Compensation Rating and Inspection Bureau of Massachusetts

2. "Aggregation of Correlated Risk Portfolio: Models and Algorithms"

   Author: Shaun Wang
   Associate Actuary
   SCOR Reinsurance Company

After a break from 9:30 a.m. to 10:00 a.m., the final General Session was held from 10:00 a.m. to 11:30 a.m.

"Dynamic Capital Adequacy Testing"

Moderator: Stephen R. Haist
Principal
Ernst & Young

Panelists: Pierre Lepage
Consulting Actuary
Tillinghast-Towers Perrin

David J. Oakden
Senior Vice President and Chief Actuary
Zurich Canada

A. David Pelletier
Executive Vice President
RGA Life Reinsurance

Mavis A. Walters officially adjourned the 1998 CAS Annual Meeting at 11:40 a.m. after closing remarks and an announcement of future CAS meetings.

Attendees of the 1998 CAS Annual Meeting

The 1998 CAS Annual Meeting was attended by 353 Fellows, 115 Associates, and 190 guests. The names of the Fellows and Associates in attendance follow:

FELLOWS

Barbara J. Addie
Kristen M. Albright
Christiane Allaire
Craig A. Allen
Richard B. Amundson
Scott C. Anderson
Robert A. Anker
Steven D. Armstrong
Timothy Atwill
Lewis V. Augustine
Barry L. Bablin
Victoria L. Bailey
Irene K. Bass
Philip A. Baum
Andrea C. Bautista
Michael J. Bednarick
Albert J. Beer
Linda L. Bell
Phillip N. Ben-Zvi
Abbe S. Bensimon
Regina M. Berens
Wayne F. Berner
James E. Biller
Annie Blais
Cara M. Blank
Barry E. Blodgett
LeRoy A. Boison
Ronald L. Bornhuetter
Charles H. Boucek
François Boulanger
Kimberly Bowen
Douglas J. Bradac
J. Scott Bradley
Nancy A. Braithwaite
Paul Braithwaite
Yaakov B. Brauner
Robert S. Briere
Margaret A. Brinkmann
Ward M. Brooks
Ron Brusky
Christopher J. Burkhalter
Mark W. Callahan
Jeanne H. Camp
Claudette Cantin
Lynn R. Carroll
Andrew R. Cartmell
Michael J. Cascio
Tania J. Cassell
Galina M. Center
Joseph S. Cheng
David R. Chernick
Gary C. Cheung
James K. Christie
Cindy C. Chu
Allan Chuck
Kasing L. Chung
Gregory J. Ciezadlo
Brian A. Clancy
Eugene C. Connell
John B. Conners
Christopher G. Cunniff
Kenneth S. Dailey
Smitesh Davé
Jeffrey W. Davis
John D. Deacon
Curtis Gary Dean
Jerome A. Degerness
Camley A. Delach
Germain Denoncourt
Claude Desilets
Behram M. Dinshaw
Pierre Dionne
James L. Dornfeld
William F. Dove
Margaret E. Doyle
Karl H. Driedger

David L. Drury
Michael C. Dubin
Kenneth Easlon
Gary J. Egnasko
Valere M. Egnasko
Nancy R. Einck
Douglas D. Eland
David M. Elkins
Thomas J. Ellefson
Paula L. Elliott
James Ely
Martin A. Epstein
Kristine M. Esposito
Philip A. Evensen
Steven R. Fallon
Dennis D. Fasking
Sylvain Fauchon
Richard I. Fein
Stephen A. Finch
Ginda K. Fisher
Russell S. Fisher
Wayne H. Fisher
Louise A. Francis
Walter H. Fransen
Kay L. Frerk
Jacqueline F. Friedland
Michael Fusco
John E. Gaines
Alice H. Gannon
David E. Gansberg
Steven A. Gapp
Robert W. Gardner
Kathy H. Garrigan
Richard Gauthier
James J. Gebhard
Margaret Wendy Germani
John F. Gibson
Michael A. Ginnelly
Mary K. Gise
Bradley J. Gleason
Spencer M. Gluck
Leonard R. Goldberg
Moshe D. Goldberg
James F. Golz
Patrick J. Grannan
John E. Green
Steven A. Green
Daniel C. Greer
Daniel E. Greer
Cynthia M. Grim
Anthony J. Grippa
David N. Hafling
Greg M. Haft
Allen A. Hall
Leigh J. Halliwell
William D. Hansen
H. D. Hanson
Robin A. Harbage
Ellen M. Hardy
Robert L. Harnatkiewicz
David C. Harrison
William N. Herr
Betty-Jo Hill
Alan M. Hines
Daniel L. Hogan
Robert J. Hopper
Paul E. Hough
David D. Hudson
Jeffrey R. Hughes
Paul Ivanovskis
Christopher D. Jacks
Peter H. James
Joseph W. Janzen
Christian Jobidon
Eric J. Johnson
Jeffrey R. Jordan
Gary R. Josephson
Jeremy M. Jump
Stephen H. Kantor
Hsien-Ming K. Keh
Brandon D. Keller
Anne E. Kelly
Steven A. Kelner
Thomas P. Kenia
Rebecca A. Kennedy
Frederick W. Kilbourne
Joseph P. Kilroy
Gerald S. Kirschner
Bradley J. Kiscaden
Frederick O. Kist
Joel M. Kleinman
John J. Kollar
Ronald T. Kozlowski
Israel Krakowski
Rodney E. Kreps
Adam J. Kreuser
Brian S. Krick
Mary D. Kroggel
Jane J. Krumrie
Jeffrey L. Kucera
Andrew E. Kudera
Howard A. Kunst
Kenneth A. Kurtzman
Edward M. Kuss

Bertrand J. LaChance
Mylene J. Labelle
Timothy J. Landick
James W. Larkin
Robert J. Larson
Francis J. Lattanzio
Pierre G. Laurin
Guy Lecours
Robert H. Lee
Thomas C. Lee
Marc-Andre Lefebvre
Scott J. Lefkowitz
Steven G. Lehmann
Winsome Leong
Pierre Lepage
Steven J. Lesser
Kenneth A. Levine
Orin M. Linden
Barry C. Lipton
Stephen P. Lowe
Robert G. Lowery
Aileen C. Lyle
Brett A. MacKinnon
Howard C. Mahler
Gary P. Maile
Donald F. Mango
Richard J. Marcks
Lawrence F. Marcus
Joseph O. Marker
Peter R. Martin
Isaac Mashitz
Jeffrey H. Mayer
Dee Dee Mays
Michael G. McCarter
Charles W. McConnell
Sean P. McDermott
Stephen J. McGee
Kelly S. McKeethan
Kathleen A. McMonigle
Michael A. McMurray
William T. Mech
Jeffrey A. Mehalic
Brian J. Melas
David L. Menning
Stephen J. Meyer
Robert S. Miccolis
David L. Miller
David L. Miller
Mary F. Miller
Michael J. Miller
Ronald R. Miller
Scott M. Miller
Neil B. Miner
Madan L. Mittal
Anne H. Moore
Matthew S. Mrozek
Raymond D. Muller
Donna S. Munt
Thomas G. Myers
Chris E. Nelson
Karen L. Nester-Schmitt
Benjamin S. Newville
Mindy Y. Nguyen
Ray E. Niswander
Mark A. O'Brien
Mary Beth O'Keefe
David J. Oakden
Marlene D. Orr
David J. Otto
Joseph M. Palmer
Jennifer J. Palo
Dmitry E. Papush
Thomas Passante
Harry T. Pearce
Steven C. Peck
Karen L. Pehrson
Brian G. Pelly
Lynne M. Peterson
Anne M. Petrides
Stephen W. Philbrick
Mark W. Phillips
Daniel C. Pickens
On Cheong Poon
Cynthia M. Potts
Robert Potvin
Jennifer K. Price
Boris Privman
Arlie J. Proctor
David S. Pugel
Richard A. Quintano
Andre Racine
Jeffrey C. Raguse
Kara L. Raiguel
Srinivasa Ramanujam
Kiran Rasaretnam
Natalie J. Rekittke
Dennis L. Rivenburgh
Steven C. Rominske
Deborah M. Rosenberg
Kevin D. Rosenstein
Gail M. Ross
Daniel G. Roth
Richard J. Roth
Chet J. Rublewski
Thomas A. Ryan
Manalur S. Sandilya

Michael B. Schenk
Matt J. Schmitt
Harold N. Schneider
Roger A. Schultz
Arthur J. Schwartz
Susanne Sclafane
Kim A. Scott
Mark R. Shapland
Edward C. Shoop
Rial R. Simons
Raleigh R. Skaggs
Gerson Smith
Lee M. Smith
Michael B. Smith
Richard A. Smith
Richard H. Snader
Keith R. Spalding
Alan M. Speert
David Spiegler
Daniel L. Splitt
Catherine E. Staats
Barbara A. Stahley
James N. Stanard
Lee R. Steeneck
Grant D. Steer
Elton A. Stephenson
Deborah L. Stone
Ilene G. Stone
Scott J. Swanay
Susan T. Szkoda
Christopher Tait
Angela E. Taylor
Karen F. Terry
Georgia A. Theocharides
Kevin B. Thompson
Margaret W. Tiller
Michael L. Toothman
Linda K. Torkelson
Christopher J. Townsend
Nancy R. Treitel
Timothy J. Ungashick
David B. Van Koevering
Gary G. Venter
Glenn M. Walker
Mavis A. Walters
Michael A. Walters
Kimberley A. Ward
Michael R. Ward
Jeffrey C. Warren
Nina H. Webb
Dominic A. Weber
John P. Welch
Hugh G. White
Wyndel S. White
William R. Wilkins
Michael J. Williams
Gregory S. Wilson
Chad C. Wischmeyer
Kirby W. Wisian
Richard G. Woll
Patrick B. Woods
John S. Wright
Floyd M. Yager
Charles J. Yesker
Alexander G. Zhu
John D. Zicarelli
Joshua A. Zirin

ASSOCIATES

Michele S. Arndt
Mohammed Q. Ashab
Rose D. Barrett
Lee M. Bowron
Richard A. Brassington
Michelle L. Busch
Robert N. Campbell
Stephanie T. Carlson
Matthew R. Carrier
Victoria J. Carter
Bethany L. Cass
Andrew K. Chu
Kuei-Hsia R. Chu
Louise Chung-Chum-Lam
Donald L. Closter
J. P. Cochran
Mary Katherine T. Dardis
Robert E. Davis
William Der
David K. Dineen
Michael E. Doyle
Kimberly J. Drennan
Nathalie Dufresne
Rachel Dutil
Carolyn F. Edmunds
Anthony D. Edwards
Wayne W. Edwards
Jane Eichmann

Dawn E. Elzinga
S. Anders Ericson
Jonathan P. Evans
Charles V. Faerber
Kristine M. Firminhac
Michelle L. Freitag
Kai Y. Fung
Jean-Pierre Gagnon
Charles E. Gegax
Bradley G. Gipson
Terry L. Goldberg
Natasha C. Gonzalez
Francis X. Gribbon
Nasser Hadidi
Eugene E. Harrison
Gary M. Harvey
Kevin B. Held
Joseph A. Herbers
Thomas E. Hettinger
Eric J. Hornick
Brett Horoff
Melissa K. Houck
Jeffrey R. Ill
Philip W. Jeffery
Susan K. Johnston
Edwin G. Jordan
Daniel R. Kamen
Mary Jo Kannon
David L. Kaufman
Jeffrey D. Kimble
Paul E. Kinson
Brandelyn C. Klenner
Andrew M. Koren
Scott C. Kurban
David W. Lacefield
Steven M. Lacke
Douglas H. Lacoss
Todd W. Lehmann
James P. Leise
Xiaoying Liang
Jason K. Machtinger
Betsy F. Maniloff
James W. Mann
Stephen P. Marsden
Scott A. Martin
Jarow G. Myers
Charles Pare
Moshe C. Pascher
Willard W. Peacock
Richard M. Pilotte
Dylan P. Place
Karen L. Queen
Kathleen M. Quinn
James E. Rech
Brenda L. Reddick
John W. Rollins
Peter A. Royek
Brian P. Rucci
Michael Sansevero
J. S. Sawyer
Michael C. Schmitz
Michael L. Scruggs
Kelvin B. Sederburg
Kelli D. Shepard-El
Meyer Shields
Carol A. Stevenson
Avivya S. Stohl
Frederick M. Strauss
Katie Suljak
Adam M. Swartz
Richard G. Taylor
David M. Terne
Diane R. Thurston
Michael J. Toth
Andy K. Tran
Michael C. Tranfaglia
Joel A. Vaag
Steven J. Vercellini
Linda M. Waite
David G. Walker
Robert J. Wallace
Robert J. Walling
William B. Westrate
Matthew M. White
David L. Whitley
Robin D. Williams
Robert F. Wolf

REPORT OF THE VICE PRESIDENT–ADMINISTRATION

This report provides a summary of CAS activities since the 1997 CAS Annual Meeting. I will first comment on these activities as they relate to the following purposes of the Casualty Actuarial Society as stated in our Constitution:

1. Advance the body of knowledge of actuarial science in applications other than life insurance;
2. Establish and maintain standards of qualifications for membership;
3. Promote and maintain high standards of conduct and competence for the members; and
4. Increase the awareness of actuarial science.

I will then provide a summary of other activities that may not relate to a specific purpose, but yet are critical to the ongoing vitality of the CAS. Finally, I will summarize the current status of our finances and key membership statistics.

The CAS call paper programs and the publication of the *Proceedings* contribute to the attainment of the first purpose. In addition to the *Proceedings*, three volumes of the *Forum* and the Spring Meeting discussion program papers were published and distributed to members in 1998:

- The 1997 *Proceedings* contained 887 pages, the greatest number of pages yet for any *Proceedings*. Included in this volume were thirteen papers and two discussions.
- The winter 1998 edition of the *Forum* included eight ratemaking call papers.
- The summer 1998 edition of the *Forum* included nine DFA call papers as well as five papers on other topics.
- The fall 1998 edition of the *Forum* included eleven reserving call papers.

- A volume titled *Dynamic Analysis of Pricing Decisions* included five submissions from the 1998 Spring Meeting discussion paper program.

Also related to purpose 1, discussion drafts of the "General Principles of Actuarial Science," a joint effort of the CAS and SOA, were released to the membership for their comments.

The Task Force on Health and Managed Care Issues presented its report to the Board of Directors in May 1998. They recommended that the CAS sponsor research on data needs for managed care issues as they affect casualty coverages and develop methods for evaluating the initial and long-term effects of managed care on Medicare and costs paid under the various property/casualty lines of insurance. A new CAS standing committee was authorized for this area.

In regards to purpose 2, a new structure was approved for the CAS examination process. Associateship will continue to require seven exams but Fellowship will require nine exams rather than the current ten. The new structure will be effective in the year 2000.

The CAS reached agreement with the SOA on the joint sponsorship and administration of the first four exams. These four exams will be given twice a year. The remaining exams will be given once a year with parts 5, 7, and 8 in the spring and 6 and 9 in the fall.

A new class of CAS membership was created: Affiliate. Affiliate members will be able to participate as active CAS members without becoming Associates or Fellows but they will not have voting rights nor be able to use the designations ACAS or FCAS. Affiliate membership recognizes that the Affiliate Member has been granted professional status as an actuary by another actuarial organization and that he or she practices in the property/casualty field.

Purpose 3 is partially achieved through a quality program of continuing education. The CAS provides these opportunities through the publication of actuarial materials and the sponsorship of meetings and seminars. This year's sessions included:

Meetings:

	Location	*Registrants*
Spring	Marco Island, Florida	555
Annual	Toronto	521

Seminars:

Title	*Location*	*Month*	*Registrants*
Ratemaking	Chicago	March	714
Emerging Technologies	Miami Beach	April	152
Reinsurance	Greenwich, CT	June	285
Dynamic Financial Analysis	Boston	July	311
Casualty Loss Reserve	Philadelphia	September	626
CIA/CAS Appointed Actuary	Toronto	September	314
Catastrophes	New Orleans	October	166
Course on Professionalism	Five locations		224

Limited Attendance Seminars:

Title	*Location*	*Month*	*Registrants*
Managing Asset Risk and Return	Washington, D.C.	April	30
Principles of Finance	Atlanta	June	23
Loss Distributions	Chicago	July	42
Reinsurance	New York	August	65
Hands-on DFA Model: Basic	Chicago	October	35

In October 1998 the CAS sponsored the first module of the online course "Financial Risk Management for Insurers." The course had 27 registrants and was taught by Steve D'Arcy, Professor of Finance at the University of Illinois. This newly developed course was the first CAS-sponsored continuing education program to be offered online via the Internet.

A new Regional Affiliate outside of North America, Casualty Actuaries of Europe, was formed to help meet the educational and professional needs of CAS members residing in Europe. The CAS Regional Affiliates provide valuable opportunities for members to participate in educational forums at less expense and travel than national meetings and seminars.

The CAS publication, *Foundations of Casualty Actuarial Science*, is being revised. Authors were selected this past year and work has begun on the rewrite of selected chapters.

The membership approved new rules of procedure for disciplinary actions presented by an investigatory body against CAS members. These rules are intended to provide fairness and due process by requiring adequate notice, an opportunity to respond, and fair and impartial decision-makers in the discipline process.

To increase the awareness of actuarial science, purpose 4, the CAS and the SOA jointly sponsored two Actuarial Career Information Fairs, one in Philadelphia and the other in New York. Attracting minority candidates to the profession is one goal of these fairs.

The CAS Academic Correspondent Program was enhanced to provide meeting and seminar fee waivers. Grants will also be made to educational institutions when an academic staff member attains a CAS member designation. The CAS intends to build stronger ties with the academic community so that students will be more aware of the property/casualty actuarial career, more courses on property/casualty actuarial science will be offered, and professors will be motivated to perform research and write articles on property/casualty actuarial science.

The CAS Web Site supports all four purposes. The *Proceedings* and *Forums* are available online for review and downloading. New research and call papers can also be accessed online. Many of the prior *Proceedings* have been added to the online library and the CAS *Bibliographies*, *The Actuarial Review*, and *Future Fellows* are all online. A text search engine is available to search all pages within the web site for a key word or phrase.

Other features have been added to the Web site including an “Advertising” section with a job posting page. The revenue from advertising will defray part of the cost of running the Web site. Students can examine their current exam status in the online database and calculate their future status using the transition rules calculator. CAS members could respond to the annual CAS Participation Survey online beginning in 1998. An e-mail mailing list of CAS membership was also activated.

New task forces were created to monitor volunteer resources, investigate nontraditional practice areas, consider mutual recognition of actuarial credentials, and review the exam and education process and procedures. An Editorial Board was created to provide advice and counsel to the staff of *The Actuarial Review*.

A detailed survey of the CAS membership was made in 1998 to gather information and opinions on a variety of topics. The results will be compiled, analyzed, and distributed in 1999.

The CAS became a member association of the International Actuarial Association, which was restructured in June 1998 from an organization with individual members. All CAS Fellows are now members of the IAA.

Joint activities with the SOA continued. The CAS is participating on the Joint CAS/SOA Task Force on Academic Ties. A joint CAS/SOA Board meeting was held on September 17, 1998 for getting to know each other and sharing ideas. Howard Bolnick, then President-Elect of the SOA, shared the idea of a "big tent" strategy for membership in actuarial organizations. Financial engineers and other practitioners who use advanced mathematical tools for management of financial risk would be welcomed into the actuarial ranks. This controversial topic will be discussed at the 1999 CAS Leadership meeting.

New members elected to the Board of Directors for next year include Charles A. Bryan, John J. Kollar, Gail M. Ross, and Michael L. Toothman. The Board appointed Russell S. Fisher to complete the last year of Alice H. Gannon's term on the Board. The membership elected Alice H. Gannon to the position of President-Elect, while Steven G. Lehmann will assume the presidency.

The Executive Council, with primary responsibility for day-to-day operations, met either by teleconference or in person at least once a month during the year. The Board of Directors elected the following Vice Presidents for the coming year:

Vice President–Administration, Curtis Gary Dean

Vice President–Admissions, Kevin B. Thompson

Vice President–Continuing Education, Abbe S. Bensimon

Vice President–Programs and Communications, David R. Chernick

Vice President–Research and Development, Robert S. Miccolis

In closing, I will provide a brief status of our membership and financial condition. Our size continued its rapid increase as we added 177 new Associates and 144 new Fellows. Our membership now stands at 3,059. In 1998 there were 6,618 registrations for CAS exams.

The CPA firm of Langan Associates was engaged to examine the CAS books for fiscal year 1998 and its findings will be reported by the Audit Committee to the Board of Directors in February 1999. The fiscal year ended with an audited Net Income from Operations of $316,827 compared to a budgeted amount of $8,069. This higher than expected net income was the result of: (1) lower than budgeted expenses, particularly office expense; (2) more exam income from higher than expected exam enrollments; and (3) more publication sales than expected.

Members' equity now stands at $2,912,962. This represents an increase in equity of $436,281 over the amount reported last year.

For 1998–1999, the Board of Directors has approved a budget of approximately \$3.9 million. Members' dues for next year will be \$280, an increase of \$10, while fees for the Subscriber Program will increase by \$15 to \$350.

Respectfully submitted,
Curtis Gary Dean
Vice President–Administration

FINANCIAL REPORT
FISCAL YEAR ENDED 9/30/98

OPERATING RESULTS BY FUNCTION

FUNCTION	*INCOME*	*EXPENSE*	*DIFFERENCE*
Membership Services	$ 1,095,602 (a)	$ 1,056,140	$ 39,462
Seminars	1,093,604	891,812	201,792
Meetings	681,724	670,797	10,927
Exams	2,509,830 (b)	2,389,915 (b)	119,915
Publications	75,630	38,550	37,080
TOTAL	$ 5,456,390	$ 5,047,214	$ 409,176 (c)

NOTES: (a) Includes income of $92,348 to adjust marketable securities to market value (SFAS 124).
(b) Includes $1,475,850 of Volunteer Services for income and expense (SFAS 116).
(c) Change in CAS Surplus net of $52,000 of interfund transfers ($50,000 to Research Fund and $2,000 to ASTIN Fund).

BALANCE SHEET

ASSETS	*9/30/97*	*9/30/98*	*DIFFERENCE*
Checking Account	$ 237,098	$ 149,088	$ (88,010)
T-Bills/Notes	2,922,852	3,436,980	514,128
Accrued Interest	49,875	49,902	27
Prepaid Expenses	31,798	74,072	42,274
Prepaid Insurance	11,467	11,184	(283)
Accounts Receivable	13,782	39,461	25,679
Textbook Inventory	14,435	12,247	(2,188)
Computers, Furniture	270,717	313,752	43,035
Less: Accumulated Depreciation	(223,531)	(254,800)	(31,269)
TOTAL ASSETS	$ 3,328,493	$ 3,831,886	$ 503,393

LIABILITIES	*9/30/97*	*9/30/98*	*DIFFERENCE*
Exam Fees Deferred	$ 338,649	$ 388,425	$ 49,776
Annual Meeting Fees Deferred	52,860	42,246	(10,614)
Seminar Fees Deferred	21,106	61,440	40,334
Accounts Payable and Accrued Expenses	372,617	372,716	99
Deferred Rent	21,744	15,384	(6,360)
Accrued Pension	44,835	38,714	(6,121)
TOTAL LIABILITIES	$ 851,811	$ 918,925	$ 67,114

MEMBERS' EQUITY

Unrestricted	*9/30/97*	*9/30/98*	*DIFFERENCE*
CAS Surplus	$ 2,150,935	$ 2,560,111	$ 409,176
Michelbacher Fund	98,425	102,249	3,824
Dorweiler Fund	3,591	2,771	(820)
CAS Trust	18,825	19,765	940
Research Fund	154,207	166,207	12,000
ASTIN Fund	31,550	43,353	11,803
Subtotal Unrestricted	2,457,533	2,894,456	436,923
Temporarily Restricted			
Scholarship Fund	7,042	6,895	(147)
Rodermund Fund	12,106	11,611	(495)
Subtotal Restricted	19,148	18,506	(642)
TOTAL EQUITY	$ 2,476,681	$ 2,912,962	$ 436,281

C. Gary Dean, 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.

CAS Audit Committee: David N. Hafling, Chairperson;
Paul Braithwaite, Anthony J. Grippa, and Richard W. Lo

1998 EXAMINATIONS—SUCCESSFUL CANDIDATES

Examinations for Parts 3B, 4A, 4B, 5A, 5B, 6, 8-United States, 8-Canada, and 10 of the Casualty Actuarial Society were held on May 4, 5, 6, 7, and 8, 1998. Examinations for Parts 3B, 4A, 4B, 5A, 5B, 7-United States, 7-Canada, and 9 of the Casualty Actuarial Society were held on October 26, 27, 28, and 29, 1998.

Examinations for Parts 1, 2, 3A, and 3C (SOA courses 100, 110, 120, and 135, respectively) are jointly sponsored by the Casualty Actuarial Society and the Society of Actuaries. Parts 1 and 2 were given in February, May, and November 1998, and Parts 3A and 3C were given in May and November of 1998. Candidates who were successful on these examinations were listed in 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 Part 1 CAS Examination.

For the February 1998 Part 1 CAS Examination, the $200 first prize winners were: Yue Che Lau, University of Michigan; Van Khanh Le, Ohio State University; and Christina Szu-Hung Liu, Simon Fraser University. The $100 second prize winners were Kin Lun Choi, Columbia University and Michael Kayne McDermid, University of Manitoba.

For the Spring 1998 Part 1 CAS Examination, the $200 first prize winner was Yi Jing, University of Science and Technology. The $100 second prize winners were: Stephen R. Griscom, Princeton University; Heng Li, Peking University; Jun Li, Peking University; Chi S. Liu, The University of Hong Kong; and Anping Wang, University of Science and Technology.

For the Fall 1998 Part 1 CAS Examination, the $200 first prize winners were Ming Cheung Choi, The University of Hong Kong; and Dong Jun Hua, Renmin University. The $100 second prize winners were: Vincent Chen, University of Western Ontario; Yiu Wai Choy, University of Hong Kong; Jeong-Suk Im, Korea University; Xiao Dong Luo, Renmin University; and Chunping Wei, University of Science & Technology of China.

The following candidates were admitted as Fellows and Associates at the 1998 CAS Spring Meeting in May. By passing Fall 1997 CAS examinations, these candidates successfully fulfilled the Society requirements for Fellowship or Associateship designations.

NEW FELLOWS

Michael K. Curry	Man-Gyu Hur	David Molyneux
Elizabeth B. DePaolo	Andre L'Esperance	Vinay Nadkarni
Steven T. Harr	Steven W. Larson	William Peter
Daniel F. Henke	Christina Link	John S. Peters
Thomas G. Hess	Michael K. McCutchan	Michael D. Price
Marie-Josée Huard	Thomas S. McIntyre	Michael J. Steward II

NEW ASSOCIATES

Mustafa Bin Ahmad	Loren R. Danielson	Daniel Eli Greer
Nancy S. Allen	Timothy A. Davis	David J. Gronski
Wendy L. Artecona	Brian H. Deephouse	Eric C. Hassel
Carl X. Ashenbrenner	Nancy K. DeGelleke	Christopher R. Heim
David S. Atkinson	Michael B. Delvaux	Chad A. Henemyer
Craig V. Avitabile	Karen D. Derstine	Melissa K. Higgins
Phillip W. Banet	Sara P. Drexler	Tina T. Huynh
Emmanuil Bardis	Tammi B. Dulberger	Susan E. Innes
Michael W. Barlow	Francois R. Dumontet	Claudine H. Kazanecki
Gina S. Binder	Mark Kelly Edmunds	Kelly Martin Kingston
Kevin M. Bingham	Brian A. Evans	James D. Kunce
James D. Buntine	Stephen C. Fiete	Carl Lambert
Alan Burns	Sarah J. Fore	Hugues Laquerre
Hayden Burrus	Mauricio Freyre	Dennis H. Lawton
Thomas J. Chisholm	Timothy J. Friers	Manuel Alberto T. Leal
Wanchin W. Chou	Bernard H. Gilden	David Leblanc-Simard
Christopher W. Cooney	Sanjay Godhwani	Bradley R. LeBlond
Jonathan S. Curlee	Daniel Cyrus Greer	Glen A. Leibowitz

Craig A. Levitz
John N. Levy
Shiu-Shiung Lin
Victoria S. Lusk
Allen S. Lynch Jr.
Stephen J. McAnena
Jennifer A. McCurry
Mark Z. McGill III
David P. Moore
Jennifer A. Moseley
Ethan Mowry
Jarow G. Myers
Seth W. Myers
Kari A. Nicholson
John E. Noble
Jason M. Nonis
Corine Nutting
Jean-François Ouellet
Kathryn A. Owsiany
Pierre Parenteau
M. Charles Parsons
Jeremy P. Pecora
Richard M. Pilotte
Glen-Roberts Pitruzzello
Christopher D. Randall
Hany Rifai
Brad E. Rigotty
Karen L. Rivara
Rebecca L. Roever
Nathan W. Root
Kimberly R. Rosen
Richard A. Rosengarten
Seth A. Ruff
Brian C. Ryder
James C. Sandor
Gary F. Scherer
Nathan A. Schwartz
Steven G. Searle
Meyer Shields
Aviva Shneider
Alastair Shore
Matthew R. Sondag
Benoit St-Aubin
Joy M. Suh
Karrie L. Swanson
Rachel R. Tallarini
Varsha A. Tantri
Glenda O. Tennis
Laura L. Thorne
Beth S. Tropp
Kris D. Troyer
Turgay F. Turnacioglu
Leslie A. Vernon
Kyle J. Vrieze
Matthew J. Wasta
Lynne K. Wehmueller
Christopher B. Wei
Scott Werfel
Dean A. Westpfahl
Thomas J. White
Vanessa C. Whitlam-Jones
Kendall P. Williams
Yoke Wai Wong
Linda Yang

The following candidates successfully completed the Parts of the Spring 1998 CAS Examinations that were held in May.

Part 3B

Sajjad Ahmad
Jennifer A. Andrzejewski
Veronique Bouchard
Ryan M. Diehl
Jonathan Palmer Evans
Genevieve Garon
Lisa N. Guglietti
Milton G. Hickman
Carole K. L. Ho
Debra Hudson
Vibha N. Jayasinghe
Brian B. Johnson
David R. Kennerud
Peter Latshaw
Tak Yam Lee
Linda L. Maier
Kevin P. McClanahan
Ryan A. Michel
Suzanne A. Mills
Gilbert Ouellet

Jorge E. Pizarro
Carmilla T. Rivera
Michael A. Sce
James M. Smieszkal
Dennis R. Unver
John Wong

Part 4A

Josh A. Abrams
Keith P. Allen
Catherine Ambrozewicz
Vagif Amstislavskiy
Brian D. Archdeacon
Farid Aziz Ibrahim
Kris Bagchi
Damian T. Bailey
Grazyna A. Bajorska
Daniel Bar-Yaacov
John M. Barish
Mark Belasco
Richard J. Bell III
Jeremy T. Benson
Jonathan P. Berenbom
Mark W. Bingener
Neil M. Bodoff
Nebojsa Bojer
Kevin E. Branson
Melissa L. Brewer
Elaine K. Brunner
Scott T. Bruns
Christopher R. Burgess
Sarah Burns
Lori Casey
Matthew J. Cavanaugh
Raji H. Chadarevian
Hao Chai
Kevin K. W. Chan
Zoe Cheung
Daisy L. Chu
Glenn A. Colby
Jeanne L. Connolly
Thomas Cosenza
Helene Crovatto
Michael B. Cunningham
Aaron T. Cushing
Jacek Czajkowski
David W. Dahlen
Lucia De Carvalho
Donna K. DiBiaso
Brian M. Donlan
Tomer Eilam
Malika El Kacemi
Jessica L. Elsinger
Melissa M. Emmendorfer
Tricia L. Evans
Matthew B. Feldman
Sean W. Fisher
Joshua L. Fishman
Feifei Ford
Jeffrey J. Fratantaro
James M. Gallagher
Angelito P. Garcia
Michael P. Gibson
Matthew J. Gillette
William G. Golush
Linda Grand
Christa Green
Jennifer T. Grimes
John W. Grove
Manuel S. Guerra Jr.
Serhat Guven
Edward Kofi Gyampo
Rena Hartstein
Gary M. Harvey
Shrinivas Havaldar
Joshua E. Hedgecorth
Todd H. Hoivik
Douglas Bruce Homer
Hyunpyo Hong
David J. Horn Jr.
Gerald K. Howard
Chih-Che Hsiao
Todd D. Hubal
Elizabeth J. Hudson
Craig D. Isaacs
Katherine Jacques
William T. Jarman
David B. Johnson
Erik A. Johnson
Jason A. Jones
Jason C. Jones
Linda M. Kane
John B. Kelly
David R. Kennerud
Susanlisa Kessler
Chung-Hun Kim
Ziv Kimmel
Brant L. Kizer

Scott M. Klabacha
Linda S. Klenk
Steve C. Klingemann
Perry A. Klingman
John E. Kollar
Andrew M. Koren
Anand S. Kulkarni
Charles B. Kullmann
Gregory Kushnir
David J. Kwak
Ting Kwok
James A. Landgrebe
Nancy E. Lanier
Michael A. Lardis
Michael L. Laufer
Damon T. Lay
Doris Lee
Patricia Lee
Stuart Saiwah Lee
Daniel Leff
Antoine Letourneau
Monika Lietz
Herman Lim
Kenneth Lin
Enkuei Liu
Xiaoquing Iris Liu
Wan Li Lu
Yu-ping Lu
Dan Mahamah
Jaclyn B. Maher
Lynn C. Malloney
Roy M. Markham
Peter K. Markiewicz
Kenneth J. Martinez
Christopher M. Maydak
John R. McCollough
Stephane McGee
Karen J. McKenna
Michael E. McKeon
Charles A. Metzger
Camilo Mohipp
Celso M. Moreira
Craig S. Mosher
Sherry L. Mueller
Matthew D. Myshrall
Christopher A. Najim
Shannon P. Newman
Vuong V. Nguyen
Carlos E. Nunez
James P. O'Donovan
Nathalie Payette
Sue L. Poduska
Robert R. P. Pouliot
Anatoly Raklyar
Arthur R. Randolph II
Bruce A. Redmond
Charity A. Rieck
Joseph L. Rizzo
Paul J. Rostand
Hal D. Rubin
Nichole M. Runnels
Gary R. Russell
Vickie J. Scherr
Thomas W. Schroeder
Elizabeth M. Scott
Anand D. Shah
Yipei Shen
Jimmy Shkolyar
Chayanna Siripirom
Jared M. Skowron
Benjamin R. Specht
Bryan V. Spero
Alexandra R. St-Onge
Esperanza Stephens
Shelley A. Stone
Stuart C. Strauss
Ju-Young Suh
Barbara Sylvain
Jonas F. Thisner
Malgorzata Timberg
Lori S. Tinsley
Melissa K. Trost
Salvatore M. Tucci
Stephen H. Underhill
William D. Van Dyke
Karen L. VanCleave
Jennifer A. Vezza
Brian A. Viscusi
Peter R. Vita
Kate L. Walsh
David J. Watson
Thomas E. Weist
Andrew T. Wiest
Duanne A. Willis
Kelvin K. Yau
Xiangfei Zeng

Part 4B

Kedar Mulgund
Rizwan Abdul Aziz
Faisal Ahmed
Rhonda K. Ahrens
Adebowale O. Ajayi
Sami Alajaji
Emilie Alary
Richard T. Alden
Michel D. Allain
Brian M. Ancharski
Pang S. Annie
Venessa L. Archibald
Stephane Arvanitis
Keith C. Bailey
Martin Bauer
Stephen H. Beal
Jennifer A. Beattie
Esther Becker
Brian R. Bedwell
Eve Belmonte
Martine Bergeron
Robert Bhatia
Steve P. Binioris
Mike Bishop
Neil M. Bodoff
Jean-Philippe Boucher
John R. Bower
Michelle S. Brandel
Roman T. Brewka
Chris Brisebois
Peter A. Brot
Elaine K. Brunner
Cheryl R. Burrows
Lori L. Burton
Fatima E. Cadle
Lianhe Cai
Isabelle Carignan
Hao Chai
Yanick Chainey
Chi Wai Chak
Gavin Brent Chambers
Alice Y. H. Chan
Beda Chan
Ming Yan Judith M. Y. Chan
Eric Charron
Sanjeev Chaudhuri
Frank H. Chechel
Tzu-Ling Chen
Henrietta H. Cheng
Tsui-Hsien J. Chien
Allison H. Chu
Chia-Chun Chu
Hannung Chu
Kam I. Chu
Peggy P. Chung
Benjamin W. Clark
Jacques Cloutier
Nathalie Cloutier
Eliezer Cohen
Sanford K. Cohn
Robert J. Collingwood
Larry Kevin Conlee
Patrick Corbin
Carmen Csillag
Wangling Cui
Arthur D. Cummings
Alan D. Dang
Michael J. Davis
Michael S. Deckert
Douglas L. Dee
Katarzyna Deja
Peter R. DeMallie
Patrick De Roy
Patrice Denis
Robert E. Dennison
Steven M. Dersom
Mark R. Desrochers
Marie Des Roches
Benjamin Diederich
Nilesh O. Dihora
Panagiotis D. Dimitriou
Jerome Dionne
Peter C. Dolan
Crisanto A. Dorado
Courtney A. Dubbs
Philippe Dunn
Melanie Dupont
Ponniah Elancheran
Daniel Ezer
Lily H. Fang
Chad A. Fix
Susan C. Flynn
David E. Forbes
Tamasin S. Ford
Michel Fournier
Pamela A. Franz
Timothy C. French
Michael J. Gaal
Patrice Gaillardetz
Alan R. Gard

Stuart G. Gelbwasser
Lyna Gendron
Mark B. Gengenbach
William J. Gerhardt
Guy Gignac
David Patrick Glenn
Jio Young Goh
Joseph E. Goldman
William G. Golush
Christopher Groendyke
Andrew J. Haider
Genevieve Hebert
Matthew T. Henry
Patrick Henry
Susan K. Himmelman
Molly K. Hitzges
Craig D. Holcomb
Xiaoyan Hu
Todd D. Hubal
Phillipe Hudon
Shyh-ho Hung
Syed T. Hussain
Elena Ilina
Pak Chung Ip
Vladimir Y. Itkin
Alain Jacob
Jennifer J. Janvrin
Lynda E. Jeffs
Suman Jiwani
William R. Johnson
Vartivar Joukakelian
Barbara L. Kanigowski
Eric A. Keener
Brennan D. Kennedy
Sean M. Kennedy
Susanlisa Kessler
Changki Kim
Jinney Kim
Yong Woon Kim
Youngsoon Kim
Yevgeniy Kirichenko
Jeff A. Kluck
Robert A. Koth
Vasileios Koutsaftis
Vincent C. Kozlowski
Matthew R. Kuczwaj
Jennifer M. Kuehl
Karen M. Kulchyski
Edward Kuo
Mun-Bin Kuo
Kevin K. Kwok
Eric Kwong
Martin Labelle
Philippe Lagace
Crystal King Hang Lam
David Larsen
Hans H. Larsen
Louise Lavoie
Sandrine Leard
Francis Letourneau
Julia Leung
Brigitte Levasseur
Daniel A. Levin
Amanda M. Levinson
Hayden Anthony Lewis
Guodong Li
Haidong Li
Harold H. Li
Jiun-tai Li
Matthew A. Lillegard
David Grant Lim
Karina G. Limsico
Hong Lin
Hsu-feng Lin
Liqian Lin
Su-Hua Lin
Caroline D. Liu
Kathleen T. Logue
Xue Qing Lu
Yu-ping Lu
Yann Lussier
Lee-Shing Ma
Kevin B. Mahoney
Grigoriy Makarov
Gregory N. Malone
Neil P. Manning
Tousignant Marjorie
Zinoviy Mazo
Brendan S. McCallum
Stephane McGee
Pantelis N. Messolonghitis
Bryan D. Miller
Mickey Moon
Harker D. Moor
Edward M. Moore
Dean E. Murray
Jordan E. Muse
Marie-Noelle Nadeau
Darren J. Nakanishi
Simon J. Nelson
Brent S. Neville
Richard U. Newell
Lester M. Y. Ng
Sean R. Nimm
Patricia Nols

Lauree J. Nuccio
Roger S. Offerman
Robert Scott Otchere-Sarfo Jr.
Eric P. Palmer
Johan Pao
Claude Paquette
Andrea L. Pass
Kathryn L. Pate
Craig F. Pedersen
Marc Pelletier
Matthew J. Perkins
Elissa D. Perl
François Perrier
David J. Petruzzellis
Terry C. Pfeifer
Jeffrey J. Pfluger
Brian S. Piccolo
Troy S. Podraza
Peter V. Polanskyj
Sean E. Porreca
Gregory T. Preble
David Previte
Christopher M. C. Pun
Amy M. Quinn
Melanie Quintin
Carl T. Rajendram
Leonid Rasin
Dean R. Reigner
Shanour Remtoulah
Isabelle Renaud
Franklin W. Reynolds
Joann C. Ribar
Keith J. Richtik
Benoit Robert
Ezra J. Robison
Nicolas Rochon
Robert C. Roddy
Ian R. Roke
Kevin D. Roll
Charles A. Romberger
Scott I. Rosenthal
Ryan S. Rowland
Normand Roy
Zahid R. Salman
Mita B. Sandilya
Ashu Sarna
Lori R. Satov
Andrew J. Schafer
Laurie Schlenkermann
Mindy B. Schmitz
Deborah A. Schultz
Darrel W. Senior
Scott A. Shaddick
Syed Afzal A. Shah
Naixiu Shen
Yipei Shen
Osman A. Shirwany
Zandria Sia
Jose David Siberon
Kenneth L. Sidikman
Gregory J. Sikora
Stuart H. Silverman
Frederic Simard
Marc-Andre Simard
Nicolas Sirois
Elizabeth C. Skiba
Walter J. Slobojun
James M. Smieszkal
Clarissa L. Smith
Jodi Smith
Richard C. Smith
Travis R. Smith
Anthony A. Solak
Scott W. Spencer
Natalie St-Jean
Karine St-Onge
Owen J. Stein
Steven J. Stender
Michael D. Stephens
Julia M. Stetter
Karen M. Strand
Lisa D. Strobel
Thomas J. Stypla
Yu-Fan Su
Brett M. Swenson
Erica W. Szeto
John T. C. Tan
Kok-How Tan
Neeza Thandi
Jean-François Therrien
Colin A. Thompson
John J. Thompson
Christopher S. Throckmorton
Daniel J. Towriss
Katie Trahan
Catherine Tremblay
Mathiew Tremblay
Rejean Tremblay
Jean-Pierre Tremblay-Canuel
Tammy Truong
Huan P. Tseng
Chung Y. Um
Valerio Valenti
William D. Van Dyke
Carole Vincent

James I. Voelker
Youcheng Wei
Mark Weinblatt
Lisa M. Weir
Barbara Wiesler
Paul G. Winnett
Julia Lynn Wirch
Jon A. Wirkkula
Ann Min-Sze Wong
Giak Diang Tan Wong
Ha Kion Wong
Randall C. Wright
Scott E. Wright
Mihoko Yamazoe
Esther Y. Yang
Huey Wen Yang
Davout Yean
Pai-Hung Yeh
Eng Kim Yeoh
Sung G. Yim
Derek Yokota
Hamilton Yuen
Kam Chuen Yuen
King Shing Yung
Michael R. Zarember
Xiangfei Zeng
Wendy Zicari

Part 5A

Muhammad Munawar Ali
Kevin J. Atinsky
Stevan S. Baloski
Tony F. Bloemer
Lee M. Bowron
Jennifer P. Capute
Patrick J. Causgrove
Brian J. Cefola
Louise Chung-Chum-Lam
Jeffrey A. Clements
Leanne M. Cornell
Richard R. Crabb
Mary Katherine T. Dardis
Donna K. DiBiaso
Crisanto A. Dorado
Nathalie Dufresne
Rebecca E. Freitag
Graham S. Gersdorff
Neil P. Gibbons
Christie L. Gilbert
Bradley G. Gipson
Natasha C. Gonzalez
Matthew R. Gorrell
Chantal Guillemette
Edward Kofi Gyampo
Hans Heldner
Scott E. Henck
Daniel D. Heyer
Xiaohu Jiang
John J. Karwath
Karen A. Kosiba
Rocky S. Latronica
Karen J. Lee
Shangjing Li
Dengxing Lin
Jin Liu
Martin Menard
William S. Ober
Jill E. Peppers
Brian S. Piccolo
Shing Chi Poon
Michael J. Quigley
Ricardo A. Ramotar
Joe Reschini
Choya A. Robinson
Ronald J. Robinson
Richard R. Sims
Adam D. Swope
Eric D. Telhiard
Craig Tien
Andy K. Tran
Andrea E. Trimble
Brian K. Turner
Lawrence A. Vann
Youcheng Wei
Daniel Westcott
Paul D. Wilbert
Dana L. Winkler
Philip Wong
Huey Wen Yang

Part 5B

Leah C. Adams
Faisal Ahmed
Michael L. Alfred
Brian M. Ancharski
Pamela G. Anderson
Anju Arora
Nancy A. Ashmore
John D. Back
Pranav D. Badheka
Kris Bagchi
Benjamin Beckman
Toby Layne Bennington
Sarah J. Billings
Neil M. Bodoff
Alison S. Carter
Kin Lun Choi
Michael B. Cunningham
Peter R. DeMallie
David E. Dela Cruz
Brian Elliott
Kyle A. Falconbury
Elizabeth J. Fethkenher
Jeffrey R. Fleischer
Mark T. Ford
Teresa M. Fox
Rebecca E. Freitag
Dustin W. Gary
Michael L. Gish
Stephanie A. Groharing
Serhat Guven
Rebecca N. Hai
Dawn Marie S. Happ
Richard A. Haugen
Hans Heldner
Daniel D. Heyer
Mark D. Heyne
Wendy L. Hopfensperger
Candace Yolande Howell
Mohammad A. Hussain
Mohamad H. Ibrahim
Craig D. Isaacs
William T. Jarman
Erik A. Johnson
Jason A. Jones
Julie A. Jordan
Betsy J. Koestler
Ting Kwok
Julie-Linda LaForce
Peta Lewin
Carrie L. Lewis
James W. Mann
Luis S. Marques
Kevin P. McClanahan
Isaac Merchant
Deborah Ann Mergens
Vadim Y. Mezhebovsky
Ross H. Michehl
Kathleen M. Miller
Celso M. Moreira
Matthew E. Morin
Matthew D. Myshrall
John F. Pagano
Jeff D. Paggi
Jorge E. Pizarro
Aaron N. Prisco
Charity A. Rieck
Erica L. Riggs
Choya A. Robinson
Ezra J. Robison
Gail S. Rohrbach-Fink
Benjamin G. Rosenblum
Richard R. Ross
Jennifer L. Rupprecht
Tammy L. Schwartz
Mandy M. Y. Seto
Paul Silberbush
Summer L. Sipes
Clarissa L. Smith
Neeza Thandi
Nicole C. Tillyer
Dominic A. Tocci
Raymond D. Trogdon
Michael S. Uchiyama
Stephen H. Underhill
Karen L. VanCleave
Karl C. Von Brockdorff
Tice R. Walker
Robert J. Wallace
Joseph C. Wenc
So Fun Wong
Joshua C. Worsham
Mihoko Yamazoe
Jacinthe Yelle
David Zambo
Lianmin Zhou

Part 6

Jason R. Abrams
Michael B. Adams
Jodie Marie Agan
Stephen A. Alexander
Genevieve L. Allen
Scott J. Altstadt
Silvia J. Alvarez
Denise M. Ambrogio
Gwendolyn Lilly Anderson
Kevin L. Anderson
Mary K. Anderson
Jonathan L. Ankney
Michele S. Arndt
Richard T. Arnold
Afrouz Assadian
Peter Attanasio
Robert D. Bachler
Maura Curran Baker
John L. Baldan
Patrick Barbeau
Linda S. Baum
Patrick Beaudoin
Marie-Eve J. Belanger
David J. Belany
Jason E. Berkey
Ellen A. Berning
Penelope A. Bierbaum
David M. Biewer
Mary Denise Boarman
Christopher D. Bohn
Veronique Bouchard
Erick A. Brandt
Jeremy James Brigham
David C. Brueckman
Robert J. Brunson
Paul E. Budde
John C. Burkett
Derek D. Burkhalter
Matthew R. Carrier
Nathalie Charbonneau
Patrick J. Charles
Todd D. Cheema
Yvonne Wai Ying Cheng
Kin Lun Choi
Andrew K. Chu
Bernadette M. Chvoy
Susan M. Cleaver
Eric J. Clymer
Carolyn J. Coe
Christian J. Coleianne
Helaina I. Connelly
Peter J. Cooper
Kiera E. Cope
Kevin A. Cormier
Sharon R. Corrigan
Tina M. Costantino
Jeffrey A. Courchene
Spencer L. Coyle
Hall D. Crowder
David F. Dahl
John E. Daniel
Robert E. Davis
Stephanie A. DeLuca
Alain P. DesChatelets
Jean-François Desrochers
Jonathan M. Deutsch
Timothy M. DiLellio
Christopher A. Donahue
Gregory L. Dunn
Louis Christian Dupuis
Sophie Duval
Carolyn F. Edmunds
Jane Eichmann
James R. Elicker
Keith A. Engelbrecht
Greg J. Engl
Tricia G. English
Laura A. Esboldt
Juan Espadas
Weishu Fan
Solomon Carlos Feinberg
Kenneth D. Fikes
Janine A. Finan
William M. Finn
Donia N. Freese
Michelle L. Freitag
Rosemary D. Gabriel
Serge Gagne
Martine Gagnon
Cynthia Galvin
Donald M. Gambardella
Anne M. Garside
Ellen M. Gavin
Amy L. Gebauer
Hannah Gee
Charles E. Gegax

Justin G. Gensler
Keith R. Gentile
Rainer Germann
Patrick J. Gilhool
Bradley G. Gipson
Todd B. Glassman
Joseph E. Goldman
Stacey C. Gotham
Stephanie A. Gould
Elizabeth A. Grande
Joseph P. Greenwood
Mark R. Greenwood
Francis X. Gribbon
Rebecca N. Hai
Brian T. Hanrahan
Dawn Marie S. Happ
Elaine M. Harbus
Harry K. Hariharan
Michael S. Harrington
Jeffery T. Hay
Qing He
James A. Heer
Kristina S. Heer
James D. Heidt
Kandace A. Heiser
Kevin B. Held
Deborah L. Herman
Amy L. Hicks
Glenn R. Hiltpold
Kurt D. Hines
Marcy R. Hirner
Patricia A. Hladun
Brook A. Hoffman
Richard M. Holtz
Kaylie Horning
Melissa K. Houck
Terrie L. Howard
Carol I. Humphrey
Christopher W. Hurst
Scott R. Hurt
Rusty A. Husted
Philip M. Imm
Michael S. Jarmusik
Weidong Wayne Jiang
Charles B. Jin
Philippe Jodin
Paul J. Johnson
Shantelle A. Johnson
Susan K. Johnston
Steven M. Jokerst
Bryon R. Jones
Mark C. Jones
Dana F. Joseph
Daniel R. Kamen
Mary Jo Kannon
Stacey M. Kidd
Jeffrey D. Kimble
Jill E. Kirby
Susan L. Klein
Steven T. Knight
Scott C. Kurban
Isabelle LaPalme
Douglas H. Lacoss
Ravikumar Lakshminarayan
Chingyee Teresa Lam
Stephen J. Langlois
Travis J. Lappe
Jean-François Larochelle
Aaron M. Larson
Khanh M. Le
Borwen Lee
Wendy R. Leferson
James P. Leise
Christian Lemay
W. Scott Lennox
Brendan M. Leonard
Diane Lesage-Cantin
Sharon Xiaoyin Li
Xiaoying Liang
Joshua Y. Ligosky
Dengxing Lin
Jing Liu
Erik F. Livingston
Richard P. Lonardo
Aviva Lubin
Jason K. Machtinger
John T. Maher
Atul Malhotra
Joshua N. Mandell
Richard J. Manship
Albert Maroun
Stephen P. Marsden
Victor Mata
James J. Matusiak Jr.
David M. Maurer
Laura S. McAnena
Timothy L. McCarthy
Ian J. McCracken
Jennifer A. McGrath
Smith W. McKee
Mitchel Merberg
Ryan A. Michel
Rebecca E. Miller
Scott A. Miller
Ain H. Milner
Matthew K. Moran

Christian Morency
John V. Mulhall
Loren J. Nickel
Sean R. Nimm
Sylvain Nolet
Brett M. Nunes
Randall W. Oja
Sheri L. Oleshko
Helen S. Oliveto
Michael A. Onofrietti
Rodrick R. Osborn
Matthew R. Ostiguy
Apryle L. Oswald
Chad M. Ott
Robert A. Painter
Cosimo Pantaleo
Moshe C. Pascher
Lorie A. Pate
Prabha Pattabiraman
Michael T. Patterson
Wendy W. Peng
Robert B. Penwick
Sylvain Perrier
Christopher K. Perry
Kevin T. Peterson
Kraig P. Peterson
Michael R. Petrarca
Andrea L. Phillips
Kristin S. Piltzecker
Donna M. Pinetti
Frank P. Pittner
Dylan P. Place
Jayne L. Plunkett
Sean E. Porreca
Brentley J. Radeloff
Leonid Rasin
Mary Elizabeth Reading
Sara Reinmann
Teresa M. Reis
Brian E. Rhoads
David C. Riek
Stephen D. Riihimaki
Arnie W. Rippener
Delia E. Roberts
Kathleen F. Robinson
Brian P. Rucci
Bryant E. Russell
Frederick D. Ryan
Salimah H. Samji
Rachel Samoil
Michelle L. Sands
James C. Santo
Frances G. Sarrel
Jeremy N. Scharnick
Jennifer A. Scher
Daniel David Schlemmer
David B. Schofield
Jeffery W. Scholl
Jonathan A. Schriber
Ronald J. Schuler
Kelvin B. Sederburg
Vladimir Shander
Scott A. Sheldon
Kelli D. Shepard-El
Michelle L. Sheppard
Maria Shlyankevich
Joseph A. Smalley
Lora L. Smith-Sarfo
John H. Soutar
Michael D. Sowka
Anya K. Sri-Skanda-Rajah
Laura B. Stein
Gary A. Sudbeck
Edward Sypher
Stephen J. Talley
Josephine L. C. Tan
Michel Theberge
Tanya K. Thielman
Robert M. Thomas II
Gary S. Traicoff
Michael C. Tranfaglia
John D. Trauffer
Nathalie Tremblay
Kieh Treavor Ty
David Uhland
Matthew L. Uhoda
Joel A. Vaag
Jennifer L. Vadney
Richard A. Van Dyke
Justin M. Van Opdorp
Steven J. Vercellini
Mark A. Verheyen
Linda M. Waite
Amy R. Waldhauer
Kristie L. Walker
Keith A. Walsh
Shaun S. Wang
Victoria K. Ward
David W. Warren
Wade T. Warriner
Kelly M. Weber
Chris J. Westermeyer
William B. Westrate
Matthew M. White
L. Alicia Williams

Robin Davis Williams
Amy M. Wixon
Karin H. Wohlgemuth
Jeffrey S. Wood
Jonathan S. Woodruff
Scott M. Woomer
Jimmy L. Wright
Yin Zhang
Steven B. Zielke

Part 8—Canada

Martin Carrier
Christopher William Cooney
Denis Dubois
Rachel Dutil
Philippe Gosselin
Patrice Jean
Elaine Lajeunesse
Thomas L. Lee
Andrew M. Lloyd
Shawn Allan McKenzie
Hany Rifai
Nitin Talwalkar

Part 8—United States

Mustafa Bin Ahmad
John Scott Alexander
John P. Alltop
Amy P. Angell
David Steen Atkinson
Craig Victor Avitabile
Keith M. Barnes
Michael J. Bednarick
Michael J. Belfatti
Brian K. Bell
Bruce J. Bergeron
Eric D. Besman
Kristen M. Bessette
John T. Binder
Mark E. Bohrer
David R. Border
Erik R. Bouvin
Rebecca Schafer Bredehoeft
Robert F. Brown
Kevin D. Burns
Stephanie T. Carlson
Allison F. Carp
Kristi Irene Carpine-Taber
Bethany L. Cass
Tania J. Cassell
Jill C. Cecchini
Hsiu-Mei Chang
Hong Chen
Brian K. Ciferri
Maryellen J. Coggins
Kathleen T. Cunningham
Jonathan Scott Curlee
Kenneth S. Dailey
Loren Rainard Danielson
Smitesh Davé
Francis L. Decker IV
Nancy K. DeGelleke
Michael Brad Delvaux
Patricia A. Deo-Campo Vuong
Mike Devine
Michael Edward Doyle
David L. Drury
Tammi Dulberger
Brandon L. Emlen
Joseph G. Evleth
Vicki A. Fendley
Stephen A. Finch
Christine M. Fleming
Ronnie S. Fowler
David I. Frank
Kay L. Frerk
Noelle Christine Fries
Donald M. Gambardella
Kathy H. Garrigan
James B. Gilbert
Emily C. Gilde
Susan I. Gildea
Moshe D. Goldberg
Jay C. Gotelaere
John W. Gradwell
Greg M. Haft
Scott T. Hallworth
Gregory Hansen

Eric Christian Hassel
Christopher Ross Heim
David E. Heppen
Ronald J. Herrig
Todd H. Hoivik
John F. Huddleston
Paul Ivanovskis
Christopher Donald Jacks
John F. Janssen
Chad C. Karls
Claudine Helene Kazanecki
James M. Kelly
Steven A. Kelner
Joseph P. Kilroy
Eleni Kourou
James Douglas Kunce
Kenneth A. Kurtzman
Steven M. Lacke
Timothy J. Landick
Robert J. Larson
Yin Lawn
Thomas C. Lee
Scott J. Lefkowitz
Neal M. Leibowitz
Diana M. S. Linehan
Robb W. Luck
Victoria S. Lusk
James P. Lynch
Daniel Patrick Maguire
Vahan A. Mahdasian
Anthony L. Manzitto
Richard J. Marcks
David E. Marra
Peter R. Martin
Bonnie C. Maxie
Jeffrey A. Mehalic
Mark F. Mercier
Richard Ernest Meuret
Stephanie J. Michalik
Michael J. Miller
Paul D. Miotke
Christopher James Monsour
Roosevelt C. Mosley
Sureena Binte Mustafa
Timothy O. Muzzey
Mindy Y. Nguyen
Michael A. Nori
Rebecca Ruth Orsi
Kathryn Ann Owsiany
Dmitry E. Papush
M. Charles Parsons
Thomas Passante
Abha B. Patel
Michael A. Pauletti
Rosemary C. Peck
Miriam E. Perkins
Anne Marlene Petrides
Ellen K. Pierce
Richard Matthew Pilotte
Jordan J. Pitz
Richard B. Puchalski
Kathleen Mary Quinn
Kiran Rasaretnam
Peter S. Rauner
Rebecca Lea Roever
John R. Rohe
Kim R. Rosen
Seth Andrew Ruff
Kevin L. Russell
Tracy A. Ryan
Joseph J. Sacala
Elizabeth A. Sander
Jason T. Sash
Matt J. Schmitt
Michael C. Schmitz
Timothy D. Schutz
M. Kate Smith
Theodore S. Spitalnick
Catherine E. Staats
Lori E. Stoeberl
Brian Tohru Suzuki
Roman Svirsky
Karrie Lynn Swanson
Adam M. Swartz
Sebastian Yuan Yew Tan
Jennifer L. Throm
Laura M. Turner
Alice M. Underwood
Timothy J. Ungashick
Eric Vaith
Robert J. Walling III
Karen E. Watson
Patricia Cheryl White
Wendy L. Witmer
Yoke Wai Wong
Stephen K. Woodard
Yuhong Yang
Charles J. Yesker
Edward J. Zonenberg

Part 10

Mustafa Bin Ahmad
Ethan David Allen
Michael E. Angelina
Lewis V. Augustine
William P. Ayres
Barry Luke Bablin
Wayne F. Berner
Barry E. Blodgett
Kimberly Bowen
Douglas J. Bradac
Betsy A. Branagan
Ron Brusky
Julie Burdick
Hugh E. Burgess
Christopher J. Burkhalter
Cindy Cin-Man Chu
Brian A. Clancy
Kay A. Cleary
Christopher G. Cunniff
Karen Barrett Daley
Karen L. Davies
Jeffrey W. Davis
Timothy Andrew Davis
John D. Deacon
Camley A. Delach
Sean R. Devlin
Kurt S. Dickmann
Margaret Eleanor Doyle
Tammy L. Dye
Kevin M. Dyke
Mark Kelly Edmunds
Wayne W. Edwards
Kristine Esposito
Alana C. Farrell
Sylvain Fauchon
Kendra M. Felisky-Watson
Chauncey E. Fleetwood
Walter H. Fransen
John E. Gaines
Gary J. Ganci
David Evan Gansberg
Margaret Wendy Germani
John E. Green
Steven A. Green
Daniel Cyrus Greer
Daniel Eli Greer
Alexander Archibold Hammett
Ellen M. Hardy
Robert L. Harnatkiewicz
William N. Herr Jr.
Christopher Todd Hochhausler
Daniel L. Hogan Jr.
Jeffrey R. Hughes
Brian L. Ingle
Joseph W. Janzen
Jeremy M. Jump
Hsien-Ming Keh
Brandon Daniel Keller
Thomas P. Kenia
Deborah M. King
James F. King
Bradley J. Kiscaden
Brandelyn C. Klenner
Brian S. Krick
Mary D. Kroggel
Alexander Krutov
Sarah Krutov
David Leblanc-Simard
Guy Lecours
Steven J. Lesser
Richard B. Lord
Robert G. Lowery
Gary P. Maile
Dee Dee Mays
Stephen J. McGee
Scott A. McPhee
Brian James Melas
Anne Hoban Moore
Matthew S. Mrozek
Raymond D. Muller
Donald R. Musante
Henry E. Newman
James L. Nutting
Mark A. O'Brien
Mary Beth O'Keefe
Christopher Edward Olson
David Anthony Ostrowski
David J. Otto
Michael G. Owen
Joseph M. Palmer
Charles Pare
Harry Todd Pearce

Lynne M. Peterson
Kathy Popejoy
Jennifer K. Price
Anthony E. Ptasznik
David S. Pugel
Kara Lee Raiguel
Ricardo A. Ramotar
Natalie J. Rekittke
Dennis L. Rivenburgh, Jr.
John W. Rollins
Richard A. Rosengarten
Daniel G. Roth
Chet James Rublewski
Thomas A. Ryan
Manalur S. Sandilya
Michael B. Schenk
Arthur J. Schwartz
Nathan Alexander Schwartz
Craig J. Scukas
Michael Shane
Alastair Charles Shore
Bret Charles Shroyer
Gerson Smith
Matthew Robert Sondag
Alan M. Speert
Ilene G. Stone
Scott J. Swanay
Christopher C. Swetonic
Christopher Tait
Georgia A. Theocharides
Jeffrey S. Trichon
Kai Lee Tse
Jeffrey A. Van Kley
Leslie Alan Vernon
Kyle Jay Vrieze
Kimberley A. Ward
Douglas M. Warner
Wyndel S. White
William Robert Wilkins
Michael J. Williams
Kirby W. Wisian
Simon Kai-Yip Wong
Yoke Wai Wong
Vincent F. Yezzi

The following candidates were admitted as Fellows and Associates at the 1998 CAS Annual Meeting in November. By passing Spring 1998 examinations, these candidates successfully fulfilled the Society requirements for Fellowship or Associateship designations.

NEW FELLOWS

John Porter Alltop
Lewis Victor Augustine
Barry Luke Bablin
Michael James Bednarick
Michael James Belfatti
Wayne F. Berner
Barry E. Blodgett
Kimberly Ann Bowen
Douglas J. Bradac
Ron Brusky
Christopher John Burkhalter
Tania Janice Cassell
Cindy C. M. Chu
Brian Arthur Clancy
Kay A. Cleary
Christopher G. Cunniff
Kenneth Scott Dailey
Smitesh Davé
Karen L. Davies
Jeffrey Wayne Davis
John David Deacon
Camley Ann Delach
Margaret E. Doyle
David L. Drury
Tammy Lynn Dye
Kristine M. Esposito
Sylvain Fauchon
Kendra Margaret Felisky-Watson
Stephen A. Finch
Walter H. Fransen

Kay L. Frerk
John Edward Gaines
David Evan Gansberg
Kathy Helene Garrigan
Margaret Wendy Germani
Moshe David Goldberg
John E. Green
Steven A. Green
Daniel Cyrus Greer
Daniel Eli Greer
Greg M. Haft
Ellen M. Hardy
Robert L. Harnatkiewicz
William Nesthus Herr Jr.
Daniel Leo Hogan Jr.
Jeffrey R. Hughes
Paul Ivanovskis
Christopher Donald Jacks
Joseph William Janzen
Jeremy M. Jump
Hsien-Ming Keh
Brandon Daniel Keller
Steven A. Kelner
Thomas Paul Kenia
Joseph P. Kilroy
Bradley James Kiscaden
Brian Scott Krick
Mary Downey Kroggel
Alexander Krutov
Kenneth Allen Kurtzman
Timothy John Landick
Robert John Larson
Guy Lecours
Thomas C. Lee
Thomas L. Lee
Scott Jay Lefkowitz
Steven Joel Lesser
Robert Glenn Lowery
Gary P. Maile
Anthony L. Manzitto
Richard Joseph Marcks
Peter Robert Martin
Dee Dee Mays
Stephen J. McGee
Jeffrey A. Mehalic
Brian James Melas
Anne Hoban Moore
Matthew Stanley Mrozek
Raymond D. Muller
Timothy O. Muzzey
Mindy Yu Nguyen
Mark A. O'Brien
Mary Beth O'Keefe
David J. Otto
Joseph Martin Palmer
Dmitry E. Papush
Thomas Passante
Abha B. Patel
Harry Todd Pearce
Lynne M. Peterson
Anne Marlene Petrides
Jennifer K. Price
David Scott Pugel
Kara Lee Raiguel
Kiran Rasaretnam
Natalie J. Rekittke
Dennis Louis Rivenburgh Jr.
Daniel Gregg Roth
Chet James Rublewski
Kevin L. Russell
Thomas A. Ryan
Elizabeth A. Sander
Manalur Sundaram Sandilya
Michael Bruce Schenk
Matt John Schmitt
Arthur J. Schwartz
Craig James Scukas
Gerson Smith
Mary Kathryn Smith
Alan M. Speert
Catherine Elaine Staats
Ilene Gail Stone
Scott Jay Swanay
Christopher Tait
Sebastian Yuan Yew Tan
Georgia A. Theocharides
Alice Underwood
Timothy John Ungashick
Jeffrey Alan Van Kley
Kimberley A. Ward
Wyndel S. White
William Robert Wilkins
Michael J. Williams
Kirby W. Wisian
Yoke Wai Wong
Charles John Yesker

NEW ASSOCIATES

Stephen Allan Alexander
Jennifer Ann Andrzejewski
Michele Segreti Arndt
Robert Daniel Bachler
Lee Matthews Bowron
John Carol Burkett
Matthew R. Carrier
Andrew K. Chu
Louise Chung-Chum-Lam
Mary Katherine Thérèse Dardis
Robert Earl Davis
Nathalie Dufresne
Carolyn F. Edmunds
Jane Eichmann
Jonathan Palmer Evans
Michelle Lynn Freitag
Donald Michael Gambardella
Charles Edward Gegax
Bradley Gordon Gipson
Natasha Cecilia Gonzalez
Francis Xavier Gribbon
Gary Michael Harvey
Kevin Blaine Held
Melissa Katherine Houck
Weidong Wayne Jiang
Susan Kay Johnston
Daniel R. Kamen
Mary Jo Kannon
Jeffrey Dale Kimble
Andrew M. Koren
Scott C. Kurban
Douglas H. Lacoss
James Peter Leise
Diane Lesage-Cantin
Xiaoying Liang
Jason Kirk Machtinger
James William Mann
Stephen Paul Marsden
John Vincent Mulhall
Moshe C. Pascher
Judith Diane Perr
Dylan Pamphilon Place
Ricardo Anthony Ramotar
Brian Paul Rucci
Asif Sardar
Kelvin Bryce Sederburg
Kelli Denise Shepard-El
John Haldane Soutar
Andrew K. Tran
Michael Charles Tranfaglia
Joel Andrew Vaag
Steven John Vercellini
Linda M. Waite
Robert Joseph Wallace
William Boyd Westrate
Matthew Michael White
L. Alicia Williams
Robin Davis Williams

The following candidates successfully completed the Parts of the Fall 1998 CAS Examinations that were held in October.

Part 3B

Kris Bagchi
Brian J. Barth
Richard J. Bell III
Jonathan P. Berenbom
Erick A . Brandt
Monica M. Buck
Daniel R. Campbell
Kin Lun Choi
Richard R. Crabb
Marc-Andre Dallaire
Bruce R. Darling
Paul B. Deemer
Peter R. DeMallie
Jean A. DeSantis
Patrick J. Dubois

Keith R. Gentile
Christopher J. Graham
Donald B. Grimm
Serhat Guven
Edward Kofi Gyampo
Christopher W. Hurst
Susan R. Johansen-Green
Jason A. Jones
Lawrence S. Katz
John H. Kerper
Nathalie M. Lavigne
Shangjing Li
Joshua Y. Ligosky
Matthew A. Lillegard
Diana M. S. Linehan
Ratsamy Manoroth
Mea Theodore Mea
Vadim Y. Mezhebovsky
Scott A. Miller
Matthew E. Morin
Stephen J. Mueller
Sean R. Nimm
Maureen D. O'Keefe
Jean-Pierre Paquet
Kristin S. Piltzecker
Carmen Racy
Neil W. Reiss
Ezra J. Robison
David G. Shafer
Duc M. Ta
Neeza Thandi
Sheng P. Tseng
Susan B. Van Horn
Shaun S. Wang
Scott M. Woomer
Xiangfei Zeng

Part 4A

Greg A. Aikey
Anju Arora
Nancy A. Ashmore
Joel E. Atkins
Peter Attanasio
Richard Audet
Pranav D. Badheka
Christine L. Berg
Andrew W. Bernstein
David M. Biewer
Jonathan E. Bransom
Gregor L. Brown
Lisa K. Buege
Ronald Cederburg
Ramses T. Celestin
Yves Charbonneau
Tsui-Hsien J. Chien
Kin Lun Choi
Byron W. Clift
John R. Cloutier
Jeffrey R. Coker
Christopher L. Cooksey
Dustin W. Curtit
Walter C. Dabrowski
Tanuja S. Dharmadhikari
Erik L. Donahue
John A. Duffy
Lawrence K. Fink
Michael J. Fiorito
Sharon L. Fochi
Dana R. Frantz
Veronique Grenon
Adrian F. Guardado
Jennifer L. Hany
Stuart J. Hayes
Joseph Hebert
Greg Helser
Mark D. Heyne
Joseph H. Hohman
Melissa S. Holt
Carissa A. Hughes
Abraham J. Israel
Kenneth L. Israelsen
David R. James
Amy B. Kamen
Hye-Sook Kang
John J. Karwath
Inga Kasatkina
Stephen F. Katz
Amy Jieseon Kim
Roman Kimelfeld
Laurie A. Knoke
David Kodama
Todd J. Kuhl
Melissa Kuykendall
Nadya Kuzkina
Eric N. Laszlo
Annie Latouche

Geraldine Marie Z. Lejano
James J. Leonard
Stephen M. Lippman
Nataliya A. Loboda
Andrea A. Lombardi
Michael H. Loretta
Paul T. Lupica
Eric A. Madia
John C. Marques
Lora K. Massino
Michael B. McCarty
James P. McCoy
Scott A. McPhee
Thomas E. Meyer
Michael E. Mielzynski
Charles W. Mitchell
Robert E. Moran
Jason L. Morgan
Timothy C. Mosler
David B. Mukerjee
Darren J. Nakanishi
Georgia A. Nelson
Lester M. Y. Ng
Julie K. Nielsen
Stoyko N. Nikolov
Alejandra S. Nolibos
Lorie A. Pate
Michael C. Petersen
Michael R. Petrarca
David M. Pfahler
Eric W. L. Ratti
William C. Reddington
Anita E. Samuelson
Mandy M. Y. Seto
Summer L. Sipes
Robert P. Siwicki
Thomas R. Slader
Emily J. Smith
Thomas S. Stadler
Christopher J. Styrsky
Lisa Liqin Sun
Lucia Tedesco
Colleen A. Timney
Lapki Tsang
Elaine Ching Tse
Choi Nai Charlies Tu
Lien K. Tu
Kevin G. Turner
Shannon C. Vecchiarello
Sebastien Y. Vignola
Eric T. Viney
Cameron J. Vogt
Hanny C. Wai
Daniel Westcott
Todd M. Wing
Philip Wong
Jeffrey S. Wood
Shawn A. Wright
Andreas Wyler
Andrew Yershov
Sung G. Yim
Yingjie Zhang

Part 4B

Angela H. A'Zary
James E. Alford
Michael L. Alfred
Vagif Amstislavskiy
Pamela G. Anderson
Brian D. Archdeacon
Joel E. Atkins
Kris Bagchi
Vicki J. Bagley
Wendy A. Barone
Suzanne Barry
David W. Batten
Craig E. Bauer
Jonathan P. Berenbom
Elliot J. Bernstein
Mark W. Bingener
Gary R. Blackwood
Marie-Josée C. Blanchet
KayAnn Blaszcyzk
Charles W. Bloss IV
Aree K. Bly
Julie K. Bohning
Brad N. Bondy
Marcus J. Bosland
Frederic Boule
Chantal Bray
Melissa L. Brewer
Elizabeth K. Brill
Jeffrey D. Brook
Lan Z. Brown
Caroline Bruniau
Rebecca L. Burton

Ryan A. Buschert
Timothy W. Caldwell
Alyssa M. Caslick
Joseph S. Cella
Kevin K. W. Chan
Michael Tsz-Kin Chan
Whye-Loon Chan
Ming-Shiu Chang
Donna Chen
Hai Fen Chen
Shu-Chuan Chen
Yung-Chih Chen
Carrie Cheung
Su-Ying Chiang
Tracy L. Child
Yiu Wai B. Chiu
Wil Chong
Alan M. Chow
Kwok Wing Choy
Shu Hung Choy
Paul L. Cohen
Chandra B. Cole
Candace A. Cooper
Ronald M. Cornwell
Stephen M. Couzens
Richard R. Crabb
Hall D. Crowder
Dermot M. Cryan
Mary Jo Curcio
Dave J. Czernicki
David W. Dahlen
Stephanie A. DeLuca
David E. Dela Cruz
Paul M. Dennee
Milind K. Desai
Manon Desrosiers
Christopher A. Donahue
Brian M. Donlan
Thierry Duchesne
Laura A. Esboldt
Ashifa Esmail
John D. Faught
Janine A. Finan
Eric J. Fitton
Sharon L. Fochi
Feifei Ford
Guy J. Foutz
Janis L. Frazer
Elinor Friedman
Po C. Fu
Andre Gagnon
Virginia K. Gammill
Joan K. C. Ganas
Dustin W. Gary
Matthew P. Gatsch
Timothy J. Geddes
Leslie A. George
Kelly C. Geragotelis
Daniel J. Gieske
Donald L. Glick
Kevin A. Gontowski
Matthew R. Gorrell
Andrew S. Greenhalgh
Donald B. Grimm
Jared M. Gross
John W. Grove
Jocelyn Guerin
Daniel R. Guilbert
Edward Kofi Gyampo
Matthew M. Ha
Michael A. Ha
Brian P. Hall
Eric D. Halpern
Kanwal Hameed
F. Keith Handley
Robert T. Hatcher
Erik L. Heiny
Kandace A. Heiser
Scott L. Herchenbach
Mark D. Heyne
Marcy R. Hirner
Kathleen Hobbs
Margaret M. Hook
William E. Hopkins
Derek R. Hoyme
Shaohe T. Huang
Kee Wai Ip
Corrine M. Iverson
Ellen H. Jacoby
Katherine Jacques
Paul T. Jakubczak
David R. James
Steven Thomas James
Donna R. Jarvis
Huahua Jiang
Brian B. Johnson
Darrell G. Johnson
Erik A. Johnson
Kristin Lynn Johnson
Samuel Johnson
Stephane Julien
Cheryl L. Jurgens
Brian J. Kasper
David R. Kennerud
Mark S. Kertzner
Ziv Kimmel
Chris L. Kinnison

Gregory L. Kissel
Darren H. Klorfine
Laurie P. Kolb
Rostislav Kongoun
Jason M. Konopik
Sonya Koo
James A. Kosinski
Frank J. Kowal
Jodi L. Krantz
Anand S. Kulkarni
Kin Chung Kung
Gregory Kushnir
Wing Wai Kwan
François Lacroix
Joel A. Lahrman
Wen Lan Lai
Hrvoje Lakota
Margaret W. Lam
Maxime Lanctot
Frederic Lapostolle
Sylvia Wy Law
Dawn M. Lawson
Marc Leblond
Doris Lee
Patricia Lee
Tak Yam Lee
William Henry Leslie IV
Wing Wah Leung
David C. Lewis
Jennifer L. Lewis
Zhiyuan Li
Eric F. Liland
Mien Seng Lim
Shiah Shiuan Lin
Christina S.H. Liu

Daniel A. Lowen
Tony Lu
Kelly A. Lysaght
Daniel K. Ma
Sally Ann MacFadden
Michael R. Mace
Deidre A. Mangan
Steven Manilov
Zhouhong Mao
Jeffrey L. Martin
Raul G. Martin
Kenneth J. Martinez
Victor Mata
Paul G. Mattson
Timothy J. McCarthy
Robert B. McCleish IV
John R. McCollough
Kerri A. McLaughlin
Isaac Merchant
Deborah Ann Mergens
Shawna A. Meyer
Paul B. Miles
Robert Mizner
David G. Moger
Lori A. Moore
Amy J. Morehouse
Joseph Muccio
Michelle L. Myers
Jason H. Natof
Anne-Marie Nawar
Norman Niami
Daisy H. M. Ning
Alejandra S. Nolibos
Julie L. Normand
Charles A. Norton
Tom E. Norwood

Francesco Nudo
James P. O'Donovan
Yanik Paquin
Felix Patry
Joy-Ann C. Payne
Bruce G. Pendergast
Holly A. Pfeiffer
Kristin S. Piltzecker
Michael A. Polonsky
Jeffrey M. Pomerantz
Samir Prasad
Elisabeth Prince
Julie-Ann Puzzo
Bujia Qiao
Jason J. Reed
Brian E. Rhoads
Danielle L. Richards
Rodney E. Rishel
Charles Rodrigue
Randall J. Rogers
Michele S. Rosenberg
Yori B. Rubinson
Srinath Sampath
Siddhartha Sankaran
Philip Santiago
Matthew D. Schafer
Leon Schmerhold
David P. M. Scollnik
Elizabeth A. Sexauer
Bintao Shi
Alan L. Shulman
Riana Sia
Janel M. Sinacori
Frederick W. Slater
Katherine E. Slover
Ada M. So

Xi Song
Harold L. Spangler Jr.
Sean B. Staggs
Ulrich Stengele
Esperanza Stephens
John D. Stiefel
Jonathan E. Stinson
Jeffery S. Stout
Wei Hua Su
Beth M. Sweeney
Mikel M. Swyers
Barbara Sylvain
Yuan Kwan Sze
David B. Thaller
Emmerik Theriault
Bernadette M. Thie
Malgorzata Timberg
Pauline P. M. To
Marie-Josée Tremblay-Canuel
Maureen B. Tresnak
Karen J. Triebe
Erik S. Troutman
Shu-Chen Tsai
Jye-Terng Tsay
Lien K. Tu
Douglas A. Turner
William H. Turner
Alexander S. Vajda
Amy W. Van Nostrand
Nancy C. Van Walleghem
Karen L. VanCleave
Jennifer A. Vezza
Cameron J. Vogt
Colleen Ohle Walker
Matthew J. Walter
Kaicheng Wang
Mingqi Wang
Ya-Feng Wang
Javanika Patel Weltig
Joseph C. Wenc
Jay W. West
Daniel Westcott
Miles C. Williams
Stephanie J. Williams
Todd M. Wing
Tammy L. Wood
Shawn A. Wright
Jun Wu
Jennifer A. Yanulavich
Kuo-Jen Yen
Andrew Yershov
Shuk-Han Lisa Yeung
Timothy M. Yi
Lai Yuen Yip
Jinsung Yoo
Yuan Yuan
Ka Chun Yue
Paul T. Zelazoski
James J. Zheng
Huawei Jamie Zhu
Christopher D. Zuiker

Part 5A

Michael B. Adams
Amy P. Angell
Lara L. Anthony
John D. Back
Jonathan P. Berenbom
Steven G. Brenk
Alison S. Carter
Aimee B. Cmar
Natalie A. Cmar
Lynn E. Cross
Alana C. Farrell
Dale A. Fethke
Todd B. Glassman
Stacey C. Gotham
Rebecca N. Hai
Dawn Marie S. Happ
Richard A. Haugen
Henry J. Konstanty
Leland S. Kraemer
Jonathan D. Leuy
Joshua Y. Ligosky
Steven R. Lindley
Jia Liu
Wing Lowe
Tony Lu
Kevin M. Madigan
Vadim Y. Mezhebovsky
Ross H. Michehl
Jason E. Mitich
Matthew E. Morin
Randall W. Oja
Gilbert Ouellet
Ajay Pahwa

Christopher A. Pett
Ezra J. Robison
James C. Santo
Avijit Sarkar
Peter A. Scourtis
Annemarie Sinclair
Thomas M. Smith
Alexandra R. St-Onge
Lisa C. Stanley
Neeza Thandi
Jo Dee Westbrook
Xiangfei Zeng

Part 5B

Muhammad Munawar Ali
Vagif Amstislavskiy
Brandie J. Andrews
Amy P. Angell
Koosh Arfa-Zanganeh
Kevin J. Atinsky
Nabila Audi
John L. Baldan
Stevan S. Baloski
Nathalie Belanger
Mark Belasco
David M. Biewer
Nebojsa Bojer
Joseph V. Bonanno Jr.
John R. Bower
Anthony P. Brown
Jeffrey A. Brueggeman
Judith E. Callahan
Thomas L. Cawley
Brian J. Cefola
Scott R. Clark
Christopher J. Cleveland
Aimee B. Cmar
Natalie A. Cmar
Cameron A. Cook
Arthur D. Cummings
Keith R. Cummings
David B. Dalton
Rich A. Davey
Scott C. Davidson
Amy L. DeHart
Stephanie A. DeLuca
Sheri L. Delaboursodiere
Christopher P. DiMartino
Diana M. Dodu
Scott H. Drab
James R. Elicker
William H. Erdman
Dale A. Fethke
William M. Finn
Kristine M. Fitzgerald
Lisa C. Fontana
Gregory A. Frankowiak
Stuart G. Gelbwasser
Christie L. Gilbert
Todd B. Glassman
William G. Golush
Boris V. Gorelik
Matthew R. Gorrell
Stacey C. Gotham
Christa Green
Manuel S. Guerra Jr.
Timothy P. Hawkins
Brandon L. Heutmaker
Troy W. Holm
David J. Horn, Jr.
Kaylie Horning
Hai Huang
Jesse T. Jacobs
Shantelle A. Johnson
William B. Johnson
Elena Y. Karzhitskaya
Susan M. Keaveny
Sarah M. Kemp
Susanlisa Kessler
Karen A. Kosiba
Vladimir A. Kremerman
Stephen J. Langlois
Jason A. Lauterbach
Amy E. LeCount
Ruth M. LeStourgeon
Geraldine Marie Z. Lejano
Jonathan D. Leuy
Jenn Yih Lian
Kenneth Lin
Jia Liu
Jin Liu
Xiaoquing Iris Liu
Todd L. Livergood
Teresa Madariaga

Roy M. Markham
Jeffrey L. Martin
Joseph W. Mawhinney
Jeffrey B. McDonald
Patricia McGahan
Ryan A. Michel
Stephanie Miller
Bilal Musharraf
Scott L. Negus
Julie K. Nielsen
Stoyko N. Nikolov
Rodrick R. Osborn
Matthew R. Ostiguy
Robin V. Padwa
Aseem Palvia
Matthew J. Perkins
James J. Rehbit
John J. Reid
Sara Reinmann
Robert C. Roddy
Michele S. Rosenberg
Ryan P. Royce
Ray M. Saathoff
Anthony D. Salido
Dionne M. Schaaffe
Bryan K. Scott
Vladimir Shander
Scott M. Shannon
Barry D. Siegman
William A. Smyth
Anthony A. Solak
Evan M. Spiegel
Jason D. Stubbs
Nicki A. Styka
Edward T. Sweeney
Adam D. Swope
Aason A. Temples
Charles A. Thayer
Christian A. Thielman
Pantelis Tomopoulos
Isabel Trepanier
Jennifer L. Vadney
Lisa M. VanDermark
Jennifer A. Vezza
Brian A. Viscusi
Qingxian Wang
Bethany R. Webb
Carolyn D. Wettstein
Stephen C. Williams
Jonelle A. Witte
Philip Wong
Xiangfei Zeng
Yingjie Zhang

Part 7—Canada

Genevieve L. Allen
Nathalie J. Auger
Veronique Bouchard
Stephane Brisson
Nathalie Charbonneau
Yvonne Wai Ying Cheng
Steven A. Cohen
Peter J. Cooper
Sophie Duval
Serge Gagne
David Gagnon
Isabelle Gingras
Lisa N. Guglietti
Julie-Linda LaForce
Jean-Sebastien Lagace
Stephane Lalancette
Isabelle LaPalme
Christian Lemay
Julie Martineau
Shawn Allan McKenzie
Christian Menard
Sylvain Nolet
Sylvain Perrier
Sylvain Renaud
Mario Richard
Ernest C. Segal

Part 7—United States

Jason R. Abrams
Anthony L. Alfieri
Silvia J. Alvarez
Brian M. Ancharski
Gwendolyn Lilly Anderson
Paul D. Anderson
Katherine H. Antonello
Amy L. Baranek
Patrick Beaudoin
Esther Becker
Saeeda Behbahany
David J. Belany
Kristen M. Bessette
John T. Binder
Mario Binetti
Neil M. Bodoff
Christopher D. Bohn
Mark E. Bohrer
Caleb M. Bonds
David R. Border
Thomas S. Botsko
Jeremy James Brigham
Karen A. Brostrom
Conni J. Brown
Paul E. Budde
Julie Burdick
Angela D. Burgess
Derek D. Burkhalter
Anthony R. Bustillo
Allison F. Carp
Daniel G. Charbonneau
Patrick J. Charles
Todd D. Cheema
Kin Lun Choi
Julia F. Chu
Jeffrey A. Clements
Jeffrey J. Clinch
Eric J. Clymer
Carolyn J. Coe
Brian R. Coleman
Larry Kevin Conlee
Sean O. Cooper
Sharon R. Corrigan
David E. Corsi
Jeffrey A. Courchene
Jose R. Couret
Julie R. Crane
John E. Daniel
Mujtaba H. Datoo
Catherine L. DePolo
Krikor Derderian
Mark R. Desrochers
Jonathan M. Deutsch
Timothy M. DiLellio
Richard J. Engelhuber
Greg J. Engl
Weishu Fan
Brian M. Fernandes
Kenneth D. Fikes
Sean P. Forbes
Ronnie S. Fowler
Mark R. Frank
Rosemary D. Gabriel
James M. Gallagher
Anne M. Garside
Amy L. Gebauer
Justin G. Gensler
Rainer Germann
Emily C. Gilde
Cary W. Ginter
Theresa Giunta
Stephanie A. Gould
Paul E. Green
Joseph P. Greenwood
Mark R. Greenwood
Rebecca N. Hai
David L. Handschke
Bryan Hartigan
Jeffery T. Hay
Qing He
Amy L. Hicks
Jay T. Hieb
Stephen J. Higgins Jr.
Glenn R. Hiltpold
Kurt D. Hines
Glenn S. Hochler
Todd H. Hoivik
Terrie L. Howard
Long-Fong Hsu
Philip M. Imm
Craig D. Isaacs
Charles B. Jin
Karen L. Jiron
Paul J. Johnson
Bryon R. Jones
Burt D. Jones
Derek A. Jones
Dana F. Joseph
Robert C. Kane
Douglas H. Kemppainen
Sean M. Kennedy
Ung Min Kim

Jennifer E. Kish
Wendy A. Knopf
Thomas F. Krause
Ravikumar
Lakshminarayan
Chingyee Teresa Lam
John B. Landkamer
Travis J. Lappe
Aaron M. Larson
Borwen Lee
Wendy R. Leferson
Brendan M. Leonard
Karen N. Levine
Sally M. Levy
Shangjing Li
Sharon Xiaoyin Li
Dengxing Lin
Erik F. Livingston
James P. Lynch
Vahan A. Mahdasian
Atul Malhotra
Joshua N. Mandell
Jason A. Martin
Meredith Martin
James J. Matusiak Jr.
Laura S. McAnena
Timothy L. McCarthy
Kevin P. McClanahan
Jennifer A. McGrath
Rasa V. McKean
Sarah K. McNair-
Grove
Kirk F. Menanson
Rebecca E. Miller
Ain H. Milner
Rodney S. Morris
Michael W. Morro
Brian C. Neitzel
John-Giang L. Nguyen
Khanh K. Nguyen
Loren J. Nickel
Michael D. Nielsen
Sean R. Nimm
Sheri L. Oleshko
Leo Martin Orth Jr.
Apryle L. Oswald
Gerard J. Palisi
Prabha Pattabiraman
Michael A. Pauletti
Fanny C. Paz-Prizant
Rosemary C. Peck
John M. Pergrossi
Christopher K. Perry
Anthony J. Pipia
Jordan J. Pitz
Thomas L. Poklen Jr.
Peter V. Polanskyj
Sean E. Porreca
William D. Rader Jr.
Leonid Rasin
Darin L. Rasmussen
Frank S. Rau
David C. Riek
Marn Rivelle
Delia E. Roberts
Kathleen F. Robinson
Joseph F. Rosta
Janelle P. Rotondi
Robert A. Rowe
Joseph J. Sacala
Frances G. Sarrel
Jason T. Sash
Jeremy N. Scharnick
Jeffery W. Scholl
Annmarie Schuster
Tina Shaw
Seth Shenghit
Michelle L. Sheppard
Helen A. Sirois
Joseph A. Smalley
Michael W. Starke
David K. Steinhilber
Karen M. Strand
Mark R. Strona
Jayme P. Stubitz
Stephen J. Talley
Josephine L. C. Tan
Robert M. Thomas II
Christopher S.
Throckmorton
Jennifer L. Throm
Gary S. Traicoff
John D. Trauffer
Andrea E. Trimble
Brian K. Turner
Peggy J. Urness
Mark A. Verheyen
Jon S. Walters
Douglas M. Warner
David W. Warren
Wade T. Warriner
Kevin E. Weathers
Kelly M. Weber
Arthur S. Whitson
William B. Wilder
Trevar K. Withers
Jonathan S. Woodruff
Perry K. Wooley
Yin Zhang
Steven B. Zielke

Part 9

Mustafa Bin Ahmad
Nancy L. Arico
Carl Xavier Ashenbrenner
Craig Victor Avitabile
Phillip Wesley Banet
Emmanuil Theodore Bardis
Michael William Barlow
Paul C. Barone
Rick D. Beam
Anna Marie Beaton
Nicolas Beaupre
Andrew S. Becker
Lisa A. Bjorkman
Michael J. Bluzer
Lesley R. Bosniack
Lori Michelle Bradley
Betsy A. Branagan
Russell J. Buckley
Hugh E. Burgess
Elliot R. Burn
Kevin D. Burns
Pamela A. Burt
Janet P. Cappers
Stephanie T. Carlson
Hsiu-Mei Chang
Eric D. Chen
Richard M. Chiarini
Christopher William Cooney
Loren Rainard Danielson
Sheri L. Daubenmier
Timothy Andrew Davis
Brian Harris Deephouse
John T. Devereux
Kevin Francis Downs
Peter F. Drogan
Wayne W. Edwards
Brandon L. Emlen
Jonathan Palmer Evans
Bruce D. Fell
Kevin Jon Fried
Noelle Christine Fries
Jean-Pierre Gagnon
Donald M. Gambardella
Christopher H. Geering
Bradley G. Gipson
Stewart H. Gleason
Sanjay Godhwani
Natasha C. Gonzalez
Robert A. Grocock
James C. Guszcza
Kenneth Jay Hammell
Christopher Ross Heim
David E. Heppen
Daniel D. Heyer
Amy L. Hoffman
Jane W. Hughes
Jason Israel
Walter L. Jedziniak
Philip W. Jeffery
Philippe Jodin
Susan K. Johnston
Sally M. Kaplan
Robert B. Katzman
Claudine Helene Kazanecki
James M. Kelly
Scott Andrew Kelly
Deborah M. King
Eleni Kourou
Claudia A. Krucher
Sarah Krutov
James Douglas Kunce
Steven M. Lacke
Marc La Palme
Yin Lawn
Dennis H. Lawton
Bradley R. LeBlond
Betty F. Lee
Kevin A. Lee
Todd William Lehmann
Neal M. Leibowitz
Bradley H. Lemons
Charles Letourneau
Janet G. Lindstrom
Richard B. Lord
Cara M. Low
Michelle Luneau
James W. Mann
William A. Mendralla
Michael J. Miller
David Patrick Moore
Sureena Binte Mustafa
Seth Wayne Myers
Kari S. Nelson

Kari A. Nicholson
Darci Z. Noonan
Rebecca Ruth Orsi
David Anthony Ostrowski
M. Charles Parsons
James Alan Partridge
Kerry S. Patsalides
John R. Pedrick
Luba O. Pesis
Jeffrey J. Pfluger
Anthony George Phillips
Richard Matthew Pilotte
Glen-Roberts Pitruzzello
Dylan P. Place
Christopher David Randall
Hany Rifai
Rebecca Lea Roever
John W. Rollins
Seth Andrew Ruff
Cindy R. Schauer
Christy Beth Schreck
Nathan Alexander Schwartz
Michael Shane
Matthew Robert Sondag
Lori E. Stoeberl
Katie Suljak
Adam M. Swartz
Christopher C. Swetonic
Nitin Talwalkar
Kai Lee Tse
Kieh Treavor Ty
Leslie Alan Vernon
Martin Vezina
Kyle Jay Vrieze
Claude A. Wagner
Tice R. Walker
Isabelle T. Wang
William B. Westrate
Vanessa Clare Whitlam-Jones
Jerelyn S. Williams
Kendall P. Williams
Laura Markham Williams
Kah-Leng Wong
Jeffrey F. Woodcock
Linda Yang

NEW FELLOWS ADMITTED IN MAY 1998

First row, from left: Thomas G. Hess, Marie-Josee Huard, David Molyneux, Elizabeth B. DePaolo, Christina Link, Michael J. Steward II. **Second row, from left:** Man-Gyu Hur, Vinay Nadkarni, **Mavis A. Walters**, Steven T. Harr, Daniel F. Henke. **Third row, from left:** Andre L'Esperance, John S. Peters, William Peter, Steven W. Larson. **New Fellows admitted in May 1998 who are not pictured:** Michael K. Curry, Michael K. McCutchan, Thomas S. McIntyre, Michael D. Price.

New Associates Admitted in May 1998

First row, from left: Kevin M. Bingham, Dean A. Westpfahl, Nathan A. Schwartz, Tammi B. Dulberger, Kyle J. Vrieze, Timothy A. Davis, Nancy S. Allen, Melissa K. Higgins. **Second row, from left:** Allen S. Lynch Jr., Christopher W. Cooney, Benoit St-Aubin, Turgay F. Turnacioglu, **CAS President Mavis A. Walters**, Jason M. Nonis, Carl X. Ashbrenner, Alan Burns, Steven G. Searle. **Third row, from left:** Brian A. Evans, Alastair Shore, Nancy K. DeGelleke, Wanchin W. Chou, Brad E. Rigotty, Leslie A. Vernon. **Fourth row, from left:** Timothy J. Friers, Bernard H. Gilden, Francois R. Dumontet, Kendall P. Williams, Thomas J. Chisholm.

NEW ASSOCIATES ADMITTED IN MAY 1998

First row, from left: Glen A. Leibowitz, Linda Yang, Gary F. Scherer, John E. Noble, Sara P. Drexler, Aviva Shneider, Jennifer A. Moseley, Beth S.Tropp. **Second row, from left:** Manuel Alberto T. Leal, Brian H. Deephouse, Thomas J. White, Phillip W. Banet, **CAS President Mavis A. Walters**, Jean-Francois Ouellet, Pierre Parenteau, John N. Levy, Emmanuil Bardis. **Third row, from left:** Matthew J. Wasta, Kimberly R. Rosen, Wendy L. Artecona, Jennifer A. McCurry, Jonathan S. Curlee, David J. Gronski, Daniel E. Greer, Craig A. Levitz. **Fourth row, from left:** Loren R. Danielson, Jeremy P. Pecora, Christopher R. Heim, Craig V. Avitabile, Stephen J. McAnena.

NEW ASSOCIATES ADMITTED IN MAY 1998

First row, from left: Daniel Cyrus Greer, Eric C. Hassel, Kathryn A. Owsiany, Karrie L. Swanson, Gina S. Binder, Michael B. Delvaux. **Second row, from left:** Karen D. Derstine, Richard A. Rosengarten, Ethan Mowry, **CAS President Mavis A. Walters**, Tina T. Huynh, Corine Nutting, Dennis H. Lawton. **Third row, from left:** Lynn K. Wehmueller, Glenda O. Tennis, Mauricio Freyre, James D. Kunce, Mark Kelly Edmunds, Matthew R. Sondag. **Fourth row, from left:** Seth A. Ruff, Sanjay Godhwani, Hayden Burrus, James C. Sandor, Scott Werfel.

New Associates Admitted in May 1998

First row, from left: Michael W. Barlow, Nathan W. Root, Rachel R. Tallarini, David P. Moore, Chad A. Henemyer, David S. Atkinson. **Second row, from left:** Rebecca L. Roever, Susan E. Innes, Varsha A. Tantri, **CAS President Mavis A. Walters**, Kris D. Troyer, Seth W. Myers. **Third row, from left:** Glen-Roberts Pitruzzello, Claudine H. Kazanecki, Vanessa C. Whitlam-Jones, M. Charles Parsons, Christopher D. Randall. **New Associates admitted in May 1998 who are not pictured:** Mustafa Bin Ahmad, James D. Buntine, Stephen C. Fiete, Sarah J. Fore, Kelly Martin Kingston, Carl Lambert, Hugues Laquerre, David Leblanc-Simard, Bradley R. LeBlond, Shiu-Shiung Lin, Victoria S. Lusk, Mark Z. McGill III, Jarow G. Myers, Kari A. Nicholson, Richard M. Pilotte, Hany Rifai, Karen L. Rivara, Brian C. Ryder, Meyer Shields, Joy M. Suh, Laura L. Thorne, Christopher B. Wei.

New Fellows Admitted November 1998

First row, from left: Mary Beth O'Keefe, Mary Downey Kroggel, Anne Marlene Petrides, Natalie J. Rekittke, **CAS President Mavis A. Walters**, Kristine M. Esposito, Ilene Gail Stone, Mindy Yu Nguyen, Margaret Wendy Germani. **Second row, from left:** Richard Joseph Marcks, Brian Arthur Clancy, Joseph Martin Palmer, Dmitry E. Papush, Mark A. O'Brien, Manalur Sundaram Sandilya, Michael J. Williams, Christopher John Burkhalter, Kimberley Ann Bowen. **Third row, from left:** Barry Luke Bablin, William Nesthus Herr Jr., William Robert Wilkins, Harry Todd Pearce, Greg M. Haft, Charles John Yesker, Hsien-Ming Keh, Michael Bruce Schenk, Robert Glenn Lowery, Kirby W. Wisian, Chet James Rublewski.

NEW FELLOWS ADMITTED NOVEMBER 1998

First row, from left: Christopher G. Cunniff, Brandon Daniel Keller, Douglas J. Bradac, Arthur J. Schwartz, **CAS President Mavis A. Walters**, Lynne M. Peterson, Jennifer K. Price, Christopher Donald Jacks, Kathy Helene Garrigan. **Second row, from left:** Kiran Rasaretnam, Timothy John Ungashick, Jeffrey Wayne Davis, David Evan Gansberg, John E. Green, Steven A. Green, Steven Joel Lesser, David Scott Pugel, Christopher Tait, Kara Lee Raiguel, Jeremy M. Jump. **Third row, from left:** Alan M. Speert, John Edward Gaines, Peter Robert Martin, Daniel Leo Hogan Jr., Joseph P. Kilroy, Timothy John Landick, Stephen J. McGee, Dennis Louis Rivenburgh Jr., Daniel Cyrus Greer.

NEW FELLOWS ADMITTED NOVEMBER 1998

First row, from left: Catherine Elaine Staats, Margaret E. Doyle, Tania Janice Cassell, Ellen M. Hardy, **CAS President Mavis A. Walters**, John David Deacon, Scott Jay Swanay, Smitesh Dave, Camley Ann Delach. **Second row, from left:** Michael James Bednarick, Kay L. Frerk, Kimberley A. Ward, Anne Hoban Moore, Bradley James Kiscaden, Kenneth Scott Dailey, Kenneth Allen Kurtzman, Guy Lecours, Raymond D. Muller, Dee Dee Mays. **Third row, from left:** Wayne F. Berner, Matthew Stanley Mrozek, Brian Scott Krick, Jeffrey A. Mehalic, Sylvain Fauchon, Jeffrey R. Hughes, Michael James Belfatti, Steven A. Kelner, Barry E. Blodgett, Paul Ivanovskis, Matt John Schmitt.

New Fellows Admitted November 1998

First row, from left: Moshe David Goldberg, Gerson Smith, Cindy C.M. Chu, Georgia A. Theocharides, **CAS President Mavis A. Walters**, Wyndel S. White, Lewis Victor Augustine, Thomas A. Ryan, Thomas Passante. **Second row, from left:** Brian James Melas, Thomas Paul Kenia, Robert John Larson, Gary P. Maile, Daniel Gregg Roth, Ron Brusky. **Third row, from left:** Scott Jay Lefkowitz, Stephen A. Finch, Daniel Eli Greer, Thomas C. Lee, David J. Otto, Joseph William Janzen, Walter H. Fransen. **New Fellows admitted in November 1998 who are not pictured:** John Porter Alltop, Kay A. Cleary, Karen L. Davies, David L. Drury, Tammy Lynn Dye, Kendra Margaret Felisky-Watson, Robert L. Harnatkiewicz, Alexander Krutov, Thomas L. Lee, Anthony L. Manzitto, Timothy O. Muzzey, Abha B. Patel, Kevin L. Russell, Elizabeth A. Sander, Craig James Scukas, Mary Kathryn Smith, Sebastian Yuan Yew Tan, Alice Underwood, Jeffrey Alan Van Kley, Yoke Wai Wong.

NEW ASSOCIATES ADMITTED IN NOVEMBER 1998

First row, from left: Kelli Denise Shepard-El, Linda M. Waite, Michele Segreti Arndt, Louise Chung-Chum-Lam, **CAS President Mavis A. Walters**, Natasha Cecilia Gonzalez, Michelle Lynn Freitag, Susan Kay Johnston, James William Mann. **Second row, from left:** Francis Xavier Gribbon, Gary Michael Harvey, Jonathan Palmer Evans, Lee Matthews Bowron, Matthew R. Carrier, Scott C. Kurban, Brian Paul Rucci, Joel Andrew Vaag, Andrew K. Tran.

New Associates Admitted in November 1998

First row, from left: Kevin Blaine Held, Carolyn F. Edmunds, Mary Jo Kannon, Melissa Katherine Houck, **CAS President Mavis A. Walters**, Robin Davis Williams, Nathalie Dufresne, Daniel R. Kamen, Andrew K. Chu. **Second row, from left:** Steven John Vercellini, Mary Katherine Thérèse Dardis, Robert Earl Davis, Matthew Michael White, Moshe C. Pascher, James Peter Leise, Jeffrey Dale Kimble, Charles Edward Gegax, Dylan Pamphilon Place, Kelvin Bryce Sederburg. **Third row, from left:** Robert Joseph Wallace, Jason Kirk Machtinger, Bradley Gordon Gipson, William Boyd Westrate, Jane Eichmann, Stephen Paul Marsden, Douglas H. Lacoss, Michael Charles Tranfaglia, Joe (Xiaoying) Liang. **New Associates admitted in November 1998 who are not pictured:** Stephen Allan Alexander, Jennifer Ann Andrzejewski, Robert Daniel Bachler, John Carol Burkett, Donald Michael Gambardella, Weidong Wayne Jiang, Andrew M. Koren, Diane Lesage-Cantin, John Vincent Mulhall, Judith Diane Perr, Ricardo Anthony Ramotar, Asif Sardar, John Haldane Soutar, L. Alicia Williams.

OBITUARIES

JAMES M. CAHILL
CHARLES DESJARDINS
K. ARNE EIDE
ROBERT C. PERRY
PAUL E. SINGER
EMIL J. STRUG
PAUL A. VERHAGE
HERBERT E. WITTICK

JAMES M. CAHILL
1905–1998

James M. Cahill died on September 28, 1998 at Inglemoor Care Center in Livingston, New Jersey. He was 92.

Born in Hartford, Connecticut, on November 16, 1905, Cahill lived in Scarsdale, New York, before moving to Ramsey, New Jersey. He lived in Ramsey for 50 years before moving to Livingston in 1995.

After graduating from Hartford Public High School in 1923, Cahill attended Trinity College in Hartford. At Trinity, Cahill was named valedictorian of the class of 1927. He also set the school's quarter-mile track record.

Cahill began his career in 1927 with the Travelers Insurance Company in Hartford. He earned his CAS Fellowship in 1929. Cahill joined the Compensation Rating Board in New York City in 1938, progressing to the National Bureau of Casualty Underwriters (NBCU) in 1944. At NBCU he served as secretary and later general manager.

Cahill played a pivotal role in consolidating the NBCU and the National Automobile Underwriters Association into the Insurance Rating Board in 1968, serving as its first general manager until his retirement in 1971.

An active member of the Casualty Actuarial Society, Cahill served as president from 1947–48, as a vice president from 1945–46, and as a member of the CAS Council from 1938–41. He wrote several *Proceedings* papers including "Product Public Liability Insurance" (*PCAS* XXI), which won the Richard Fondiller Prize in 1934–35. Cahill also participated as a panelist in CAS meeting sessions.

During his CAS presidency, Cahill took a great interest in mentoring mathematics students by initiating apprentice programs. In his presidential address on November 19, 1948, he predicted an increased need for casualty actuaries and encouraged members to mentor young people who were considering careers as casualty actuaries.

Cahill is survived by two daughters, Susan Ramsey of Wayne, New Jersey, and Barabara Melendez of Phoenix, Arizona; four grandchildren; and seven great grandchildren. His wife, Mildred (nee Potter) Cahill, predeceased him in 1993.

CHARLES DESJARDINS
1962–1998

Charles Desjardins died February 20, 1998 at the age of 35.

Born March 11, 1962, Québec City, Québec, Desjardins graduated from Laval University in Montréal in 1983 with a degree in actuarial science.

In 1983, Desjardins worked for Fireman's Fund Insurance Company of Canada, and from 1983 to 1985 he worked in the ratemaking department for Wellington Insurance Company. While at Wellington he was responsible for implementing and updating the company's loss reserving databases and evaluating its IBNR needs. Desjardins moved to Commercial Union in Toronto, Ontario in 1985 as an actuarial assistant. Desjardins became an Associate of the Casualty Actuarial Society in 1989.

Christopher J. Townsend (FCAS 1986), a colleague of Desjardins at Commercial Union, recalled his clever wit. In particular, Townsend remembered Desjardins' definition of a large loss as being a "failure to meet plan." Although Desjardins left the actuarial field in 1992, he continued as a member of Ontario Conference of Casualty Actuaries and the Casualty Actuarial Society.

A close friend of Desjardins, Andrew Hamilton, described him as the "life of every party." "At every gathering, you would be sure to see him, fashionably dressed, talking a mile a minute and surrounded by people," said Hamilton. Desjardins was a thoughtful man who never forgot birthdays or anniversaries and whose gifts were always beautifully appropriate to the occasion.

"Charles was a brave man," said Hamilton. "Living with a terminal illness was both difficult and debilitating, but Charles remained cheerful, optimistic, and concerned about others. Few of his friends knew how ill he was because he hid it so well. He was always asking about his friends, rather than dwelling on his own problems. We miss him."

Desjardins is survived by his parents, Jeanne and Raymond Desjardins of Sillery, Québec; and a brother, Jacques of Québec City.

KNUT ARNE EIDE
1912–1997

Knut Arne Eide died November 8, 1997. He was 85.

Eide was born July 10, 1912 in Fertile, Minnesota, the oldest of eight children. He graduated from Fertile High School in 1929 and, from 1929 to 1933, he attended Luther College in Decorah, Iowa, earning a bachelors degree in liberal arts. While at Luther College, Eide was active in the Boy Scouts as a troop leader and in the college band. Eide continued to support the college's music program many years after he graduated.

During the summer of 1937, Eide did some graduate course work at the University of Minnesota in Minneapolis. He attended the University of Iowa in Iowa City and earned a master's degree in 1939.

Eide joined the Army in 1942 and began officer training. During the World War II he served in the Quartermaster Corps in Europe. After his wartime service, Eide served in the Army Reserves for more than 20 years, retiring as a lieutenant colonel.

Former SOA president Gilbert Fitzhugh (FCAS 1935) sponsored Eide when he became an Associate of the Casualty Actuarial Society in 1954. Eide was a member of the American Academy of Actuaries and earned Fellowship in the CAS and the SOA in 1959 and 1967, respectively. He served as the editor of the SOA's *Statistical Bulletin* for many years.

During the 1950s, Eide began a long career with Metropolitan Life Insurance Company in New York City. He began as an actuarial clerk and by his retirement in 1977 he was assistant vice president.

Eide left New York in 1977 to retire in Fertile. During his retirement Eide was active in the community serving as president of the Lion's Club, member of the Fair Meadow Nursing Home board, substitute math teacher for the local high school, and as a member of the Concordia Lutheran Church. In 1992, Eide was honored with the Citizen of the Years Award by the Fertile Community Club. The award is given to individuals exhibiting outstanding community service over several years.

Eide is survived by two brothers, Alf of Fertile, and Arvid of Morris, Minnesota; and three sisters, Anna Pagenhart of Rochester, Minnesota, Alfhild O'Malley of Minneapolis, and Agnes Cranmer of Vista, California. Two brothers, Roald and Anders, predeceased him.

ROBERT C. PERRY
1910–1997

Robert C. Perry died December 7, 1997 in Tucson, Arizona, at the age of 87.

He was born October 12, 1910 in Bloomington, Illinois, and raised in Downs, Illinois, where he attended public school. In 1932, Perry graduated from the University of Illinois.

For 43 years, Perry was employed in the home office of the State Farm Insurance Company in Bloomington, Illinois. Throughout his years with the company, he served as vice president and actuary, executive vice president, and as vice chairman and secretary. Perry retired to Tucson in 1976.

Perry became an Associate of the Casualty Actuarial Society in 1947 and a Fellow of the Society of Actuaries in 1948. He was also a Fellow of the Life Office Management Institute, a Chartered Life Underwriter, and a member of the American Academy of Actuaries. Perry was active in international life insurance trade associations and was past chairman of the Life Office Management Association. He was also a member of the Canadian Institute of Actuaries and of the Chicago Actuarial Club.

Among his other community activities, Perry was a 32-degree Mason and Shriner, a former member of Bloomington Chamber of Commerce and the Skyline Chamber Commerce, and a member of Theta Chi Fraternity.

Perry's wife of 61 years, Harriett, preceded him in death in 1995.

PAUL E. SINGER
1923–1998

Paul E. Singer died March 21, 1998. He was 74.

Singer was born April 26, 1923 in Chicago, Illinois, and lived there for much of his life. He attended the University of Chicago during the 1940s, graduating with a degree in liberal arts. He married Jean Henkel June 19, 1948, and together they had four children.

Singer worked for several years for Continental Casualty in Chicago, which eventually became CNA Insurance. While working for Continental Casualty, Singer earned his Associateship in the Casualty Actuarial Society in 1963. He then went on to become assistant vice president of National Insurance Group in 1964, an actuary with Continental National American Group in 1967, vice president and actuary for CNA Insurance in 1975, and CNA vice president in 1976. From 1977 until his retirement in 1985, Singer was president of Illinois State Medical Insurance Services, Inc. After retirement, Singer also lent his expertise as an expert witness.

From the mid-1960s until well into the 1970s, Singer was an active member of the CAS, participating in several CAS meetings as a member of various panels. Some of the panels included "Accident and Health Development" (1965), "Government Medical Assistance Programs" (1967), "National Health Insurance" (1971), and "Recognition of Anticipated Investment Income" (1979). Singer was also active in CAS committees, serving on the Committee on Social Insurance (1965–67), the Committee to Study Forms of Amalgamation (1971), and the Committee on Financial Reporting (1974–78). "He was really proud to be a member [of the Casualty Actuarial Society]," said Singer's daughter, Margaret Piper. "His pride in it was obvious."

A bicycling and bridge enthusiast, Singer enjoyed crossword, double crostic, and mathematical puzzles. Singer and his wife spent many years living in Michigan after he retired.

Singer's survivors include two daughters, Margaret of Wheaton, Illinois, and Elise of Philadelphia, Pennsylvania; and eight grandchildren. Singer's wife Jean and two sons, Martin and William, preceded him in death.

EMIL J. STRUG
1924–1998

Emil J. Strug died at his home in Stoughton, Massachusetts on March 14, 1998 after a four-year battle with cancer. He was 73.

Strug was born September 2, 1924 in Mendon, Massachusetts, and attended Cathedral High School. During World War II, Strug served as first lieutenant and navigator with the 44th Bomb Group of the 8th Air Corps. He flew missions over Europe from his station in Norwich, England.

After the war, Strug attended Boston College, graduating in 1950 with a degree in mathematics. That same year he married Eleanor M. Misiewicz on September 10. The couple had four children.

After working briefly for Liberty Mutual Fire Insurance Company in Boston, Massachusetts, Strug began a thirty-year career with what would become Blue Cross/Blue Shield in Boston. He began as assistant actuary and later served as manager of actuarial-statistical, and associate vice president and associate actuary for the company. In 1988, he became vice president and treasurer for Benefit Management Association, Inc. in Rockland, Massachusetts. He retired in 1991.

Strug was an active member of the Casualty Actuarial Society and frequently attended CAS meetings, during which he forged strong friendships with his fellow members. “He had some very good times with the group at CAS meetings,” said Strug’s wife Eleanor. During the evenings at CAS meetings, Strug’s friends and colleagues would often perform musical numbers and plays or would gather around a piano to sing.

A frequent contributor to the *Proceedings*, Strug’s papers include “Joint Underwriting as a Reinsurance Problem” (1972), “Determining Ultimate Claim Liabilities for Health Insurance

Coverages" (1980), and "The Pricing of Medi-Gap Coverage" (1983).

Strug became an Associate of the Casualty Actuarial Society in 1959 and a Fellow in 1970. He was also a member of the American Academy of Actuaries, the International Actuarial Association, the 44th Bomb Group Association, and Veterans of Foreign Wars.

He is survived by his wife, Eleanor; three sons, Stephen J. of Halifax, Massachusetts, Christopher G. of Westboro, Massachusetts, and Scott D. of White Plains, New York; a daughter, Susan M. Keshian of North Andover, Massachusetts; a sister, Laura Grant of Randolph, Massachusetts; and nine grandchildren. The family requested that donations be made to Dana Farber Cancer Institute in Boston.

PAUL A. VERHAGE
1938–1998

Paul A. Verhage died September 25, 1998. He was 59.

Born November 8, 1938 in Kankakee, Illinois, Verhage attended Northwest High School in Sheboygan, Wisconsin and the University of Wisconsin in Milwaukee. He was a member of the American Academy of Actuaries and became an Associate of the Casualty Actuarial Society in 1962 and a Fellow in 1965.

While attending church in Vesper, Wisconsin, Verhage met his future wife, Patricia. The two were married on September 24, 1960. Together they had two sons.

After graduation, Verhage worked briefly for Northwestern Mutual Life in 1960 and then began working as an actuarial technician for Hardware Mutuals in Stevens Point, Wisconsin, which eventually became Sentry Insurance Group. Verhage worked for Sentry over 33 years, most recently as senior vice president and chief financial officer. In his last position with Sentry he was responsible for investment management, accounting, treasury, internal audit, and strategic planning.

After Sentry, Verhage became senior manager and consulting actuary for Arthur Andersen's property/casualty actuarial consulting group in Milwaukee, Wisconsin.

Verhage was active in the CAS as a *Proceedings* author (Volume LIII) and as a committee member. He served on the two sections of the CAS Education and Examination Committee from 1966 to 1975. He was also involved in the CAS Committee to Identify Interest Areas in 1971 and the Finance Committee from 1975 to 1977.

John H. Muetterties (FCAS 1956), consulting actuary for MBA Inc. in Mountain Lakes, New Jersey, hired Verhage out of college for Hardware Mutuals. "He was a very bright individual," said Muetterties. "I'd rate him in the upper five percent of casualty actuaries." Muetterties told how Verhage, who had no background in electronics, took one of his company's first computers and successfully upgraded it. "That was a big step in the 1960s!" said Muetterties.

A modest and quiet family man, Verhage stood well over six feet tall. He was dedicated to his work and liked traveling with family and friends. In particular, he and his wife enjoyed a trip to London with their two sons and their daughters-in-law.

Verhage was also interested in genealogy. He was intrigued with discovering the genealogical tie to a close friend who shared his surname. While traveling in Holland with their wives, the two men traced their ancestries back to the year 1698 to find a common link.

Verhage is survived by his wife Patricia; two sons Peter of Stevens Point and Paul Jr. of North Liberty, Iowa; mother Minnie Verhage of Sheboygan Falls, Wisconsin; sister Annette Kashney also of Sheboygan Falls; and one grandchild. In lieu of flowers, the family requested that donations be given to the American Cancer Society.

HERBERT E. WITTICK
1903–1998

Herbert E. Wittick was born June 25, 1903 in Peoria, Illinois. He attended Bradley Polytechnic Institute in Peoria from 1916 to 1920 and the University of Illinois from 1920 to 1924, earning a bachelor of science degree. Wittick became an Associate of the Casualty Actuarial Society in 1929 and a Fellow in 1931. He was also a Fellow of the Canadian Institute of Actuaries.

From 1924 through 1931, Wittick worked at Standard Accident Insurance Company in Detroit, Michigan, in the company's compensation and liability underwriting department. In 1932, Wittick moved to Canada to work for Pilot Insurance Company in Toronto, Ontario. He would work for Pilot for over 36 years until his retirement in 1969. While at Pilot, he held the positions of assistant secretary, secretary, assistant general manager, and finally vice president and general manager.

Wittick was active in the CAS—attending meetings, serving as a committee member, and writing *Proceedings* papers. From 1964 to 1968, he served on the CAS Publicity Committee. A *Proceedings* author in 1958 and 1964, Wittick wrote "The Canadian Merit Rating Plan for Individual Automobile Risks" (*PCAS* XLV) and "Estimating the Cost of Accident Insurance as a Part of Automobile Liability Insurance" (*PCAS* LI). Wittick was also a panelist for the 1966 meeting session, "Management and the Actuary."

INDEX TO VOLUME LXXXV

INDEX—CONTINUED

INDEX—CONTINUED

Page

INDEX—CONTINUED

INDEX—CONTINUED

INDEX—CONTINUED

Page

INDEX—CONTINUED

Errata for Discussion by Howard Mahler of "Retrospective Rating: 1997 Excess Loss Factors"

At the bottom of page 320, the equation for $\hat{R}(100)$ is incorrect.

This example of simple dispersion is an example of a mixture with five pieces.

The excess ratio of the mixture is a weighted average of individual excess ratios, with the weights the product of the means and the probabilities for each piece of the mixture.[1]

If the probability of each piece of a mixture is p_i, $\sum p_i = 1$, the mean of each piece of the mixture is m_i, and R_i is the excess ratio for each piece of the mixture, then $\hat{R}(L) = \sum p_i\, m_i\, R_i(L)$.

If each loss is divided by for example .75, then after development, the excess ratio at L is the same as the original excess ratio at .75 L.[2]

$R_i(L)$ is the excess ratio when the losses have all been divided by r_i.

Thus $R_i(L) = R(r_i\, L)$.

In the example on page 320, each mean is proportional to 1/divisor = $1/r_i$, and each probability is the same at 1/5. Thus the weights are: $(1/5)(1/r_i)$.

The sum of the weights is: $\sum (1/5)(1/r_i) = (1/5)(1/.75 + 1/.833 + 1/1 + 1/1.25 + 1/1.5) = 1$.[3]

Thus $\hat{R}(L) = \sum (1/5)(1/r_i)\, R(r_i\, L) = (1/5) \sum R(r_i\, L)\, /r_i$.

Therefore, the corrected equation at the bottom of page 320 is:

$$\hat{R}(100) = (1/5)\{R(75)/.75 + R(83.3)/.833 + R(100)/1 + R(125)/1.25 + R(150)/1.50\}$$
$$= (1/5)\{.6009/.75 + .5817/.833 + .5582/1 + .5384/1.25 + .5191/1.50\} = 0.5669.$$

Similarly, the corrected equation at the top of page 321 is:

$$\hat{R}(5000) = (1/5)\{R(3750)/.75 + R(4165)/.833 + R(5000)/1 + R(6250)/1.25 + R(7500)/1.50\}$$
$$= (1/5)\{.0157/.75 + .0070/.833 + 0/1 + 0/1.25 + 0/1.50\} = 0.0059.$$

---

[1] See page 154 of "Workers Compensation Excess Ratios: An Alternate Method of Estimation" by Mahler.

[2] If each loss is multiplied by 1/.75 = 1.333, this is mathematically the same as uniform inflation of 33.3%. Thus we can get the excess ratio after development, by taking the original excess ratio at the deflated value of L/1.333 = .75 L. Increasing the sizes of loss, increases the excess ratio over a fixed limit.

[3] Mahler chose these loss divisors so that the total expected losses are unaffected.

At page 324, some of the numerical values shown in the computation of $R_3(2000)$ are mixed up, although the final value is correct at 0.384 as shown.
It should have read:

$$R_3(2000) = (1.04167)(0.9999980) - (0.04167)(0.0057148)$$
$$+ (0.1667)(0.0026029) - (0.8333)(0.999995)$$
$$+ (0.1761)(0.999987) - (0.1761)(0.0011302) = 0.384.$$

Also, in Table 1 the excess ratios were computed for Gamma loss divisors with shape parameter 16.67 and inverse scale parameter 15.67. However, the text at page 323 refers to Gamma loss divisors with shape parameter s = 18.67 and inverse scale parameter l = 17.67; this distribution of loss divisors corresponds to a mean loss development of 1 and a variance of loss development of 0.060, matching the simple dispersion example.

Using the intended Gamma parameters of s = 18.67 and l = 17.67 changes the excess ratios in Table 1 slightly, although the pattern remains the same.

The values in the simple dispersion column of Table 1 at page 320 are revised in a similar manner to that for 5000.
The values in Gamma dispersion column of Table 1 at page 320 are revised based on a shape parameter of s = 18.67 and inverse scale parameter of l = 17.67.

Corrected Table 1
Excess Ratios

LIMIT	No Development	Simple Dispersion	Gamma Dispersion
50	.6888	.6949	.6939
100	.5582	.5669	.5673
500	.3012	.3080	.3069
1,000	.1606	.1705	.1709
2,000	.0904	.0931	.0927
3,000	.0402	.0462	.0453
4,000	.0100	.0194	.0182
5,000	.0000	.0059	.0062
6,000	.0000	.0007	.0020
7,000	.0000	.0000	.0006
8,000	.0000	.0000	.0002
9,000	.0000	.0000	.0001
10,000	.0000	.0000	.0000

As can be seen in corrected Table 1, the simple dispersion effect raises the excess ratios, especially at the higher limits.[4]

At page 326, the formula near the bottom of page should have λ in place of X:

$R(L) = (\lambda\, l/L)^{s-1}\, U(s-1, s+1-\alpha, \lambda\, l/L)$.

The statement in the third paragraph of page 327 of the Discussion is backwards.
It should have read:
As the shape parameter of the Pareto, α, gets smaller, the losses have a heavier tail and the multiplicative impact of the dispersion on the excess ratios at high limits decreases.

At page 331, the first equation needs parentheses around the b - a:
$F(y) = (y/(b-a))\, \{E[R\, ;\, b/y] - E[R\, ;\, a/y]\}$.

[4] It can be demonstrated that when dispersion has no overall effect, loss dispersion either increases an excess ratio or keeps it the same. In most practical applications, the excess ratio will be increased by loss dispersion.

At page 332, lines 6 and 7, the dr is missing from the integrals:

$$\frac{y}{b-a}\left\{\left[\int_0^{b/y} r\, h(r)\, dr + \frac{b}{y}(1-H(b/y))\right] - \left[\int_0^{a/y} r\, h(r)\, dr + \frac{a}{y}(1-H(a/y))\right]\right\}$$

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXXIV

RETROSPECTIVE RATING: 1997 EXCESS LOSS FACTORS

WILLIAM R. GILLAM AND JOSE R. COURET

DISCUSSION BY HOWARD C. MAHLER

INTRODUCTION

This discussion will present some of the mathematical aspects of the effect of dispersion of loss development on excess ratios. It will be shown how the formulas developed in "Retrospective Rating: 1997 Excess Loss Factors" fit into this more general mathematical framework.

THE PROBLEM

Even if one included average loss development beyond fifth report in the estimation of excess ratios, there are at least two phenomena that would affect excess ratios that are not being considered. First, the different sizes of claims may have varying expected amounts of development. If larger claims had higher average development, this would raise the excess ratios for higher limits.

Secondly, there is a "dispersion" effect. Assume we have two claims of $1 million each that are expected on average to develop by 10%. It makes a difference whether we assume we'll have two claims each at $1.1 million or one claim at $1 million and one claim at $1.2 million. The ratio excess of $1.1 million will differ in the two cases.[1]

It is assumed for simplicity that there is no development on average; alternatively, any average development has already been

[1] In the former case it is zero, since there are no dollars excess of $1.1 million. In the latter case it is 0.1/2.2, since there are $1.2–$1.1 million = $.1 million dollars excess of $1.1 million, and total losses of $2.2 million.

incorporated into the size of loss distribution. However, some individual claims will develop more than average while others will develop less than average. In total, the average development factor is assumed to be unity.

SIMPLE EXAMPLE

Assume we have a piece-wise linear size of accident distribution such that:[2]

$$\begin{aligned} F(0) &= 0 \\ F(100) &= .90 \\ F(1{,}000) &= .99 \\ F(5{,}000) &= 1.00. \end{aligned}$$

Any size of loss distribution can be approximated sufficiently well by such an "ogive."[3] For actual applications one would have many more intervals, but this example will illustrate the principles involved.

The probability density function is:

$$f(x) = \begin{cases} .009 & 0 < x \leq 100 \\ .0001 & 100 < x \leq 1{,}000 \\ .0000025 & 1{,}000 < x \leq 5{,}000 \\ 0 & x > 5{,}000. \end{cases}$$

One can compute the average size of claim as the sum of three integrals of $xf(x)$:

$$\begin{aligned} E[X] &= \int_0^{100} (.009)x\,dx + \int_{100}^{1000} (.0001)x\,dx \\ &\quad + \int_{1000}^{5000} (.0000025)x\,dx \\ &= 45 + 49.5 + 30 = 124.5. \end{aligned}$$

[2]Assume everything is in units of thousands of dollars. Thus, 5,000 actually corresponds to $5 million.

[3]See Hogg and Klugman [3].

The excess ratio at limit L can be computed as:

$$R(L) = \int_L^\infty (x-L) f(x)\,dx / E[X].$$

In this case we can compute the numerator as a sum of three terms:

$$\begin{aligned}
\int_L^\infty (x-L) f(x)\,dx \\
&= (\text{If } L < 100) \int_L^{100} (x-L)(.009)\,dx \\
&\quad + (\text{If } L < 1000) \int_{\max[100,L]}^{1000} (x-L)(.0001)\,dx \\
&\quad + (\text{If } L < 5000) \int_{\max[1000,L]}^{5000} (x-L)(.0000025)\,dx.
\end{aligned}$$

If $L < 100$, let

$$\begin{aligned}
R_1(L) &= \int_L^{100} (x-L)\,dx \Big/ \int_0^{100} x\,dx \\
&= \text{excess ratio at } L \text{ if losses are uniformly distributed on the interval } [0,100].
\end{aligned}$$

Note that $R_1(L) = 0$ if $L \geq 100$. Then the first term above is

$$R_1(L) \int_0^{100} .009x\,dx = R_1(L) E_1[X],$$

where $E_1[X] = \int_0^{100} (.009)x\,dx$ is the contribution to the overall mean from claims in the first interval. Then, working similarly with the other two intervals:

$$\int_0^\infty (x-L) f(x)\,dx = R_1(L)E_1[X] + R_2(L)E_2[X] + R_3(L)E_3[X],$$

$$R(L) = \frac{R_1(L)E_1[X] + R_2(L)E_2[X] + R_3(L)E_3[X]}{E_1[X] + E_2[X] + E_3[X]}.$$

Thus, the overall excess ratio can be expressed as a weighted average of excess ratios each computed as if the losses were uniformly distributed on an interval. The weights are the contributions to the overall mean of the claims in each interval. In this case, the weights are 45, 49.5, and 30 or 45/124.5, 49.5/124.5, and 30/124.5.

For example, for a limit of 70, the individual excess ratios are:[4] .09, .8727, and .9767. The weighted average is

$$R(70) = \frac{(45)(.09) + (49.5)(.8727) + (30)(.9767)}{124.5} = .6149.$$

Further, if the losses were uniform from 100 to 1000 then the excess ratio would be:

$$\frac{1}{900}\int_{100}^{1000}(x-70)\,dx \Big/ \frac{1}{900}\int_{100}^{1000} x\,dx = (550-70)/550$$
$$= 480/550 = .8727.$$

Table 1 shows the excess ratios for this simple example for several limits. As can be seen, in the absence of any loss development, the ratio excess of 5,000 is zero; there are no losses above 5,000.

SIMPLE DISPERSION

Assume for simplicity that each accident has an equal likelihood of developing in a manner such that it is divided[5] by either: .75, .833, 1, 1.25, or 1.5. Then the average development is

$$\frac{1}{5}\left(\frac{1}{.75} + \frac{1}{.833} + \frac{1}{1} + \frac{1}{1.25} + \frac{1}{1.5}\right) = 1.$$

---

[4]For losses distributed uniformly on $[a,b]$, for $b > L > a$, $R(L) = (b-L)^2/(b^2-a^2)$; for $L < a$, $R(L) = 1 - 2L/(b+a)$; for $L > b$, $R(L) = 0$. For the interval [0,100] we have the first case. For the intervals [100,1,000] and [1,000,5,000] we have the second case.

[5]Loss development *divisors* are used in order to match the presentation in "Retrospective Rating: 1997 Excess Loss Factors." Loss development multipliers or factors could have been used equally well for the presentation.

TABLE 1

EXCESS RATIOS*

LIMIT	No Development	Simple Dispersion**	Gamma Dispersion***
50	.6888	.6813	.6945
100	.5582	.5597	.5684
500	.3012	.2932	.3076
1,000	.1606	.1632	.1721
2,000	.0904	.0856	.0930
3,000	.0402	.0394	.0459
4,000	.0100	.0156	.0190
5,000	.0000	.0045	.0069
6,000	.0000	.0005	.0024
7,000	.0000	.0000	.0008
8,000	.0000	.0000	.0003
9,000	.0000	.0000	.0001
10,000	.0000	.0000	.0000

*For simple piece-wise linear distribution with $F(0) = 0$, $F(100) = .9$, $F(1000) = .99$, $F(5000) = 1$.
**For five possibilities, see text. Mean development = 1; Variance of development = .060.
***For a gamma loss divisor with $\alpha = 16.67$, $\lambda = 15.67$, see text. Mean development = 1; Variance of development = .060.

Thus, the total expected losses are unaffected. The variance of the development is .060.

We can compute excess ratios for each of the five possibilities and average the results together. If all the accidents were divided by 1.25; i.e., multiplied by .8, then a limit of 100 is equivalent to a limit of 125 without any development. So the excess ratio for 100 for the developed losses can be computed as $R(125)$ for the original distribution.[6]

Thus, to compute the excess ratio for the developed losses for a limit of 100:

$$\begin{aligned}\hat{R}(100) &= \tfrac{1}{5}(R(75) + R(83.3) + R(100) + R(125) + R(150))\\ &= \tfrac{1}{5}(.6009 + .5817 + .5582 + .5384 + .5191) = .5597.\end{aligned}$$

[6]If each of the accidents are divided by 1.25, then the ratio excess of a limit of 100 declines from .5582 to .5384. Reducing the size of accidents reduces the excess ratio over any fixed limit.

Similarly, we can compute the excess ratio for the developed losses for a limit of 5,000:

$$\begin{aligned}\hat{R}(5000) &= \tfrac{1}{5}(R(3750) + R(4165) + R(5000)\\ &\quad + R(6250) + R(7500))\\ &= \tfrac{1}{5}(.0157 + .0070 + 0 + 0 + 0) = .0045.\end{aligned}$$

So the dispersion effect has now produced some losses excess of 5,000, without affecting the total expected losses.

As can be seen in Table 1, the dispersion effect raises the excess ratios for higher limits and alters those for lower limits. While this example could be changed to include more than 5 possibilities, the essence of the dispersion effect has been captured. However, if the possibilities were more dispersed around the mean; i.e., if the variance of the development were greater, then the impact of the dispersion would be greater.

CONTINUOUS LOSS DIVISORS APPLIED TO A UNIFORM DISTRIBUTION ON AN INTERVAL

What if, rather than five possible loss divisors, one had a continuous probability distribution?

Assume:

1. Losses are distributed uniformly on the interval $[a,b]$.
2. Losses will develop with loss divisors r given by a distribution $H(r)$, with density $h(r)$.[7]

Then, as shown in Appendix A, the distribution function for the developed losses y, is given by:

$$F(y) = [y/(b-a)]\{E(R;b/y) - E(R;a/y)\},$$

[7] It is assumed that $\int_0^\infty (h(r)/r)\,dr$ is finite, so that the overall loss development is finite. In the case where H is a gamma distribution, this requirement means that the shape parameter s must be greater than one.

where $E[R;L]$ is the limited expected value of the distribution of loss divisors, at a limit L.

Appendix A also shows that the density function can be written in a number of forms, as summarized below:

$$\begin{aligned}
f(y) &= \frac{1}{b-a}\int_{a/y}^{b/y} rh(r)\,dr \\
&= \frac{1}{b-a}\{E[R;b/y]-(b/y)(1-H(b/y))-E[R;a/y] \\
&\qquad + (a/y)(1-H(a/y))\} \\
&= \frac{1}{b-a}\{E[R;b/y]-E[R;a/y]\} \\
&\quad + \frac{1}{y(b-a)}\{bH(b/y)-aH(a/y)\}-\frac{1}{y}.
\end{aligned}$$

Further, Appendix A describes how one can use the density function and distribution function to calculate the excess ratio of the developed losses, as follows:

$$\begin{aligned}
R(L) = \frac{1}{b^2-a^2}\Bigg\{ & b^2\int_0^{b/L} h(r)/r\,dr - a^2\int_0^{a/L} h(r)/r\,dr \\
& + 2aLH(a/L) - 2bLH(b/L) \\
& + L^2\int_{a/L}^{b/L} rh(r)\,dr \Bigg\} \Bigg/ \int_0^{\infty} h(r)/r\,dr.
\end{aligned}$$

GAMMA DISPERSION APPLIED TO THE UNIFORM DISTRIBUTION

Assume that the loss divisor r is distributed according to a gamma distribution[8] with parameter s and l:

$$h(r) = \frac{l^s r^{s-1} e^{-lr}}{\Gamma(s)},$$

where $\Gamma(n) = (n-1)!$.

[8]Then the loss multipliers are distributed according to an inverse gamma. We assume $s > 1$, so that the overall loss development is finite.

Then, as shown in Appendix B, based on the general formula in Appendix A, if the losses are uniformly distributed on the interval $[a,b]$, after development the excess ratio for the limit L is given by:[9]

$$\begin{aligned} R(L) = {} & \frac{b^2}{b^2-a^2}\Gamma(s-1;lb/L) - \frac{a^2}{b^2-a^2}\Gamma(s-1;la/L) \\ & + \frac{2L(s-1)}{(b^2-a^2)l}\{a\Gamma(s;la/L) - b\Gamma(s;lb/L)\} \\ & + \frac{L^2(s-1)s}{(b^2-a^2)l^2}\{\Gamma(s+1;lb/L) - \Gamma(s+1;la/L)\}, \end{aligned}$$

where $\Gamma(s;y) = 1/\Gamma(s)\int_0^y t^{s-1}e^{-t}\,dt$ is the incomplete gamma function.

One can apply this "gamma dispersion" effect to a piece-wise linear size of accident distribution, such as in the prior example.

The mean development is the mean of an inverse gamma, $l/(s-1)$. For this discussion, the average development is unity, so we take $l = s-1$. The variance of the development is the variance of an inverse gamma, $l^2/\{(s-1)^2(s-2)\}$. For $l = s-1$, the variance is $1/s-2$. Thus, if one takes $s = 18.67$, (and $l = 17.67$) then the variance of the development is $1/16.67 = .060$, which matches that in the simple dispersion example. However, the gamma allows extreme possibilities (with a small probability), so one gets a somewhat different behavior than in the simple dispersion example.

As seen in Table 1, using the gamma dispersion for very high limits (7,000 and above) yields excess ratios that are now positive rather than zero. Gamma dispersion can have a particularly significant impact on very high limits, particularly if the variance is large.

---

[9]These are the formulas developed and shown in "Retrospective Rating: 1997 Excess Loss Factors."

Each excess ratio is computed as a weighted average of the excess ratios computed for losses uniformly distributed on each of the three assumed intervals. For example, for a limit of 2,000, the excess ratio for losses distributed uniformly from 1,000 to 5,000, with gamma dispersion with $s = 18.67$ and $l = 17.67$ is given by the formula from Appendix B:

$$
\begin{aligned}
R_3&(2000)\\
&= (1.04167)\Gamma(17.67;44.175) - (.04167)\Gamma(17.67;8.835)\\
&\quad + (.1667)\Gamma(18.67;8.835) - (.8333)\Gamma(18.67;44.175)\\
&\quad + (.1761)\Gamma(19.67;44.175) - (.1761)\Gamma(19.67;8.835)\\
&= (1.04167)(.9999980) - (.04167)(.0057148)\\
&\quad + (.1667)(.0011302) - (.8333)(.999987)\\
&\quad + (.1761)(.999949) - (.1761)(.0026)\\
&= .384.
\end{aligned}
$$

Similarly, for losses uniform from 100 to 1,000, $R_2(2000) = .00008$. For losses uniform from 0 to 100, $R_1(2000) = 10^{-19}$. Taking a weighted average, using weights of 45, 49.5, and 30, one obtains $R(2000) = .093$, as displayed in Table 1.

Note that the gamma distribution used in this example has a large value of s, the shape parameter. Therefore, the distribution of loss divisors is close to normal.[10] The distribution of loss divisors has a skewness of $2/\sqrt{s}$, which is only .49. The skewness of the distribution of loss multipliers is that of an inverse gamma: $4\sqrt{(s-2)}/(s-3) = 1.12$. If one were to take a different form of distribution with a larger skewness one would have a larger chance of extreme results. Therefore, in the case of very high limits, the excess ratios would be even larger.

[10]The distribution of loss multipliers is close to an inverse normal distribution.

DISTRIBUTION OF DEVELOPED LOSSES

The particular situations examined so far are a special case of a more general framework. As shown in Appendix C, if losses at latest report are distributed via $G(x)$ and the loss divisors r are distributed independently of x via density function $h(r)$,[11] then the distribution for the developed losses y is given by:

$$F(y) = \int_0^\infty G(yr)h(r)\,dr.$$

GAMMA LOSS DIVISORS APPLIED TO AN EXPONENTIAL DISTRIBUTION

For example, if $G(x)$ is an exponential distribution $G(x) = 1 - e^{-\lambda x}$ and the loss divisors are gamma distributed $h(r) = l^s r^{s-1} e^{-lr}/\Gamma(s)$, then

$$\begin{aligned} F(y) &= 1 - \frac{l^s}{\Gamma(s)} \int_0^\infty r^{s-1} e^{-lr} e^{-\lambda yr}\,dr \\ &= 1 - \frac{l^s}{\Gamma(s)} \frac{\Gamma(s)}{(l + \lambda y)^s} = 1 - \left(\frac{(l/\lambda)}{(l/\lambda) + y}\right)^s. \end{aligned}$$

Thus $F(y)$ has a Pareto distribution as per Hogg and Klugman [3], with shape parameter s and scale parameter l/λ. Thus, the excess ratio of the developed losses is that of a Pareto distribution:

$$R(L) = \left(\frac{(l/\lambda)}{(l/\lambda) + L}\right)^{s-1}.$$

MATHEMATICAL RELATION TO MIXED DISTRIBUTIONS

The calculation of the distribution of the developed losses is the same as that used to calculate the mixed distribution in the inverse gamma-exponential conjugate prior.[12] (An inverse gamma

[11] With $\int_0^\infty (h(r)/r)\,dr$ finite.

[12] See Herzog [2], or Venter [4]. The mixed distribution in the case of an inverse gamma—Exponential conjugate prior is a Pareto distribution, as obtained above.

distributed multiplier is the same as a gamma distributed divisor.) In general, if the loss multipliers and the loss distribution form any of the well known pairs[13] of prior distributions of the scale parameters of the conditional distributions and conditional distributions, then the developed losses will be given by the mixed distribution. For example, as shown in Venter [4], a Weibull loss distribution and a transformed gamma loss divisor[14] would produce a Burr distribution of developed losses. Thus, there are a number of mathematically convenient examples that might approximate a particular real world application.

GAMMA LOSS DIVISORS APPLIED TO PARETO LOSSES

Since the Pareto distribution is often used to model losses (or at least the larger losses), it would be valuable to be able to apply the concept of loss divisors to the Pareto distribution.

As shown in Appendix C, one can develop the mathematics of applying gamma loss divisors to losses distributed by a Pareto distribution with parameters α and λ: $F(x) = 1 - (\lambda/(\lambda + x))^{\alpha}$. As derived in Appendix C, the excess ratio for the developed losses is given by:

$$R(L) = \left(\frac{Xl}{L}\right)^{s-1} U(s-1, s+1-\alpha, \lambda l/L),$$

where U is a confluent hypergeometric function.[15]

It is also shown in Appendix C that when the average development is unity[16] then the excess ratios of the developed losses can be approximated by replacing λ in the Pareto by $\lambda' = \lambda(s-1)/(s-(\alpha/2)-1)$.

---

[13] Such as shown in Venter [4]. Venter displays a list of conjugate priors, but for the current application there is no requirement that it be a conjugate prior situation.

[14] An inverse transformed gamma loss multiplier.

[15] See Appendix D and *Handbook of Mathematical Functions* [1].

[16] Also, we need the shape parameter of the gamma, s, to be greater than $\alpha + 1$.

Table 2 and Figure 1 compare the excess ratios for a Pareto with $\alpha = 3.5$ and $\lambda = 1{,}000$, for the developed losses[17] with a gamma divisor with $s = 6$ and $l = 5$, and for an approximating Pareto with $\alpha = 3.5$ and $\lambda = 1{,}000(s-1)/(s-(\alpha/2)-1) = 1{,}538$. The excess ratios for the developed losses are larger than those for the undeveloped losses. The approximation using the rescaled Pareto yields excess ratios that are too high for the lower limits, but it does an excellent job of approximating the excess ratios for higher limits.

As shown in Appendix C, in the tail, the loss development[18] multiplies the excess ratios by a factor of approximately:

$$(s-1)^{\alpha-1}\Gamma(s-\alpha)/\Gamma(s-1) \approx \left((s-1)\Big/\left(s-\frac{\alpha}{2}-1\right)\right)^{\alpha-1}.$$

In this example, this factor is: $5^{2.5}\Gamma(2.5)/\Gamma(5) = 3.1$. Figure 2 shows how this adjustment factor varies as the shape parameters of the Pareto and gamma vary. As the shape parameter of the Pareto, α, gets smaller, the losses have a heavier tail and the impact of the dispersion increases. As the coefficient of variation of the gamma[19] increases, the impact of the dispersion increases.

In general, as the coefficient of variation of the loss divisors increases, the impact of the dispersion increases. As the coefficient of variation approaches zero, we approach the situation where each claim develops by the average amount and there is no effect of dispersion. Thus, in order to apply this technique, one of the key inputs would be the coefficient of variation of the loss divisors.

CONCLUSIONS

The effect of the dispersion of loss development beyond the latest available report can be incorporated into the calculation of

---

[17]Then $R(L) = (\lambda l/L)^{s-1}U(s-1, s+1-\alpha, \lambda l/L) = (L/5000)^{-5}U(5, 3.5, 5000/L)$.

[18]For gamma dispersion with $l = s-1$ so the average development is unity.

[19]The coefficient of variation is the standard deviation divided by the mean. For the gamma distribution with shape parameter s, the coefficient of variation is $1/\sqrt{s}$.

TABLE 2

EXCESS RATIOS

LIMIT	Undeveloped Losses Pareto ($\alpha = 3.5, \lambda = 1{,}000$)	Developed Losses*	Approximating Developed Pareto ($\alpha = 3.5, \lambda = 1{,}538$)
500	.3629	.3960	.4947
1,000	.1768	.2152	.2859
2,500	.0436	.0668	.0895
5,000	.0113	.0211	.0268
10,000	.0025	.0055	.0065
25,000	.00029	.00076	.00081
50,000	.00005	.00015	.00015
100,000	.000010	.000029	.000028

*Assuming gamma loss divisor, with $s = 6$ and $l = 5$. $R(L) = (5000/L)^{2.5}U(5, 1.5; 5000/L)$.

FIGURE 1

EXCESS RATIOS

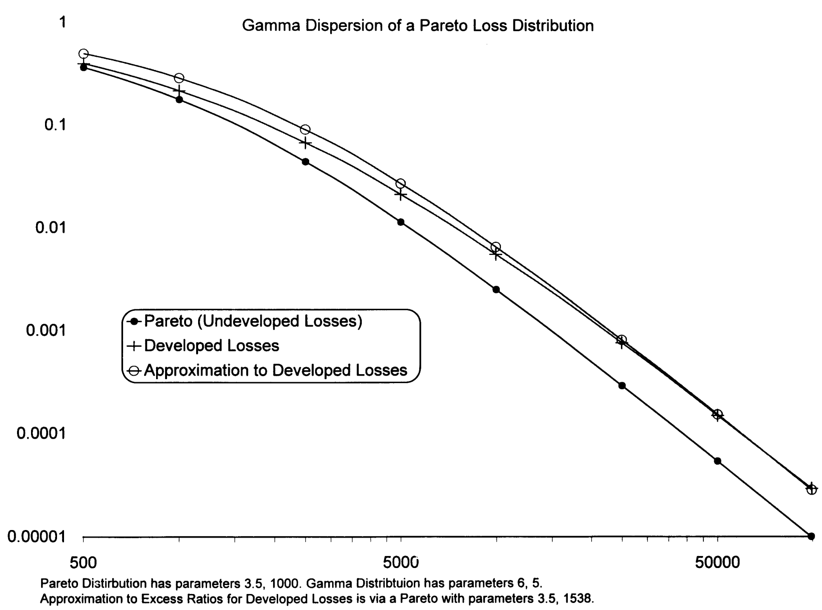
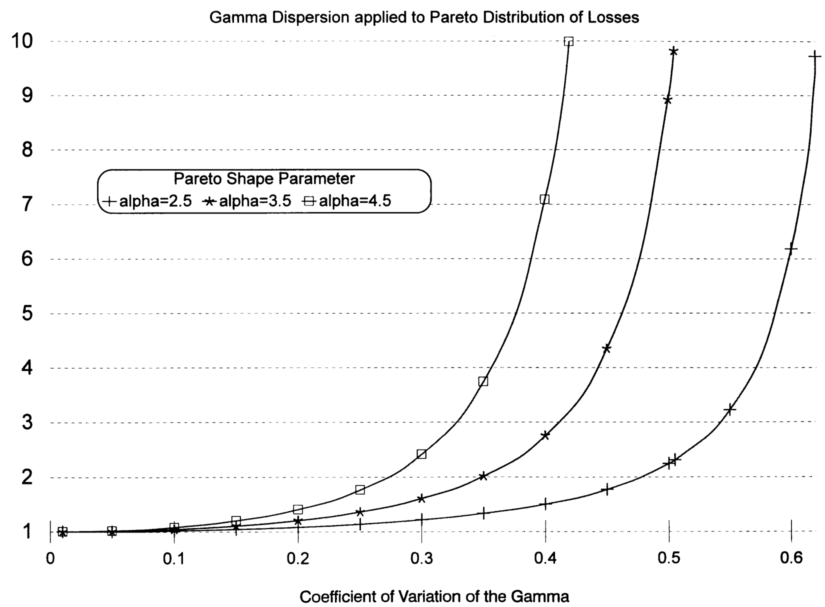


FIGURE 2

ADJUSTMENT FACTOR TO APPLY TO EXCESS RATIOS



excess ratios. In the case of loss dispersion which is (approximately) independent of size of loss, for many special cases one can calculate the distribution of the developed losses in closed form. In these cases, the excess ratios follow directly.

In other situations, one can approximate the loss distribution via a piece-wise linear distribution and then apply the effects of dispersion to each interval. Since on each interval the piece-wise linear approximation is a uniform distribution, one can apply the formulas developed in Appendix A. Then one can weight together the excess ratios for the developed losses from the individual intervals in order to get the excess ratio for all developed losses.

REFERENCES

[1] Abramowitz, Milton and Irene A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 1964, U.S. Department of Commerce.

[2] Herzog, Thomas N., *Introduction to Credibility Theory*, 1994, Mad River Books.

[3] Hogg, Robert V. and Stuart A. Klugman, *Loss Distributions*, John Wiley & Sons, 1984, New York.

[4] Venter, Gary G., "Credibility," Chapter 7 in *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990, pp. 375–483.

APPENDIX A

LOSS DIVISORS APPLIED TO A UNIFORM DISTRIBUTION ON AN INTERVAL

Assume:

Losses are distributed uniformly on the interval $[a,b]$. Losses will develop with loss divisors r given by a distribution $H(r)$ and density $h(r)$.

Then:

The distribution function for the developed losses y, is given by:

$$F(y) = (y/b-a)\{E[R;b/y] - E[R;a/y]\},$$

where $E[R;L]$ is the limited expected value of the distribution of loss divisors, at a limit L.

Proof:

The developed losses y are the ratio of the undeveloped losses x and the loss divisor r; $y = x/r$ or $x = yr$. Thus since x is uniform on $[a,b]$,[20] the conditional distribution of y given r is:

$$F(y \mid r) = \begin{cases} 0 & yr \le a \\ \dfrac{ry-a}{b-a} & a \le yr \le b \\ 1 & yr \ge b. \end{cases}$$

The unconditional distribution of y can be computed by integrating the conditional distribution of y given r times the assumed density function of r:

[20]For the uniform distribution on $[a,b]$, $F(x) = 0$ if $x \le a$, $F(x) = (x-a)/(b-a)$ if $a \le x \le b$, and $F(x) = 1$ if $x \ge b$.

$$
\begin{aligned}
F(y) &= \int_{r=0}^{\infty} F(y \mid r) h(r)\, dr \\
&= \int_{a/y}^{b/y} \left(\frac{ry-a}{b-a}\right) h(r)\, dr + \int_{b/y}^{\infty} h(r)\, dr \\
&= \frac{y}{b-a} \int_{a/y}^{b/y} r h(r)\, dr + \frac{a}{b-a}\left(H\left(\frac{a}{y}\right) - H\left(\frac{b}{y}\right)\right) \\
&\quad + 1 - H(b/y) \\
&= \frac{y}{b-a}\left\{\int_{0}^{b/y} r h(r)\, dr - \int_{0}^{a/y} r h(r)\, dr + \frac{a}{y} H\left(\frac{a}{y}\right) - \frac{a}{y}\right. \\
&\qquad \left. + \frac{b}{y} - \frac{b}{y} H\left(\frac{b}{y}\right)\right\} \\
&= \frac{y}{b-a}\left\{\left[\int_{0}^{b/y} r h(r) + \left(\frac{b}{y}\right)(1 - H(b/y))\right]\right. \\
&\qquad \left. - \left[\int_{0}^{a/y} r h(r) + \left(\frac{a}{y}\right)(1 - H(a/y))\right]\right\} \\
&= \frac{y}{b-a}\left\{E\left[R; \frac{b}{y}\right] - E\left[R; \frac{a}{y}\right]\right\}.
\end{aligned}
$$

Similarly, we can get the density function $f(y)$. For conditional density at y given r is:

$$
f(y \mid r) = \begin{cases} 0 & yr \leq a \\ \dfrac{r}{b-a} & a \leq yr \leq b \\ 0 & yr \geq b. \end{cases}
$$

The unconditional density at y can be computed by integrating the conditional density at y given r times the assumed density function of r:

$$f(y) = \int_0^\infty f(y \mid r) h(r)\,dr = \int_{a/y}^{b/y} \frac{r}{b-a} h(r)\,dr = \frac{1}{b-a}\int_{a/y}^{b/y} r h(r)\,dr.$$

We can put this type of integral in terms of limited expected values, since

$$E[R;r] = \int_0^r r h(r)\,dr + r(1-H(r))$$

$$f(y) = \frac{1}{b-a}\{E[R;b/y] - (b/y)(1-H(b/y)) - E[R;a/y] + (a/y)(1-H(a/y))\}$$

$$= \frac{1}{b-a}\{E[R;b/y] - E[R;a/y]\} + \frac{1}{y(b-a)}\{bH(b/y) - aH(a/y)\} - \frac{1}{y}.$$

One can use the density function and distribution function to calculate the excess ratio of the developed losses. The numerator of this excess ratio is the (developed) losses excess of L:

$$\int_L^\infty (y-L) f(y)\,dy = \int_L^\infty y f(y)\,dy - L(1-F(L)).$$

Since $f(y) = 1/(b-a)\int_{a/y}^{b/y} r h(r)\,dr$ we have

$$\int_L^\infty y f(y)\,dy = \frac{1}{b-a}\int_{y=L}^\infty y \int_{r=a/y}^{b/y} r h(r)\,dr\,dy.$$

Switching the order of integration:

$$\int_L^\infty yf(y)\,dy = \frac{1}{b-a}\int_{r=0}^{a/L}\int_{y=a/r}^{b/r} yrh(r)\,dy\,dr$$

$$+\frac{1}{b-a}\int_{r=a/L}^{b/L}\int_{y=L}^{b/r} yrh(r)\,dy\,dr$$

$$=\frac{1}{2(b-a)}\int_{r=0}^{a/L}\left(\frac{b^2}{r^2}-\frac{a^2}{r^2}\right)rh(r)\,dr$$

$$+\frac{1}{2(b-a)}\int_{r=a/L}^{b/L}\left(\frac{b^2}{r^2}-L^2\right)rh(r)\,dr$$

$$=\frac{b^2}{2(b-a)}\int_{r=0}^{b/L} h(r)/r\,dr-\frac{a^2}{2(b-a)}\int_{r=0}^{a/L} h(r)/r\,dr$$

$$-\frac{L^2}{2(b-a)}\int_{r=a/L}^{b/L} rh(r)\,dr.$$

In the course of deriving the form of the distribution function we had

$$F(y)=\frac{y}{b-a}\int_{a/y}^{b/y} rh(r)\,dr+1+\frac{a}{b-a}H\left(\frac{a}{y}\right)-\frac{b}{b-a}H\left(\frac{b}{y}\right).$$

Thus

$$1-F(L)=$$

$$\frac{b}{b-a}H(b/L)-\frac{a}{b-a}H(a/L)-\frac{L}{b-a}\int_{a/L}^{b/L} rh(r)\,dr.$$

Thus combining the terms, the numerator of the excess ratio is:

$$\int_L^\infty yf(y)\,dy - L(1-F(L))$$

$$= \frac{b^2}{2(b-a)}\int_0^{b/L} h(r)/r\,dr - \frac{a^2}{2(b-a)}\int_0^{a/L} h(r)/r\,dr$$

$$+ \frac{aL}{(b-a)}H(a/L) - \frac{bL}{b-a}H(b/L)$$

$$+ \frac{L^2}{2(b-a)}\int_{a/L}^{b/L} rh(r)\,dr.$$

The denominator of the excess ratio is:[21]

$$\int_0^\infty yf(y)\,dy = \lim_{L\to 0}\int_L^\infty yf(y)\,dy$$

$$= \frac{b^2-a^2}{2(b-a)}\int_0^\infty h(r)/r\,dr$$

$$= \frac{b+a}{2}\int_0^\infty h(r)/r\,dr.$$

Combining the numerator and denominator, the excess ratio (of the developed losses) at L is:

$$R(L) = \frac{1}{b^2-a^2}\left\{b^2\int_0^{b/L} h(r)/r\,dr - a^2\int_0^{a/L} h(r)/r\,dr\right.$$

$$+ 2aLH(a/L) - 2bLH(b/L)$$

$$\left. + L^2\int_{a/L}^{b/L} rh(r)\,dr\right\}\bigg/\int_0^\infty h(r)/r\,dr.$$

[21] The denominator of the excess ratio is the mean of the developed losses. It is equal to the product of the mean undeveloped losses $(b+a)/2$, and the average loss development $\int_0^\infty h(r)/r\,dr$.

APPENDIX B

GAMMA LOSS DIVISORS APPLIED TO LOSSES UNIFORM ON AN INTERVAL

For the situation discussed in Appendix A but for the specific case where the distribution of the loss divisors, $h(r)$, is a gamma distribution with parameters s and l:

$$\begin{aligned}
\int_0^x h(r)r\,dr &= \int_0^x l^s e^{-lr} r^s/\Gamma(s)\,dr = (l^s/\Gamma(s))\int_0^x e^{-lr} r^s\,dr \\
&= (l^s/\Gamma(s))(\Gamma(s+1)/l^{s+1})\Gamma(s+1;lx) \\
&= (s/l)\Gamma(s+1;lx) \\
H(x) &= \int_0^x h(r)\,dr = \int_0^x l^s e^{-lr} r^{s-1}/\Gamma(s)\,dr = \Gamma(s;lx) \\
\int_0^x h(r)/r\,dr &= \int_0^x l^s e^{-lr} r^{s-2}/\Gamma(s)\,dr \\
&= (l^s/\Gamma(s))(\Gamma(s-1)/l^{s-1})\Gamma(s-1;lx) \\
&= \frac{l}{s-1}\Gamma(s-1;lx) \\
\int_0^\infty h(r)/r\,dr &= (l/s-1)\Gamma(s-1;\infty) = l(s-1).
\end{aligned}$$

Thus, using the formula from Appendix A, the excess ratio of the developed losses for limit L is in this case:

$$\begin{aligned}
R(L) &= \frac{b^2}{b^2-a^2}\Gamma(s-1;lb/L) - \frac{a^2}{b^2-a^2}\Gamma(s-1;la/L) \\
&\quad + \frac{2L(s-1)}{(b^2-a^2)l}\{a\Gamma(s;la/L) - b\Gamma(s;lb/L)\} \\
&\quad + \frac{L^2(s-1)s}{(b^2-a^2)l^2}\{\Gamma(s+1;lb/L) - \Gamma(s+1;la/L)\}.
\end{aligned}$$

For the loss divisors given by a gamma distribution with parameters s and l, we can plug in the limited expected value for the gamma distribution in terms of the incomplete gamma function:[22]

$$E[R;r] = \frac{s}{l}\Gamma(s+1;lr) + r[1-\Gamma(s;lr)].$$

Thus using the formula derived in Appendix A:

$$\begin{aligned} F(y) &= \frac{y}{b-a}\left\{E\left[R;\frac{b}{y}\right] - E\left[R;\frac{a}{y}\right]\right\} \\ &= \frac{ys}{l(b-a)}\left\{\Gamma\left(s+1;\frac{lb}{y}\right) - \Gamma\left(s+1;\frac{la}{y}\right)\right\} \\ &\quad + 1 + \frac{a}{b-a}\Gamma\left(s;\frac{la}{y}\right) - \frac{b}{b-a}\Gamma\left(s;\frac{lb}{y}\right). \end{aligned}$$

Also using the formula derived in Appendix A, the probability density function is given by:

$$\begin{aligned} f(y) &= \frac{1}{b-a}\int_{a/y}^{b/y} rh(r)\,dr \\ &= \frac{s}{(b-a)l}\{\Gamma(s+1;lb/y) - \Gamma(s+1;la/y)\}. \end{aligned}$$

[22] See Hogg and Klugman [3, page 226].

APPENDIX C

Gamma Loss Divisors Applied to a Pareto Distribution

Assume:

Losses are distributed (at latest report) on $(0,\infty)$ via a distribution function $G(x)$. Losses will develop with loss divisors r given by a density function $h(r)$.[23] (The distribution of r is independent of x.)

Then:

The distribution function for the developed losses y, is given by:

$$F(y) = \int_0^\infty G(yr)h(r)\,dr.$$

Proof:

The developed losses y are the ratio of the undeveloped losses x and the loss divisor r; $y = x/r$ or $x = yr$.

Given a value for r, the developed losses are less than y if the undeveloped losses are less than yr. Thus:

$$F(y \mid r) = G(yr).$$

Integrating over all possible values of r we have

$$F(y) = \int_0^\infty F(y \mid r)h(r)\,dr = \int_0^\infty G(yr)h(r)\,dr.$$

In the specific case where r follows a gamma distribution with parameters s and l and the undeveloped losses follow a Pareto distribution with parameters α and λ:

$$G(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha},$$

$$h(r) = l^s r^{s-1} e^{-lr}/\Gamma(s).$$

[23] It is assumed that $\int_0^\infty (h(r)/r)\,dr$ is finite, so that the average loss development is finite.

Then the distribution function for the developed losses is

$$\begin{aligned}F(y) &= \int_0^\infty h(r)G(yr)\,dr\\ &= \int_0^\infty \left(1-\left(\frac{\lambda}{\lambda+yr}\right)^{\alpha}\right) l^s r^{s-1}e^{-lr}/(\Gamma(s))\,dr\\ &= \int_0^\infty l^s r^{s-1}e^{-lr}/\Gamma(s)\,dr - \frac{\lambda^\alpha l^s}{\Gamma(s)}\int_0^\infty r^{s-1}e^{-lr}(\lambda+yr)^{-\alpha}\,dr.\end{aligned}$$

The first integral is unity,[24] while the second integral can be put in terms of confluent hypergeometric functions.[25]

Let $q = (y/\lambda)r$, then the second integral becomes

$$\begin{aligned}&\frac{\lambda^{s-\alpha}}{y^s}\int_{q=0}^\infty q^{s-1}e^{-\lambda lq/y}(1+q)^{-\alpha}\,dq\\ &\qquad = \frac{\lambda^{s-\alpha}}{y^s}\Gamma(s)U(s,s+1-\alpha;\lambda l/y)\end{aligned}$$

where U is a confluent hypergeometric function such that[26]

$$U(a,b;z) = (1/\Gamma(a))\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}\,dt.$$

Thus the distribution function of the developed losses is:

$$\begin{aligned}F(y) &= 1 - \frac{\lambda^\alpha l^s}{\Gamma(s)}\frac{\lambda^{s-\alpha}}{y^s}\Gamma(s)U(s,s+1-\alpha;\lambda l/y)\\ &= 1-\left(\frac{\lambda l}{y}\right)^s U(s,s+1-\alpha;\lambda l/y).\end{aligned}$$

Similarly one can compute the density function of the developed losses. Differentiating the distribution function one gets:

$$f(y) = \int_0^\infty rg(yr)h(r)\,dr.$$

[24] It is the cumulative distribution function of the gamma distribution at infinity.
[25] See Appendix D and *Handbook of Mathematical Functions* [1].
[26] See Equation 13.2.5 in *Handbook of Mathematical Functions* [1].

In the specific case where h is gamma and g is Pareto it turns out that the density of the developed losses is:

$$f(y) = \frac{s\alpha l^s \lambda^s}{y^{s+1}} U(s+1, s+1-\alpha; l\lambda/y).$$

This can be obtained either by substituting the specific form of y and h into the above integral or by differentiating $F(y)$, and making use of the facts that[27]

$$\frac{d}{dz}U(a,b;z) = -aU(a+1,b+1;z),$$

$$U(a-1,b;z) - zU(a,b+1;z) = (a-b)U(a,b;z),$$

$$\begin{aligned}
f(y) = \frac{d}{dy}F(y) &= \frac{d}{dy}\left(1 - \frac{\lambda^s l^s}{y^s} U\left(s, s+1-\alpha; \frac{\lambda l}{y}\right)\right) \\
&= \frac{\lambda^s l^s s}{y^{s+1}} U\left(s, s+1-\alpha; \frac{\lambda l}{y}\right) \\
&\quad - \frac{\lambda^s l^s}{y^s}\left(\frac{\lambda l}{y^2}\right) U'\left(s, s+1-\alpha; \frac{\lambda l}{y}\right) \\
&= \frac{s\lambda^s l^s}{y^{s+1}}\{U(s, s+1-\alpha; \lambda l/y) \\
&\qquad - (\lambda l/y)U(s+1, s+2-\alpha; \lambda l/y)\} \\
&= \frac{s\lambda^s l^s}{y^{s+1}}\{(s+1) - (s+1-\alpha)\}U(s+1, s+1-\alpha; \lambda l/y) \\
&= \frac{s\alpha\lambda^s l^s}{y^{s+1}} U(s+1, s+1-\alpha; \lambda l/y).
\end{aligned}$$

One can use the density function and distribution function to calculate the excess ratio of the developed losses. The numerator of this excess ratio is the total (developed) losses excess of L:

$$\int_L^\infty (y-L)f(y)\,dy = \int_L^\infty y f(y)\,dy - L(1-F(L)),$$

[27] See Equations 13.4.21 and 13.4.18 in *Handbook of Mathematical Functions* [1].

$$\int_L^\infty yf(y)dy = s\alpha\lambda^s l^s \int_L^\infty y^{-s} U(s+1, s+1-\alpha; \lambda l/y)dy.$$

Letting $z = \lambda l/y$ this integral becomes

$$-s\alpha \int_0^{\lambda l/L} z^s U(s+1, s+1-\alpha; z)\frac{\lambda l}{-z^2}dz$$

$$= \lambda l s\alpha \int_0^{\lambda l/L} z^{s-2} U(s+1, s+1-\alpha; z)dz.$$

Using the theorem from Appendix D:

$$\int_L^\infty yf(y)dy = \frac{\lambda l s\alpha z^{s-1}}{s\alpha}\left[U(s, s+1-\alpha; z) + \frac{U(s-a, s+1-\alpha; z)}{(s-1)(\alpha-1)}\right]_{Z=0}^{\lambda l/L}$$

$$= \lambda l \left(\frac{\lambda l}{L}\right)^{s-1}\left(U\left(s, s+1-\alpha; \frac{\lambda l}{L}\right) + \frac{U(s-1, s+1-\alpha; \lambda l/L)}{(s-1)(\alpha-1)}\right).$$

Now

$$L(1-F(L)) = \frac{\lambda^s l^s}{L^{s-1}} U(s, s+1-\alpha; \lambda l/L).$$

Thus the numerator of the excess ratio is

$$\int_L^\infty yf(y)dy + L(1-F(L))$$

$$= \frac{\lambda^s l^s}{(s-1)(\alpha-1)L^{s-1}} U(s-1, s+1-\alpha; \lambda l/L).$$

The denominator of the excess ratio is the total (developed) losses or the mean of the undeveloped losses times the mean loss development. The former is the mean of the Pareto or $\lambda/(\alpha-1)$. The latter is the mean of the inverse gamma or $l/(s-1)$.

Thus the excess ratio is:

$$R(L) = \left(\frac{\lambda l}{L}\right)^{s-1} U(s-1, s+1-\alpha, \lambda l/L).$$

Note that this compares to the excess ratio for the undeveloped losses (given by a Pareto) of $(\lambda/\lambda + L)^{\alpha-1}$. For z small:[28]

$$U(a,b,z) \approx z^{1-b}\Gamma(b-1)/\Gamma(a) \quad b > 2.$$

Thus for large limits L, and $s > \alpha + 1$

$$\begin{aligned} R(L) &= \left(\frac{\lambda l}{L}\right)^{s-1} U(s-1, s+1-\alpha, \lambda l/L) \\ &\approx \left(\frac{\lambda l}{L}\right)^{s-1} \frac{\Gamma(s-\alpha)}{\Gamma(s-1)} \left(\frac{\lambda l}{L}\right)^{\alpha-s} \\ &= \left(\frac{\lambda l}{L}\right)^{\alpha-1} \Gamma(s-\alpha)/\Gamma(s-1). \end{aligned}$$

For the Pareto for large limits

$$R(L) = (\lambda/(\lambda+L))^{\alpha-1} \approx (\lambda/L)^{\alpha-1}.$$

Thus the ratio of the excess ratios for the developed and the undeveloped losses is approximately: $l^{\alpha-1}\Gamma(s-\alpha)/\Gamma(s-1)$. If the mean development is unity, then $l = s-1$. Then this ratio is:

$$\begin{aligned} &(s-1)^{\alpha-1}/\{(s-\alpha-1)\cdots(s-2)\} \\ &\approx \left((s-1)\Big/\left(s-\frac{\alpha}{2}-1\right)\right)^{\alpha-1}. \end{aligned}$$

Since for a Pareto for large limits $R(L) \approx \lambda^{\alpha-1}/L^{\alpha-1}$ if one adjusts λ by multiplying by a factor of $(s-1)/(s-(\alpha/2)-1)$, then one will approximately multiply the excess ratios by the desired adjustment factor.

[28] See Equation 13.5.6, *Handbook of Mathematical Functions* [1].

APPENDIX D

CONFLUENT HYPERGEOMETRIC FUNCTIONS[29]

There are a number of related functions referred to as confluent hypergeometric functions. They can be usefully thought of as generalizations of the beta and gamma functions. They can be thought of as two parameter distributions. Let:

$$M(a,b,z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_0^1 e^{zt}t^{a-1}(1-t)^{b-a-1}\,dt,$$
$$U(a,b,z) = \frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}\,dt.$$

Then M can be computed using the following power series in z:

$$M(a,b;z) = 1 + \frac{az}{b} + \frac{a(a+1)z^2}{b(b+1)(2!} + \frac{a(a+1)(a+2)z^3}{b(b+1)(b+2)(3!)} + \cdots.$$

U can be computed as a combination of two values of M:

$$U(a,b;z) = \frac{\pi}{\sin\pi b}\left(\frac{M(a,b,z)}{\Gamma(1+a-b)\Gamma(b)} - \frac{z^{1-b}M(1+a-b,2-b,z)}{\Gamma(a)\Gamma(2-b)}\right).$$

U is related to the incomplete gamma function:

$$U(1-a,1-a;x) = e^x\Gamma(a;x).$$

Among the facts used in Appendix C are:

$$\frac{d}{dz}U(a,b;z) = -aU(a+1,b+1;z),$$
$$U(a-1,b;z) - zU(a,b+1;z) = (a-b)U(a,b;z).$$

For z small and $b > 2$, $U(a,b;z) \approx z^{1-b}\Gamma(b-1)/\Gamma(a)$.

---

[29]See *Handbook of Mathematical Functions* [1].

THEOREM

$$\int z^{a-3} U(a,b,z)\,dz$$

$$= -\frac{z^{a-2}}{(a-1)(b-a)} \frac{U(a-1,b,z) - U(a-2,b,z)}{\{(a-2)(b+1-a)\}}.$$

Given:

$$\frac{dU(a,b,z)}{dz} = -aU(a+1,b+1,z), \qquad \text{and}$$

$$zU(a,b+1,z) - U(a-1,b,z) = (b-a)U(a,b,z).$$

Proof:

Let

$$\nu = z^{a-2}(U(a-1,b,z) - U(a-2,b,z)/\{(a-2)(b+1-a)\})$$

$$\frac{d\nu}{dz} = (a-2)\nu/z + z^{a-2}\frac{-(a-1)U(a,b+1,z) + U(a-1,b+1,z)}{(b+1-a)}$$

$$= z^{a-3}\{(a-2)U(a-1,b,z) - U(a-2,b,z)/(b+1-a)$$

$$- (a-1)zU(a,b+1,z)$$

$$+ zU(\alpha-1,b+1,z)/(b+1-\alpha)\}$$

$$= z^{a-3}\{(zU(\alpha-1,b+1,z) - U(\alpha-2,b,z))/(b+1-\alpha)$$

$$- (a-1)(zU(a,b+1,z) - U(\alpha-1,b,z)) - U(\alpha-1,b,z)\}$$

$$= z^{a-3}(U(a-1,b,z) - (a-1)(b-a)U(a,b,z) - U(a-1,b,z))$$

$$= -z^{a-3}(a-1)(b-a)U(a,b,z). \qquad \text{Q.E.D.}$$

Errata for Discussion by Howard Mahler of "Retrospective Rating: 1997 Excess Loss Factors"

At the bottom of page 320, the equation for $\hat{R}(100)$ is incorrect.

This example of simple dispersion is an example of a mixture with five pieces.

The excess ratio of the mixture is a weighted average of individual excess ratios, with the weights the product of the means and the probabilities for each piece of the mixture.[1]
If the probability of each piece of a mixture is p_i, $\sum p_i = 1$, the mean of each piece of the mixture is m_i, and R_i is the excess ratio for each piece of the mixture, then $\hat{R}(L) = \sum p_i\, m_i\, R_i(L)$.

If each loss is divided by for example .75, then after development, the excess ratio at L is the same as the original excess ratio at .75 L.[2]
$R_i(L)$ is the excess ratio when the losses have all been divided by r_i.
Thus $R_i(L) = R(r_i\, L)$.

In the example on page 320, each mean is proportional to 1/divisor = $1/r_i$, and each probability is the same at 1/5. Thus the weights are: $(1/5)(1/r_i)$.
The sum of the weights is: $\sum (1/5)(1/r_i) = (1/5)(1/.75 + 1/.833 + 1/1 + 1/1.25 + 1/1.5) = 1$.[3]

Thus $\hat{R}(L) = \sum (1/5)(1/r_i)\, R(r_i\, L) = (1/5) \sum R(r_i\, L)\, /r_i$.

Therefore, the corrected equation at the bottom of page 320 is:

$$\hat{R}(100) = (1/5)\{R(75)/.75 + R(83.3)/.833 + R(100)/1 + R(125)/1.25 + R(150)/1.50\}$$
$$= (1/5)\{.6009/.75 + .5817/.833 + .5582/1 + .5384/1.25 + .5191/1.50\} = .5669.$$

Similarly, the corrected equation at the top of page 321 is:

$$\hat{R}(5000) = (1/5)\{R(3750)/.75 + R(4165)/.833 + R(5000)/1 + R(6250)/1.25 + R(7500)/1.50\}$$
$$= (1/5)\{.0157/.75 + .0070/.833 + 0/1 + 0/1.25 + 0/1.50\} = .0059.$$

[1] See page 154 of "Workers Compensation Excess Ratios: An Alternate Method of Estimation" by Mahler.
[2] If each loss is multiplied by 1/.75 = 1.333, this is mathematically the same as uniform inflation of 33.3%.
Thus we can get the excess ratio after development, by taking the original excess ratio at the deflated value of L/1.333 = .75 L. Increasing the sizes of loss, increases the excess ratio over a fixed limit.
[3] Mahler chose these loss divisors so that the total expected losses are unaffected.

The values in the simple dispersion column of Table 1 at page 320 are revised in a similar manner.

Corrected Table 1
Excess Ratios

LIMIT	No Development	Simple Dispersion	Gamma Dispersion
50	.6888	.6949	.6945
100	.5582	.5669	.5684
500	.3012	.3080	.3076
1,000	.1606	.1705	.1721
2,000	.0904	.0931	.0930
3,000	.0402	.0462	.0459
4,000	.0100	.0194	.0190
5,000	.0000	.0059	.0069
6,000	.0000	.0007	.0024
7,000	.0000	.0000	.0008
8,000	.0000	.0000	.0003
9,000	.0000	.0000	.0001
10,000	.0000	.0000	.0000

As can be seen in corrected Table 1, the simple dispersion effect raises the excess ratios, especially at the higher limits.[4]

At page 326, the formula near the bottom of page should have λ in place of X:

$R(L) = (\lambda \ell/L)^{s-1} U(s-1, s+1-\alpha, \lambda \ell/L)$.

[4] It can be demonstrated that when dispersion has no overall effect, loss dispersion either increases an excess ratio or keeps it the same. In most practical applications, the excess ratio will be increased by loss dispersion.

Errata for Discussion by Howard Mahler of “Retrospective Rating: 1997 Excess Loss Factors”

Appended: 6 July 2009

On page 331, the first formula, for *F(y)*, the term $y/b - a$ should be $y/(b - a)$.