VOLUME LXXXVI NUMBERS 164 AND 165

PROCEEDINGS

OF THE

Casualty Actuarial Society

ORGANIZED 1914

1999

VOLUME LXXXVI

Number 164—May 1999

Number 165—November 1999

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 risks 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 three classes of members—Fellows, Associates, and Affiliates. Both Fellowship and Associateship require successful completion of examinations, held 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. Affiliates are qualified actuaries who practice in the general insurance field and wish to be active in the CAS but do not meet the qualifications to become a Fellow or Associate.

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 (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, 1999

EXECUTIVE COUNCIL*

STEVEN G. LEHMANN . . . . *President*
ALICE H. GANNON . . . . *President-Elect*
CURTIS GARY DEAN . . . . *Vice President–Administration*
KEVIN B. THOMPSON . . . . *Vice President–Admissions*
ABBE S. BENSIMON . . . . *Vice President–Continuing Education*
DAVID R. CHERNICK . . . *Vice President–Programs & Communications*
ROBERT S. MICCOLIS . . . . . *Vice President–Research & Development*

THE BOARD OF DIRECTORS

*Officers**

STEVEN G. LEHMANN . . . . *President*
ALICE H. GANNON . . . . *President-Elect*

Immediate Past President†

MAVIS A. WALTERS 1999

Elected Directors†

SHOLOM FELDBLUM 1999
RUSSELL S. FISHER 1999
DAVID N. HAFLING 1999
RICHARD J. ROTH JR. 1999
PAUL BRAITHWAITE 2000
JEROME A. DEGERNESS 2000
MICHAEL FUSCO 2000
STEPHEN P. LOWE 2000
CHARLES A. BRYAN 2001
JOHN J. KOLLAR 2001
GAIL M. ROSS 2001
MICHAEL L. TOOTHMAN 2001

* Term expires at the 1999 Annual Meeting. All members of the Executive Council are Officers. The Vice President–Administration also serves as the Secretary and Treasurer.

† Term expires at Annual Meeting of year given.

1999 PROCEEDINGS
CONTENTS OF VOLUME LXXXVI

Page

1999 PROCEEDINGS
CONTENTS OF VOLUME LXXXVI

1999 PROCEEDINGS
CONTENTS OF VOLUME LXXXVI

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*.

Editorial Committee, *Proceedings* Editors

Volume LXXXVI, Part 1 **No. 164**

PROCEEDINGS
May 16, 17, 18, 19, 1999

LEVELS OF DETERMINISM IN WORKERS COMPENSATION REINSURANCE COMMUTATIONS

GARY BLUMSOHN

Abstract

When commuting workers compensation reinsurance claims, the standard method is to project the future value of the claims using stated assumptions for future medical usage, medical inflation, cost-of-living adjustments, and investment income. The actuary selects a best estimate for each variable, and assumes this deterministic number will be realized in the future. To account for the date of death being stochastic, a mortality table is used to model the future lifetime.

By assuming deterministic values for future medical usage, medical inflation, cost-of-living adjustments, and investment income, the calculation ignores the possibilities of higher or lower values. It is shown that these do not generally balance out, and that the standard method produces biased results. In low reinsurance layers, the *commutation amount is overstated, and in high layers it is understated. By removing deterministic assumptions from the calculation, bias is removed from the results. The paper gives a detailed, realistic, example to illustrate this.*

It is impossible to eliminate all determinism, but it is often appropriate to judgmentally adjust the answers to account for this. In discussing this, the paper draws parallels to the work of economists on "genuine uncertainty."

The implications of the paper reach beyond the narrow realm of workers compensation reinsurance commutations. The most obvious implications are for workers compensation reserving, but the essential message applies to pricing and reserving of any excess insurance and reinsurance: deterministic assumptions often lead to biased results.

ACKNOWLEDGEMENT

The author is grateful to Eric Brosius, Sholom Feldblum, Joe Gilles, Richard Homonoff, Tony Neghaiwi, Jill Petker, John Rathgeber, Lee Steeneck, Mike Teng, Bryan Ware, Wendy Witmer, and the anonymous referees of the Committee on Review of Papers for providing comments on earlier drafts of this paper. Remaining mistakes are, of course, my responsibility.

1. INTRODUCTION

Excess reinsurance for workers compensation generally pays out over many decades. While workers compensation claims are usually reported to the insurer soon after the accident, and the insurer may soon report them to the reinsurer, the loss payments are slow, being made over the lifetime of the injured worker or even the lifetime of uninjured dependents. Consequently, even for reinsurance with a relatively modest retention, it can take many years to breach the retention, and many more years to exhaust a layer. Gary Venter [17] has estimated that it takes, on average, over 30 years to pay half the ultimate claim amount.

At some point after an excess reinsurance treaty ends, but before the losses have been fully paid, it is common to commute either the reinsurance treaty or the individual reinsured claims. The commutation entails having the reinsurer pay the ceding company a flat amount, in exchange for canceling future liabilities. This saves costs for both parties, since the expense of reporting on claims to the reinsurer and the cost of paying these claims are eliminated. It allows the parties to shut their reinsurance files and spend their time on more profitable activities.

The actuarial techniques for evaluating workers compensation commutations differ from the techniques generally used in commutations of other lines of business. With workers compensation (and in some other cases, like unlimited medical benefits for no-fault auto) the population of claims is generally known at the time of the commutation—there is very little lag in claims being reported to the primary company. Also, the amount of the payments does not depend on some future court verdict. The payments are based on a fixed annual indemnity amount, subject, in some states, to an annual cost-of-living adjustment (COLA), and on the actual medical payments incurred by the claimant. In the case of permanent-total disability cases, these payments often continue for the rest of the claimant's life. Since the losses are so closely tied to the claimant's life span, it is natural to use the mortality techniques more generally associated with life actuaries than with their property/casualty brethren.

While the actuarial techniques in these calculations are by now well-accepted, this paper will argue that the results are systematically biased and can be improved upon. The life-table techniques generally assume that mortality is stochastic, but that other variables (amount of medical care, inflation rates, investment yields) are deterministic. These deterministic variables can be stripped away, much as earlier actuaries stripped away the assumption of deterministic mortality. By doing this, we improve the accuracy of our calculations and eliminate some biases.

Though this paper will express the issues in terms of commutations, the issues are similar when doing excess workers compensation case reserving using life-table methods. And, as will be noted later, the same issues find their way into most actuarial work.

2. LIFE-TABLE TECHNIQUES

Method 1: Totally Deterministic Calculation

The simplest method for performing the calculation is to assume the claimant will live to his life expectancy and then calculate the present value of the future stream of payments for this time. This method, though simple and appealing, is wrong. As actuaries are well aware, and as will be discussed in detail later, assuming a deterministic life-span leads to systematically incorrect results.

Method 2: Stochastic Date Of Death

The actuarial literature contains several papers that discuss the calculation of reserves for long-term workers compensation cases, and the calculation of a commutation value only differs in minor respects from the calculation of a reserve.[1] Actuaries

[1] The classic paper is Ronald Ferguson's *Actuarial Note on Workmen's Compensation Loss Reserves* [8], which applied life-table methods to excess indemnity reserves. He did not address the issue of the medical portion of the reserve. Richard Snader [15] applied similar methods to long-term medical claims. A recent valuable addition to the literature is by Lee Steeneck [16], who uses an analysis very close to "Method 2" discussed in this paper. Another approach is given in Venter and Gillam [18].

and, to a lesser extent, the wider insurance community, generally accept that the right way to reserve these claims is through the life-table techniques routinely used by life actuaries. Life-table techniques take into account the probabilities of the claimant dying either earlier or later than his life expectancy, rather than assuming he lives to his life expectancy and then dies.

Using a life table to make the number of payments stochastic, rather than deterministic, is a crucial advance in the accuracy of the calculation. A life-table approach allows for the possibility that a claimant may live to age 95, and hence pierce reinsurance layers that would not have been pierced if he had died at his life expectancy. In other words, if the claimant lives to his life expectancy of, say, 75, a retention of $5 million may not be breached. But if he lives another 10 years, to 85, the total payments in the additional 10 years of life may be enough to breach the $5 million retention, and if he lives to 95, it may breach a $10 million retention. The probabilities of living to these ages, and thus breaching higher layers, must be reflected in the commutation price.

Put another way, there will be a positive commutation amount in layers that we do not expect to get hit. The commutation is effectively a purchase of reinsurance by the reinsurer, covering the possibility of the claimant breaching the retention. There need not be a guarantee that the retention will be breached in order for the expected losses in the layer to be positive.

Assumptions

In doing the commutation calculation, the actuary needs to make a number of assumptions:[2]

---

[2] In practice, some reinsurance contracts have commutation clauses in which the parties have negotiated some of the parameters at the time the contract is drawn up. For example, the clause may specify what mortality table to use and what interest rate to use in discounting the future payments.

- An appropriate *mortality table* must be selected.
- For workers compensation, the indemnity amount is generally known, but it may be subject to *cost-of-living adjustments*, which depend usually on movements in the average weekly wage in the state.
- The amount of medical expenses must be estimated for each year in the future. This is usually done in two steps: first, estimate the future *annual medical expense* in today's dollars, and, second, estimate future *medical price inflation*, to convert today's dollars into tomorrow's dollars.
- The *rate at which to discount* future dollar payments to present value.

Once assumptions have been chosen, the calculations can be performed, and the parties can agree on an amount for settlement.[3]

3. LEVELS OF DETERMINISM

The life-table method ignores fluctuations in other key variables. Just as it is wrong to assume a claimant's life-span is fixed, so it is wrong to assume that medical usage and inflation

---

[3] This paper will not address the crucial impact of income tax. In the calculations, one must account for taxes without the commutation, compared to taxes with the commutation.

i) If the claim is not commuted, the reinsurer carries a reserve on its books. For tax purposes, this reserve is discounted by the IRS discount factors, and the unwinding of the discount is counted into the incurred losses of the company each year. On the other hand, the investment income earned on the reserve is taxable.

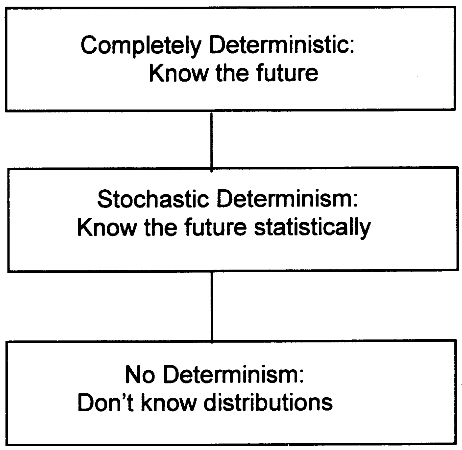
ii) If the claim is commuted, the reinsurer takes down the reserves and puts up a paid loss. If the reserve exceeds the paid loss (as it frequently does, because statutory accounting demands undiscounted, or tabularly-discounted, reserves) the reinsurer's profit rises by the difference between the reserve and the paid loss. This profit is taxable.

The ceding company has the reverse entries on its books.

The tax benefits or costs are as important as any other cash flows, but they are beyond the scope of this paper. For a detailed discussion of the tax effects, see Connor and Olsen [5].

FIGURE 1

LEVELS OF DETERMINISM



are fixed. Assuming a deterministic life-span leads to inaccurate calculations. Likewise, assuming deterministic medical care and inflation will lead to inaccurate calculations. A deterministic life span implies that high layers of reinsurance will not be hit, when they do, in fact, have a chance of getting hit if the claimant lives long enough. Likewise, deterministic medical care and deterministic inflation understate the costs to the highest reinsurance layers.

Actuarial calculations can contain varying levels of determinism, and this can be represented as shown in Figure 1.

At the "completely deterministic" level, our calculations assume we know what the future will bring. This is the viewpoint of typical loss development work: we look at the historical loss development patterns, select patterns to represent the future, and develop the losses to ultimate. We assume that the selected patterns represent loss emergence in the future, and we make no allowance for deviations from the selected patterns. In many uses, this approach is perfectly reasonable. In others, and particularly in dealing with excess reinsurance, it can generate misleading results.

The next stopping point on the continuum of determinism is what I call "stochastic determinism." Here we do not assume that we know what the future will be, but we assume that we know the statistical distributions of the relevant variables. For example, Ferguson [8] pointed out that we do not know when a workers compensation claimant will die, but we have mortality tables that tell us the probability of dying at any given age. Using these probabilities, Ferguson showed, generates a more accurate answer to the required reserve for an excess workers compensation claim.

Note, though, that the typical actuarial approach to workers compensation cases does not have all variables stochastic. For example, the rate of medical inflation, cost-of-living adjustments, investment yields, and the annual real amount of medical expenses are assumed to be fixed. The typical approach (to be labeled "Method 2" later in this paper) is partway between complete determinism and stochastic determinism. The calculations in Section 4 of this paper will shift the approach further towards stochastic determinism.

At the end of the continuum is "no determinism," which is where we assume that we do not know even the distributions that underlie what will happen in the future. We can imagine various scenarios occurring in the future, but we cannot assess the probabilities. We know, for example, that doctors might find a way to surgically fix the damage to a quadriplegic, and thus get him back to work and end his workers compensation claim. But we do not know the probability of this happening. This is obviously the hardest level to deal with from an actuarial standpoint.

We will return later to a more detailed discussion of these various levels of determinism. At this point it is sufficient to notice that Ferguson's paper stripped away some determinism from the workers compensation calculation by making mortality stochastic. To add even greater accuracy, we need to strip away more determinism.

4. A COMPREHENSIVE EXAMPLE

This section gives a realistic example of how one can remove more determinism from the model. The calculations are significantly more complex than the standard life-table method. However, using computers, the problems are not insurmountable, and the results are significantly less biased.

The Data

Suppose we are commuting the following claim:[4]

- Joe Soap has been permanently and totally disabled since 1993. On January 1, 1998, the effective date of the commutation, he will turn 35 years old.
- Through 12/31/97, the primary company has paid out \$300,000 in medical expenses and \$70,000 in indemnity payments.[5] This is an unusually large claim, but by no means unheard of. A smaller claim would not affect any of the conclusions.
- In 1997, Mr. Soap received indemnity payments at the rate of \$20,000 per year, but these are subject to a cost-of-living adjustment that is effective on January 1 of each year, based on

[4]A similar example was used in a previous version of this paper (Blumsohn [1]). Some items have been updated to incorporate more recent data. Substantive changes from the previous version will be noted.

[5]For simplicity, the example ignores ALAE, which is usually covered by reinsurance and should be included in the calculations. ALAE is usually relatively small in workers compensation, and including it would not change any of the principles discussed in this paper.

the increase in the state average weekly wage over the previous year.

- The best estimate of his future medical expenses is $70,000 per year, in 1997 dollars. These will increase with medical inflation.
- We assume that Joe's mortality follows that for the overall male population, as shown in the 1990 US census (Exhibit 1). Based on this mortality table, his life expectancy is 39.6 years.[6]
- We project future inflation of 4.11% per year.[7] For convenience, we assume that changes in the state average weekly

---

[6]Depending on the claimant's condition, one may use impaired-mortality tables. Note that, contrary to the usual intuition, workers compensation lifetime-pension cases do not, overall, appear to have higher mortality rates than the general population. Gillam [9] shows that at some ages, the mortality of workers compensation claimants is even lower than the general population. Gillam's technique weights each claimant equally. That may not be the optimal approach, since some claims are bigger than others. In particular, many of the really big claims are for people who are extremely badly injured and require, say, 24-hour attendant care. One might speculate that a dollar-weighted average of mortality could be found to be significantly worse than the general population.

By using the 1990 census table, we ignore mortality improvements: as medical care improves, mortality rates have historically dropped. By ignoring mortality improvements, we are implicitly assuming Joe Soap has impaired mortality.

[7]The 4.11% rate is the average of Consumer Price Index changes from 1935 to 1997, using data from the US Bureau of Labor Statistics. Using this average is a matter of convenience, rather than a matter of believing that it is a good predictor of future inflation. The data, though not a predictor of future inflation, give an idea of long-term inflationary movements.

The earlier version of this paper (Blumsohn [1]) used 4.2%, based on data from 1935 to 1995. Steeneck ([16], p. 252), when faced with projecting indemnity inflation into the indefinite future, selects 4.0% as his annual rate.

The author admits to cringing at the spurious accuracy implied in publishing an inflation average to two decimal places. Past inflation is a poor way of predicting future inflation, and there's no scientific way to project inflation decades into the future to even the nearest whole percent, never mind two decimals. We are reminded of Gauss's comment that "Lack of mathematical culture is revealed nowhere so conspicuously as in meaningless precision in numerical computations." (Quoted in Coddington [4, p. 160].) However, the problem is that we are trying to contrast various methods of doing the computations, and this requires keeping the assumptions and arithmetic in the methods as consistent as possible, to avoid obscuring the main message by implicitly switching assumptions. The only way to transparently do this was to use more decimal places than are meaningful.

wage follow the overall price inflation in the economy. (Generally, wages actually rise faster than prices over the long run because of productivity improvements.)

- Our best guess of future medical inflation is 5.25% per year.[8] Exhibit 2 shows historical changes in the CPI and medical CPI.

- Selecting an appropriate discount rate is somewhat tricky. The future cash flows are highly uncertain, and the uncertainty arises from two principal places:

  i) *Mortality*: We do not know how long the claimant will live. However, if the insurer and reinsurer both write reasonably large books of workers compensation, the mortality risks of the individual claimants will be diversified away.

  ii) *Inflation*: Wage inflation affects cost-of-living adjustments and medical inflation affects medical payments. This risk cannot be diversified by writing a large book because all claimants are subject to the highs and lows of inflation together.

In setting its investment strategy, the insurer would be wise to hedge against inflation by buying investments that rise when inflation is high—for example common stocks. (See Feldblum [7].) This strategy is particularly appealing for excess workers compensation, where the payouts are extremely slow, so the year-to-year volatility of stock prices are less of a concern.

[8]As with CPI changes, this average is based on changes in the medical component of the CPI from 1935 to 1997. The earlier version of this paper used data from 1935 to 1995 and had average medical inflation of 5.36% per year. As with the CPI, we are using this number for illustrative purposes, rather than as a considered prediction of future medical inflation. Steeneck [16, p. 252] projects annual medical inflation of 5.5%.

Starting in 1997, another inflation hedge was introduced in the market, namely, inflation-indexed Treasuries. Like other Treasuries, they are considered "risk-free" in the sense of not having default risk, and, unlike other Treasuries, they hedge against inflation as well.[9]

For discounting, we will use inflation-indexed Treasuries. At January 1, 1998, these had a real yield (above inflation) of about 3.75%. In general, discounting should be based on a rate below the investment yield, with the risk adjustment accounting for the riskiness in the flows being discounted (Butsic [3]). I will assume that a reasonable risk adjustment for excess workers compensation is 2.5 percentage points. In other words, we will discount at a real yield of 1.25% (= 3.75% − 2.5%).

As noted above, inflation is assumed to be 4.11% per year. Discounting at a real yield of 1.25% thus entails adding 1.25% to the assumed inflation of 4.11%, to get a discount rate of 5.36% per annum.[10]

- The primary insurer has purchased reinsurance in a number of layers, as shown in Table 1.

---

[9]The hedge for excess layers of workers compensation is imperfect because:

i) They are indexed to the CPI, whereas the workers compensation risk is based on changes to the state average weekly wage (for COLAs) and the medical component of the CPI. The CPI is only a proxy for these.

ii) Excess reinsurance layers suffer a leveraged effect from inflation. For example, suppose a reinsurer covers a layer of $1 million excess of $1 million, and there's a $1.1 million claim, with no inflation. In that case, the reinsurer will pay $100,000. If there's 10% inflation, raising the claim to $1.21 million, the reinsurer's portion more than doubles, to $210,000. (Of course, if the claim without inflation were, say, $3 million, inflating it to $3.3 million would have no effect on the reinsurance layer. This does not affect the general point that excess layers are typically more sensitive to inflation than are ground-up layers.)

[10]The earlier version of the paper assumed the risk-adjusted discount rate was exactly equal to the inflation rate.

TABLE 1

REINSURANCE LAYERS

Layer 1	$130,000 excess of $370,000
Layer 2	$500,000 excess of $500,000
Layer 3	$1 million excess of $1 million
Layer 4	$3 million excess of $2 million
Layer 5	$5 million excess of $5 million
Layer 6	$5 million excess of $10 million
Layer 7	$5 million excess of $15 million
Layer 8	$10 million excess of $20 million
Layer 9	$10 million excess of $30 million
Layer 10	$10 million excess of $40 million
Layer 11	$10 million excess of $50 million
Layer 12	$10 million excess of $60 million
Layer 13	$10 million excess of $70 million
Layer 14	$10 million excess of $80 million
Layer 15	$10 million excess of $90 million
Layer 16	Unlimited excess of $100 million

The first layer is somewhat artificial: since $370,000 has already been paid by the end of 1997, the layer will pay from the first dollar in 1998. This allows us to look at the value of all future payments. Also, the top layer is somewhat unusual. Reinsurers do not usually sell unlimited layers. However, it will be instructive to see the value of reinsurance on the unlimited top layer.

Method 1: Totally Deterministic Calculation

Though actuaries would not use a totally deterministic method (i.e., one that assumes Joe lives exactly to his life expectancy and then dies) it is interesting to see what result this produces. Exhibit 3 shows this calculation, and Table 2 summarizes the results.

Total payments are $11.2 million, exhausting the lowest five layers and part of the sixth. The lack of payments in higher layers

TABLE 2

RESULTS OF COMMUTATION CALCULATIONS USING METHOD 1

Layer (in $000's)	Nominal Payments (in $000's)	Present Value of Payments (in $000's)
130 xs 370	130	125
500 xs 500	500	413
1,000 xs 1,000	1,000	612
3,000 xs 2,000	3,000	1,092
5,000 xs 5,000	5,000	970
5,000 xs 10,000	1,606	217
Higher Layers	0	0
Total, All Layers	**11,236**	**3,430**

implies these layers will not be breached, and no commutation payment is needed.

This method ignores the chances of dying before or after one's life expectancy. We correct this by using a life-table approach, following Ferguson [8].

Method 2: Stochastic Date of Death

In Method 2, a mortality table models Joe's life span, as shown in Exhibit 4. Table 3 compares the commutation amounts from Methods 1 and 2.

Comparison of Method 2 Versus Method 1

Several points are worth noting:

- Using Method 2, twelve layers have non-zero commutation amounts, compared to only six layers in Method 1. This is because Method 2 recognizes that people can live beyond their life expectancies. If the person lives to the outer reaches of the mortality table, say to 110, many more layers will be breached. The highest layer reached is $10 million excess of $60 million,

TABLE 3

COMPARISON OF RESULTS OF COMMUTATION CALCULATIONS FOR METHODS 1 AND 2

Layer (in $000's)	Expected Nominal Payments (in $000's)		Expected Present-Value Payments (in $000's)	
	Method 1	Method 2[11]	Method 1	Method 2
130 xs 370	130.0	129.7	124.9	124.6
500 xs 500	500.0	494.9	413.2	409.1
1,000 xs 1,000	1,000.0	970.4	611.7	594.1
3,000 xs 2,000	3,000.0	2,725.1	1,092.4	998.6
5,000 xs 5,000	5,000.0	3,703.0	970.4	729.8
5,000 xs 10,000	1,605.9	2,574.7	217.1	311.2
5,000 xs 15,000	0.0	1,607.4	0.0	139.8
10,000 xs 20,000	0.0	1,359.7	0.0	86.5
10,000 xs 30,000	0.0	293.0	0.0	13.2
10,000 xs 40,000	0.0	39.2	0.0	1.4
10,000 xs 50,000	0.0	3.1	0.0	0.1
10,000 xs 60,000	0.0	0.1	0.0	0.0
Higher layers	0.0	0.0	0.0	0.0
Total, All Layers	**11,235.9**	**13,900.4**	**3,429.8**	**3,408.3**

compared to only the $5 million excess of $10 million layer using Method 1.[12]

- For all layers combined, which translates to the value of all future amounts payable to the claimant, the nominal total from Method 1 ($11.2 million) is considerably lower than the nominal total from Method 2 ($13.9 million). However, the present value from Method 1 ($3.43 million) is about the same as the

[11] "Nominal" payments for Method 2 are discounted for mortality, but not for the time-value of money.

[12] Exhibit 4 in fact shows that the maximum possible loss for Method 2 is $74 million, which is one layer higher than is reflected in the text. The tiny probability of this happening means that the expected losses in the layers above $70 million are below $1,000, and thus do not show up on Table 3. In other words, the numbers are different, even though rounding makes them look the same.

present value from Method 2 ($3.41 million). How can we explain this?

i) *Nominal total from Method 2 considerably greater than Method 1* The easiest way to explain the relation between the nominal totals is by analogy to a more familiar idea involving annuities, namely, that the present value of a life annuity is less than the present value of an annuity certain for the person's life expectancy. (Bowers [2], pp. 149–150 (example 5.13) and p. 158 (exercise 5.45).) In other words, the expected cost of paying someone $1 per year for life is less than the cost of paying $1 per year for a guaranteed period equal to the person's life expectancy. The intuition is that if you pay for the person's actual lifetime, there's a chance of living beyond the life expectancy, and those payments are discounted at a higher rate than the earlier payments. By contrast, the annuity certain ignores the possibility of these higher discounts.

   How does this relate to the nominal payments from Method 2 being much greater than Method 1? In our situation, we have inflation affecting the payments in two ways: the indemnity amounts are increased by the annual cost-of-living increase, and the medical amounts are increased by the annual medical inflation. If the claimant lives to, say, 95 years old, there will be many years of inflation increasing the annual payments beyond the inflation contemplated in Method 1, which halts at the life expectancy. Thus, without inflation, the nominal amounts from Methods 1 and 2 would be identical; with inflation, the nominal amount from Method 1 will be lower than that for Method 2.

ii) *Present value of Method 2 almost the same as Method 1* Without inflation, the payments would be the same each year. Then, as noted above, the present value of Method 1 (an annuity certain for the life expectancy) would exceed the present value for Method 2 (a life annuity). When there is inflation, things are more complicated. The issue is whether the effect of the additional inflation beyond the life expectancy outweighs the effect of the additional discounting. Depending on the rates used for inflation and discounting, the present value of Method 2 could be either higher or lower than the present value of Method 1. Though the total present values for Methods 1 and 2 are close, the amounts in particular layers differ considerably.

- On the layers that are pierced by Method 1, the commutation value from Method 2 is lower than the value from Method 1. For example, on the $500,000 excess $500,000 layer, the value under Method 1 is $413,200, while under Method 2 it's $409,100. This is because Method 1 assumes the amounts are paid for certain, and discounts only for the time-value of money. By contrast, Method 2 recognizes that the claimant may die early, so the amounts may not be paid. Of course, in the layers not pierced in Method 1, the commutation value for Method 2 is always higher.
- We can make no general statement about whether a commutation calculated using Method 1 will produce a total amount, for all layers combined, that is greater than or less than the total for Method 2. For example, if the primary company buys reinsurance on only very low layers, Method 1 will tend to be higher. If it buys reinsurance only on high layers, Method 2 will tend to be higher.

Determinism and Risk

Once a claim has been commuted, the cedent takes the risk of future losses. If the claimant lives to a ripe old age, the primary company will suffer a loss—it would have been better off not to have commuted. That's not a problem: insurance is about taking risks. The commutation calculation measured the mortality risk, and included it in the commutation price. Though the primary company may not be happy to have to pay higher than expected losses, the mortality risk has been priced into the commutation amount.

But there are other risks faced by the ceding company that have not been priced into the commutation amount. Medical inflation is one such example.

The assumed rate of medical inflation is often a contentious issue in commutation negotiations. The parties may argue over whether we should use the average for the past decade (currently about 6%), a longer term average (also about 6% if we average back to World War II), or an econometrician's projection for medical inflation for the next decade. In many cases we are projecting inflation for 70 years or more, so we cannot expect our numbers to be perfect. But often the parties find a number on which they can agree—let us assume it is 5.25%, and let us assume this number is, indeed, the future long-term average medical inflation rate. If the parties use Method 2 with 5.25% medical inflation, and agree on the amount, the ceding company appears to have been compensated for future inflation.

But the ceding company has not, in fact, been compensated for future inflation. It has been compensated for a fixed 5.25% future inflation. It faces the risk that 2 or 3 years hence, there will be high medical inflation, say 20% or 25% per year, for 3 or 4 years, after which medical inflation will drop back to its long-term average. This period of abnormally high medical inflation will quickly erode the retention, which is in nominal dollars, and breach the excess layers much more quickly than the commutation calculation assumes.

There is, similarly, a chance that medical inflation for the next few years will be lower than the long term average, and high medical inflation may not occur for another 60 years. Over the course of the 70 years, one might expect things to even out. So, the skeptic may ask, why should we care? If, on average, it evens out, and if a company does a large number of commutations

TABLE 4

MEDICAL AMOUNT PAYABLE EACH YEAR

Year	Scenario 1: 5% inflation each year	Scenario 2: 20% inflation in year 1; 0% in all other years	Scenario 3: 20% inflation in year 4; 0% in all other years
0	100.00	100.00	100.00
1	105.00	120.00	100.00
2	110.25	120.00	100.00
3	115.76	120.00	100.00
4	121.55	120.00	120.00
Total	**552.56**	**580.00**	**520.00**

over a large number of years, the overall result will be about right.

The problem is that it will not be "about right," as things do not average out in the long run. Just as Method 1 gave biased results, so Method 2, by assuming certain inputs are deterministic, gives biased results.

The Effects of Variable Inflation

To see why things do not average out, let us examine the effects of variable inflation more closely. Consider, on Table 4, an average inflation rate of 5% per year in each of 3 scenarios, and assume the pre-inflation amount payable per year is $100.

Inflation early on (scenario 2) raises the nominal dollar amounts in all future years, causing the total nominal amount to be higher. If there is reinsurance on these payments, the reinsurance retention would be breached earlier, and perhaps a layer will be breached that would not otherwise have been breached. The average inflation over the three scenarios is the same, but

Scenario 2 results in more dollars of medical expenses, and Scenario 3 results in fewer dollars of medical expenses.[13]

For a given average inflation rate, the path of inflation over the life of the claim will affect the future payments: high inflation early on will result in higher amounts; low inflation early on will result in lower amounts. While the total amount over all layers of reinsurance may roughly average out to be the same when present-valued, the amounts within the various layers will differ significantly.

If there is high inflation early on, the reinsurance retention will be breached earlier than expected. There is thus a greater chance that the claimant will still be alive to receive the payment. This greater possibility of payment directly affects the commutation calculation.

The standard commutation calculation fails to include certain risks, and thus neglects to price them. Method 2 assumes mortality is stochastic, but that medical inflation is deterministic. It also assumes wage inflation (and hence cost-of-living adjustments, in states that have them), investment income, and the annual medical usage of the claimant are deterministic. Analogous to Method 1 overstating the lower layers and understating the higher layers, Method 2 will generally bias the commutation amount upwards for lower layers and downwards for higher layers. ("Higher" and "lower" is relative to the size of an individual claim.) Making each of these factors stochastic removes some of the bias in the calculation.

Method 3: Stochastic economic factors and medical costs

Method 3 incorporates several additional random variables into the calculation:

---

[13]Lee Steeneck pointed out that it might be more appropriate to use a geometric mean of inflation in this example, rather than an arithmetic mean. Doing so would somewhat complicate the example, without changing the point being made.

- Inflation is not constant over time. It fluctuates, with the year-to-year rates correlated. [A note on terminology: by "inflation," with no modifier, we mean inflation relating to the overall economy, most popularly measured by the CPI. When referring specifically to price rises for medical care, we will refer to "medical inflation."]
- Medical inflation, while roughly tracking the ups and downs of general inflation, will not be the same as inflation, or even some constant difference from inflation.
- Investment yields fluctuate from year to year, but, like inflation, years are correlated.
- The annual medical payment to the claimant will not be a constant real amount each year. As the claimant's health changes, this amount will change. The claimant may take a turn for the worse, and require $200,000 of hospitalization one year; or he may have a stable period where his medical expense is a lot lower than projected.

Each of these variables needs to be modeled. The specific ways they have been modeled here is not the only way it could be done. The details of the example are less important than the general point being made, namely, that additional fluctuations need to be taken into account.

Inflation

Inflation was modeled using an autoregressive process of the following form:

$$\begin{aligned}&\text{Inflation rate}_{\text{Year } t}\\&\quad= \text{Long-term average inflation rate}\\&\qquad+ \alpha[\text{Inflation rate}_{\text{Year } (t-1)}\\&\qquad\quad- \text{Long-term average inflation rate}]\\&\qquad+ \text{error}_{\text{Year } t}\end{aligned}$$

Daykin, et al. [6, pp. 218–225], discusses this model, and a number of other inflation models that may better fit the data. In the interests of simplicity, this model was chosen. The model starts with a known inflation rate for 1997, and simulates a series of future paths of inflation.

Using least-squares fitting of inflation data from the Bureau of Labor Statistics from 1935–1995, the following parameters were obtained:

$$\text{Long-term average inflation} = 4.11\% \text{ per year}$$
$$\alpha = 0.511.$$

The error term was modeled using a lognormal distribution. Since the error can be positive or negative, but a lognormal is only defined for positive variables, I shifted the lognormal. The best fit was obtained from a shifted lognormal with parameters $\mu = -2.76$ and $\sigma = 0.501$. To ensure a zero mean for the error term, the lognormal was shifted by the mean of this distribution, or about 0.0718. Exhibit 5 shows the derivation of these parameters.

This inflation variable was used to model the cost-of-living adjustment to the indemnity payments. COLAs are usually tied to changes in the state average weekly wage, and wage inflation was assumed to be the same as overall price inflation—a convenient simplification, not necessarily correct. Since most COLAs are capped, the COLA was assumed to not exceed 5% in any year. It was also assumed that if inflation is negative, the indemnity amount would not drop. Since COLAs are lagged a year, it was assumed that the COLA in 1998 is based on 1997 inflation, etc.

Medical Inflation

Medical inflation may be higher or lower than inflation, but they are linked: if the inflation rate were 20% for a sustained period, one would not expect medical inflation to remain at 2%.

The selected model of medical inflation is tied to the overall inflation rate, but with a degree of error allowed. The model is:

$$\begin{aligned}
&\text{Medical Inflation}_{\text{Year } t}\\
&\quad = \text{Inflation}_{\text{Year } t}\\
&\qquad + \beta[\text{Medical inflation}_{\text{Year } (t-1)} - \text{Inflation}_{\text{Year } (t-1)}]\\
&\qquad + [\text{long-term average medical inflation}\\
&\qquad\quad - \text{long-term average inflation}]\\
&\qquad + \text{error}_{\text{Year } t}
\end{aligned}$$

The error term is assumed to be normally distributed, with a mean of zero.[14]

The longest available data series was used to get these parameters. The Bureau of Labor Statistics has medical CPI numbers back to 1935. From 1935 to 1997, average medical inflation was 1.14 percentage points higher than average inflation. This is what was used for the third term of the above expression. We are assuming the long-term trend will continue, although, there is of course no guarantee of this.

The fitted β was 0.38, and the error term was normally distributed with a mean of 0 and a standard deviation of 0.027. Exhibit 6 shows the derivation.

Investment Yields

As noted above, the firm is assumed to invest in inflation-indexed Treasuries, to hedge the inflation risk.[15] These currently have a real yield of about 3.75%. For discounting purposes, a

[14]The inflation model had a lognormal error term, but the medical inflation model has a normal error term. The author had a strong feeling that the error for inflation was skewed, whereas it is less obvious, both from the data and intuitively, that the difference between overall inflation and medical inflation (which largely drives the medical inflation model) is skewed.

[15]It is beyond the scope of the paper to address the question of whether discounting should be based on the firm's (either the reinsurer's or reinsured's) actual investments, or whether it should be based on market discount rates.

2.5 percentage point risk adjustment was made to the rate, thus discounting at 1.25 percentage points above the inflation rate.

For example, if the annual CPI in a particular year is 5.3%, as generated by the autoregressive model discussed above, the discounting for that year would be at 6.55%.

Even if inflation is negative, one would not expect interest rates to drop below some threshold (e.g., 2.5%), so the risk-adjusted discount rate was assumed to not go below zero, i.e., the rate for discounting was set at the greater of the inflation rate plus 1.25% and zero.

Medical Services Used By Claimant

Medical usage will fluctuate from year to year, but we would expect the services from year to year to be correlated. For example, if a claimant has surgery this year, the costs of post-operative treatment may keep the costs higher than average in the next year. One can model this process using an autoregressive model, similar to the one for inflation:

$$\begin{aligned}\text{Medical amount}_{\text{Year } t} &\\ &= \text{Long-term average medical amount}\\ &\quad + \gamma[\text{Medical amount}_{\text{Year } (t-1)}\\ &\qquad - \text{long-term average medical amount}]\\ &\quad + \text{error}_{\text{Year } t}\end{aligned}$$

The long-term average medical amount for this case is, by assumption, \$70,000. Empirically, there does not appear to be a very strong link between last year's medical amount and this year's, so $\gamma = 0.05$ was used. The error term was modeled using a lognormal distribution with $\mu = 10.80089$ and $\sigma = 0.75$. The mean of this lognormal is \$65,000, so the distribution was shifted by 65,000 to ensure the error term has a mean of zero.

TABLE 5

COMPARISON OF RESULTS FROM METHODS 1, 2, AND 3

Layer (in $000's)	Expected Nominal Payments (in $000's)			Expected Present-Value Payments (in $000's)		
	Method 1	Method 2	Method 3	Method 1	Method 2	Method 3
130 xs 370	130	130	130	125	125	125
500 xs 500	500	495	494	413	409	415
1,000 xs 1,000	1,000	970	969	612	594	609
3,000 xs 2,000	3,000	2,725	2,705	1,092	999	1,031
5,000 xs 5,000	5,000	3,703	3,643	970	730	766
5,000 xs 10,000	1,606	2,575	2,591	217	311	344
5,000 xs 15,000	0	1,607	1,788	0	140	175
10,000 xs 20,000	0	1,360	2,093	0	87	152
10,000 xs 30,000	0	293	1,047	0	13	55
10,000 xs 40,000	0	39	558	0	1	23
10,000 xs 50,000	0	3	316	0	0	11
10,000 xs 60,000	0	0	188	0	0	6
10,000 xs 70,000	0	0	117	0	0	3
10,000 xs 80,000	0	0	75	0	0	2
10,000 xs 90,000	0	0	49	0	0	1
Unlimited xs $100MM	0	0	120	0	0	2
Total, All Layers	**11,236**	**13,900**	**16,881**	**3,430**	**3,408**	**3,719**

Running the Model

Each of these parameters was then put into a simulation model. By simulating inflation, medical inflation, and the annual medical amount, one gets a set of input parameters for each simulation. These parameters are then run through the same model as is used in Method 2. The difference is that each time it is run through with different parameters, so that instead of getting a single present value of the future payments, we get a distribution. (Exhibit 7 shows a single simulation from this distribution.)

The means of these distributions, for each layer, are shown on Table 5, compared with the results for Methods 1 and 2.

It is worth noting a few things regarding these results:

- Unlike Methods 1 and 2, Method 3 hits all the reinsurance layers. A less deterministic approach recognizes that higher layers are exposed to loss. Thus, layers that might otherwise have been thought to have no possibility of a loss, are shown to have some commutation value.
- The total nominal value of Method 3 is higher than the nominal value of Method 2 (and Method 2 is higher than Method 1, as discussed earlier).

  This is largely explained by the treatment of inflation. The medical and indemnity amounts paid in some future period depend on the products of (1 + inflation rate) for all prior periods. For example, the amount paid in period 3 depends on what inflation was in periods 1 and 2. The inflation rates are not independent from period to period: the autocorrelation model ensures that they are positively correlated. With positive correlation, the expected value of the product is greater than the product of the expected values, making the overall nominal payments for Method 3 higher than the payments in Method 2.[16]
- The overall present value factor for Method 2 is 25% (= 3,408 ÷ 13,900), but the present value factor for Method 3 is only 22% (= 3,719 ÷ 16,881). In other words, Method 3 has, on average, a steeper discount applied to it.

  This is partly because the year-to-year discount factors (like the inflation factors) are correlated, implying a higher average discount. Also, high medical inflation is correlated with high discount factors, so the higher nominal payments caused by high inflation are more heavily discounted.
- The relationship between the present values of Methods 2 and 3 is complex, largely because the assumptions are not con-

[16] $E(XY) = E(X)E(Y) + \mathrm{cov}(X,Y)$. Thus, if X and Y are positively correlated, the expected value of the product exceeds the product of the expected values.

sistent between the two methods. Yes, we tried to make them consistent, but the differences in the assumptions become clear once we examine them more carefully.

Consider the indemnity cost-of-living adjustments. In Method 2 we used 4.11% for the cost-of-living adjustment. In Method 3, inflation varies stochastically, with a mean of 4.11%. But our cost-of-living adjustment rules were that it couldn't be above 5%, or below 0%. In Method 3, the average inflation rate is 4.11%, but the average cost-of-living adjustment is about 2.9% because it is sometimes capped. A similar, though smaller, discrepancy occurs in the discount rate, due to assuming that the discount rate cannot be negative.

In general, the relationship between the present values of Methods 2 and 3 will depend on the particular assumptions, and how they interact with the various caps and correlations.

- The present value factor for Method 3 losses declines sharply in the higher layers. For example, for the $5 million excess of $5 million layer, the present value is $766,000, compared to the nominal value of $3,643,000. This translates to a present value factor of 21%. By contrast, in the $10 million excess of $90 million layer, the present value factor is only 2%.
- In the lowest layers, the nominal value of Method 1 is higher than Method 2, and Method 2 is higher than Method 3.[17] This

[17] On the earlier table, the nominal values for Methods 2 and 3 look the same in the low layers, but the numbers in the table are rounded. If the complete numbers had been shown, the nominal values in the low layers would be systematically less (though admittedly by a small amount) for Method 3 than for Method 2:

	Nominal Value (in $000's)	
Layer	Method 2	Method 3
1	129.74	129.69
2	494.88	494.44
3	970.39	968.63
4	2,725.08	2,704.59

is because Method 1 implies these layers will be hit for certain, whereas Methods 2 and 3 recognize that the claimant could die before the layer is penetrated. In addition, Method 3 recognizes that there could be years of unusually low claim amounts, so that it may take longer than expected to breach the retention. This reduces the commutation amount in two ways:

i) The longer it is until the retention is breached, the greater the chance of the claimant dying before breaching the retention.

ii) The longer it is until the retention is breached, the steeper the effect of discounting.

In higher layers, which have a lower probability of being penetrated, this situation reverses itself: Method 3 gives higher results than Method 2. The upper layers are most vulnerable to a period of sustained high inflation or high claim levels. Methods 1 and 2 assume inflation and claim levels are fixed, so they do not contemplate any chance of sustained high inflation or claim levels.

- For the lower layers, where the chances are good that the claimant will live long enough to breach them, Method 2 gives similar results to Method 3. But as the layers get higher, the Method 2 number gets lower and lower as a percentage of Method 3, as shown in Table 6.

5. ARE THERE FURTHER LEVELS OF DETERMINISM?

We have shown that the commutation calculation is significantly affected by making a variety of variables non-deterministic. Have we now stripped away all determinism? Put another way, is Method 3 "the perfect" commutation calculation, or is there further determinism that remains?

There is, indeed, further determinism. This paper has shown how we can strip away determinism in the levels of inflation,

TABLE 6

METHOD 2 RESULT AS PERCENTAGE OF METHOD 3 RESULT

Layer	Nominal	Present Value
1	100%	100%
2	100%	99%
3	100%	98%
4	101%	97%
5	102%	95%
6	99%	90%
7	90%	80%
8	65%	38%
9	28%	24%
10	7%	4%
11	1%	0%
Higher Layers	0%	0%

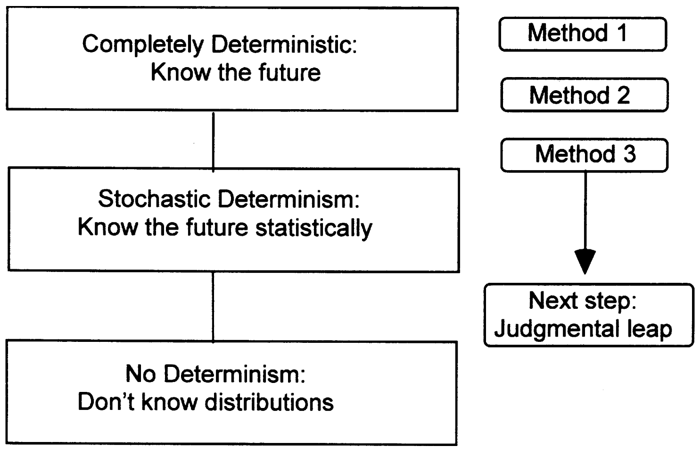
medical utilization, etc. But to measure the paths for these variables, we have relied on statistical measures on past data. Clearly, the historical data may not be valid predictors of the future. For example, the paper assumes that the best predictor of medical inflation is the last 60 years of medical CPI information. One can plausibly argue that what drove medical inflation in the 1930s and 1940s was completely different from what drove it in the 1970s and 1980s, and different from what will drive it in the second half of the twenty-first century. It is quite possible that the drivers of inflation will change periodically over the course of the claimant's lifetime. We have assumed that we know what the future path of medical inflation will be, at the level of a statistical model. But the parameters of the model are deterministic, and so is the structure of the model.

This same issue applies to other variables. For example, advances in medical care could affect the medical utilization for the claimant's condition—and perhaps render the assumed mortality table inappropriate.

In other words, the parameters of our stochastic models could shift, or the model structure itself could change. Method 3 is

FIGURE 2

METHODS 1, 2, AND 3 IN PERSPECTIVE



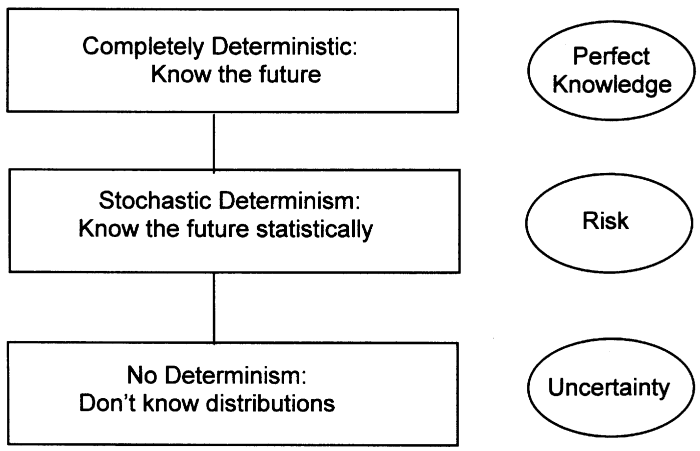
closer to being "stochastically deterministic" than Method 2 is, but it still contains determinism.

The problem is that this next level of determinism is not easily subject to measurement, and hence is not amenable to quantification by the usual actuarial methods. But not being able to quantify does not mean we can simply ignore something.

To put things in perspective, we return now to the graphic introduced at the start of the paper. As Figure 2 makes clear, Method 1 is completely deterministic, Method 2 is somewhat less deterministic, and Method 3 is even less deterministic. But, note carefully that Method 3 is not completely at the level of stochastic determinism, though it is close. There are still various items in Method 3 that are deterministic—for example, mortality rates are assumed to be given. Also, we assume that the parameters of our inflation and interest-rate generators are constant, whereas we could make those parameters themselves stochastic. There are doubts as to whether there is much use in adding these further

FIGURE 3

THE ECONOMIC PERSPECTIVE ON LEVELS OF DETERMINISM



stochastic elements, but the simple point is that Method 3 is not at the level of pure stochastic determinism.

The arrow on Figure 2 shows where we likely need to go after Method 3. The next step requires jumping over the level of pure stochastic determinism and going directly to those items that we cannot measure. Before discussing this, it will be useful to take a brief tour of how economists have viewed some of these issues.

6. THE ECONOMICS OF UNCERTAINTY

The earlier graphic is useful for showing how the ideas in this paper relate to how economists think about risk and uncertainty. Figure 3 repeats the earlier graphic, but now adds some ovals on the right that relate the actuarial ideas to the way economists think about uncertainty.

Many familiar economic models, notably that of perfect competition, assume that people have perfect knowledge. This corresponds with the one end of our continuum: in a completely deterministic calculation, the actuary proceeds as if he or she knows exactly what the future will be.

Moving away from perfect knowledge, economists distinguish between "risk" and "uncertainty."[18] Risk includes things that can be measured statistically, and uncertainty includes things that cannot be measured, but which might occur. For example, if one bets on a fair coin coming up heads, one is facing a risk. But if one bets on the chance of intelligent life being found on an as-yet-undiscovered planet, one faces uncertainty—we have no way of measuring the associated probabilities.

Furthermore, there are events for which we not only do not know the probabilities, but we don't even realize that the event can happen. For example, no actuary pricing liability insurance in the 1930s could even have imagined the wave of asbestos litigation that would hit those policies decades later. This lack of knowledge is sometimes referred to as "sheer ignorance" or "genuine uncertainty."

The economist's idea of risk corresponds to what we called "stochastic determinism": the future is known statistically. And the economist's notion of uncertainty corresponds to what we have called "no determinism."

In practice, most mainstream economics incorporates risk but ignores uncertainty. It is rare to find an economist who deals seriously with uncertainty. And this is, perhaps, for the same reason that one finds so little discussion of this in the actuarial literature—namely, that it is very difficult to include genuine uncertainty in "rigorous" work. Dealing with uncertainty is difficult, and cannot be made numerically precise. Nevertheless, we need to acknowledge that uncertainty is inherent in what we are doing, and that we are fooling ourselves if we believe that our results are perfectly accurate. This applies to both economists and actuaries.

---

[18] The classic reference on risk and uncertainty is Knight [11]. For more recent discussions of the economics of uncertainty, see O'Driscoll and Rizzo [13] and Kirzner [10].

A focus on uncertainty is mainly found outside the mainstream of economics, and is closely associated with the "Austrian" school of economic thought.[19] Their emphasis is on the role of *sheer ignorance* in the economy:

> For the Austrian approach, imperfect information is seen as involving an element which cannot be fitted at all into neoclassical models, that of "sheer" (i.e., unknown) ignorance…[S]heer ignorance differs from imperfect information in that the discovery which reduces sheer ignorance is necessarily accompanied by the element of *surprise*—one had not hitherto realized one's ignorance. (Kirzner, [10, p. 62])

For the Austrians, uncertainty is an inescapable part of human decision-making. We cannot avoid uncertainty, and the fact that it is difficult for economists to quantify and precisely model is not a reason to ignore it.

7. RISK AND UNCERTAINTY IN INSURANCE

Most insurance problems consist of a mixture of risk and uncertainty. Insurers are good at dealing with risk. By measuring the probabilities of loss and pooling risk, we can work to eliminate risk and make losses more stable in the aggregate. It is far more difficult to deal with uncertainty.[20]

---

[19]The "Austrian" school's roots were with Carl Menger at the University of Vienna in the late nineteenth century. Perhaps the best-known Austrian in contemporary times has been Friedrich Hayek, who won the Nobel Prize for Economics in 1974. Today, the main concentrations of "Austrians" are at American universities, most notably New York University and George Mason University. For an introduction to Austrian thought, see Kirzner [10].

Uncertainty is also a concern of some other non-mainstream schools, especially the Post-Keynesians. Some economists, notably George Shackle [14] and Ludwig Lachmann [12] are considered by some to straddle the divide between the Austrians and the Post-Keynesians. For a discussion of Shackle's views on uncertainty, see Coddington [4].

[20]Readers may be tempted to equate the term "risk" with "process risk," and "uncertainty" with "parameter risk." It is advisable to avoid this temptation. Risk, in the sense used by economists, includes both process risk and parameter risk, at least when parameter risk is narrowly defined as the risk of misestimating a parameter due to having a too-small

In this paper, we have been measuring risk: we have only dealt with those things that can be measured. (Insofar as they cannot be modeled well, there are elements of uncertainty.) The next level of determinism consists of uncertainty.

While we cannot easily measure the effect of uncertainty, we can make some qualitative statements about its effects on commutations. Just as removing earlier layers of determinism increased the commutation amount in the higher layers, so removing yet another layer of determinism will increase the commutation amount in higher layers, and higher layers that would not otherwise have been pierced will have some commutation value.

Consider, for example, the inflation model postulated in the example in this paper. There is a real, but very small, chance that the model will generate years where inflation will run above, say, 100% a year as the result of a random blip in the model. In reality, if hyperinflation at that level occurs, it will be more likely to be a result of a structural change in the economy rather than a random event. Since this type of structural change was not included in the data used to fit the model, it is not contemplated in the resulting commutation amount.

Put another way, a completely deterministic model assumes the future will be like the past. Our inflation model, while not completely deterministic, assumes that *fluctuations in* future inflation will be like the past. While this may be more realistic than a completely deterministic model, it is not necessarily true.

All the other variables in the commutation are subject to similar uncertainty: mortality rates might plummet as cures are found for cancer and heart disease; or mortality rates might soar, as a

---

sample size. Narrowly defined in this way, parameter risk can be diversified away, just like any other risk.

In popular usage, parameter risk has acquired a more elastic meaning, to include such things as having an incorrectly structured model. (Uncertainty about the structure of the model is sometimes separated from parameter risk, and called "specification risk" or "model risk.") This is much closer to the economist's notion of uncertainty, and is impossible to quantify. Models that quantify parameter risk almost always have a narrower notion of parameter risk in mind, and so it is confusing to equate uncertainty with parameter risk. Furthermore, uncertainty has connotations of the underlying structure of the economy changing over time, and this is not contemplated by parameter risk.

new virus kills half the population. The annual medical usage might drop, if a cure is found for the claimant's ailment, which was previously thought to be permanent. Or the cost of medical care might soar as a new drug is discovered that greatly improves the claimant's quality of life, at twice the cost. What if the government takes over the entire health-care system, and insurers are no longer responsible for medical costs?

We can dream up many different situations that will change what insurers owe to claimants. We can put probabilities on none of these, and we also know that there are many possibilities that we may not even think of, until they actually happen.

In commutations, it is common to ignore this uncertainty, and to commute some of the very high layers without payment. This is unwarranted. Commuting reinsurance is really a matter of pricing future possibilities, and reinsurers do not give away free layers, even if they have only a remote chance of being hit. For example, suppose I want to buy workers compensation reinsurance for a layer of $1 million excess of $800 million. (To avoid catastrophe issues, let us assume the reinsurance is per claim, not per occurrence.) There has never been a workers compensation claim that large, or even remotely close to it. Yet, would a reinsurer be willing to give the layer away free, even assuming they have no costs to service the contract? Of course they won't. Reinsurers recognize the remote possibility of having to pay on this contract, and they need to charge for that risk. The risk is remote, but remote does not mean non-existent. The chance of the layer being hit is not measurable, but not measurable does not mean zero.

8. THE DILEMMA OF THE "AUSTRIAN" ACTUARY

The dilemma of an actuary who recognizes ubiquitous uncertainty described by the Austrian economists is illustrated by a supposed comment of Lord Kelvin that "If you cannot measure, your knowledge is meager and unsatisfactory."[21]

[21] Coddington [4, p. 160] notes that there is no record of Kelvin ever having said exactly this, but it is inscribed on the facade of the Social Science Research Building at the University of Chicago.

As actuaries, we are paid to advise people on the numbers. In the case of a commutation, we are paid to decide whether a particular commutation offer is reasonable. If we are presented with a commutation offer, we can recommend that it be accepted or rejected. But saying "I don't know because the future is uncertain and I can't measure that" won't help. The dilemma of the "Austrian" actuary is that he recognizes that his knowledge *is* "meager and unsatisfactory," but he has to make a recommendation nevertheless.

One way of handling the dilemma is to take the advice of Frank Knight, who commented that the meaning for social scientists of Kelvin's remark is that "If you cannot measure, measure anyhow."[22] But "measuring anyhow" just leads to ignoring things that cannot be measured. If you have no reason to believe that these unmeasurables will bias your answers one way or another, that doesn't matter. But in many cases, especially when dealing with excess reinsurance, the unmeasurables will frequently bias the answers.

We must recognize that we will have to judgmentally adjust our answers for the unmeasurables. Judgmental adjustments are often uncomfortable, because they are hard to justify when attacked by others. But we have no choice other than to make our best judgments and explain the uncertainty of what we are doing.

9. POSSIBLE WAYS TO "MEASURE" THE UNMEASURABLE

When making judgmental adjustments, we are not completely without guidance. For a workers compensation commutation, here are some ways to check one's judgments:

Check 1: How much difference does the uncertainty make?

The first issue is to check the level of uncertainty, and the effects it can have. In the Joe Soap example discussed at length

[22]Quoted in Coddington [4, p. 160].

above, the different reinsurance layers have very different levels of uncertainty. One would expect that the lowest two or three layers will be breached fairly quickly, if the claimant survives. Even fairly dramatic changes in inflation and mortality rates will have relatively little impact on the numbers. The real impact of uncertainty is on the upper layers, where decades of compounded inflation, investment yields, changes in medical practice, and the claimant's condition come together to make the results of the calculations very fuzzy.

In the lower layers, Method 2 gives reasonable results. For medium layers, Method 3, unadjusted, may be reasonable. For higher layers, Method 3 results may need to be judgmentally increased, with the higher the layer, the higher the increase.

Check 2: What would it take to breach the layer?

For high layers, one can ask what it would take to breach the layer. For example, if it would take sustained medical inflation of 25% per annum to breach the layer, one would probably feel that this possibility is remote. But if it would take medical inflation of 10% per annum, which is considerably more likely, it should get a bigger charge. One can do similar reasonability checks for other parameters.

Check 3: What does the market charge?

We can get useful information from finding out what the market charges. To get useful information from market prices, we do not need to assume that the market price is exactly at its equilibrium level. The market price, as some consensus of supply and demand, provides a reality check.

There are, of course, no large, liquid markets for workers compensation commutations, but that doesn't mean there is no available information. A commutation is nothing more than reinsurance pricing, albeit for accidents that have already happened a number of years ago. It is quite reasonable to look at the reinsurance market for help.

For example, we generally find that the higher the layer being covered, the higher the risk load for the layer. [This higher risk load might be expressed in different ways—for example, a lower discounted loss ratio, or a "capacity charge" for layers that are seen to have a remote chance of being breached—but, in essence, these are all just risk loads.] With a commutation, we can look at the market structure of risk loads by layer, and use those to develop commutation risk loads for corresponding layers.

10. OTHER LINES OF BUSINESS; PRICING AND RESERVING, TOO

The issues discussed in this paper apply more broadly than just to workers compensation commutations. A commutation calculation for a general liability treaty would usually develop the expected losses to ultimate, and commute based on the discounted value of those losses. But this ignores risks that are transferred back to the ceding company in the commutation. For example, a general liability treaty being commuted in 1978 would have relieved the reinsurer for liability for environmental claims that were generated by the Superfund law, which passed a couple of years later. It was unknown, at the time of the commutation, that the cedent was giving up coverage for this risk, but it was not unknown that the cedent was taking the risk of some such change in the future. Just as a company selling general liability reinsurance will not give away remote layers free of charge, so the commutation should not be free for these layers either.

And it is not just commutations that are affected by determinism. It applies to regular pricing and reserving work as well. The clearest example would be the reserving of workers compensation reinsurance, where the methods used in this paper can be directly applied. But for pricing and reserving of any excess insurance or reinsurance, it is important to keep in mind the problems of determinism. If we simply assume the future will turn out to be what was expected, or that the future will follow the patterns of the past, we are bound to be led astray. The scary part of writing insurance is the uncertainty of what the future will bring. The uncertainty cannot be quantified, but we must not stick our heads in the sand and assume that if something cannot be quantified, it doesn't exist.

REFERENCES

[1] Blumsohn, Gary, "Levels of Determinism in Workers' Compensation Reinsurance Commutations," Casualty Actuarial Society *Forum*, Spring 1997, pp. 53–114.

[2] Bowers, Newton L., et al., *Actuarial Mathematics*, Society of Actuaries, 1986.

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

[4] Coddington, Alan, "Creaking Semaphore and Beyond: A Consideration of Shackle's 'Epistemics and Economics'," *British Journal of the Philosophy of Science* 26, 1975, pp. 151–163.

[5] Connor, Vincent, and Richard Olsen, "Commutation Pricing in the Post Tax-Reform Era," *PCAS* LXXVIII, 1991, pp. 81–109.

[6] Daykin, C. D., T. Pentikäinen, and M. Pesonen, *Practical Risk Theory for Actuaries*, Chapman and Hall, 1994.

[7] Feldblum, Sholom, "Asset Liability Matching for Property/Casualty Insurers," "Valuation Issues" Casualty Actuarial Society Special Interest Seminar, 1989, pp. 117–154.

[8] Ferguson, Ronald E., "Actuarial Note on Workmen's Compensation Loss Reserves," *PCAS* LVIII, 1971, pp. 51–57.

[9] Gillam, William R., "Injured Worker Mortality," *PCAS* LXXX, 1993, pp. 34–54.

[10] Kirzner, Israel M., "Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach," *Journal of Economic Literature* XXXV, March 1997, pp. 60–85.

[11] Knight, Frank H., *Risk, Uncertainty, and Profit*, University of Chicago Press, 1921.

[12] Lachmann, Ludwig M., "From Mises to Shackle: An Essay," *Journal of Economic Literature* 14, 1976, pp. 54–62.

[13] O'Driscoll, Gerald P., and Mario J. Rizzo, *The Economics of Time and Ignorance*, Basil Blackwell, 1985.

[14] Shackle, G. L. S., *Epistemics and Economics*, Cambridge: Cambridge University Press, 1972.

[15] Snader, Richard H., "Reserving Long Term Medical Claims," *PCAS* LXXIV, 1987, pp. 322–353.

[16] Steeneck, Lee R., "Actuarial Note on Workmen's Compensation Loss Reserves—25 Years Later," Casualty Actuarial Society *Forum*, Summer 1996, pp. 245–271.

[17] Venter, Gary G., "Workers Compensation Excess Reinsurance—The Longest Tail?," NCCI *Issues Report*, 1995, pp. 18–20.

[18] Venter, Gary G., and William R. Gillam, "Simulating Serious Workers' Compensation Claims," Casualty Actuarial Society Discussion Paper Program, 1986, pp. 226–258.

EXHIBIT 1

1990 US LIFE TABLE (MALES)

Age	$l(x)$	Life Expectancy	Age	$l(x)$	Life Expectancy	Age	$l(x)$	Life Expectancy
0	100,000.0	71.8	37	94,585.0	37.8	74	54,249.0	9.9
1	98,969.0	71.6	38	94,316.0	36.9	75	51,519.0	9.4
2	98,894.0	70.6	39	94,038.0	36.0	76	48,704.0	8.9
3	98,840.0	69.7	40	93,753.0	35.1	77	45,816.0	8.4
4	98,799.0	68.7	41	93,460.0	34.2	78	42,867.0	7.9
5	98,765.0	67.7	42	93,157.0	33.3	79	39,872.0	7.5
6	98,735.0	66.8	43	92,840.0	32.4	80	36,848.0	7.1
7	98,707.0	65.8	44	92,505.0	31.6	81	33,811.0	6.7
8	98,680.0	64.8	45	92,147.0	30.7	82	30,782.0	6.3
9	98,657.0	63.8	46	91,764.0	29.8	83	27,782.0	5.9
10	98,638.0	62.8	47	91,352.0	28.9	84	24,834.0	5.5
11	98,623.0	61.8	48	90,908.0	28.1	85	21,962.0	5.2
12	98,608.0	60.8	49	90,429.0	27.2	86	19,216.8	4.9
13	98,586.0	59.9	50	89,912.0	26.4	87	16,607.4	4.5
14	98,547.0	58.9	51	89,352.0	25.5	88	14,157.7	4.2
15	98,485.0	57.9	52	88,745.0	24.7	89	11,889.0	3.9
16	98,397.0	57.0	53	88,084.0	23.9	90	9,819.5	3.7
17	98,285.0	56.0	54	87,363.0	23.1	91	7,962.6	3.4
18	98,154.0	55.1	55	86,576.0	22.3	92	6,326.9	3.2
19	98,011.0	54.2	56	85,719.0	21.5	93	4,915.0	2.9
20	97,863.0	53.3	57	84,788.0	20.7	94	3,723.5	2.7
21	97,710.0	52.3	58	83,777.0	20.0	95	2,743.0	2.5
22	97,551.0	51.4	59	82,678.0	19.2	96	1,958.3	2.3
23	97,388.0	50.5	60	81,485.0	18.5	97	1,349.7	2.1
24	97,221.0	49.6	61	80,194.0	17.8	98	894.0	1.9
25	97,052.0	48.7	62	78,803.0	17.1	99	566.2	1.8
26	96,881.0	47.8	63	77,314.0	16.4	100	340.6	1.6
27	96,707.0	46.9	64	75,729.0	15.8	101	193.2	1.5
28	96,530.0	45.9	65	74,051.0	15.1	102	102.4	1.3
29	96,348.0	45.0	66	72,280.0	14.5	103	50.1	1.2
30	96,159.0	44.1	67	70,414.0	13.8	104	22.3	1.1
31	95,962.0	43.2	68	68,445.0	13.2	105	8.9	1.0
32	95,758.0	42.3	69	66,364.0	12.6	106	3.1	0.9
33	95,545.0	41.4	70	64,164.0	12.0	107	0.9	0.8
34	95,322.0	40.5	71	61,847.0	11.5	108	0.2	0.7
35	95,089.0	39.6	72	59,419.0	10.9	109	0.0	0.5
36	94,843.0	38.7	73	56,885.0	10.4	110	0.0	

Source: Vital Statistics of the United States, 1990 [US Department of Health and Human Services, 1994].
Note that the published tables extend only to age 85; beyond 85, the numbers are extrapolations.

EXHIBIT 2

INFLATION: CONSUMER PRICE INDEX AND MEDICAL CONSUMER PRICE INDEX

	Index at December		Annual Inflation			Index at December		Annual Inflation	
Year	CPI	Medical CPI	CPI	Medical CPI	Year	CPI	Medical CPI	CPI	Medical CPI
1935	13.8	10.2			1967	33.9	28.9	3.0%	6.3%
1936	14.0	10.2	1.4%	0.0%	1968	35.5	30.7	4.7%	6.2%
1937	14.4	10.3	2.9%	1.0%	1969	37.7	32.6	6.2%	6.2%
1938	14.0	10.3	– 2.8%	0.0%	1970	39.8	35.0	5.6%	7.4%
1939	14.0	10.4	0.0%	1.0%	1971	41.1	36.6	3.3%	4.6%
1940	14.1	10.4	0.7%	0.0%	1972	42.5	37.8	3.4%	3.3%
1941	15.5	10.5	9.9%	1.0%	1973	46.2	39.8	8.7%	5.3%
1942	16.9	10.9	9.0%	3.8%	1974	51.9	44.8	12.3%	12.6%
1943	17.4	11.4	3.0%	4.6%	1975	55.5	49.2	6.9%	9.8%
1944	17.8	11.7	2.3%	2.6%	1976	58.2	54.1	4.9%	10.0%
1945	18.2	12.0	2.2%	2.6%	1977	62.1	58.9	6.7%	8.9%
1946	21.5	13.0	18.1%	8.3%	1978	67.7	64.1	9.0%	8.8%
1947	23.4	13.9	8.8%	6.9%	1979	76.7	70.6	13.3%	10.1%
1948	24.1	14.7	3.0%	5.8%	1980	86.3	77.6	12.5%	9.9%
1949	23.6	14.9	– 2.1%	1.4%	1981	94.0	87.3	8.9%	12.5%
1950	25.0	15.4	5.9%	3.4%	1982	97.6	96.9	3.8%	11.0%
1951	26.5	16.3	6.0%	5.8%	1983	101.3	103.1	3.8%	6.4%
1952	26.7	17.0	0.8%	4.3%	1984	105.3	109.4	3.9%	6.1%
1953	26.9	17.6	0.7%	3.5%	1985	109.3	116.8	3.8%	6.8%
1954	26.7	18.0	−0.7%	2.3%	1986	110.5	125.8	1.1%	7.7%
1955	26.8	18.6	0.4%	3.3%	1987	115.4	133.1	4.4%	5.8%
1956	27.6	19.2	3.0%	3.2%	1988	120.5	142.3	4.4%	6.9%
1957	28.4	20.1	2.9%	4.7%	1989	126.1	154.4	4.6%	8.5%
1958	28.9	21.0	1.8%	4.5%	1990	133.8	169.2	6.1%	9.6%
1959	29.4	21.8	1.7%	3.8%	1991	137.9	182.6	3.1%	7.9%
1960	29.8	22.5	1.4%	3.2%	1992	141.9	194.7	2.9%	6.6%
1961	30.0	23.2	0.7%	3.1%	1993	145.8	205.2	2.7%	5.4%
1962	30.4	23.7	1.3%	2.2%	1994	149.7	215.3	2.7%	4.9%
1963	30.9	24.3	1.6%	2.5%	1995	153.5	223.8	2.5%	3.9%
1964	31.2	24.8	1.0%	2.1%	1996	158.6	230.6	3.3%	3.0%
1965	31.8	25.5	1.9%	2.8%	1997	161.3	237.1	1.7%	2.8%
1966	32.9	27.2	3.5%	6.7%					
							Average	**4.11%**	**5.25%**

Source: US Department of Labor, Bureau of Labor Statistics.

EXHIBIT 3

PART 1—PAGE 1

COMPLETELY DETERMINISTIC COMMUTATION CALCULATION

Parameters:

(A)	Evaluation Date:	1/1/98
(B)	Age at evaluation date:	35
(C)	Annual indemnity payment	20,000
(D)	Annual medical payment: (at mid-1997 price levels)	70,000
(E)	Indemnity paid to date	70,000
(F)	Medical paid to date	300,000
(G)	Life expectancy:	39.6
(H)	Cost-of-Living Adjustment:	4.11%
(I)	Medical Inflation Rate:	5.25%
(J)	Annual Discount Rate:	5.36%

Year	(1) Cost of Living Adjustment	(2) Indemnity Payment	(3) Medical Inflation	(4) Medical Payment	(5) Total Payment (2) + (4)	(6) Cumulative Total Payment Cumulative of (5)
1997 and prior		70,000		300,000	370,000	370,000
1998	4.11%	20,822	5.25%	73,675	94,497	464,497
1999	4.11%	21,678	5.25%	77,543	99,221	563,718
2000	4.11%	22,569	5.25%	81,614	104,183	667,900
2001	4.11%	23,496	5.25%	85,899	109,395	777,295
2002	4.11%	24,462	5.25%	90,408	114,870	892,166
2003	4.11%	25,467	5.25%	95,155	120,622	1,012,788
2004	4.11%	26,514	5.25%	100,150	126,665	1,139,452
2005	4.11%	27,604	5.25%	105,408	133,012	1,272,465
2006	4.11%	28,738	5.25%	110,942	139,681	1,412,145
2007	4.11%	29,920	5.25%	116,767	146,686	1,558,831
2008	4.11%	31,149	5.25%	122,897	154,046	1,712,878
2009	4.11%	32,429	5.25%	129,349	161,778	1,874,656
2010	4.11%	33,762	5.25%	136,140	169,902	2,044,558
2011	4.11%	35,150	5.25%	143,287	178,437	2,222,995
2012	4.11%	36,595	5.25%	150,810	187,404	2,410,400
2013	4.11%	38,099	5.25%	158,727	196,826	2,607,226
2014	4.11%	39,664	5.25%	167,061	206,725	2,813,951
2015	4.11%	41,295	5.25%	175,831	217,126	3,031,077

EXHIBIT 3

PART 1—PAGE 2

Year	(1) Cost of Living Adjustment	(2) Indemnity Payment	(3) Medical Inflation	(4) Medical Payment	(5) Total Payment (2) + (4)	(6) Cumulative Total Payment Cumulative of (5)
2016	4.11%	42,992	5.25%	185,062	228,054	3,259,131
2017	4.11%	44,759	5.25%	194,778	239,537	3,498,668
2018	4.11%	46,598	5.25%	205,004	251,602	3,750,270
2019	4.11%	48,514	5.25%	215,767	264,280	4,014,551
2020	4.11%	50,508	5.25%	227,094	277,602	4,292,152
2021	4.11%	52,583	5.25%	239,017	291,600	4,583,753
2022	4.11%	54,745	5.25%	251,565	306,310	4,890,063
2023	4.11%	56,995	5.25%	264,772	321,767	5,211,830
2024	4.11%	59,337	5.25%	278,673	338,010	5,549,840
2025	4.11%	61,776	5.25%	293,303	355,079	5,904,919
2026	4.11%	64,315	5.25%	308,702	373,017	6,277,935
2027	4.11%	66,958	5.25%	324,909	391,867	6,669,802
2028	4.11%	69,710	5.25%	341,966	411,676	7,081,478
2029	4.11%	72,575	5.25%	359,920	432,495	7,513,973
2030	4.11%	75,558	5.25%	378,815	454,373	7,968,346
2031	4.11%	78,663	5.25%	398,703	477,367	8,445,713
2032	4.11%	81,897	5.25%	419,635	501,532	8,947,245
2033	4.11%	85,263	5.25%	441,666	526,928	9,474,173
2034	4.11%	88,767	5.25%	464,853	553,620	10,027,793
2035	4.11%	92,415	5.25%	489,258	581,673	10,609,466
2036	4.11%	96,213	5.25%	514,944	611,158	11,220,624
2037	4.11%	60,101	5.25%	325,187	385,288	11,605,912
Total		2,060,654		9,545,258		

Future payments = 11,605,912 – 370,000 = 11,235,912

EXHIBIT 3

PART 2—PAGE 1

Year	(6) Cumulative Total Payment Cumulative of (5)	(7)	(8)	(9)	(10)	(11)	(12)
		Incremental Payments By Layer					
		$500,000 xs $370,000	$500,000 xs $500,000	$1 million xs $1 million	$3 million xs $2 million	$5 million xs $5 million	$5 million xs $10 million
1997 and prior	370,000	0	0	0	0	0	0
1998	464,497	94,497	0	0	0	0	0
1999	563,718	35,503	63,718	0	0	0	0
2000	667,900	0	104,183	0	0	0	0
2001	777,295	0	109,395	0	0	0	0
2002	892,166	0	114,870	0	0	0	0
2003	1,012,788	0	107,834	12,788	0	0	0
2004	1,139,452	0	0	126,665	0	0	0
2005	1,272,465	0	0	133,012	0	0	0
2006	1,412,145	0	0	139,681	0	0	0
2007	1,558,831	0	0	146,686	0	0	0
2008	1,712,878	0	0	154,046	0	0	0
2009	1,874,656	0	0	161,778	0	0	0
2010	2,044,558	0	0	125,344	44,558	0	0
2011	2,222,995	0	0	0	178,437	0	0
2012	2,410,400	0	0	0	187,404	0	0
2013	2,607,226	0	0	0	196,826	0	0
2014	2,813,951	0	0	0	206,725	0	0
2015	3,031,077	0	0	0	217,126	0	0
2016	3,259,131	0	0	0	228,054	0	0

2017	3,498,668	0	0	0	239,537	0	0
2018	3,750,270	0	0	0	251,602	0	0
2019	4,014,551	0	0	0	264,280	0	0
2020	4,292,152	0	0	0	277,602	0	0
2021	4,583,753	0	0	0	291,600	0	0
2022	4,890,063	0	0	0	306,310	0	0
2023	5,211,830	0	0	0	109,937	211,830	0
2024	5,549,840	0	0	0	0	338,010	0
2025	5,904,919	0	0	0	0	355,079	0
2026	6,277,935	0	0	0	0	373,017	0
2027	6,669,802	0	0	0	0	391,867	0
2028	7,081,478	0	0	0	0	411,676	0
2029	7,513,973	0	0	0	0	432,495	0
2030	7,968,346	0	0	0	0	454,373	0
2031	8,445,713	0	0	0	0	477,367	0
2032	8,947,245	0	0	0	0	501,532	0
2033	9,474,173	0	0	0	0	526,928	0
2034	10,027,793	0	0	0	0	525,827	27,793
2035	10,609,466	0	0	0	0	0	581,673
2036	11,220,624	0	0	0	0	0	611,158
2037	11,605,912	0	0	0	0	0	385,288
Total		130,000	500,000	1,000,000	3,000,000	5,000,000	1,605,912

EXHIBIT 3

PART 2—PAGE 2

	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
		Discounted Value by Layer						
Year	Present Value Factor	\$500,000 xs \$370,000	\$500,000 xs \$500,000	\$1 million xs \$1 million	\$3 million xs \$2 million	\$5 million xs \$5 million	\$5 million xs \$10 million	All Layers Combined
1997 and prior								
1998	0.9742	92,062	0	0	0	0	0	92,062
1999	0.9247	32,829	58,918	0	0	0	0	91,746
2000	0.8776	0	91,434	0	0	0	0	91,434
2001	0.8330	0	91,124	0	0	0	0	91,124
2002	0.7906	0	90,817	0	0	0	0	90,817
2003	0.7504	0	80,917	9,596	0	0	0	90,513
2004	0.7122	0	0	90,212	0	0	0	90,212
2005	0.6760	0	0	89,913	0	0	0	89,913
2006	0.6416	0	0	89,617	0	0	0	89,617
2007	0.6089	0	0	89,324	0	0	0	89,324
2008	0.5780	0	0	89,034	0	0	0	89,034
2009	0.5486	0	0	88,746	0	0	0	88,746
2010	0.5207	0	0	65,261	23,200	0	0	88,461
2011	0.4942	0	0	0	88,178	0	0	88,178
2012	0.4690	0	0	0	87,898	0	0	87,898
2013	0.4452	0	0	0	87,621	0	0	87,621
2014	0.4225	0	0	0	87,346	0	0	87,346
2015	0.4010	0	0	0	87,073	0	0	87,073
2016	0.3806	0	0	0	86,803	0	0	86,803
2017	0.3613	0	0	0	86,536	0	0	86,536

2018	0.3429	0	0	0	86,270	0	0	86,270
2019	0.3254	0	0	0	86,007	0	0	86,007
2020	0.3089	0	0	0	85,747	0	0	85,747
2021	0.2932	0	0	0	85,488	0	0	85,488
2022	0.2783	0	0	0	85,232	0	0	85,232
2023	0.2641	0	0	0	29,034	55,944	0	84,979
2024	0.2507	0	0	0	0	84,727	0	84,727
2025	0.2379	0	0	0	0	84,478	0	84,478
2026	0.2258	0	0	0	0	84,230	0	84,230
2027	0.2143	0	0	0	0	83,985	0	83,985
2028	0.2034	0	0	0	0	83,742	0	83,742
2029	0.1931	0	0	0	0	83,502	0	83,502
2030	0.1832	0	0	0	0	83,263	0	83,263
2031	0.1739	0	0	0	0	83,026	0	83,026
2032	0.1651	0	0	0	0	82,791	0	82,791
2033	0.1567	0	0	0	0	82,559	0	82,559
2034	0.1487	0	0	0	0	78,195	4,133	82,328
2035	0.1411	0	0	0	0	0	82,099	82,099
2036	0.1340	0	0	0	0	0	81,872	81,872
2037	0.1271	0	0	0	0	0	48,988	48,988
Total		124,890	413,209	611,704	1,092,436	970,442	217,093	3,429,774

EXHIBIT 4

PART 1—PAGE 1

METHOD 2: STOCHASTIC MORTALITY (OTHER INPUTS DETERMINISTIC)

Parameters:

(A)	Evaluation Date:	1/1/98
(B)	Current Age:	35
(C)	Annual Indemnity Payment	20,000
(D)	Annual Medical Payment (at mid-1997 price levels)	70,000
(E)	Indemnity Paid to Date	70,000
(F)	Medical Paid to Date:	300,000
(G)	Cost-of-Living Adjustment	4.11%
(H)	Medical Inflation Rate:	5.25%
(I)	Annual Discount Rate:	5.36%

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Year	Cost of Living Adjustment	Indemnity Payment	Medical Inflation	Medical Payment	Total Payment (2) + (4)	Cumulative Total Payment Cum. of (5)	Probability of claimant living to mid-year	Present Value Factor	Discount for mortality & investment income (7) × (8)
1997 and prior		70,000		300,000	370,000	370,000			
1998	4.11%	20,822	5.25%	73,675	94,497	464,497	0.999	0.974	0.973
1999	4.11%	21,678	5.25%	77,543	99,221	563,718	0.996	0.925	0.921
2000	4.11%	22,569	5.25%	81,614	104,183	667,900	0.993	0.878	0.872
2001	4.11%	23,496	5.25%	85,899	109,395	777,295	0.990	0.833	0.825
2002	4.11%	24,462	5.25%	90,408	114,870	892,166	0.987	0.791	0.781
2003	4.11%	25,467	5.25%	95,155	120,622	1,012,788	0.984	0.750	0.739
2004	4.11%	26,514	5.25%	100,150	126,665	1,139,452	0.981	0.712	0.699

2005	4.11%	27,604	5.25%	105,408	133,012	1,272,465	0.978	0.676	0.661
2006	4.11%	28,738	5.25%	110,942	139,681	1,412,145	0.975	0.642	0.625
2007	4.11%	29,920	5.25%	116,767	146,686	1,558,831	0.971	0.609	0.591
2008	4.11%	31,149	5.25%	122,897	154,046	1,712,878	0.967	0.578	0.559
2009	4.11%	32,429	5.25%	129,349	161,778	1,874,656	0.963	0.549	0.528
2010	4.11%	33,762	5.25%	136,140	169,902	2,044,558	0.958	0.521	0.499
2011	4.11%	35,150	5.25%	143,287	178,437	2,222,995	0.954	0.494	0.471
2012	4.11%	36,595	5.25%	150,810	187,404	2,410,400	0.948	0.469	0.445
2013	4.11%	38,099	5.25%	158,727	196,826	2,607,226	0.943	0.445	0.420
2014	4.11%	39,664	5.25%	167,061	206,725	2,813,951	0.936	0.423	0.396
2015	4.11%	41,295	5.25%	175,831	217,126	3,031,077	0.930	0.401	0.373
2016	4.11%	42,992	5.25%	185,062	228,054	3,259,131	0.923	0.381	0.351
2017	4.11%	44,759	5.25%	194,778	239,537	3,498,668	0.915	0.361	0.330
2018	4.11%	46,598	5.25%	205,004	251,602	3,750,270	0.906	0.343	0.311
2019	4.11%	48,514	5.25%	215,767	264,280	4,014,551	0.897	0.325	0.292
2020	4.11%	50,508	5.25%	227,094	277,602	4,292,152	0.886	0.309	0.274
2021	4.11%	52,583	5.25%	239,017	291,600	4,583,753	0.875	0.293	0.257
2022	4.11%	54,745	5.25%	251,565	306,310	4,890,063	0.863	0.278	0.240
2023	4.11%	56,995	5.25%	264,772	321,767	5,211,830	0.850	0.264	0.225
2024	4.11%	59,337	5.25%	278,673	338,010	5,549,840	0.836	0.251	0.210
2025	4.11%	61,776	5.25%	293,303	355,079	5,904,919	0.821	0.238	0.195
2026	4.11%	64,315	5.25%	308,702	373,017	6,277,935	0.805	0.226	0.182
2027	4.11%	66,958	5.25%	324,909	391,867	6,669,802	0.788	0.214	0.169
2028	4.11%	69,710	5.25%	341,966	411,676	7,081,478	0.769	0.203	0.157
2029	4.11%	72,575	5.25%	359,920	432,495	7,513,973	0.750	0.193	0.145
2030	4.11%	75,558	5.25%	378,815	454,373	7,968,346	0.730	0.183	0.134
2031	4.11%	78,663	5.25%	398,703	477,367	8,445,713	0.709	0.174	0.123
2032	4.11%	81,897	5.25%	419,635	501,532	8,947,245	0.686	0.165	0.113
2033	4.11%	85,263	5.25%	441,666	526,928	9,474,173	0.663	0.157	0.104
2034	4.11%	88,767	5.25%	464,853	553,620	10,027,793	0.638	0.149	0.095

EXHIBIT 4

PART 1—PAGE 2

METHOD 2: STOCHASTIC MORTALITY (OTHER INPUTS DETERMINISTIC)

Year	(1) Cost of Living Adjustment	(2) Indemnity Payment	(3) Medical Inflation	(4) Medical Payment	(5) Total Payment (2) + (4)	(6) Cumulative Total Payment Cum. of (5)	(7) Probability of claimant living to mid-year	(8) Present Value Factor	(9) Discount for mortality & investment income (7) × (8)
2035	4.11%	92,415	5.25%	489,258	581,673	10,609,466	0.612	0.141	0.086
2036	4.11%	96,213	5.25%	514,944	611,158	11,220,624	0.584	0.134	0.078
2037	4.11%	100,168	5.25%	541,979	642,146	11,862,770	0.556	0.127	0.072
2038	4.11%	104,285	5.25%	570,433	674,717	12,537,487	0.527	0.121	0.064
2039	4.11%	108,571	5.25%	600,380	708,951	13,246,438	0.497	0.115	0.057
2040	4.11%	113,033	5.25%	631,900	744,933	13,991,372	0.466	0.109	0.051
2041	4.11%	117,679	5.25%	665,075	782,754	14,774,125	0.435	0.103	0.045
2042	4.11%	122,515	5.25%	699,991	822,507	15,596,632	0.403	0.098	0.040
2043	4.11%	127,551	5.25%	736,741	864,292	16,460,924	0.372	0.093	0.035
2044	4.11%	132,793	5.25%	775,420	908,213	17,369,136	0.340	0.088	0.030
2045	4.11%	138,251	5.25%	816,129	954,380	18,323,517	0.308	0.084	0.026
2046	4.11%	143,933	5.25%	858,976	1,002,909	19,326,426	0.277	0.079	0.022
2047	4.11%	149,848	5.25%	904,073	1,053,921	20,380,347	0.246	0.075	0.019
2048	4.11%	156,007	5.25%	951,536	1,107,544	21,487,890	0.217	0.072	0.016
2049	4.11%	162,419	5.25%	1,001,492	1,163,911	22,651,801	0.188	0.068	0.013
2050	4.11%	169,095	5.25%	1,054,070	1,223,165	23,874,966	0.162	0.065	0.010
2051	4.11%	176,044	5.25%	1,109,409	1,285,453	25,160,420	0.137	0.061	0.008
2052	4.11%	183,280	5.25%	1,167,653	1,350,933	26,511,352	0.114	0.058	0.007

2053	4.11%	190,813	5.25%	1,228,955	1,419,767	27,931,120	0.094	0.055	0.005
2054	4.11%	198,655	5.25%	1,293,475	1,492,130	29,423,250	0.075	0.052	0.004
2055	4.11%	206,820	5.25%	1,361,382	1,568,202	30,991,452	0.059	0.050	0.003
2056	4.11%	215,320	5.25%	1,432,855	1,648,175	32,639,626	0.045	0.047	0.002
2057	4.11%	224,170	5.25%	1,508,080	1,732,249	34,371,876	0.034	0.045	0.002
2058	4.11%	233,383	5.25%	1,587,254	1,820,637	36,192,513	0.025	0.042	0.001
2059	4.11%	242,975	5.25%	1,670,585	1,913,560	38,106,073	0.017	0.040	0.001
2060	4.11%	252,961	5.25%	1,758,290	2,011,252	40,117,324	0.012	0.038	0.001
2061	4.11%	263,358	5.25%	1,850,601	2,113,959	42,231,283	0.008	0.036	0.000
2062	4.11%	274,182	5.25%	1,947,757	2,221,939	44,453,222	0.005	0.035	0.000
2063	4.11%	285,451	5.25%	2,050,015	2,335,465	46,788,688	0.003	0.033	0.000
2064	4.11%	297,183	5.25%	2,157,640	2,454,823	49,243,511	0.002	0.031	0.000
2065	4.11%	309,397	5.25%	2,270,916	2,580,314	51,823,825	0.001	0.029	0.000
2066	4.11%	322,113	5.25%	2,390,140	2,712,253	54,536,078	0.0004	0.028	0.000
2067	4.11%	335,352	5.25%	2,515,622	2,850,974	57,387,052	0.0002	0.027	0.000
2068	4.11%	349,135	5.25%	2,647,692	2,996,827	60,383,879	0.0001	0.025	0.000
2069	4.11%	363,485	5.25%	2,786,696	3,150,181	63,534,059	0.00002	0.024	0.000
2070	4.11%	378,424	5.25%	2,932,997	3,311,421	66,845,481	0.00001	0.023	0.000
2071	4.11%	393,977	5.25%	3,086,980	3,480,957	70,326,438	0.000001	0.022	0.000
2072	4.11%	410,170	5.25%	3,249,046	3,659,216	73,985,653	0.0000002	0.020	0.000

EXHIBIT 4

PART 2—PAGE 1

Year	(10) $130,000 xs $370,000	(11) $500,000 xs $500,000	(12) $1 million xs $1 million	(13) $3 million xs $2 million	(14) $5 million xs $5 million	(15) $5 million xs $10 million
	Incremental Payments by Layer					
1997 and prior						
1998	94,497	0	0	0	0	0
1999	35,503	63,718	0	0	0	0
2000	0	104,183	0	0	0	0
2001	0	109,395	0	0	0	0
2002	0	114,870	0	0	0	0
2003	0	107,834	12,788	0	0	0
2004	0	0	126,665	0	0	0
2005	0	0	133,012	0	0	0
2006	0	0	139,681	0	0	0
2007	0	0	146,686	0	0	0
2008	0	0	154,046	0	0	0
2009	0	0	161,778	0	0	0
2010	0	0	125,344	44,558	0	0
2011	0	0	0	178,437	0	0
2012	0	0	0	187,404	0	0
2013	0	0	0	196,826	0	0
2014	0	0	0	206,725	0	0
2015	0	0	0	217,126	0	0
2016	0	0	0	228,054	0	0
2017	0	0	0	239,537	0	0
2018	0	0	0	251,602	0	0
2019	0	0	0	264,280	0	0
2020	0	0	0	277,602	0	0
2021	0	0	0	291,600	0	0
2022	0	0	0	306,310	0	0
2023	0	0	0	109,937	211,830	0
2024	0	0	0	0	338,010	0
2025	0	0	0	0	355,079	0
2026	0	0	0	0	373,017	0
2027	0	0	0	0	391,867	0
2028	0	0	0	0	411,676	0
2029	0	0	0	0	432,495	0
2030	0	0	0	0	454,373	0
2031	0	0	0	0	477,367	0
2032	0	0	0	0	501,532	0
2033	0	0	0	0	526,928	0
2034	0	0	0	0	525,827	27,793
2035	0	0	0	0	0	581,673
2036	0	0	0	0	0	611,158
2037	0	0	0	0	0	642,146
2038	0	0	0	0	0	674,717
2039	0	0	0	0	0	708,951
2040	0	0	0	0	0	744,933
2041	0	0	0	0	0	782,754
2042	0	0	0	0	0	225,875
2043	0	0	0	0	0	0
2044	0	0	0	0	0	0

(16)	(17)	(18)	(19)	(20)	(21)	(22)
Incremental Payments by Layer						
$5 million xs $15 million	$10 million xs $20 million	$10 million xs $30 million	$10 million xs $40 million	$10 million xs $50 million	$10 million xs $60 million	$10 million xs $70 million
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
596,632	0	0	0	0	0	0
864,292	0	0	0	0	0	0
908,213	0	0	0	0	0	0

EXHIBIT 4

PART 2—PAGE 2

	(10)	(11)	(12)	(13)	(14)	(15)
	Incremental Payments by Layer					
Year	$130,000 xs $370,000	$500,000 xs $500,000	$1 million xs $1 million	$3 million xs $2 million	$5 million xs $5 million	$5 million xs $10 million
2045	0	0	0	0	0	0
2046	0	0	0	0	0	0
2047	0	0	0	0	0	0
2048	0	0	0	0	0	0
2049	0	0	0	0	0	0
2050	0	0	0	0	0	0
2051	0	0	0	0	0	0
2052	0	0	0	0	0	0
2053	0	0	0	0	0	0
2054	0	0	0	0	0	0
2055	0	0	0	0	0	0
2056	0	0	0	0	0	0
2057	0	0	0	0	0	0
2058	0	0	0	0	0	0
2059	0	0	0	0	0	0
2060	0	0	0	0	0	0
2061	0	0	0	0	0	0
2062	0	0	0	0	0	0
2063	0	0	0	0	0	0
2064	0	0	0	0	0	0
2065	0	0	0	0	0	0
2066	0	0	0	0	0	0
2067	0	0	0	0	0	0
2068	0	0	0	0	0	0
2069	0	0	0	0	0	0
2070	0	0	0	0	0	0
2071	0	0	0	0	0	0
2072	0	0	0	0	0	0
	130,000	500,000	1,000,000	3,000,000	5,000,000	5,000,000

(16)	(17)	(18)	(19)	(20)	(21)	(22)
Incremental Payments by Layer						
$5 million xs $15 million	$10 million xs $20 million	$10 million xs $30 million	$10 million xs $40 million	$10 million xs $50 million	$10 million xs $60 million	$10 million xs $70 million
954,380	0	0	0	0	0	0
1,002,909	0	0	0	0	0	0
673,574	380,347	0	0	0	0	0
0	1,107,544	0	0	0	0	0
0	1,163,911	0	0	0	0	0
0	1,223,165	0	0	0	0	0
0	1,285,453	0	0	0	0	0
0	1,350,933	0	0	0	0	0
0	1,419,767	0	0	0	0	0
0	1,492,130	0	0	0	0	0
0	576,750	991,452	0	0	0	0
0	0	1,648,175	0	0	0	0
0	0	1,732,249	0	0	0	0
0	0	1,820,637	0	0	0	0
0	0	1,913,560	0	0	0	0
0	0	1,893,927	117,324	0	0	0
0	0	0	2,113,959	0	0	0
0	0	0	2,221,939	0	0	0
0	0	0	2,335,465	0	0	0
0	0	0	2,454,823	0	0	0
0	0	0	756,489	1,823,825	0	0
0	0	0	0	2,712,253	0	0
0	0	0	0	2,850,974	0	0
0	0	0	0	2,612,948	383,879	0
0	0	0	0	0	3,150,181	0
0	0	0	0	0	3,311,421	0
0	0	0	0	0	3,154,519	326,438
0	0	0	0	0	0	3,659,216
5,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	3,985,653

EXHIBIT 4

PART 2—PAGE 3

	(23)	(24)	(25)	(26)	(27)	(28)
	Commutation Value by Layer, Discounted for Both Mortality and Investment Income. Columns are derived by multiplying the corresponding column from Exhibit 4, pages 3 and 4, by Column 9, from pages 1 and 2. For example, Column 23 = Column 10 × Column 9					
Year	$500,000 xs $0	$500,000 xs $500,000	$1 million xs $1 million	$3 million xs $2 million	$5 million xs $5 million	$5 million xs $10 million
1997 and prior						
1998	91,943	0	0	0	0	0
1999	32,699	58,685	0	0	0	0
2000	0	90,820	0	0	0	0
2001	0	90,250	0	0	0	0
2002	0	89,677	0	0	0	0
2003	0	79,655	9,446	0	0	0
2004	0	0	88,522	0	0	0
2005	0	0	87,936	0	0	0
2006	0	0	87,340	0	0	0
2007	0	0	86,729	0	0	0
2008	0	0	86,100	0	0	0
2009	0	0	85,451	0	0	0
2010	0	0	62,544	22,234	0	0
2011	0	0	0	84,079	0	0
2012	0	0	0	83,352	0	0
2013	0	0	0	82,593	0	0
2014	0	0	0	81,797	0	0
2015	0	0	0	80,962	0	0
2016	0	0	0	80,080	0	0
2017	0	0	0	79,147	0	0
2018	0	0	0	78,158	0	0
2019	0	0	0	77,111	0	0
2020	0	0	0	76,002	0	0
2021	0	0	0	74,825	0	0
2022	0	0	0	73,573	0	0
2023	0	0	0	24,684	47,561	0
2024	0	0	0	0	70,835	0
2025	0	0	0	0	69,348	0
2026	0	0	0	0	67,783	0
2027	0	0	0	0	66,145	0
2028	0	0	0	0	64,435	0
2029	0	0	0	0	62,653	0
2030	0	0	0	0	60,795	0
2031	0	0	0	0	58,854	0
2032	0	0	0	0	56,823	0
2033	0	0	0	0	54,703	0
2034	0	0	0	0	49,860	2,635
2035	0	0	0	0	0	50,208
2036	0	0	0	0	0	47,844
2037	0	0	0	0	0	45,408
2038	0	0	0	0	0	42,910
2039	0	0	0	0	0	40,359
2040	0	0	0	0	0	37,764
2041	0	0	0	0	0	35,138
2042	0	0	0	0	0	8,924

(29)	(30)	(31)	(32)	(33)	(34)	(35)
$5 million xs $15 million	$10 million xs $20 million	$10 million xs $30 million	$10 million xs $40 million	$10 million xs $50 million	$10 million xs $60 million	$10 million xs $70 million
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
23,571	0	0	0	0	0	0

EXHIBIT 4

PART 2—PAGE 4

	(23)	(24)	(25)	(26)	(27)	(28)
	Commutation Value by Layer, Discounted for Both Mortality and Investment Income. Columns are derived by multiplying the corresponding column from Exhibit 4, pages 3 and 4, by Column 9, from pages 1 and 2. For example, Column 23 = Column 10 × Column 9					
Year	\$500,000 xs \$0	\$500,000 xs \$500,000	\$1 million xs \$1 million	\$3 million xs \$2 million	\$5 million xs \$5 million	\$5 million xs \$10 million
2043	0	0	0	0	0	0
2044	0	0	0	0	0	0
2045	0	0	0	0	0	0
2046	0	0	0	0	0	0
2047	0	0	0	0	0	0
2048	0	0	0	0	0	0
2049	0	0	0	0	0	0
2050	0	0	0	0	0	0
2051	0	0	0	0	0	0
2052	0	0	0	0	0	0
2053	0	0	0	0	0	0
2054	0	0	0	0	0	0
2055	0	0	0	0	0	0
2056	0	0	0	0	0	0
2057	0	0	0	0	0	0
2058	0	0	0	0	0	0
2059	0	0	0	0	0	0
2060	0	0	0	0	0	0
2061	0	0	0	0	0	0
2062	0	0	0	0	0	0
2063	0	0	0	0	0	0
2064	0	0	0	0	0	0
2065	0	0	0	0	0	0
2066	0	0	0	0	0	0
2067	0	0	0	0	0	0
2068	0	0	0	0	0	0
2069	0	0	0	0	0	0
2070	0	0	0	0	0	0
2071	0	0	0	0	0	0
2072	0	0	0	0	0	0
	124,642	409,088	594,069	998,595	729,794	311,190
	Overall Total = 3,408,316					

(29)	(30)	(31)	(32)	(33)	(34)	(35)
$5 million xs $15 million	$10 million xs $20 million	$10 million xs $30 million	$10 million xs $40 million	$10 million xs $50 million	$10 million xs $60 million	$10 million xs $70 million
29,848	0	0	0	0	0	0
27,214	0	0	0	0	0	0
24,609	0	0	0	0	0	0
22,052	0	0	0	0	0	0
12,502	7,060	0	0	0	0	0
0	17,169	0	0	0	0	0
0	14,898	0	0	0	0	0
0	12,762	0	0	0	0	0
0	10,777	0	0	0	0	0
0	8,959	0	0	0	0	0
0	7,320	0	0	0	0	0
0	5,868	0	0	0	0	0
0	1,694	2,911	0	0	0	0
0	0	3,530	0	0	0	0
0	0	2,636	0	0	0	0
0	0	1,912	0	0	0	0
0	0	1,342	0	0	0	0
0	0	855	53	0	0	0
0	0	0	589	0	0	0
0	0	0	365	0	0	0
0	0	0	214	0	0	0
0	0	0	118	0	0	0
0	0	0	18	43	0	0
0	0	0	0	29	0	0
0	0	0	0	12	0	0
0	0	0	0	4	1	0
0	0	0	0	0	2	0.00
0	0	0	0	0	0	0.00
0	0	0	0	0	0	0.01
0	0	0	0	0	0	0.01
139,796	86,507	13,185	1,358	88	3	0.02

EXHIBIT 5

FITTING OF AUTO-REGRESSIVE MODEL FOR CPI

Model: Inflation rate = average inflation + α(last year's inflation – average inflation) + error term where error term is represented by a shifted lognormal

$$\alpha = 0.511$$

α is chosen to minimize the sum of the squared errors in Col. 5

Year	(1) CPI at December	(2) Annual % Increase in CPI	(3) Least-Squares Fit of Inflation Model*	(4) Error**	(5) Squared Error***	(6) Error + 0.07	(7) log(error + 0.07)
1935	13.8						
1936	14.0	1.4%					
1937	14.4	2.9%	2.8%	0.00105	0.00000	0.07105	(2.64431)
1938	14.0	−2.8%	3.5%	(0.06249)	0.00390	0.00751	(4.89101)
1939	14.0	0.0%	0.6%	(0.00593)	0.00004	0.06407	(2.74771)
1940	14.1	0.7%	2.0%	(0.01297)	0.00017	0.05703	(2.86421)
1941	15.5	9.9%	2.4%	0.07553	0.00570	0.14553	(1.92739)
1942	16.9	9.0%	7.1%	0.01949	0.00038	0.08949	(2.41361)
1943	17.4	3.0%	6.6%	(0.03666)	0.00134	0.03334	(3.40113)
1944	17.8	2.3%	3.5%	(0.01224)	0.00015	0.05776	(2.85142)
1945	18.2	2.2%	3.2%	(0.00938)	0.00009	0.06062	(2.80321)
1946	21.5	18.1%	3.2%	0.14973	0.02242	0.21973	(1.51537)
1947	23.4	8.8%	11.3%	(0.02436)	0.00059	0.04564	(3.08692)
1948	24.1	3.0%	6.5%	(0.03534)	0.00125	0.03466	(3.36215)
1949	23.6	−2.1%	3.5%	(0.05614)	0.00315	0.01386	(4.27884)
1950	25.0	5.9%	1.0%	0.04980	0.00248	0.11980	(2.12189)
1951	26.5	6.0%	5.0%	0.00958	0.00009	0.07958	(2.53094)
1952	26.7	0.8%	5.1%	(0.04321)	0.00187	0.02679	(3.61990)

1953	26.9	0.7%	2.4%	(0.01648)	0.00027	0.05352	(2.92768)
1954	26.7	−0.7%	2.4%	(0.03138)	0.00098	0.03862	(3.25387)
1955	26.8	0.4%	1.6%	(0.01257)	0.00016	0.05743	(2.85721)
1956	27.6	3.0%	2.2%	0.00782	0.00006	0.07782	(2.55331)
1957	28.4	2.9%	3.5%	(0.00638)	0.00004	0.06362	(2.75477)
1958	28.9	1.8%	3.5%	(0.01731)	0.00030	0.05269	(2.94341)
1959	29.4	1.7%	2.9%	(0.01181)	0.00014	0.05819	(2.84398)
1960	29.8	1.4%	2.9%	(0.01535)	0.00024	0.05465	(2.90674)
1961	30.0	0.7%	2.7%	(0.02035)	0.00041	0.04965	(3.00281)
1962	30.4	1.3%	2.4%	(0.01021)	0.00010	0.05979	(2.81690)
1963	30.9	1.6%	2.7%	(0.01048)	0.00011	0.05952	(2.82140)
1964	31.2	1.0%	2.9%	(0.01881)	0.00035	0.05119	(2.97215)
1965	31.8	1.9%	2.5%	(0.00584)	0.00003	0.06416	(2.74642)
1966	32.9	3.5%	3.0%	0.00465	0.00002	0.07465	(2.59489)
1967	33.9	3.0%	3.8%	(0.00739)	0.00005	0.06261	(2.77080)
1968	35.5	4.7%	3.6%	0.01156	0.00013	0.08156	(2.50644)
1969	37.7	6.2%	4.4%	0.01775	0.00032	0.08775	(2.43327)
1970	39.8	5.6%	5.2%	0.00393	0.00002	0.07393	(2.60458)
1971	41.1	3.3%	4.9%	(0.01590)	0.00025	0.05410	(2.91699)
1972	42.5	3.4%	3.7%	(0.00274)	0.00001	0.06726	(2.69912)
1973	46.2	8.7%	3.8%	0.04955	0.00245	0.11955	(2.12406)
1974	51.9	12.3%	6.5%	0.05879	0.00346	0.12879	(2.04954)
1975	55.5	6.9%	8.3%	(0.01377)	0.00019	0.05623	(2.87830)
1976	58.2	4.9%	5.6%	(0.00690)	0.00005	0.06310	(2.76297)
1977	62.1	6.7%	4.5%	0.02205	0.00049	0.09205	(2.38546)
1978	67.7	9.0%	5.4%	0.03583	0.00128	0.10583	(2.24588)
1979	76.7	13.3%	6.6%	0.06676	0.00446	0.13676	(1.98950)
1980	86.3	12.5%	8.8%	0.03714	0.00138	0.10714	(2.23358)
1981	94.0	8.9%	8.4%	0.00518	0.00003	0.07518	(2.58791)
1982	97.6	3.8%	6.6%	(0.02739)	0.00075	0.04261	(3.15569)
1983	101.3	3.8%	4.0%	(0.00177)	0.00000	0.06823	(2.68482)
1984	105.3	3.9%	3.9%	0.00001	0.00000	0.07001	(2.65914)
1985	109.3	3.8%	4.0%	(0.00230)	0.00001	0.06770	(2.69263)
1986	110.5	1.1%	4.0%	(0.02854)	0.00081	0.04146	(3.18299)

EXHIBIT 5

(Continued)

Year	(1) CPI at December	(2) Annual % Increase in CPI	(3) Least-Squares Fit of Inflation Model*	(4) Error**	(5) Squared Error***	(6) Error + 0.07	(7) log(error + 0.07)
1987	115.4	4.4%	2.6%	0.01862	0.00035	0.08862	(2.42338)
1988	120.5	4.4%	4.3%	0.00143	0.00000	0.07143	(2.63905)
1989	126.1	4.6%	4.3%	0.00378	0.00001	0.07378	(2.60660)
1990	133.8	6.1%	4.4%	0.01721	0.00030	0.08721	(2.43943)
1991	137.9	3.1%	5.1%	(0.02066)	0.00043	0.04934	(3.00906)
1992	141.9	2.9%	3.6%	(0.00676)	0.00005	0.06324	(2.76082)
1993	145.8	2.7%	3.5%	(0.00745)	0.00006	0.06255	(2.77173)
1994	149.7	2.7%	3.4%	(0.00740)	0.00005	0.06260	(2.77105)
1995	153.5	2.5%	3.4%	(0.00839)	0.00007	0.06161	(2.78699)
1996	158.6	3.3%	3.3%	0.00014	0.00000	0.07014	(2.65720)
1997	161.3	1.7%	3.7%	(0.02006)	0.00040	0.04994	(2.99696)
Average		4.11%		0.00023	0.00106	0.07023	(2.76199)
Std. Dev.				0.03284		0.03284	0.50068

*Column 3 is calculated as: [Avg. of Col. 2] + α[Value of Col. 3 for previous yr − Avg. of Col. 2].

**Column 4 is calculated as: {Col. 2 − Col. 3}.

***Column 5 is calculated as: $\{\text{Col. } 4\}^2$.

Shifted lognormal to model the error term is calculated by fitting a lognormal to Col. 6, the error term, plus a shift of 0.07, which ensures that all the error terms are positive. The lognormal is fitted using the method of moments to the underlying normal distribution (rather than directly to the lognormal), yielding:

$$\mu = -2.7620$$

$$\sigma = 0.5007$$

EXHIBIT 6

FITTING OF MODEL FOR MEDICAL INFLATION

Model: Medical inflation$_t$ = inflation$_t$ + β(Medical inflation$_{t-1}$ − Inflation$_{t-1}$) + (Average medical inflation − average inflation) + error$_t$

$$\beta = 0.380$$

β is chosen to minimize the sum of the squared errors in column 6

Year	(1) Medical CPI at December	(2) Annual % Increase in Medical CPI	(3) Annual % Increase in Overall CPI	(4) Least-Squares Fit of Medical Inflation Model*	(5) Error**	(6) Squared Error***
1935	10.2					
1936	10.2	0.0%	1.4%			
1937	10.3	1.0%	2.9%	3.4%	−2.46%	0.00061
1938	10.3	0.0%	−2.8%	−2.4%	2.35%	0.00055
1939	10.4	1.0%	0.0%	2.2%	−1.22%	0.00015
1940	10.4	0.0%	0.7%	2.2%	−2.22%	0.00049
1941	10.5	1.0%	9.9%	10.8%	−9.83%	0.00967
1942	10.9	3.8%	9.0%	6.8%	−2.95%	0.00087
1943	11.4	4.6%	3.0%	2.1%	2.48%	0.00061
1944	11.7	2.6%	2.3%	4.1%	−1.42%	0.00020
1945	12.0	2.6%	2.2%	3.5%	−0.95%	0.00009
1946	13.0	8.3%	18.1%	19.4%	−11.06%	0.01223
1947	13.9	6.9%	8.8%	6.2%	0.68%	0.00005
1948	14.7	5.8%	3.0%	3.4%	2.35%	0.00055
1949	14.9	1.4%	−2.1%	0.1%	1.25%	0.00016
1950	15.4	3.4%	5.9%	8.4%	−5.02%	0.00252
1951	16.3	5.8%	6.0%	6.2%	−0.31%	0.00001
1952	17.0	4.3%	0.8%	1.8%	2.46%	0.00061

EXHIBIT 6

(Continued)

Year	(1) Medical CPI at December	(2) Annual % Increase in Medical CPI	(3) Annual % Increase in Overall CPI	(4) Least-Squares Fit of Medical Inflation Model*	(5) Error**	(6) Squared Error***
1953	17.6	3.5%	0.7%	3.2%	0.30%	0.00001
1954	18.0	2.3%	−0.7%	1.5%	0.82%	0.00007
1955	18.6	3.3%	0.4%	2.7%	0.67%	0.00005
1956	19.2	3.2%	3.0%	5.2%	−2.02%	0.00041
1957	20.1	4.7%	2.9%	4.1%	0.56%	0.00003
1958	21.0	4.5%	1.8%	3.6%	0.90%	0.00008
1959	21.8	3.8%	1.7%	3.9%	−0.09%	0.00000
1960	22.5	3.2%	1.4%	3.3%	−0.08%	0.00000
1961	23.2	3.1%	0.7%	2.5%	0.60%	0.00004
1962	23.7	2.2%	1.3%	3.4%	−1.24%	0.00015
1963	24.3	2.5%	1.6%	3.1%	−0.56%	0.00003
1964	24.8	2.1%	1.0%	2.4%	−0.39%	0.00002
1965	25.5	2.8%	1.9%	3.5%	−0.65%	0.00004
1966	27.2	6.7%	3.5%	4.9%	1.73%	0.00030
1967	28.9	6.3%	3.0%	5.4%	0.85%	0.00007
1968	30.7	6.2%	4.7%	7.1%	−0.85%	0.00007
1969	32.6	6.2%	6.2%	7.9%	−1.72%	0.00030
1970	35.0	7.4%	5.6%	6.7%	0.66%	0.00004
1971	36.6	4.6%	3.3%	5.1%	−0.51%	0.00003
1972	37.8	3.3%	3.4%	5.0%	−1.76%	0.00031
1973	39.8	5.3%	8.7%	9.8%	−4.50%	0.00203
1974	44.8	12.6%	12.3%	12.2%	0.39%	0.00001

1975	49.2	9.8%	6.9%	8.2%	1.66%	0.00028
1976	54.1	10.0%	4.9%	7.1%	2.86%	0.00082
1977	58.9	8.9%	6.7%	9.8%	−0.91%	0.00008
1978	64.1	8.8%	9.0%	11.0%	−2.15%	0.00046
1979	70.6	10.1%	13.3%	14.4%	−4.22%	0.00178
1980	77.6	9.9%	12.5%	12.5%	−2.54%	0.00064
1981	87.3	12.5%	8.9%	9.1%	3.43%	0.00118
1982	96.9	11.0%	3.8%	6.3%	4.67%	0.00218
1983	103.1	6.4%	3.8%	7.7%	−1.26%	0.00016
1984	109.4	6.1%	3.9%	6.1%	0.03%	0.00000
1985	116.8	6.8%	3.8%	5.8%	1.00%	0.00010
1986	125.8	7.7%	1.1%	3.4%	4.34%	0.00188
1987	133.1	5.8%	4.4%	8.1%	−2.28%	0.00052
1988	142.3	6.9%	4.4%	6.1%	0.83%	0.00007
1989	154.4	8.5%	4.6%	6.7%	1.77%	0.00031
1990	169.2	9.6%	6.1%	8.7%	0.87%	0.00008
1991	182.6	7.9%	3.1%	5.5%	2.39%	0.00057
1992	194.7	6.6%	2.9%	5.9%	0.74%	0.00005
1993	205.2	5.4%	2.7%	5.3%	0.09%	0.00000
1994	215.3	4.9%	2.7%	4.8%	0.10%	0.00000
1995	223.8	3.9%	2.5%	4.5%	−0.58%	0.00003
1996	230.6	3.0%	3.3%	5.0%	−1.96%	0.00038
1997	237.1	2.8%	1.7%	2.7%	0.09%	0.00000
Average		5.25%	4.1%		−0.39%	0.00074
					2.71% = Std. Dev. of errors	0.04505 = Sum of square errors

Average difference between medical inflation and inflation (i.e., avg. of Col. 2 − avg. of Col. 3) = 1.14%.
*Column 4 is calculated as Col. 3 for previous year + β[Col. 2 for previous year − Col. 3 for previous year] + [Avg. of Col. 2 − Avg. of Col. 3].
**Column 5 = Column 2 − Column 4.
***Column 6 = $\{\text{Column } 5\}^2$.
β is fitted to minimize the sum of column 6.

EXHIBIT 7

PART 1—PAGE 1

ONE SIMULATION FROM METHOD 3
STOCHASTIC MORTALITY, INFLATION, MEDICAL INFLATION, AND INVESTMENT YIELDS

Parameters:

(A)	Evaluation Date:	1/1/98
(B)	Current Age:	35
(C)	Annual Indemnity Payment	20,000
(D)	Annual Medical Payment (at mid-1997 price levels)	Varies
(E)	Indemnity Paid to Date	70,000
(F)	Medical Paid to Date:	300,000
(G)	Cost-of-Living Adjustment	Varies
(H)	Medical Inflation Rate:	Varies
(I)	Annual Discount Rate:	Varies

Year	(1) Cost of Living Adjustment	(2) Indemnity Payment	(3) Medical Inflation	(4) Medical Payment	(5) Total Payment (2) + (4)	(6) Cumulative Total Payment Cum. of (5)	(7) Probability of claimant Living to Mid-year	(8) Present Value Factor	(9) Discount for Mortality & Investment Income (7) × (8)
1997 and prior		70,000		300,000	370,000	370,000			
1998	4.2%	20,843	4.22%	153,624	174,467	544,467	0.999	0.9744	0.9732
1999	4.1%	21,691	6.02%	46,557	68,248	612,716	0.996	0.9336	0.9299
2000	2.2%	22,167	2.74%	154,490	176,657	789,373	0.993	0.9168	0.9107

2001	0.0%	22,167	11.60%	26,456	48,624	837,996	0.990	0.8804	0.8719
2002	5.0%	23,275	9.33%	104,053	127,329	965,325	0.987	0.8175	0.8073
2003	5.0%	24,439	2.17%	28,936	53,375	1,018,700	0.984	0.7862	0.7739
2004	0.0%	24,439	−0.47%	95,726	120,166	1,138,866	0.981	0.7807	0.7661
2005	0.0%	24,439	−0.48%	275,599	300,038	1,438,904	0.978	0.7723	0.7554
2006	0.4%	24,540	5.79%	78,292	102,832	1,541,736	0.975	0.7660	0.7465
2007	0.0%	24,540	10.62%	135,876	160,416	1,702,151	0.971	0.7461	0.7244
2008	4.2%	25,559	7.01%	91,516	117,075	1,819,227	0.967	0.7179	0.6943
2009	1.2%	25,870	3.07%	153,420	179,289	1,998,516	0.963	0.6926	0.6669
2010	3.6%	26,805	6.35%	281,942	308,747	2,307,263	0.958	0.6681	0.6403
2011	1.2%	27,133	3.92%	47,016	74,150	2,381,413	0.954	0.6497	0.6195
2012	2.0%	27,663	4.97%	95,496	123,159	2,504,572	0.948	0.6292	0.5966
2013	2.1%	28,236	2.39%	86,667	114,903	2,619,476	0.943	0.6146	0.5794
2014	0.2%	28,284	1.54%	82,222	110,506	2,729,982	0.936	0.6012	0.5630
2015	1.8%	28,797	−2.26%	74,619	103,416	2,833,398	0.930	0.5810	0.5402
2016	2.6%	29,554	1.48%	130,737	160,291	2,993,689	0.923	0.5581	0.5149
2017	3.1%	30,462	3.73%	595,604	626,065	3,619,754	0.915	0.5365	0.4907
2018	2.5%	31,220	2.09%	280,560	311,780	3,931,534	0.906	0.5259	0.4765
2019	0.0%	31,220	−0.36%	73,362	104,582	4,036,116	0.897	0.5173	0.4638
2020	1.8%	31,780	12.41%	241,678	273,458	4,309,574	0.886	0.4811	0.4264
2021	5.0%	33,369	19.56%	121,747	155,116	4,464,690	0.875	0.4206	0.3682
2022	5.0%	35,038	15.70%	78,638	113,675	4,578,366	0.863	0.3574	0.3085
2023	5.0%	36,790	2.77%	152,619	189,408	4,767,774	0.850	0.3136	0.2666
2024	5.0%	38,629	1.02%	134,724	173,353	4,941,127	0.836	0.2955	0.2470
2025	1.8%	39,310	−2.82%	104,389	143,699	5,084,826	0.821	0.2899	0.2380
2026	0.0%	39,310	−2.30%	73,213	112,523	5,197,350	0.805	0.2887	0.2323
2027	0.0%	39,310	2.10%	332,860	372,170	5,569,519	0.788	0.2878	0.2267
2028	0.0%	39,310	2.59%	296,832	336,143	5,905,662	0.769	0.2853	0.2195

EXHIBIT 7

PART 1—PAGE 2

ONE SIMULATION FROM METHOD 3
STOCHASTIC MORTALITY, INFLATION, MEDICAL INFLATION, AND INVESTMENT YIELDS

Year	(1) Cost of Living Adjustment	(2) Indemnity Payment	(3) Medical Inflation	(4) Medical Payment	(5) Total Payment (2) + (4)	(6) Cumulative Total Payment Cum. of (5)	(7) Probability of claimant Living to Mid-year	(8) Present Value Factor	(9) Discount for Mortality & Investment Income (7) × (8)
2029	0.0%	39,310	−3.51%	84,297	123,607	6,029,269	0.750	0.2835	0.2127
2030	0.0%	39,310	5.22%	487,497	526,807	6,556,076	0.730	0.2790	0.2037
2031	2.0%	40,103	3.22%	153,476	193,580	6,749,656	0.709	0.2736	0.1939
2032	0.0%	40,103	12.65%	233,858	273,961	7,023,617	0.686	0.2639	0.1811
2033	5.0%	42,109	10.66%	139,118	181,227	7,204,844	0.663	0.2461	0.1631
2034	5.0%	44,214	6.31%	253,736	297,950	7,502,795	0.638	0.2329	0.1485
2035	2.3%	45,228	6.52%	119,357	164,585	7,667,380	0.612	0.2243	0.1372
2036	2.9%	46,533	20.05%	120,464	166,996	7,834,377	0.584	0.2030	0.1186
2037	5.0%	48,859	10.89%	529,686	578,545	8,412,922	0.556	0.1799	0.1001
2038	5.0%	51,302	8.96%	970,521	1,021,823	9,434,745	0.527	0.1679	0.0885
2039	4.5%	53,620	2.71%	284,077	337,697	9,772,442	0.497	0.1614	0.0802
2040	1.0%	54,155	8.04%	293,634	347,789	10,120,231	0.466	0.1511	0.0705
2041	5.0%	56,863	11.12%	694,986	751,848	10,872,079	0.435	0.1349	0.0587
2042	5.0%	59,706	9.13%	481,244	540,950	11,413,029	0.403	0.1224	0.0494
2043	5.0%	62,691	5.21%	1,250,236	1,312,927	12,725,956	0.372	0.1153	0.0428

2044	2.8%	64,466	7.24%	668,178	732,644	13,458,599	0.340	0.1111	0.0377
2045	2.2%	65,885	13.38%	490,791	556,676	14,015,275	0.308	0.1042	0.0321
2046	5.0%	69,179	5.43%	967,270	1,036,449	15,051,724	0.277	0.0976	0.0270
2047	2.6%	70,949	5.70%	2,033,827	2,104,775	17,156,500	0.246	0.0934	0.0230
2048	3.9%	73,682	7.81%	637,227	710,908	17,867,408	0.217	0.0895	0.0194
2049	2.3%	75,366	0.57%	1,858,383	1,933,749	19,801,158	0.188	0.0871	0.0164
2050	0.7%	75,896	3.92%	1,383,577	1,459,473	21,260,631	0.162	0.0849	0.0137
2051	2.1%	77,460	10.91%	923,621	1,001,081	22,261,712	0.137	0.0816	0.0112
2052	3.6%	80,212	4.92%	1,120,646	1,200,858	23,462,570	0.114	0.0791	0.0090
2053	0.4%	80,529	1.24%	958,304	1,038,833	24,501,403	0.094	0.0780	0.0073
2054	0.0%	80,529	3.62%	1,195,236	1,275,765	25,777,168	0.075	0.0759	0.0057
2055	3.1%	83,039	−1.81%	1,381,415	1,464,454	27,241,622	0.059	0.0740	0.0044
2056	0.0%	83,039	7.28%	1,128,922	1,211,961	28,453,583	0.045	0.0730	0.0033
2057	0.4%	83,359	6.51%	2,148,141	2,231,500	30,685,083	0.034	0.0719	0.0024
2058	0.4%	83,668	12.41%	1,726,074	1,809,742	32,494,825	0.025	0.0684	0.0017
2059	5.0%	87,851	7.14%	1,323,207	1,411,058	33,905,883	0.017	0.0639	0.0011
2060	4.0%	91,399	5.37%	1,560,112	1,651,512	35,557,395	0.012	0.0614	0.0007
2061	1.6%	92,839	9.29%	999,046	1,091,885	36,649,280	0.008	0.0603	0.0005
2062	0.0%	92,839	7.37%	541,530	634,369	37,283,649	0.005	0.0587	0.0003
2063	3.5%	96,131	9.95%	1,645,362	1,741,493	39,025,142	0.003	0.0549	0.0002
2064	5.0%	100,937	4.86%	1,090,053	1,190,990	40,216,132	0.002	0.0513	0.0001
2065	3.9%	104,845	4.80%	2,549,822	2,654,666	42,870,799	0.001	0.0487	0.0000
2066	4.3%	109,377	11.04%	1,219,660	1,329,038	44,199,836	0.0004	0.0465	0.0000
2067	2.7%	112,317	8.81%	3,720,340	3,832,657	48,032,494	0.0002	0.0451	0.0000
2068	1.2%	113,665	−2.52%	1,892,894	2,006,560	50,039,053	0.0001	0.0445	0.0000
2069	0.0%	113,665	0.52%	1,166,240	1,279,905	51,318,959	0.00002	0.0443	0.0000
2070	0.0%	113,665	−0.14%	1,513,593	1,627,258	52,946,217	0.00001	0.0442	0.0000
2071	0.0%	113,665	3.76%	11,045,559	11,159,225	64,105,442	0.000001	0.0434	0.0000
2072	2.2%	116,181	0.46%	5,311,459	5,427,640	69,533,082	0.0000002	0.0424	0.0000

EXHIBIT 7

PART 2—PAGE 1

	(10)	(11)	(12)	(13)	(14)	(15)
	Incremental Payments by Layer					
Year	$130,000 xs $370,000	$500,000 xs $500,000	$1 million xs $1 million	$3 million xs $2 million	$5 million xs $5 million	$5 million xs $10 million
1997 and prior						
1998	130,000	44,467	0	0	0	0
1999	0	68,248	0	0	0	0
2000	0	176,657	0	0	0	0
2001	0	48,624	0	0	0	0
2002	0	127,329	0	0	0	0
2003	0	34,675	18,700	0	0	0
2004	0	0	120,166	0	0	0
2005	0	0	300,038	0	0	0
2006	0	0	102,832	0	0	0
2007	0	0	160,416	0	0	0
2008	0	0	117,075	0	0	0
2009	0	0	179,289	0	0	0
2010	0	0	1,484	307,263	0	0
2011	0	0	0	74,150	0	0
2012	0	0	0	123,159	0	0
2013	0	0	0	114,903	0	0
2014	0	0	0	110,506	0	0
2015	0	0	0	103,416	0	0
2016	0	0	0	160,291	0	0
2017	0	0	0	626,065	0	0
2018	0	0	0	311,780	0	0
2019	0	0	0	104,582	0	0
2020	0	0	0	273,458	0	0
2021	0	0	0	155,116	0	0
2022	0	0	0	113,675	0	0
2023	0	0	0	189,408	0	0
2024	0	0	0	173,353	0	0
2025	0	0	0	58,873	84,826	0
2026	0	0	0	0	112,523	0
2027	0	0	0	0	372,170	0
2028	0	0	0	0	336,143	0
2029	0	0	0	0	123,607	0
2030	0	0	0	0	526,807	0
2031	0	0	0	0	193,580	0
2032	0	0	0	0	273,961	0
2033	0	0	0	0	181,227	0
2034	0	0	0	0	297,950	0
2035	0	0	0	0	164,585	0
2036	0	0	0	0	166,996	0
2037	0	0	0	0	578,545	0
2038	0	0	0	0	1,021,823	0
2039	0	0	0	0	337,697	0
2040	0	0	0	0	227,558	120,231
2041	0	0	0	0	0	751,848
2042	0	0	0	0	0	540,950
2043	0	0	0	0	0	1,312,927

(16)	(17)	(18)	(19)	(20)	(21)
Incremental Payments by Layer					
$5 million xs $15 million	$10 million xs $20 million	$10 million xs $30 million	$10 million xs $40 million	$10 million xs $50 million	$10 million xs $60 million
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

EXHIBIT 7

PART 2—PAGE 2

	(10)	(11)	(12)	(13)	(14)	(15)
	Incremental Payments by Layer					
Year	$130,000 xs $370,000	$500,000 xs $500,000	$1 million xs $1 million	$3 million xs $2 million	$5 million xs $5 million	$5 million xs $10 million
2044	0	0	0	0	0	732,644
2045	0	0	0	0	0	556,676
2046	0	0	0	0	0	984,725
2047	0	0	0	0	0	0
2048	0	0	0	0	0	0
2049	0	0	0	0	0	0
2050	0	0	0	0	0	0
2051	0	0	0	0	0	0
2052	0	0	0	0	0	0
2053	0	0	0	0	0	0
2054	0	0	0	0	0	0
2055	0	0	0	0	0	0
2056	0	0	0	0	0	0
2057	0	0	0	0	0	0
2058	0	0	0	0	0	0
2059	0	0	0	0	0	0
2060	0	0	0	0	0	0
2061	0	0	0	0	0	0
2062	0	0	0	0	0	0
2063	0	0	0	0	0	0
2064	0	0	0	0	0	0
2065	0	0	0	0	0	0
2066	0	0	0	0	0	0
2067	0	0	0	0	0	0
2068	0	0	0	0	0	0
2069	0	0	0	0	0	0
2070	0	0	0	0	0	0
2071	0	0	0	0	0	0
2072	0	0	0	0	0	0
	130,000	500,000	1,000,000	3,000,000	5,000,000	5,000,000

(16)	(17)	(18)	(19)	(20)	(21)
Incremental Payments by Layer					
$5 million xs $15 million	$10 million xs $20 million	$10 million xs $30 million	$10 million xs $40 million	$10 million xs $50 million	$10 million xs $60 million
0	0	0	0	0	0
0	0	0	0	0	0
51,724	0	0	0	0	0
2,104,775	0	0	0	0	0
710,908	0	0	0	0	0
1,933,749	0	0	0	0	0
198,842	1,260,631	0	0	0	0
0	1,001,081	0	0	0	0
0	1,200,858	0	0	0	0
0	1,038,833	0	0	0	0
0	1,275,765	0	0	0	0
0	1,464,454	0	0	0	0
0	1,211,961	0	0	0	0
0	1,546,417	685,083	0	0	0
0	0	1,809,742	0	0	0
0	0	1,411,058	0	0	0
0	0	1,651,512	0	0	0
0	0	1,091,885	0	0	0
0	0	634,369	0	0	0
0	0	1,741,493	0	0	0
0	0	974,858	216,132	0	0
0	0	0	2,654,666	0	0
0	0	0	1,329,038	0	0
0	0	0	3,832,657	0	0
0	0	0	1,967,506	39,053	0
0	0	0	0	1,279,905	0
0	0	0	0	1,627,258	0
0	0	0	0	7,053,783	4,105,442
0	0	0	0	0	5,427,640
5,000,000	10,000,000	10,000,000	10,000,000	10,000,000	9,533,082

EXHIBIT 7

PART 3—PAGE 1

	(22)	(23)	(24)	(25)	(26)	(27)
	Commutation Value by Layer, Discounted for Both Mortality and Investment Income Columns are derived by multiplying the corresponding column from Exhibit 4, pages 3 and 4, by Column 9, from pages 1 and 2. For example, Column 24 = Column 10 × Column 9					
Year	$500,000 xs $0	$500,000 xs $500,000	$1 million xs $1 million	$3 million xs $2 million	$5 million xs $5 million	$5 million xs $10 million
1997 and prior						
1998	126,511	43,274	0	0	0	0
1999	0	63,463	0	0	0	0
2000	0	160,878	0	0	0	0
2001	0	42,396	0	0	0	0
2002	0	102,789	0	0	0	0
2003	0	26,835	14,472	0	0	0
2004	0	0	92,054	0	0	0
2005	0	0	226,638	0	0	0
2006	0	0	76,768	0	0	0
2007	0	0	116,211	0	0	0
2008	0	0	81,284	0	0	0
2009	0	0	119,565	0	0	0
2010	0	0	950	196,744	0	0
2011	0	0	0	45,935	0	0
2012	0	0	0	73,479	0	0
2013	0	0	0	66,570	0	0
2014	0	0	0	62,213	0	0
2015	0	0	0	55,868	0	0
2016	0	0	0	82,533	0	0
2017	0	0	0	307,205	0	0
2018	0	0	0	148,548	0	0
2019	0	0	0	48,501	0	0
2020	0	0	0	116,604	0	0
2021	0	0	0	57,106	0	0
2022	0	0	0	35,074	0	0
2023	0	0	0	50,503	0	0
2024	0	0	0	42,825	0	0
2025	0	0	0	14,011	20,187	0
2026	0	0	0	0	26,140	0
2027	0	0	0	0	84,370	0
2028	0	0	0	0	73,781	0
2029	0	0	0	0	26,296	0
2030	0	0	0	0	107,320	0
2031	0	0	0	0	37,540	0
2032	0	0	0	0	49,623	0
2033	0	0	0	0	29,549	0
2034	0	0	0	0	44,253	0
2035	0	0	0	0	22,578	0
2036	0	0	0	0	19,813	0
2037	0	0	0	0	57,891	0
2038	0	0	0	0	90,393	0
2039	0	0	0	0	27,092	0
2040	0	0	0	0	16,034	8,472
2041	0	0	0	0	0	44,133
2042	0	0	0	0	0	26,704

(28)	(29)	(30)	(31)	(32)	(33)
$5 million xs $15 million	$10 million xs $20 million	$10 million xs $30 million	$10 million xs $40 million	$10 million xs $50 million	$10 million xs $60 million
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

EXHIBIT 7

PART 3—PAGE 2

	(22)	(23)	(24)	(25)	(26)	(27)
	Commutation Value by Layer, Discounted for Both Mortality and Investment Income Columns are derived by multiplying the corresponding column from Exhibit 4, pages 3 and 4, by Column 9, from pages 1 and 2. For example, Column 24 = Column 10 × Column 9					
Year	$500,000 xs $0	$500,000 xs $500,000	$1 million xs $1 million	$3 million xs $2 million	$5 million xs $5 million	$5 million xs $10 million
2043	0	0	0	0	0	56,253
2044	0	0	0	0	0	27,654
2045	0	0	0	0	0	17,864
2046	0	0	0	0	0	26,577
2047	0	0	0	0	0	0
2048	0	0	0	0	0	0
2049	0	0	0	0	0	0
2050	0	0	0	0	0	0
2051	0	0	0	0	0	0
2052	0	0	0	0	0	0
2053	0	0	0	0	0	0
2054	0	0	0	0	0	0
2055	0	0	0	0	0	0
2056	0	0	0	0	0	0
2057	0	0	0	0	0	0
2058	0	0	0	0	0	0
2059	0	0	0	0	0	0
2060	0	0	0	0	0	0
2061	0	0	0	0	0	0
2062	0	0	0	0	0	0
2063	0	0	0	0	0	0
2064	0	0	0	0	0	0
2065	0	0	0	0	0	0
2066	0	0	0	0	0	0
2067	0	0	0	0	0	0
2068	0	0	0	0	0	0
2069	0	0	0	0	0	0
2070	0	0	0	0	0	0
2071	0	0	0	0	0	0
2072	0	0	0	0	0	0
	126,511	439,635	727,941	1,403,719	732,859	207,656
	Overall Total = 3,813,435					

(28)	(29)	(30)	(31)	(32)	(33)
$5 million xs $15 million	$10 million xs $20 million	$10 million xs $30 million	$10 million xs $40 million	$10 million xs $50 million	$10 million xs $60 million
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
1,396	0	0	0	0	0
48,369	0	0	0	0	0
13,781	0	0	0	0	0
31,742	0	0	0	0	0
2,731	17,315	0	0	0	0
0	11,188	0	0	0	0
0	10,837	0	0	0	0
0	7,573	0	0	0	0
0	7,274	0	0	0	0
0	6,402	0	0	0	0
0	4,021	0	0	0	0
0	3,779	1,674	0	0	0
0	0	3,058	0	0	0
0	0	1,568	0	0	0
0	0	1,196	0	0	0
0	0	506	0	0	0
0	0	177	0	0	0
0	0	269	0	0	0
0	0	78	17	0	0
0	0	0	104	0	0
0	0	0	24	0	0
0	0	0	28	0	0
0	0	0	5	0.11	0
0	0	0	0	1.18	0
0	0	0	0	0.41	0
0	0	0	0	0.38	0.22
0	0	0	0	0.00	0.04
98,019	68,389	8,526	178	2.07	0.26
	Overall Total = 3,813,435				

CALIFORNIA WORKERS COMPENSATION BENEFIT UTILIZATION

A STUDY OF CHANGES IN FREQUENCY AND SEVERITY IN RESPONSE TO CHANGES IN STATUTORY WORKERS COMPENSATION BENEFIT LEVELS

WARD BROOKS

Abstract

Traditionally, workers compensation insurance ratemaking in California assumed that the utilization of benefits was independent of changes in statutory benefit levels. This assumption was retained for many years in the face of growing evidence that changes in statutory benefits indirectly affected the utilization of those benefits. Because the overall level of benefit utilization is a function of many factors, however, it was difficult to isolate which changes in utilization resulted from changes in statutory benefits and which resulted from changes in economic or social variables, randomness, or other factors. This paper explores and attempts to quantify the causal link between changes in statutory benefit levels and changes in the utilization of workers compensation benefits.

ACKNOWLEDGEMENT

The author would like to thank Mr. Dave Bellusci, Mr. William Kahley, and the members and attendees of the Actuarial Committee of the Workers' Compensation Insurance Rating Bureau of California for their guidance, and Mr. Liam O'Connor for his assistance, which was invaluable on this project.

1. INTRODUCTION

Historically the Workers' Compensation Insurance Rating Bureau of California (the Bureau) has assumed frequency will not change in response to benefit level changes and severity will change by exactly the change in benefits.[1] If benefits are increased 10%, we expect no change in frequency and a 10% increase in severity, all other things being equal. However, if benefits are increased 10% and frequency increases 1% in response, then we say we have observed a 1% change in *frequency benefit utilization*, again, all other things being equal. If severity increases 12%, perhaps because durations have increased as workers stay on claim longer, then we say we have observed a 2% change in *severity benefit utilization*.

If we chronically over- or underestimate changes in frequency or severity by failing to recognize changes in utilization, then this error will be reflected in the residual trend component of the ratemaking process. We should be able to increase the accuracy of the ratemaking process by quantifying changes in benefit utilization and incorporating them into our on-leveling procedure, thereby removing them from the residual trend. The accuracy of both our on-leveling and trend procedures will be improved as well as our understanding of the workers compensation system.

Some changes are administrative rather than statutory. When we refer to statutory benefit levels, we mean both those promulgated by statute and those effected administratively.[2] Each

[1]For the purposes of this paper, a change in *benefit utilization* means an indirect effect of the benefit change. That is, a change in frequency or severity that is related to the change in benefit level but not measured by the direct effect. The direct effect is measured by the Bureau's benefit level change estimate. Note that this definition is broader than that used for utilization in other contexts. For an overview of workers compensation ratemaking, including the role of benefit change estimates and their potential indirect effects, the reader should consult Feldblum [1]. In particular, Sections 5.C and 10 will be helpful to the reader not familiar with the issue of the indirect effects of benefit changes.

[2]As an example of an administrative change, in 1997 California's Division of Industrial Relations (DIR) revised the official Permanent Disability Rating Schedule (PDRS). The PDRS is used to evaluate an injured worker's loss of functional work capacity and culminates in the assignment of a permanent disability rating. The injured worker's weekly indemnity benefit is based on this permanent disability rating according to a schedule promulgated in California statute. The estimated impact of the DIR's revisions became controversial, highlighting the fact that these estimated cost impacts are just that, *estimates*. Sometimes they are revised ex post facto, as more information becomes available.

year the Bureau evaluates the expected impact of legislative and administrative changes on the cost of benefits. For the more common changes, the Bureau uses a model to estimate the impact. For the less common changes, the Bureau typically conducts a special study. In both cases, the estimated impact is used in the Bureau's pure premium ratemaking to adjust historical accident year indemnity losses to a current or prospective level. This estimated impact is for direct effects only.[3] It assumes there will be no change in benefit utilization. In economics parlance, it assumes that the utilization of benefits is inelastic.

Finally, we note that benefit utilization is internal to the workers compensation system. Changes in costs that result from changes in statutory benefits are a matter of public policy. California legislators and the administrators of the California workers compensation system routinely solicit the Bureau's estimated cost impacts for proposed changes. Public policy decision making will be enhanced if actuaries can estimate both the expected direct and indirect fiscal impacts of proposed changes in benefits.

2. HISTORY

In 1996 the Bureau's Governing Committee directed the Bureau to conduct a study to determine an appropriate loading in pure premium rates for changes in benefit utilization. The Bureau had commissioned two prior studies: Meyer [2] in 1991 and Appel [3] in 1992. Based on these studies, the Bureau incorporated into its pure premium ratemaking an adjustment to losses to reflect expected changes in utilization resulting from benefit level changes. The California Commissioner of Insurance, however, questioned the accuracy and method of incorporation of this utilization adjustment in his October 13, 1995 decision (Ruling

[3] For indemnity costs, this is no longer true. An earlier version of this paper was accepted by the California Department of Insurance as the basis for an adjustment to losses to reflect expected changes in utilization resulting from benefit level changes. This adjustment has been incorporated in the Bureau's filing for pure premium rates effective January 1, 1998.

No. 287). The Commissioner directed that a more in-depth study of utilization be undertaken before such an adjustment would be acceptable in pure premium ratemaking. This paper documents the findings of that study.

3. METHODOLOGY

The goal of this study is to quantify changes in frequency and severity that occur in response to changes in benefit levels. The model design selected assumes that the indirect effects of benefit changes are a function of the direct effects. That is, changes in benefit utilization are assumed to be a function of the Bureau's estimated changes in benefit levels. We will attempt to quantify this relationship using multivariate regression supplemented by nonparametric techniques where appropriate. Following is an outline of the methodology we will use to investigate indemnity frequency utilization. We will discuss medical frequency utilization along the way. Severity utilization will be discussed in a later section.

We will start by surveying graphically the candidate dependent and independent variables. We will look at the level of each variable over time and its annual percentage changes. We will then look at the correlations among variables. Here we are looking for combinations of the independent variables that are highly correlated with the dependent variable but not highly correlated with each other. We want to avoid highly correlated independent variables in a regression to avoid multicollinearity with its attendant risk of unstable and distorted least-squares estimates. It will happen that we will encounter a group of highly correlated independent variables that we wish to retain in the model. We will apply a special transformation, principal components extraction, to retain the explanatory variance while removing the multicollinearity. We will discuss this further at that time.[4]

---

[4]Readers wanting a review or more information on analysis of variance, multicollinearity, transformations, analysis of residuals, and other topics in regression analysis should see Miller [4].

The first correlations we will consider are the standard Pearson Product Moment Correlations. (These are the familiar correlations obtained using the appropriate function in Lotus or Excel.) The Pearson Product Moment Correlation between two variables assumes each is drawn from a normally distributed population. The significance of the Pearson correlation is only as strong as this assumption is valid. Because of this, we will also look at a nonparametric statistic, the Spearman Rank Correlation Coefficient. This statistic relies on much weaker assumptions. Intuitively, we will be most comfortable when these two measures of correlation are in agreement. Before proceeding, let us consider the common interpretation when these statistics are not in agreement.

If there is a significant correlation indicated by the nonparametric statistic but not the parametric statistic, then we propose that a correlation exists, but that it cannot be precisely measured. If there is a significant correlation indicated by the parametric statistic but not the nonparametric statistic, then we propose that the parametric statistic is erroneous, probably because of a violation of the underlying assumptions, though sometimes because of an outlier.[5]

Following this examination of the variables (Exhibits 1 through 4), a series of candidate regression models will be postulated. Each will be regressed and we will diagnose each model (Exhibit 5). We will first look to see if the coefficients make sense. We will compare the models' relative performance, adjusted for degrees of freedom. We will test each model for bias and the normality of its residuals. For the better models we will look more closely at performance and the appropriateness of the model's specification (Exhibit 6).

Following this, for the best models we will look at projected performance in practice (Exhibits 7 through 10, and 12). We will

---

[5]Readers interested in more information on nonparametric statistics should see Ferguson [5] or Siegel [6].

do some sensitivity testing on our most novel variable (Exhibit 11). Finally, we will present the best model with confidence intervals for our point estimates. The best model will be presented along with three other models as a form of sensitivity testing of our economic variables (Exhibit 13).

Before proceeding to the main analysis, a technical aside is in order. During the following discussion the reader might wonder if a transformation of the data relating to workers compensation reporting bases was considered. It was. But to cut down on the volume of analysis to be presented and discussed, we will deal with this issue here, summarily.

Reporting Bases

In California, workers compensation rate level indications are based on calendar-accident year data while classification relativities are based on policy year data. Variables that are collected outside of the workers compensation system—economic variables, for example—are generally on a calendar year basis. Therefore, variables of interest may be on different reporting bases. Because there is a timing difference between variables with different reporting bases, the correlations between variables can be affected. This is essentially the same issue as whether there is a lagged correlation between two variables; here the lag would be due to the timing difference of the reporting bases.

To eliminate this lag, we explored transforming calendar year variables into policy year variables. For example, suppose premiums are written and losses occur uniformly over a year. (We used more exact distributions for our transformations.) Also, suppose real gross state product increased 0.01% in 1982 and 4.93% in 1983. Then policy year 1982 real gross state product increased 2.47% [(0.0001 + 0.0493)/2]. It turned out, however, that matching variables' reporting bases delivered inferior results. This implies that a slight lag exists between the calendar year events and their policy year manifestation.[6] That is, there is a higher correlation between a calendar year 1982 economic event and a policy year 1982 (not transformed) event, with an implicit six-month lag, than between the calendar year event and the policy year event transformed to match average dates of occurrence.

The policy year variables used in this paper are developed from the Bureau's Unit Statistical Reporting (USR) system. Incurred claim counts and exposures are defined per the California Workers' Compensation Uniform Statistical Reporting Plan. Frequencies are developed from the USR data in Appendix A; severities are developed in Appendix I. The benefit level variables, which are used to adjust historical losses to a current or projected benefit level, are calendar-accident year.

4. THE VARIABLES

We begin the analysis of indemnity frequency utilization by reviewing all available candidate variables. We preface this section by noting the importance of accounting for all significant factors that affect indemnity frequency. In the end, we would like to have accounted for as much variation as possible and we would like the variation unaccounted for to be purely random noise. We do not want any significant factors to be omitted from the final regression model. If they are omitted, then the model is misspecified. This misspecification may bias the estimates or lead to erroneous conclusions about the confidence we have in the estimates.

The variables considered in the analysis are presented graphically in Exhibit 1. The top graph of each part of Exhibit 1 displays the value of each variable over time. The bottom graph shows the annual percentage change in the original variable. A tabular presentation of the variables and their annual percentage

[6]The average date of occurrence for both calendar year and accident year variables is about July 1st. The average date of occurrence of a policy year variable is December 31st.

changes is presented in Exhibit 2. Following is a discussion of each variable.

Indemnity Claim Frequency

This is the dependent variable—our first target.

All frequencies are policy year claims per million dollars of reported payroll, adjusted to a 1987 wage level. Claim counts were taken from the Bureau's USR system at third report level. Payrolls were adjusted to a 1987 wage level using average wages developed from the California Statistical Abstract (Appendix A).

Part 1 of Exhibit 1 shows the history of indemnity claim frequency from policy year 1961 through 1994.

Medical-Only Claim Frequency

Medical-only claim frequency has exhibited a persistent long-term downward trend for over three decades (Exhibit 1, Part 2). This trend is counter-intuitive, as we would expect indemnity and medical-only claim frequencies to move together. There is a wide range of speculation regarding the causes of this trend. Suspect causes include changes in medical-only reporting patterns, the decreasing hazardousness of the California insured mix of business, or an increasing tendency for all claims to have an indemnity component. In any case, since medical-only claims represent less than 5% of workers compensation costs and there is a lack of consensus about this long-term trend's causation, no attempt was made to model medical-only utilization.

Total Claim Frequency

Total claim frequency (Exhibit 1, Part 3) was not analyzed. Total claim frequency is dominated by medical-only claims, which in policy year 1992 outnumbered indemnity claims by roughly two-to-one.

Indemnity Benefit Level

This is the key independent variable. The coefficient on this variable will measure frequency benefit utilization. If benefit level estimates are accurate and unbiased, then our a priori expectation is that the coefficient on this variable will be zero if no utilization effect is present. Absent a utilization effect, a change in benefit level will produce no change in frequency. If the coefficient was 0.3 and significant, then in response to a 10% benefit level increase we would expect a 3% ($0.3 \times 10\%$) increase in frequency. The null hypothesis is that this coefficient equals zero. If we can reject this hypothesis, then we can conclude a utilization effect is present and that the coefficient measures it as a function of the benefit level change.

Because the indemnity benefit level variable is key, it is critical that it be as accurate as possible and, perhaps more importantly, be unbiased. The process for quantifying the cost impact of benefit level changes was discussed earlier. Clearly, if the process is biased, we could inadvertently capture this bias in our model and falsely conclude there is a utilization effect where there is only systematic bias in our estimates of legislative changes. Some preliminary analysis suggested that historical benefit level estimates were indeed biased, and the Bureau revised its law amendment evaluation models to remove the bias.[7]

[7]What was this bias? It was related to the Bureau's prior use of an average wage level intended to reflect the insured population. This has been replaced by an average wage level intended to reflect the expected insured *claimant* population, based on the Bureau's Individual Case Report data. This change addressed the fact that the average wage and wage distribution of the population of insured workers and the population of insured claimants are different. The latter is a subset of the former. The author has experimented with projecting the distribution of *insured* wages by fitting insured claimant wage distributions for successively higher levels of permanent partial disability. The underlying assumption here—though unproven—is that a primary cause of the difference between the insured and claimant wage distributions is self-selection and that the effect of self-selection diminishes with the seriousness of injury. Further improvements in the procedure to evaluate legislative changes may be possible by quantifying the relationship between the insured and claimant wage distributions as a function of benefit levels. Also, we note that the Bureau's evaluation methodology and the tables underlying the calculations were substantially the same throughout the period under study, so no bias was introduced by a change in methodology.

The calendar year indemnity benefit level history, revised to correct the bias discussed above, is presented in Exhibit 1, Part 4 and developed in Appendix B.

Medical Benefit Level

The medical benefit level index captures changes in California's Official Medical Fee Schedule and an index of hospital inflation costs. Unlike indemnity benefit level changes, however, a great many other factors affect medical costs in addition to the costs of medical procedures and hospital costs. Examples include the advent of managed care and the development of new technologies, such as magnetic resonance imaging and new arthroscopic surgery techniques. Indeed, these other factors are widely believed to have dominated changes in medical costs over the last several decades. For the task at hand, it may be impossible to isolate utilization effects out of this larger body of factors.

The calendar year medical benefit level history is presented in Exhibit 1, Part 5 and developed in Appendix B.

Total Benefit Level

The total benefit level combines the indemnity and medical benefit levels, weighted by their respective partial pure premiums. The calendar year total benefit level history is presented in Exhibit 1, Part 6 and developed in Appendix B.

Economic Variables

The general state of the economy is important in workers compensation. As an economy nears capacity, employees work longer hours, less skilled workers are pulled into the production cycle and the opportunity cost of safety measures may increase. As a result, claim frequency per worker varies with the economic cycle. We considered three economic variables in our analysis: aggregate employment, real gross state product, and the unemployment rate. The economic variables are shown graphically in

Exhibit 1, Parts 7 through 9. Each variable is specific to California, and its development is presented in Appendix C. These variables, which are broad measures of the robustness of the state's economy and labor market, serve to quantify changes in utilization that are a natural consequence of the economic cycle.

We note that the importance of economic influences in workers compensation systems is an on-going area of research. In this paper, we assume a priori that economic variables should be considered in the model.

Hazardousness Indices

The prior utilization studies commissioned by the Bureau examined only a subset of classifications. Only 50 classes were analyzed over a 22-year period in the 1992 study. Unfortunately, the selected classes may not be representative of the mix of business throughout the experience period. Changes in the mix of business may explain some of the changes in the overall utilization level over time. So, as California shifted from a predominantly manufacturing economy to a service economy over the last several decades, the level of hazardousness shifted concurrently. In 1970, for example, manufacturing classifications accounted for 16.9% of total workers compensation payroll; in 1990, 13.6%. The clerical standard classification 8810 grew from 20.7% of payroll in 1970 to 28.5% in 1990. To capture this phenomenon, we examined the entire insured population of classifications.

Additionally, two indices were developed to measure changes in the hazardousness of the insured California workers compensation population from policy year to policy year. The first index, the indemnity frequency hazardousness index, captures changes in frequency attributable to changes in the mix of business. The second index, the pure premium hazardousness index, captures changes in frequency and severity attributable to changes in the mix of business. These indices are developed in Appendix D.

These indices capture the subtle, long-term transformation of the California economy's level of hazardousness (Exhibit 1, Parts 10 and 11). Both illustrate the growing dominance of the service sector in the California economy. Because manufacturing is both more highly cyclical and more hazardous, the insured population's hazardousness fluctuates with the state's economic cycles. Throughout the period studied, the indemnity pure premium index fell sharply with the onset of recessions. This relationship may change in the future if the relative frequency and severity of claims among economic sectors changes.

Annual changes in these indices, however, were not highly correlated with annual changes in indemnity frequency (Exhibit 4, Parts 7 and 8). Indeed, indemnity frequency persistently increased over the period studied in spite of the decreasing hazardousness of the insured population. This does not mean the hazardousness indices are invalid or inaccurate. The hazardousness indices capture a long-term trend, while we are looking at annual changes. Further, the divergent trends in hazardousness due to changes in the mix of business and in indemnity frequency merely suggest there are other factors that are pushing indemnity frequency from different directions. In any complex system there may be a variety of forces that push in different directions at the same time. Though annual changes in the hazardousness indices did not prove relevant in the final model, we have included them here for their relevance to the utilization phenomenon and to introduce the concept of a metric for changes in mix of business.

Litigation Rates

Discretion makes benefit utilization possible and litigiousness is commonly considered a proxy for discretion in the workers compensation system. Benefit utilization exists because workers can exercise some discretion in the filing of workers compensation claims. In a textbook world, benefit utilization might not exist. No one would use workers compensation instead of vacation time, health insurance or unemployment insurance. Highly paid workers would not opt to use sick pay and health insurance benefits instead of workers compensation benefits.[8] But in the real world, many workers are presented with the choice to utilize their workers compensation benefit, or not, and this discretionary act is anecdotally correlated with litigation. To examine this, a variable measuring litigiousness was developed.

From 1972 to 1992 (except 1990) the California Workers' Compensation Institute (CWCI) collected information on the number of Applications for Adjudication filed with the Workers' Compensation Appeals Board (Appendix E). The CWCI ratioed the number of applications to the total number of claims to arrive at a litigation rate. This litigation rate might serve as a proxy for litigiousness. The denominator of this ratio, however, includes medical-only claims, which are rarely litigated. A ratio to indemnity claims would be a better measure. The litigation rate history, adjusted to an indemnity claim basis, is presented in Exhibit 1, Part 12. When the litigation rate is adjusted to an indemnity claim basis, the marked upward trend in the litigation rate disappears and the rate is fairly flat.

This result was surprising. The phenomenon of medical-only claims decreasing as a share of total claims is the obvious mathematical "cause" of the flattening of the litigation rate. When earlier years are adjusted to account for the lesser share of indemnity claims to total, the litigation rate for indemnity claims soars. The level of litigation suggested by this data is much higher than for other states. Some of this magnitude may be due to peculiarities associated with the survey method or California's adjudication process. Nevertheless, this data suggest the level of litigiousness in California not only is high, but also has been so for several decades. Still more surprising, changes in the litigation rate proved to be negatively correlated with changes in

[8]The higher a worker's income over the maximum benefit, the lower the percentage of pre-injury income workers compensation benefits replace. The benefit, therefore, decreases as a worker's income increases, and at some point may actually present an additional burden.

indemnity frequency, a result counter to our a priori expectation. This raised uncertainty as to whether this variable is accurately measuring litigiousness or some other phenomenon. Because of this uncertainty, this variable was dropped from consideration in the analysis.

Ratio of Cumulative Injuries to Total Indemnity Claims

This is the ratio of incurred claims coded as cumulative injury as defined by the Unit Statistical Reporting system to total incurred indemnity claims for each policy year.[9] Note that this ratio does not necessarily rise or fall with changes in the frequency or absolute number of cumulative injury or total indemnity claims. Cumulative injuries never comprised more than 10% of indemnity claims. Therefore, it is not appreciably correlated with indemnity frequency by definition. This variable is probably a more direct measure of changes in the discretionary element than litigiousness because cumulative injury claims have a higher degree of discretion available. For example, if you have an accident on the job, a nasty cut say, you are more likely to be seen and sent to the human resources department to fill out a form. But initiating a carpal tunnel or stress claim is much more within a worker's sole control. Note that in the presence of a benefit level variable we expect the ratio to capture discretion *unrelated* to changes in benefit levels.

The ratio of cumulative injuries to total indemnity claims is presented in Exhibit 1, Part 13 and developed in Appendix F.

Principal Components of Economic Variables

The economic variables are highly correlated among themselves. The Pearson Product Moment Correlation between annual changes in real gross state product (rGSP) and aggregate employment (AggE) is 0.655; between rGSP and the unemployment rate

[9] This variable was suggested by Mr. James J. Gebhard, FCAS, MAAA, following the failure of the litigiousness proxy.

(Unemp), −0.892; between AggE and Unemp, −0.677. If regression is to be used, these correlations are too high to use more than one variable without risking multicollinearity—that is, the linear dependence of the independent variables. If independent variables in a model are linearly dependent, then least squares estimates tend to be unstable and may be far from their expected values. To extract any additional explanatory information lost by using only one economic variable while not introducing multicollinearity, the principal components of the economic variables were formed. Principal components are the uncorrelated linear combinations of the subject variables that maximize variability.[10]

The first and second principal components of two sets of economic variables were formed. The first set was annual changes in rGSP and AggE. The second set was annual changes in rGSP, AggE and Unemp. The principal components are presented in Exhibit 1, Parts 14 through 17. Their development is presented in Appendix G.

Self-Insurance Share Index

A complicating issue in virtually all analyses of the California workers compensation market is the changing composition of the insured population. The data collected by the Bureau represents only the insured population. When an employer exits the insured market by self-insuring, his experience under self-insurance is lost to the Bureau while his insured history cannot be isolated from the Bureau's historical experience. The reverse is true when an employer returns to the insured market from self-insurance. Clearly, the comings and goings of employers has the potential to distort the insured experience. This is particularly true when large groups of employers with unique experience come and go en masse.

[10]For more information on principal components see Chapter 8 of Johnson [7]. This is also a good general reference for multivariate regression.

This problem is neither unique to this analysis, nor to California. In fact, the potential exists for changes in the self-insured population to affect aggregate pure premium ratemaking. As an example, if a group of risks with poorer experience than the aggregate begins to exit the insured market over a period of time, an improving loss ratio will be picked up by the residual trend procedure. Not knowing that the improvement is due to a change in the mix of insureds, the trend might be forecast to continue beyond the time the insured population has stabilized. To address this problem, a variable was developed to measure changes in the self-insured market.

The self-insurance share index was developed to capture annual changes in self-insurance costs as a share of total California workers compensation costs. This variable is developed from information reported by the state and federal governments and the Bureau and compiled by the Social Security Administration. This variable is presented in Exhibit 1, Part 18; the development is presented in Appendix H. This variable captures only changes in the net volume of the self-insured market. Qualitative changes are not captured (i.e., whether the experience of the self-insured market is improving or deteriorating, absolutely or relatively).

There is no appreciable correlation between annual changes in the self-insurance share index and indemnity frequency (Exhibit 3 and Exhibit 4, Part 15). On this basis, we conclude that change in the level of self-insurance is not a candidate independent variable nor likely to affect the analysis.

5. THE MODELS

We first examined the correlations among the variables. The Pearson Product Moment Correlations among the variables' annual changes and the significance of these correlations are summarized in Exhibit 3. In all cases, the analysis was conducted on the least common denominator of years for a given set of subject variables. Note that the analysis was on the annual changes in these variables—not their absolute levels. For example, the annual change in the unemployment rate is an independent variable—not the unemployment rate itself. Further references to variables will mean their annual percentage changes unless otherwise stated.

The candidate variables were tested for normality (using Kolmogorov–Smirnov). All variables except the changes in indemnity and total benefit levels, which are clearly skewed, passed tests for normality. Note that interpretation of the significance of the Pearson Product Moment Correlation between two variables assumes both to be distributed normally and that our key independent variable is not.

Exhibit 4 presents a graph of each candidate independent variable against indemnity frequency as well as the regression of indemnity frequency on the independent variable and the Spearman Rank Correlation Coefficients. The normality assumption is not required of the Spearman Rank Correlation Coefficient. For the benefit level changes, Exhibit 4 also presents regressions with a dummy variable. The dummy variable is 1 for years with an indemnity benefit change and 0 otherwise. Introduction of the dummy variable did not improve the amount of variation explained by benefit changes alone. Note, however, that the nonparametric Spearman Rank Correlation is strong and highly significant.

We examined these variables to select candidates for multivariate regression. As discussed above, candidates should be reasonably correlated with frequency but not highly correlated with other variables in the model. From a review of the information in Exhibits 3 and 4, and other exploratory analysis, we chose models with the following structure.

Y-Intercept

Models with or without a constant term.

Benefit Level

Calendar year indemnity benefit level changes, total benefit level changes, or indemnity and medical benefit level changes separately. The coefficient on the benefit level variable measures frequency utilization. We will conclude there is no utilization effect if this variable is not significantly different from zero.

Economic Variable

We considered models with the following economic variables:

1. Real gross state product (rGSP);
2. Aggregate employment (AggE);
3. Real gross state product and aggregate employment (for comparison purposes only);
4. The first principal component of rGSP and AggE;
5. The first and second principal components of rGSP and AggE;
6. The first principal component of rGSP, AggE and the unemployment rate (Unemp);
7. The first and second principal components of rGSP, AggE and Unemp.

Ratio of Cumulative Injury Claims to Total Indemnity Claims

Models with or without the cumulative injury index.

A simple multivariate linear structure was selected, as no strong nonlinear or lagged patterns were present. We next performed multivariate regressions using Manugistic's STATGRAPHICS Plus (1995) statistical software. Kalmia's WinSTAT, Version 3.1 (1995) was also used for certain diagnostic tests and to confirm results obtained using STATGRAPHICS Plus.

6. THE RESULTS

Eighty-four multivariate regressions are possible with the selected variables. A summary of selected statistics for these eighty-four models is presented in Exhibit 5. Part 1 of Exhibit 5 summarizes all models using the indemnity benefit level; Part 2 summarizes all models using the total benefit level; Part 3, the indemnity and medical benefit levels separately. For the better models (as judged by R^2 adjusted for degrees of freedom), the indemnity benefit level consistently outperforms both the total and component benefit level models. This is not surprising, because, as discussed above, the medical benefit level measures only a narrow component of medical benefit costs and the connection between changes in medical costs and indemnity benefit utilization is tenuous.

The models are ordered by adjusted R^2 on each part of Exhibit 5. The mean residual error is presented for each model. This indicates whether or not the model is biased. We want a model whose mean residual error is very close to zero. The normality of the residual errors for each model was tested using the Kolmogorov–Smirnov and Shapiro–Wilks tests. A low p-value on these tests means we can conclude the residuals are not distributed normally. The primary concern is that the residuals are skew. A low p-value on the skewness test would indicate a model's residuals are more skew than the normal distribution's. A low p-value on the kurtosis test would indicate a model's residuals are not as kurtotic as a normal distribution. A few models fail ($p < 0.10$) both the Shapiro–Wilks and kurtosis tests—but neither the Kolmogorov–Smirnov nor skewness tests. These models' residuals are more highly kurtotic than a normal distribution's. This is not bad—it means the actual data are more tightly distributed about the fitted line than if they were normally distributed.

The seven models with the highest adjusted R^2 include the cumulative injury index variable and a constant term. The regression output for these seven models is presented in Exhibit

6. All seven models are significant based on an analysis of variance. The model with the highest adjusted R^2 explains 91.4% of the variance in annual changes in indemnity claim frequency. However, the second principal component of this model is not significant at a 90% or higher confidence level. The model excluding this term (with the second highest adjusted R^2) explains 88.7% of the variance and all terms are significant at a 95% confidence level. This model, Model 2, includes the indemnity benefit level, a constant term, the first principal component of rGSP, AggE and Unemp, and the cumulative injury index.

Three other models have terms that are all significant at a 95% confidence level, each differing in the choice of economic variable. The fifth model includes the first principal component of rGSP and AggE. The sixth model includes AggE. The seventh model includes rGSP. These models explain 86.1%, 84.2% and 82.9% of the variance, respectively, as compared to the second model, which explains 88.7%. Exhibits 7 through 10 present a graphical analysis of each of the four models (Models 2, 5, 6 and 7).

The graph on Part 1 of Exhibits 7 through 10 shows the actual and fitted annual percentage changes. Part 2 of each exhibit demonstrates application of the model to predict annual frequency changes presuming we have past or estimated frequency information. That is, Part 2 is analogous to the graph on Part 1, but with a one, two or three period projection interval. For example, in the first graph of Part 2 of Exhibit 7, if we are projecting policy year 1997 we must know or have estimated the indemnity frequency for policy year 1996 and the benefit level changes and economic variable changes for 1997. The second graph, again projecting policy year 1997, assumes we have the frequency for policy year 1995 and the benefit level and economic variable changes for 1996 and 1997. These graphs illustrate how the fitted models would perform in practice. Part 3 of Exhibits 7 through 10 parallels Part 2, but for the level of indemnity claim frequency—not the annual changes in it.

These results are promising. A large portion of the annual variation in indemnity frequency is explained. The overall models are highly significant (based on an analysis of variance) and all the variables in the models are significant at a 95% level of confidence. The estimates of the coefficient on the indemnity benefit level range from 0.221 to 0.321, with the estimate for the most powerful model squarely inside this range at 0.262. So our best estimates using a variety of economic variables fall within a fairly narrow range.

One weakness of these results is the limited time frame of observation. Only sixteen years of data were available concurrently for the included variables. This limitation was imposed by the cumulative injury index, which was available beginning with policy year 1977. A key concern here is the number of economic cycles over which the economic variables were observed. With economic variables we would like to include several economic cycles to have greater confidence in our findings. To examine what impact this limitation may have had, we look now to the same models, but exclude the cumulative injury index.

Models Excluding the Ratio of Cumulative Injuries to Total Indemnity Claims

Thirty years of data are available for models including the indemnity benefit level, a constant term and the economic variables presented in Exhibits 7 through 10. Selected results for these regressions appear on Exhibit 5, Part 1 and the regression output is included in Exhibit 11. Although the models explain only 18.8% to 20.3% of the total variation (adjusted for the degrees of freedom), all four are significant at the 95% confidence level based on an analysis of variance. The coefficients on the indemnity benefit level range from 0.287 to 0.330. This range overlaps considerably the range of the models that include the cumulative injury index. Additionally, these coefficients are significant at the 90% confidence level in two models and the 95% confidence level in the other two.

Clearly, the introduction of the cumulative injury index does not significantly affect the estimated indemnity benefit level coefficient. The estimates would be only a few points higher without this variable. The cumulative injury index does, however, explain over 60% of the variance and allows us to be confident our utilization estimates are not distorted due to a misspecified model with a large portion of unaccounted-for variance.

Interpretation of the Negative Constant Term

The constant term in the final model is statistically significant. It is also negative, implying that, all other things equal, indemnity frequency will fall 3.58% per year. Why might this be?

Note that the coefficient on the first principal component of the three economic variables is negative. It happens here that a negative first principal component corresponds to an expanding economy while a positive first principal component corresponds to a recessionary economy.

Consider the median value of the first principal component over the fifteen-year fitting range. This value corresponds to 1989 and is −4.7881 (Exhibit 2, Part 2). In 1989 California's real gross state product grew 3.8%, aggregate employment grew 3.6% and the unemployment rate fell to 5.1% from 5.3% the prior year. The increase in frequency for 1989 due to the state of the economy is about 1.03% [-0.214998×-4.7881]. Indeed, 1989 seems representative of what we might expect for long-term economic growth.

But long-term, frequency, which is a rate and not an absolute number, cannot increase without bound. If it did, at some point our model would project every insured to file a claim on average! If our future were a series of 1989s without end, we would project annual increases of 1.03% in frequency, without end. Clearly the model would be misspecified. To balance the economic variable, the model must have some offset for the long-term level of economic growth. This offset is reflected in the constant term.

The situation with the indemnity benefit level is similar. In California, statutory benefit levels are not indexed to inflation. To maintain the real (inflation adjusted) value of indemnity benefits, periodic increases must be made. Over the years, we expect some portion of benefit level increases reflect adjustments to maintain purchasing power. But these adjustments have been made sporadically. In the intervening years, the *real* purchasing power of indemnity benefits is decreasing. It is being deflated by inflation. If frequency is sensitive to changes in *real* benefit levels, then we expect frequency to decline on average during the years when real benefit levels are falling (i.e., in years when benefit level changes are less than inflation). This phenomenon is reflected in the constant term.

Finally, as discussed above in the development of the hazardousness indices, the mix of business in California has been changing over the last several decades. Although annual changes in hazardousness did not predict annual changes in indemnity frequency, this does not mean the long-term trend in hazardousness is absent from our model. Both the average and median change in indemnity frequency as measured by the indemnity frequency hazardousness index are about -0.75% per year over 1978–1992. This long-term trend is reflected in the constant term.

Returning to our fitted models, Exhibit 12 presents additional performance information for the seven models in Exhibit 6. The average absolute error and adjusted R^2 are presented for the fitted model and the projection interval models. The relative performance of the projection interval models is consistent with the performance of the original models. The accuracy of the models does not deteriorate excessively with the increasing projection interval.

These results indicate that we can be highly confident that an indemnity frequency benefit utilization response exists and is statistically significant. Our estimates of this response are remarkably stable over different time periods, a variety of economic variables, and the inclusion or exclusion of a variable to capture changes in the non-benefit-related discretionary element in the workers compensation system.

7. APPLICATION

Exhibit 13 presents the indemnity frequency benefit utilization point estimates and confidence intervals for the four models in Exhibits 7 through 10. The best estimate of indemnity frequency benefit utilization, Model 2's estimate, is from Exhibit 7. The model indicates that indemnity frequency would increase 2.6% in response to a 10% increase in the indemnity benefit level. The model is linear and might be interpreted also as implying that a 10% decrease in the indemnity benefit level would produce a 2.6% decrease in indemnity frequency. However, no benefit level decreases were included in the parameterization of the models, so any conclusions about the utilization response to benefit level decreases would be extrapolating beyond the data, with its attendant risks.

We should stress that the Bureau's goal here was quantifying the utilization effect—*not* forecasting the future level of indemnity frequency. Although the models developed here can be used to project future levels of indemnity frequency (and we tested their performance to do so), the Bureau's first concern was with the benefit level coefficient to estimate expected utilization effects. We examined whole models under the theory that our confidence would be higher if both the whole and its parts were sound and because a regression approach is always sounder when most of the variance is explained by the model.

8. SEVERITY

Two analyses parallel to the above analysis of indemnity frequency were performed for indemnity severity—one using calendar year benefit level changes and one using policy year benefit level changes. Exhibit 14 graphically presents indemnity severity and real indemnity severity (adjusted to a 1982-84 level using the California Consumer Price Index). Exhibit 15 tabulates the value of each variable and its annual percentage changes. Exhibit 16 shows the Pearson Product Moment Correlations among the variables. Exhibit 17 shows a graph of the indemnity benefit level against indemnity severity and real indemnity severity as well as the regression of the severities on the indemnity benefit level and the Spearman Rank Correlation Coefficients.

Note that while the Pearson Product Moment Correlations appear respectable, the nonparametric correlations are small and insignificant. Nor do the graphs reveal any relationship between changes in severities and changes in indemnity benefit levels. The lack of any nonparametric correlation suggests that the parametric statistics are spurious. This is bolstered by our visual inspection.

Because we can find no correlation with our target independent variable—benefit level changes—our analysis stops here. This does not mean, however, that we could not build a model for changes in severity that are a function of economic or other factors. Since we are reasonably confident that our approach will not work here, today, with this data, we have tried to do no more. We do not imply more could not be done. Remember, our goal was to quantify changes in utilization as a function of changes in benefit levels—not to create a model for severity.

This situation highlights a common trap in regression analyses. Had we not looked at the dependent variable and target independent variable graphically and used a nonparametric test, it might have seemed appropriate to cobble together a model with a deceptively satisfying R^2. In fact, one can be put together. Would the model have passed an analysis of variance or would the t-statistics on the individual parameters have been significant? Perhaps. Would we have examined the mean residual error for bias or tested the residuals for normality? Hopefully.

To summarize, we found no relationship between changes in calendar year indemnity benefit levels and changes in indemnity severities. As discussed earlier in the text, we also looked at the policy year transformation of the indemnity benefit levels to confirm that the results were not a result of a poor matching between the dependent and independent variables.[11] Using policy year changes, we were able to develop models with high adjusted R^2, though they were very skew and, for the better models, the coefficients on the benefit level changes were not significantly different from zero. We also explored adding the self-insurance share index. This variable never reached statistical significance in any of the regressions.

9. CONCLUSION

We found no evidence of a benefit utilization effect for either medical costs or indemnity severity. The lack of correlation for medical costs did not surprise us. The delivery of medical benefits in the California workers compensation market has been in a state of flux for some time and will likely continue to be so in the near future. Because of this, isolating medical benefit utilization will likely be very challenging, if even possible, at present.

We were surprised to find no correlation between changes in indemnity severity, real or nominal, and changes in indemnity benefit levels. We had been conditioned by anecdotal evidence to expect a relationship. But we found none. A difference in statistical approach and rigor may be involved. We remind the reader of the importance of the visual inspection and nonparametric tests in rejecting the seemingly significant parametric findings. Also,

[11] These results were presented at the March 31, 1997 Actuarial Committee meeting of the Workers' Compensation Insurance Rating Bureau of California. They are not reproduced here but are available from the author or the Bureau.

the experimental design assumed that indirect effects could be modeled on the direct effects. Perhaps there is a relationship, but it is just too complex for a linear model. Or perhaps there was simply too much noise in California over this period of time. Our findings are, of course, temporal and local and we do not imply a relationship might not exist in the future or in other states. Nevertheless, seeing how we cannot support a severity utilization effect may be as important to our understanding as finding one, though perhaps not as gratifying.

We have developed two metrics which measure changes in hazardousness due to changes in mix of business—the indemnity frequency hazardousness index and the indemnity pure premium hazardousness index. As discussed above, although annual changes in hazardousness did not predict annual changes in indemnity frequency, this does not mean the long-term trend in hazardousness is absent from our model. This long-term trend is reflected in the constant term, and our metric has allowed us to quantify this trend. The hazardousness index may have other applications and may yet prove to be a significant variable in a model of a future, more stable economy and workers compensation system.

We have succeeded in developing a sound model of indemnity claim frequency. We can be highly confident that an indemnity frequency benefit utilization response exists and is statistically significant. This response is remarkably stable over different time periods, a variety of economic variables, and the inclusion or exclusion of a variable that captures changes in the non-benefit-related discretionary element in the workers compensation system. Our estimate of the utilization response to changes in indemnity benefit levels does not differ significantly from those of prior studies, yet the model has improved on the accuracy of the estimate and the level of confidence in the pure premium ratemaking adjustment. While there is still much to be learned, we are pleased to have made one solid step forward to a better understanding of workers compensation benefit utilization.

REFERENCES

[1] Feldblum, Sholom, "Workers Compensation Ratemaking," Casualty Actuarial Society Part 6 Study Note, September 1993.

[2] Meyer, Robert E., "A Study of Workers' Compensation Benefit Utilization," submitted to the Workers' Compensation Insurance Rating Bureau of California, October 1991.

[3] Appel, David, and David Durbin, "Impact of Economic Conditions on Workers' Compensation Benefit Utilization," Report to the Workers' Compensation Insurance Rating Bureau, August 1992.

[4] Miller, Robert B., and Dean W. Wichern, *Intermediate Business Statistics: Analysis of Variance, Regression, and Time Series*, New York: Holt, Rinehart and Winston, 1977.

[5] Ferguson, George, *Nonparametric Trend Analysis*, Montreal: McGill University Press, 1965. Available from UMI Books on Demand, Ann Arbor, Michigan.

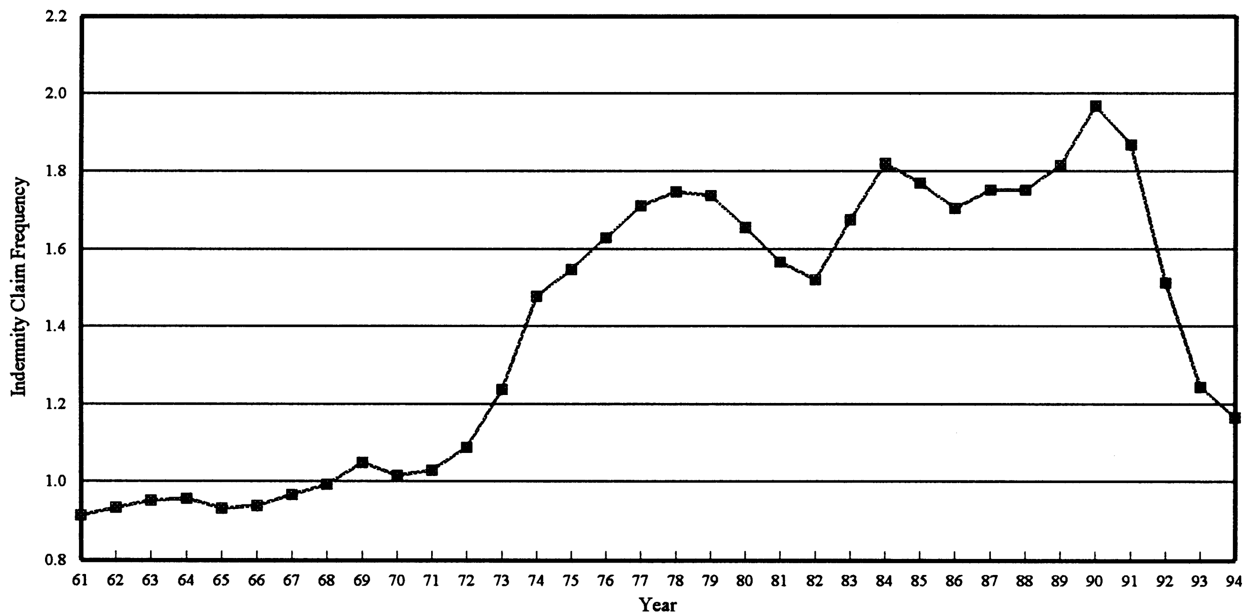
[6] Siegel, Sidney, and John N. Castellan, Jr., *Nonparametric Statistics for the Behavioral Sciences*, Second Edition, Boston: McGraw Hill, 1988.

[7] Johnson, Richard A., and Dean W. Wichern, *Applied Multivariate Statistical Analysis*, Third Edition, Englewood Cliffs, New Jersey: Prentice Hall, 1992.

EXHIBIT 1

PART 1—PAGE 1

INDEMNITY CLAIM FREQUENCY



Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 - 1994 at latest report level as of 11/12/96.
Frequency is claims per $1 Million Payroll at 1987 wage level - see Appendix A.

EXHIBIT 1

PART 1—PAGE 2

ANNUAL PERCENT CHANGE: INDEMNITY CLAIM FREQUENCY

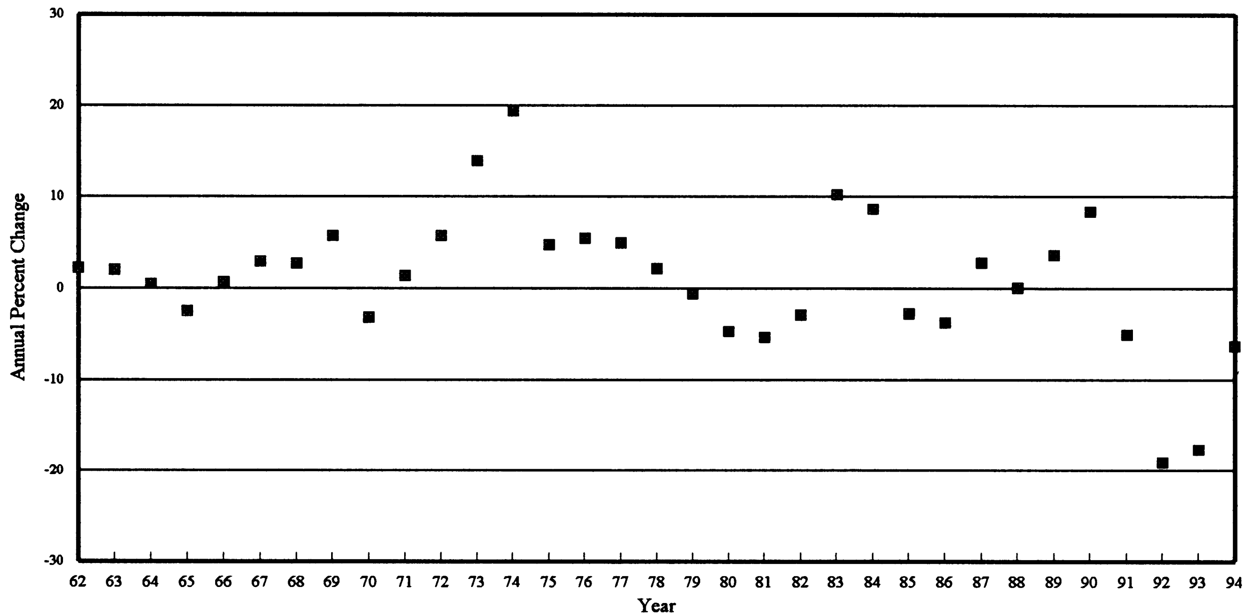
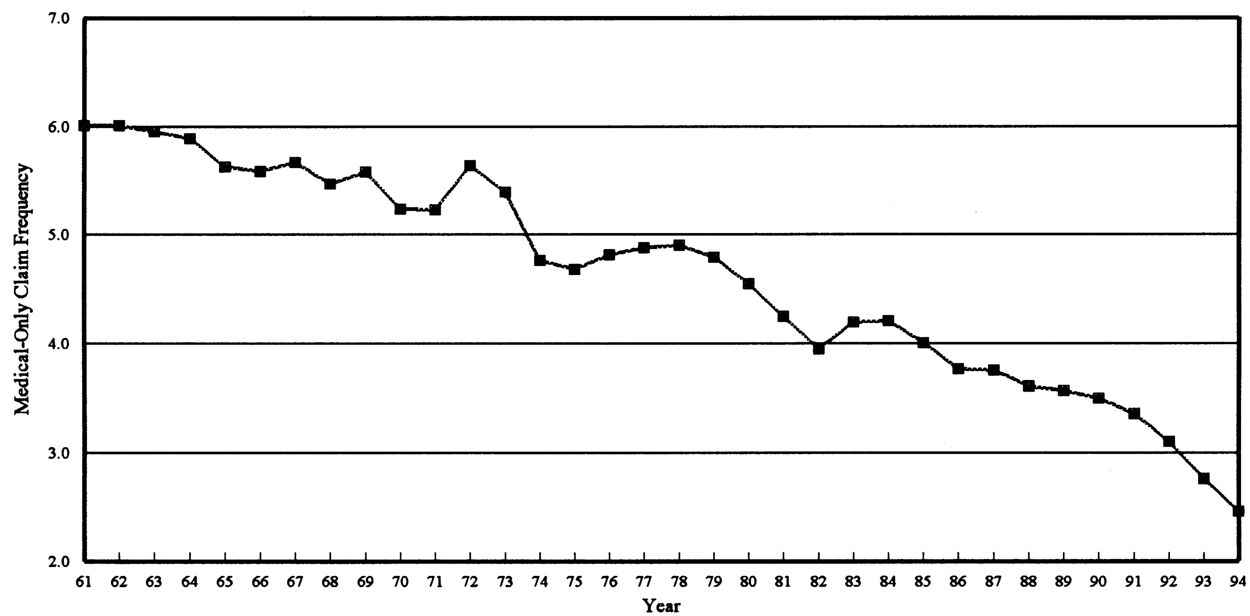


EXHIBIT 1

PART 2—PAGE 1

MEDICAL-ONLY CLAIM FREQUENCY



Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 - 1994 at latest report level as of 11/12/96.
Frequency is claims per $1 Million Payroll at 1987 wage level - see Appendix A.

EXHIBIT 1

PART 2—PAGE 2

ANNUAL PERCENT CHANGE: MEDICAL-ONLY CLAIM FREQUENCY

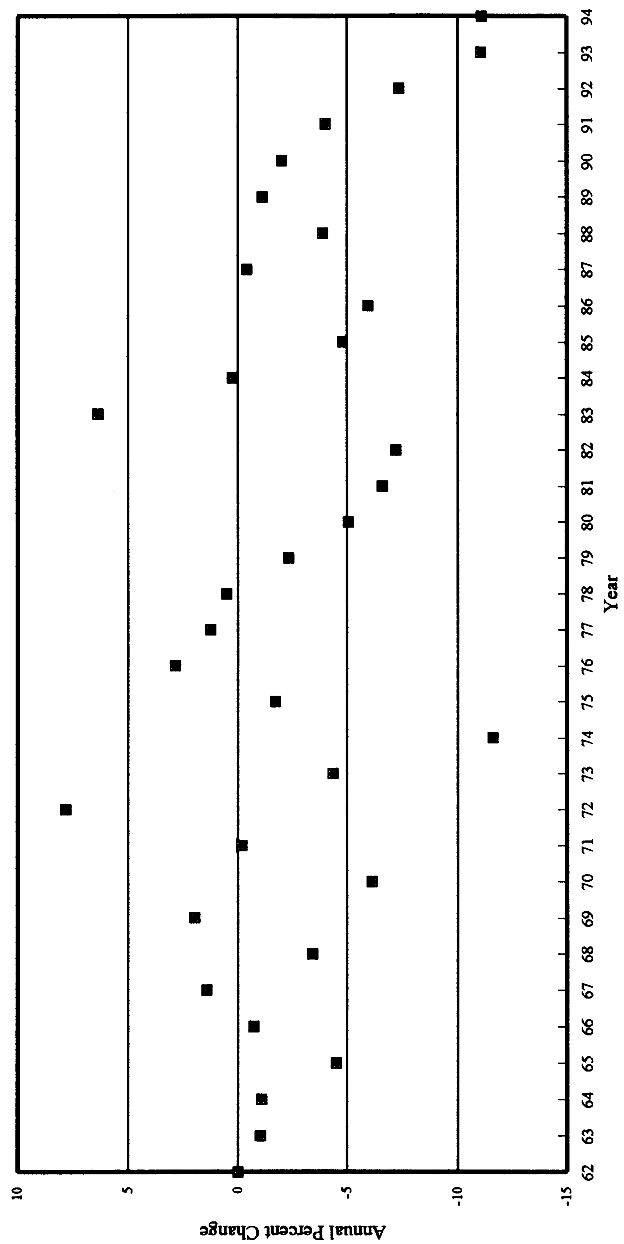
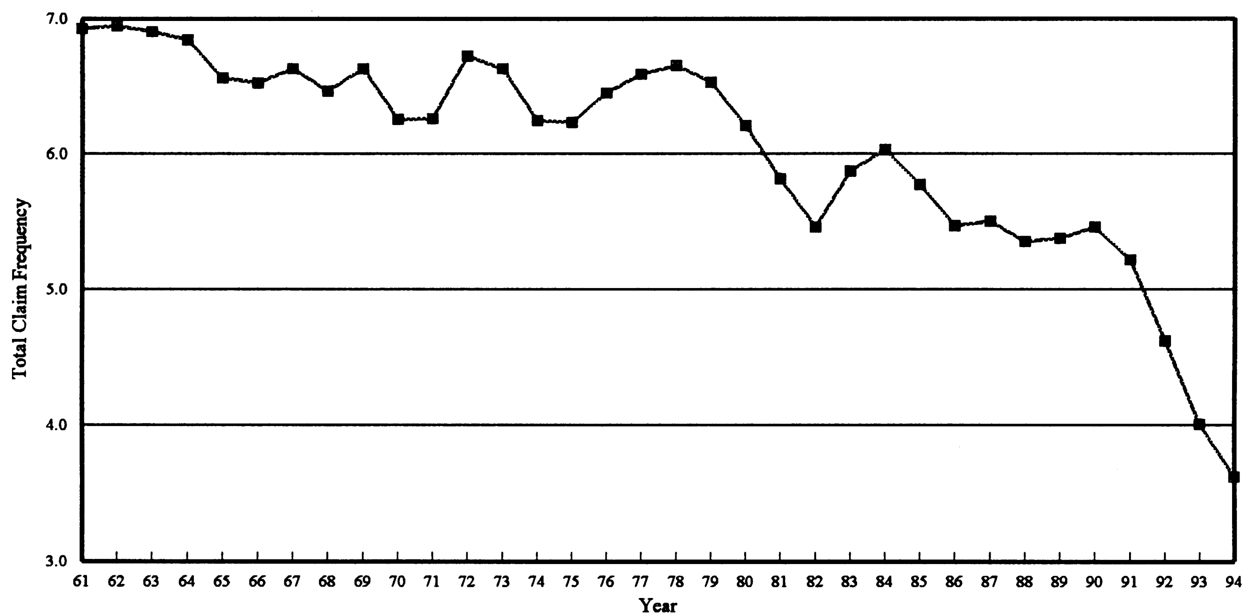


EXHIBIT 1

PART 3—PAGE 1

TOTAL CLAIM FREQUENCY



Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 - 1994 at latest report level as of 11/12/96.
Frequency is claims per $1 Million Payroll at 1987 wage level - see Appendix A.

EXHIBIT 1

PART 3—PAGE 2

ANNUAL PERCENT CHANGE: TOTAL CLAIM FREQUENCY

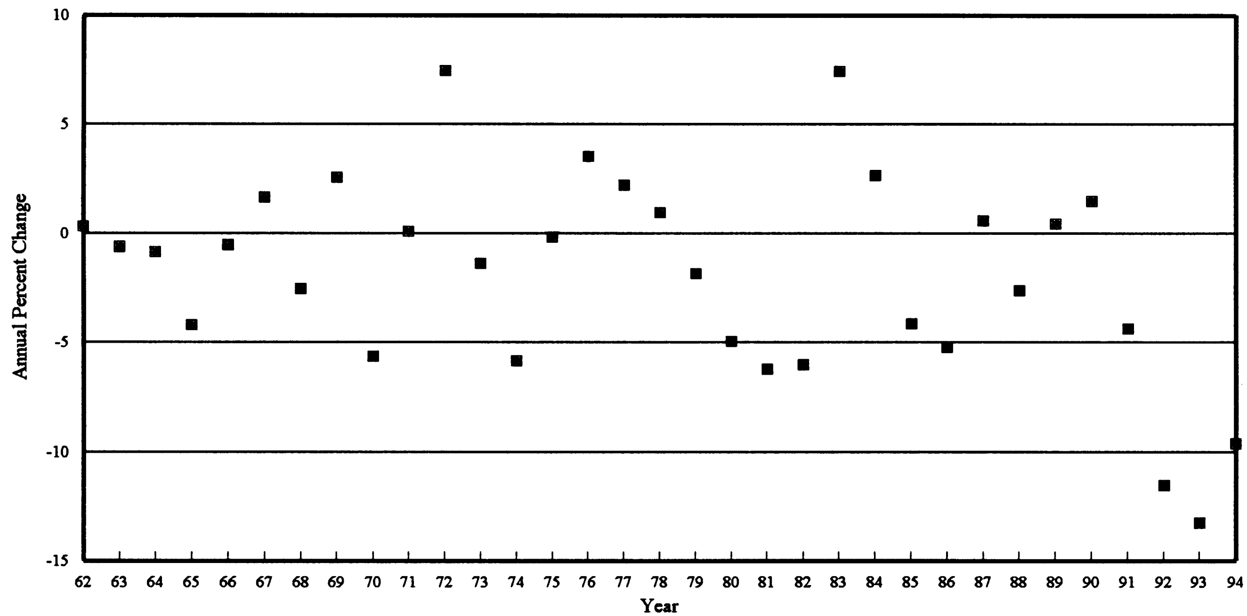
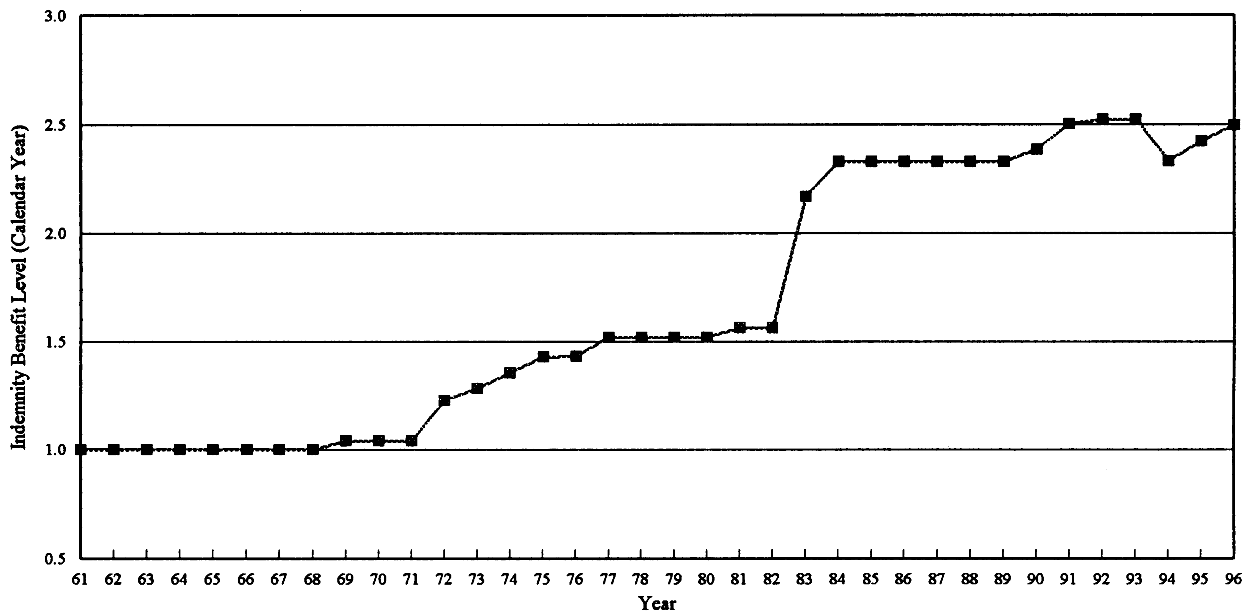


EXHIBIT 1

PART 4—PAGE 1

INDEMNITY BENEFIT LEVEL



Source: W.C.I.R.B. of California Analysis of Legislative Benefit Level Changes - see Appendix B.

EXHIBIT 1

PART 4—PAGE 2

ANNUAL PERCENT CHANGE: INDEMNITY BENEFIT LEVEL

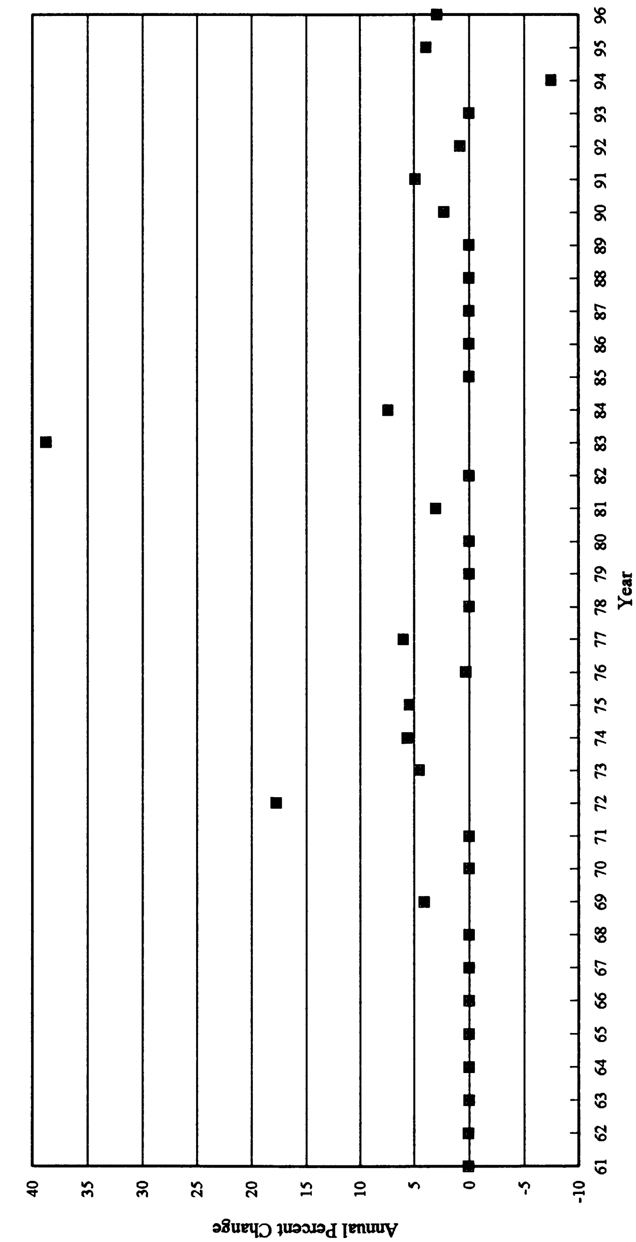
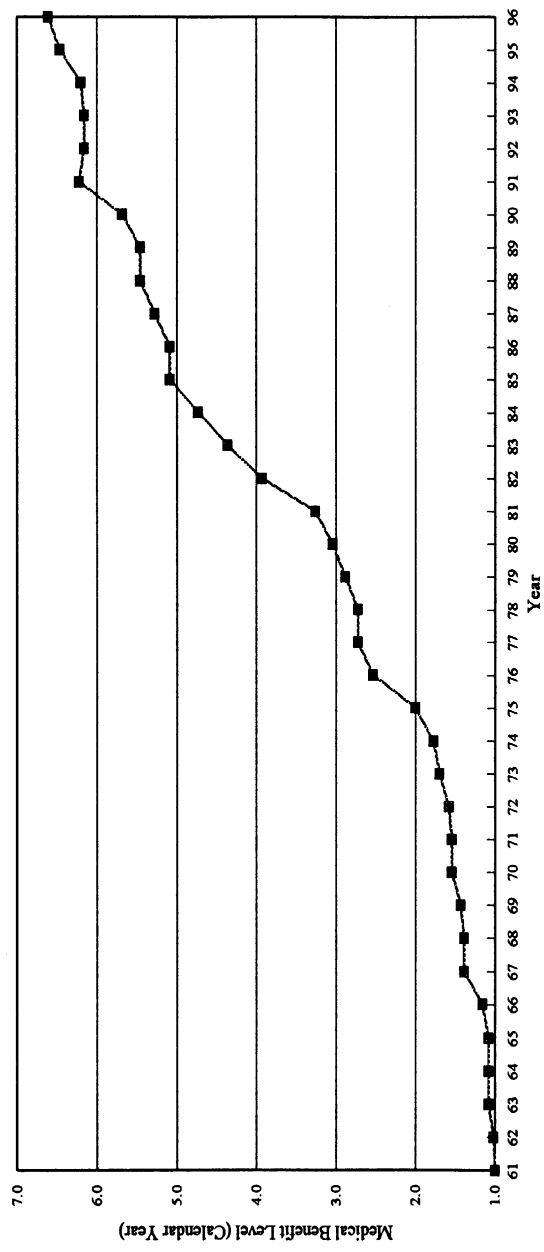


EXHIBIT 1

PART 5—PAGE 1

MEDICAL BENEFIT LEVEL



Source: W.C.I.R.B. of California Analysis of Legislative Benefit Level Changes - see Appendix B.

EXHIBIT 1

PART 5—PAGE 2

ANNUAL PERCENT CHANGE: MEDICAL BENEFIT LEVEL

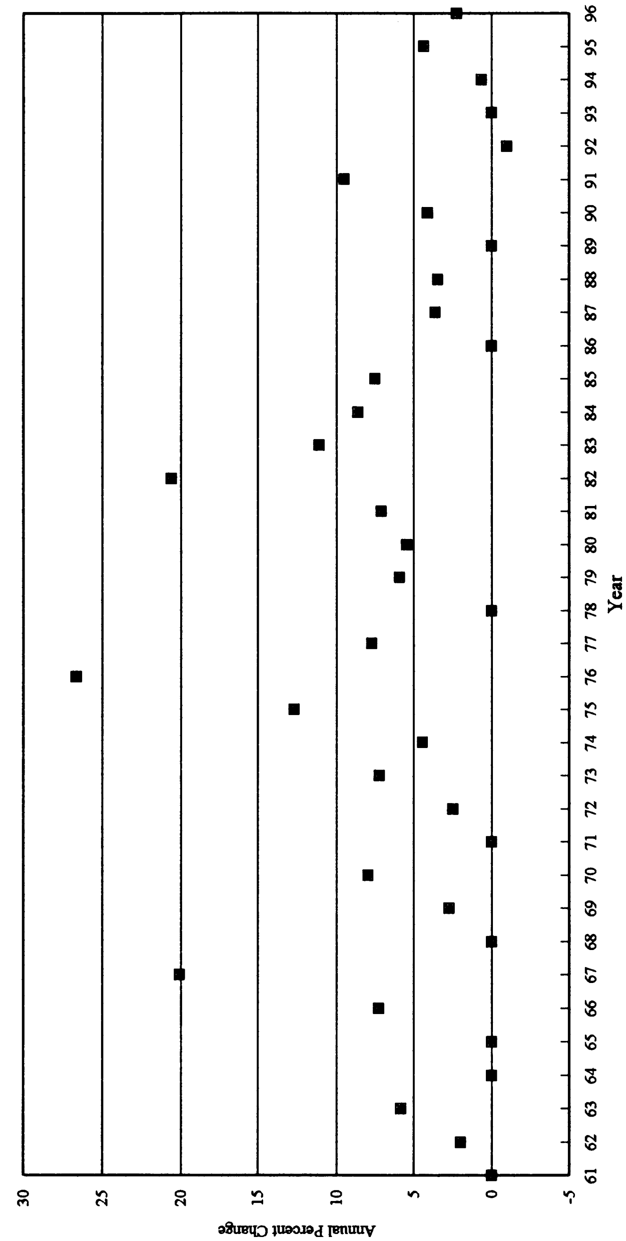
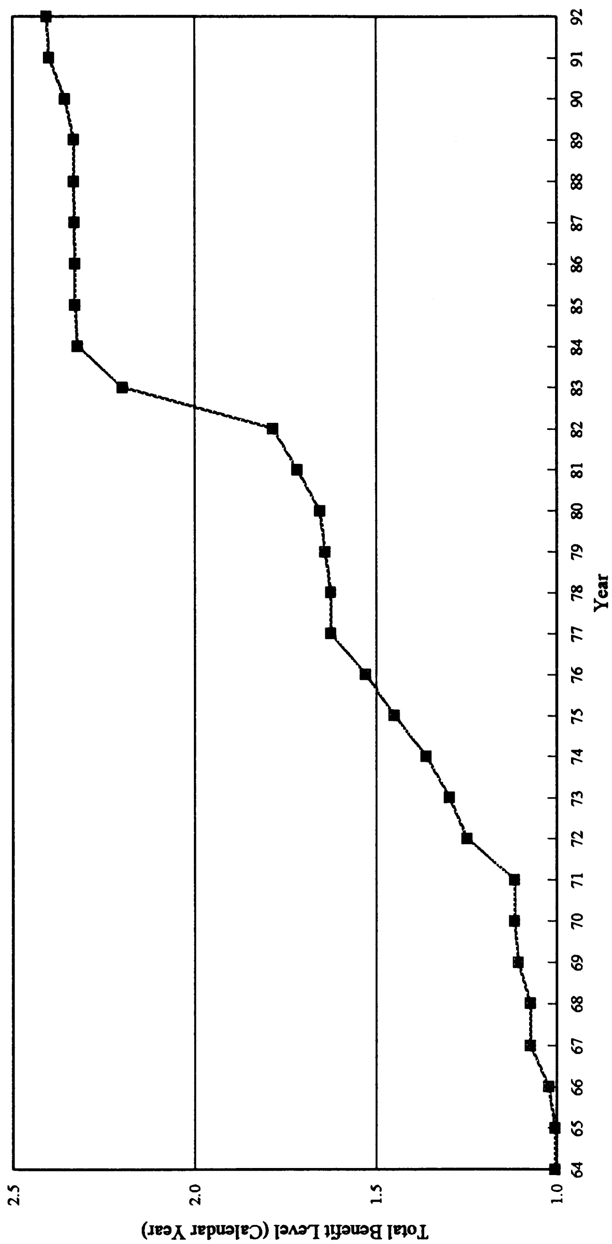


EXHIBIT 1

PART 6—PAGE 1

TOTAL BENEFIT LEVEL



Source: W.C.I.R.B. of California Analysis of Legislative Benefit Level Changes - see Appendix B.

EXHIBIT 1

PART 6—PAGE 2

ANNUAL PERCENT CHANGE: TOTAL BENEFIT LEVEL

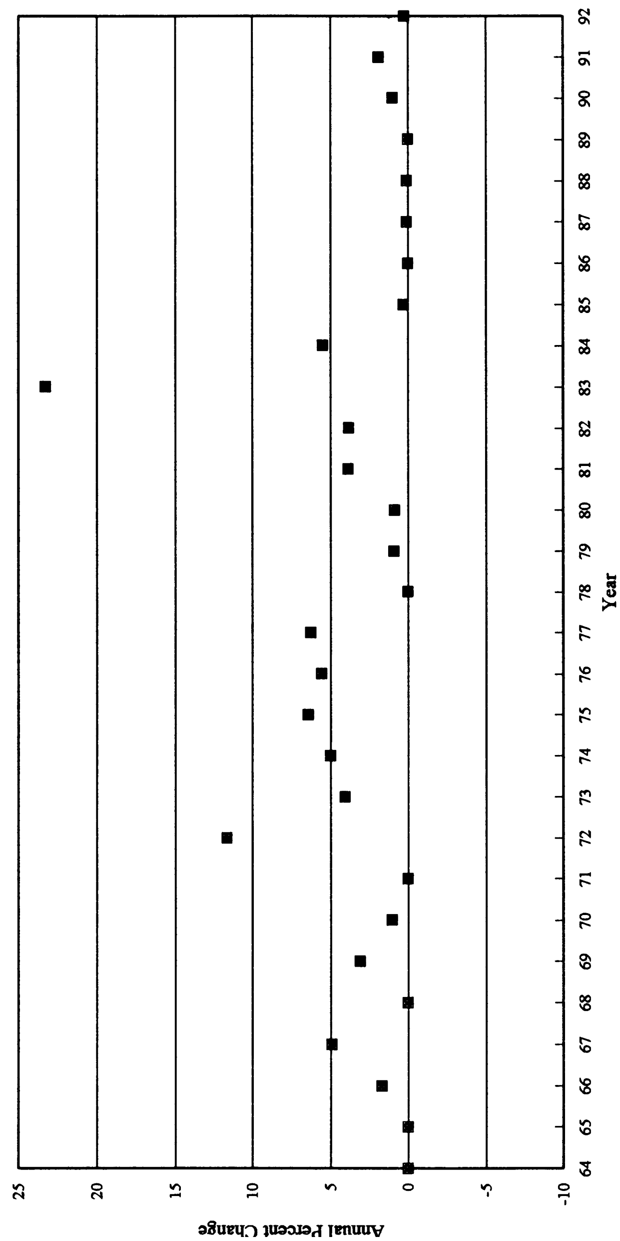
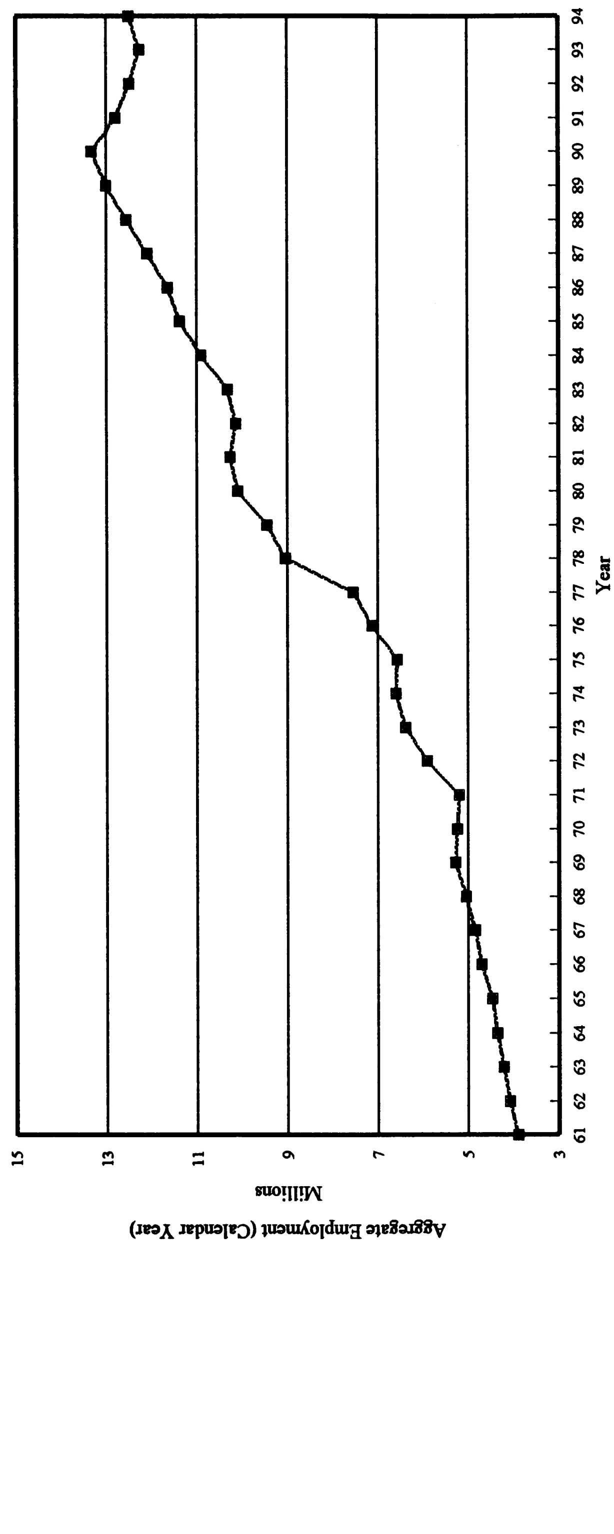


EXHIBIT 1

PART 7—PAGE 1

CALIFORNIA AGGREGATE EMPLOYMENT



Source: 1995 California Statistical Abstract - (Average Monthly Employees Covered by Unemployment Insurance)
See Appendix C, Part 1.

EXHIBIT 1

PART 7—PAGE 2

ANNUAL PERCENT CHANGE: CALIFORNIA AGGREGATE EMPLOYMENT

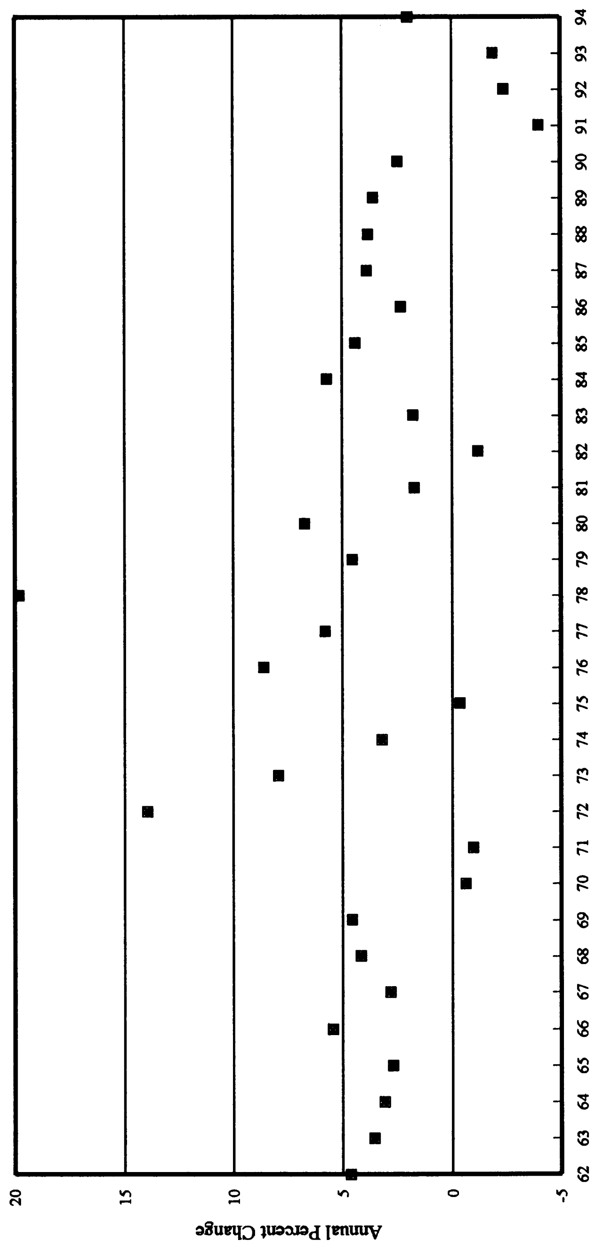
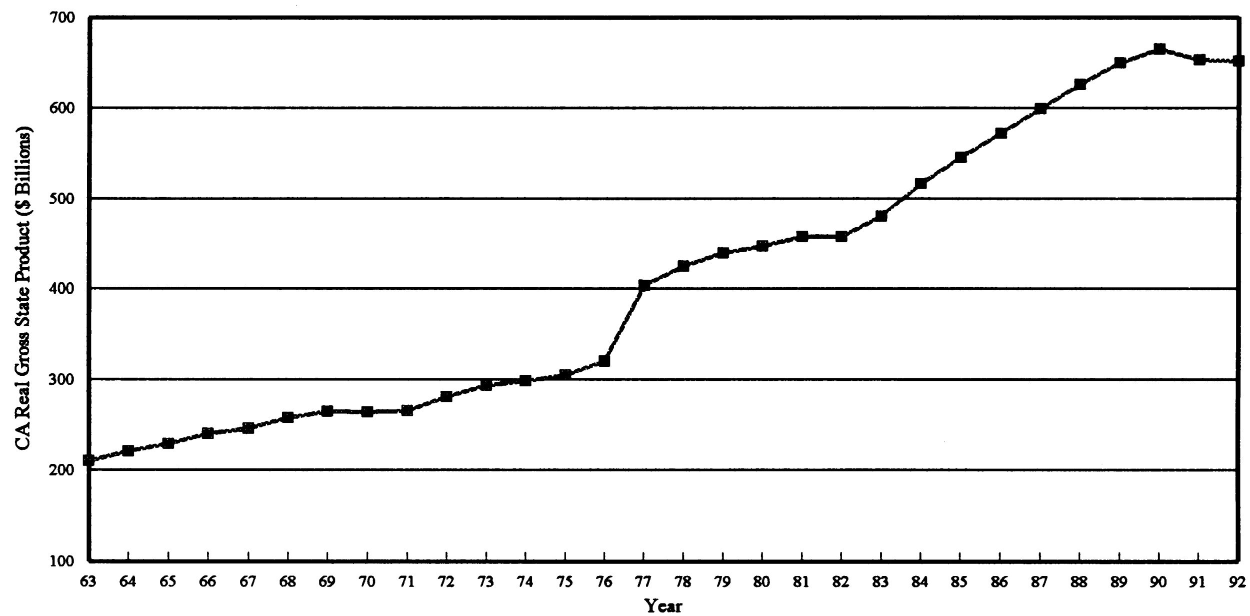


EXHIBIT 1

PART 8—PAGE 1

CALIFORNIA REAL GROSS STATE PRODUCT



Source: U.S. Department of Commerce, Bureau of Economic Analysis - (1995 California Statistical Abstract)
See Appendix C, Part 2.

EXHIBIT 1

PART 8—PAGE 2

ANNUAL PERCENT CHANGE: CALIFORNIA REAL GROSS STATE PRODUCT

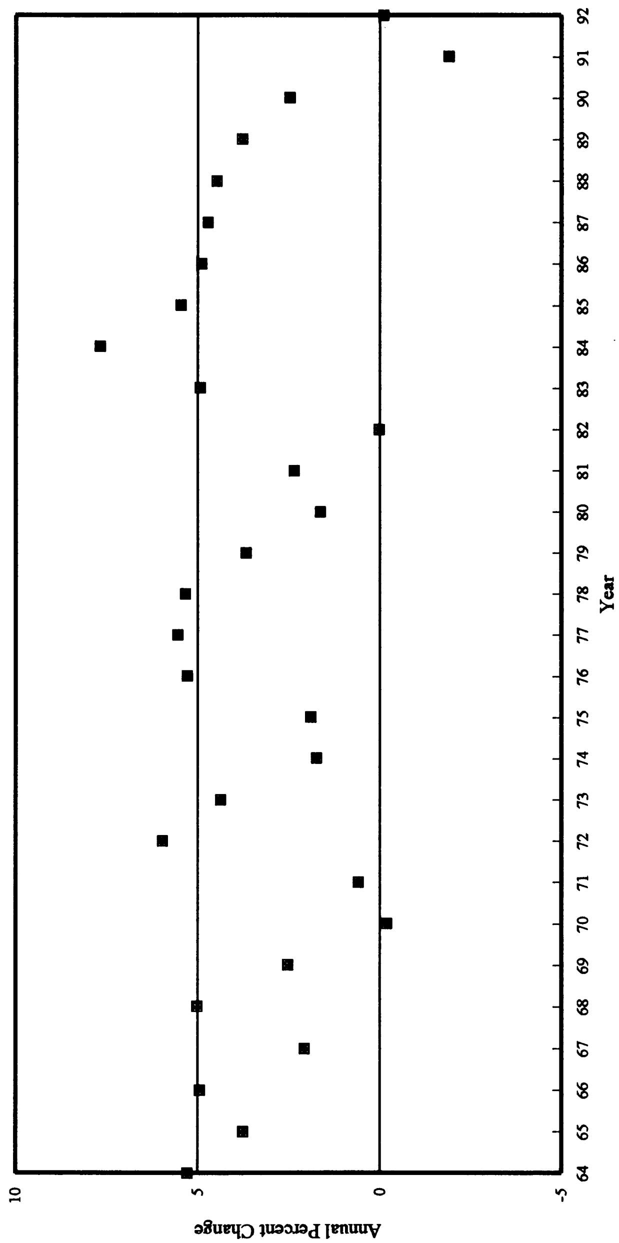
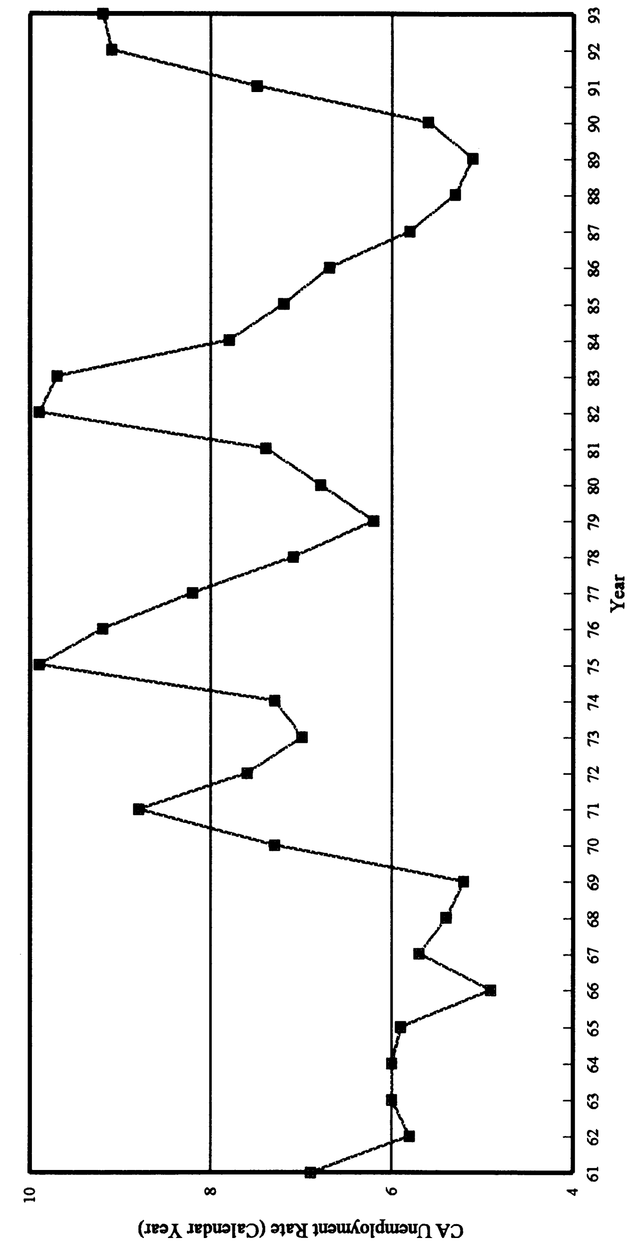


EXHIBIT 1

PART 9—PAGE 1

CALIFORNIA UNEMPLOYMENT RATE



Source: 1995 California Statistical Abstract - see Appendix C, Part 3.

EXHIBIT 1

PART 9—PAGE 2

ANNUAL PERCENT CHANGE: CALIFORNIA UNEMPLOYMENT RATE

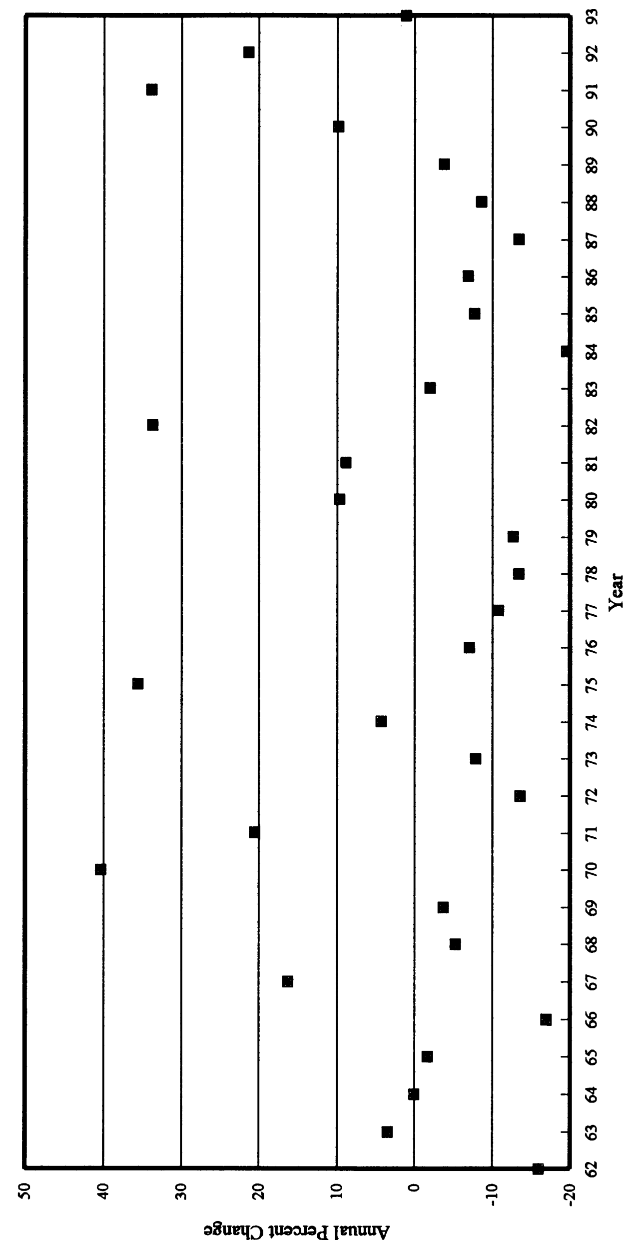
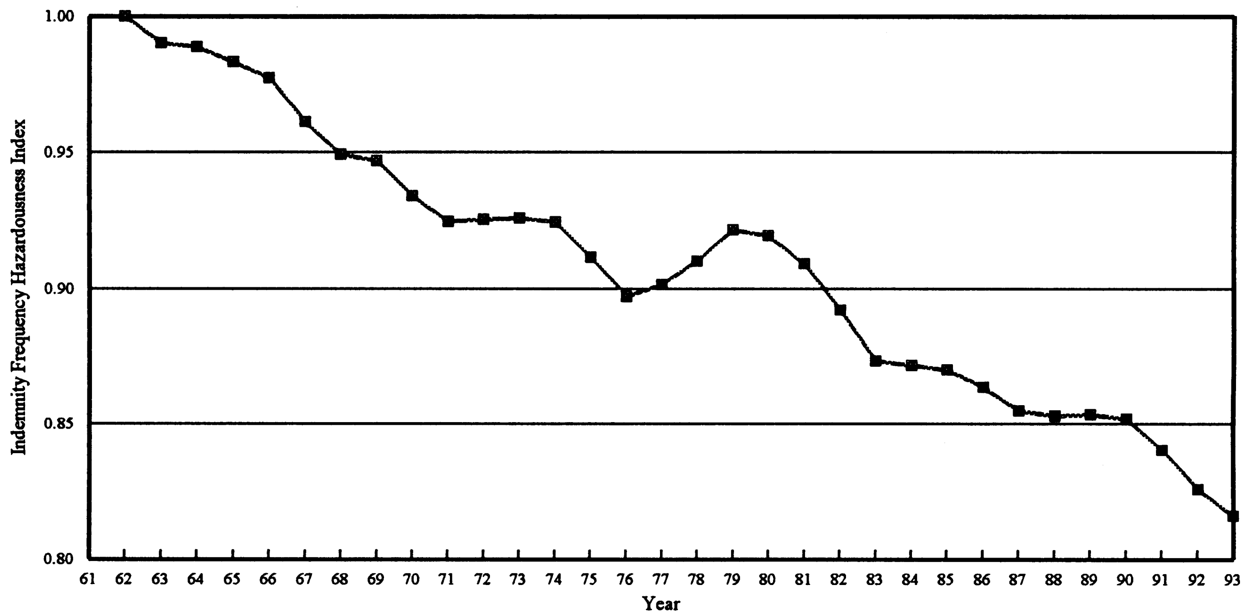


EXHIBIT 1

PART 10—PAGE 1

INDEMNITY FREQUENCY HAZARDOUSNESS INDEX



Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 at 2nd and 1993 at 1st report - see Appendix D, Part 1.

EXHIBIT 1

PART 10—PAGE 2

ANNUAL PERCENT CHANGE: INDEMNITY FREQUENCY HAZARDOUSNESS INDEX

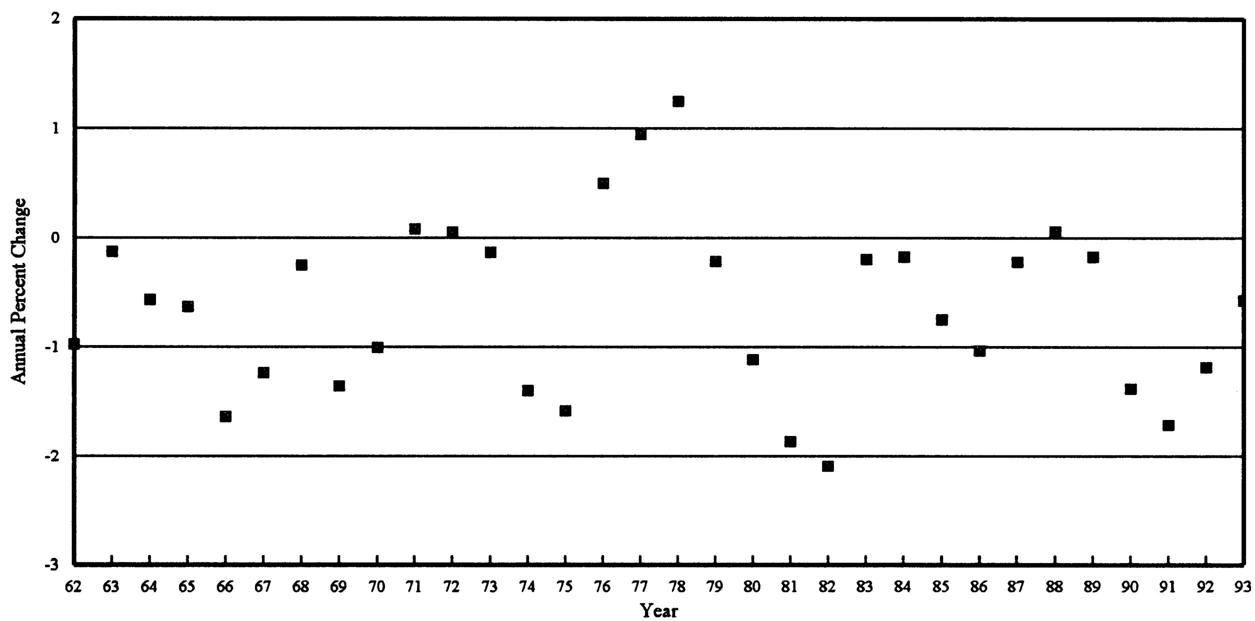
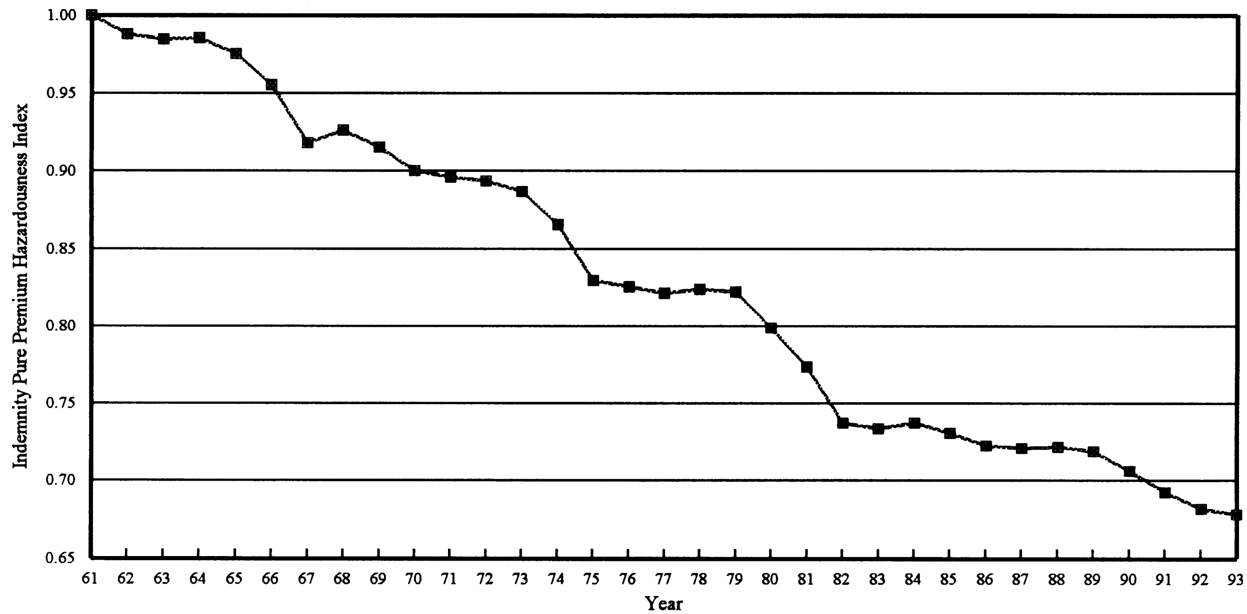


EXHIBIT 1

PART 11—PAGE 1

INDEMNITY PURE PREMIUM HAZARDOUSNESS INDEX



Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 at 2nd and 1993 at 1st report - see Appendix D, Part 2.

EXHIBIT 1

PART 11—PAGE 2

ANNUAL PERCENT CHANGE: INDEMNITY PURE PREMIUM HAZARDOUSNESS INDEX

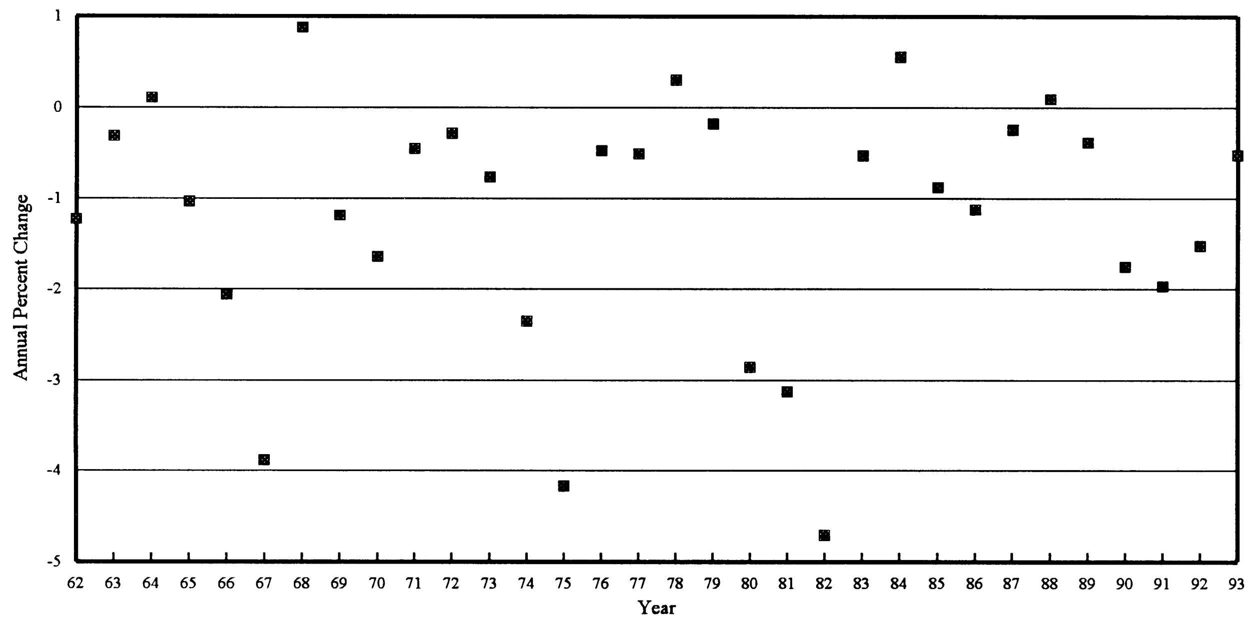
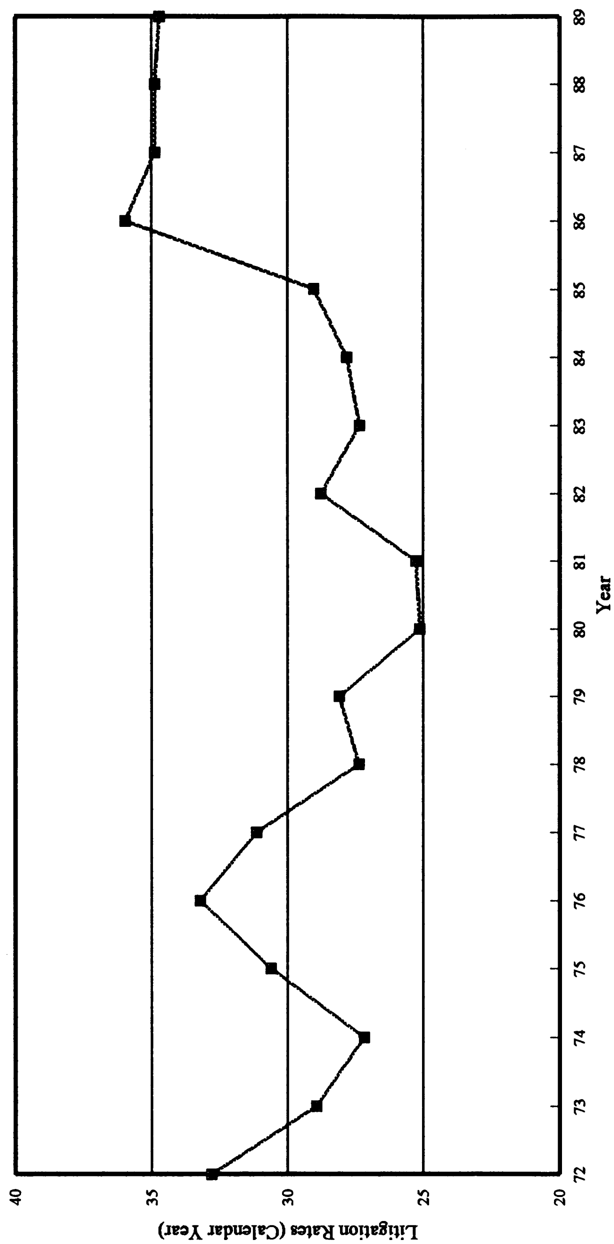


EXHIBIT 1

PART 12—PAGE 1

LITIGATION RATES



Source: C.W.C.I. Litigation Incidence Survey - see Appendix E, Part 1.

EXHIBIT 1

PART 12—PAGE 2

ANNUAL PERCENT CHANGE: LITIGATION RATES

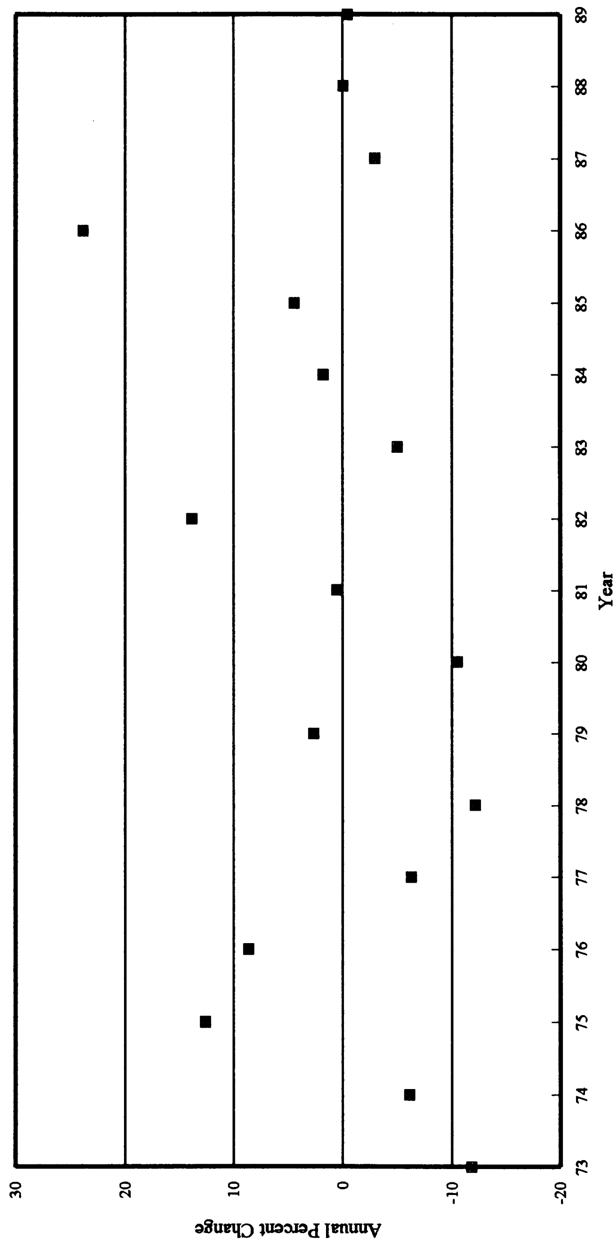
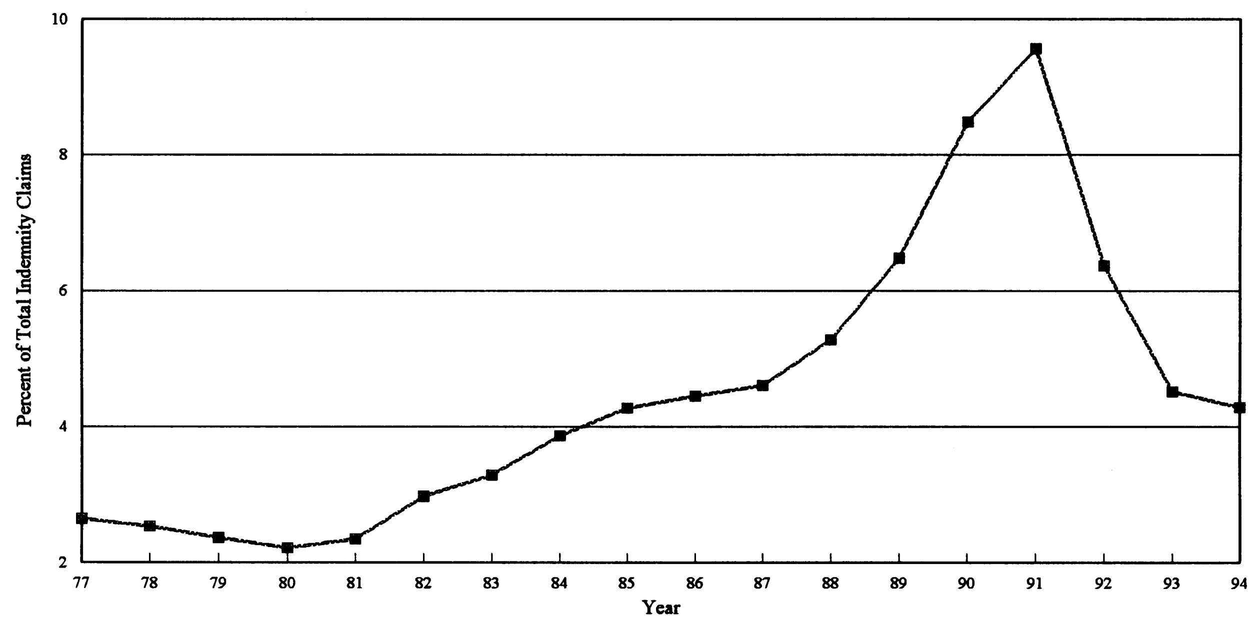


EXHIBIT 1

PART 13—PAGE 1

RATIO OF CUMULATIVE INJURIES TO TOTAL INDEMNITY CLAIMS



Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at most current report level as of 4/22/97 - see Appendix F.

EXHIBIT 1

PART 13—PAGE 2

ANNUAL PERCENT CHANGE: RATIO OF CUMULATIVE INJURIES TO TOTAL INDEMNITY CLAIM

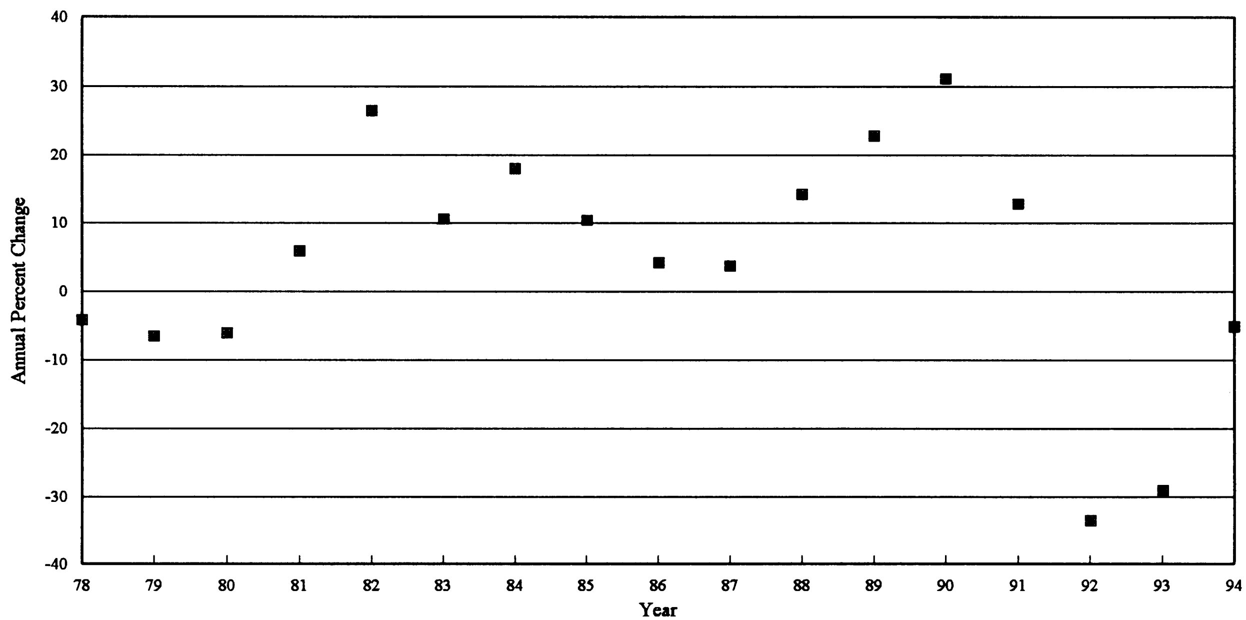
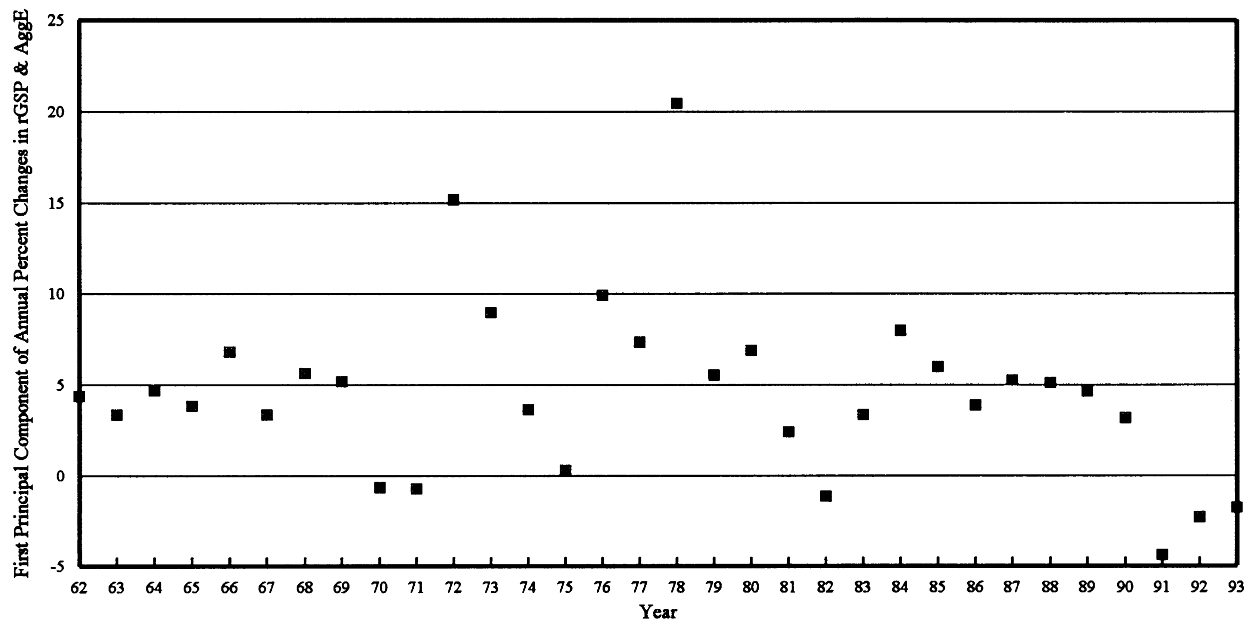


EXHIBIT 1

PART 14

FIRST PRINCIPAL COMPONENT OF ANNUAL PERCENT CHANGES IN rGSP & AggE

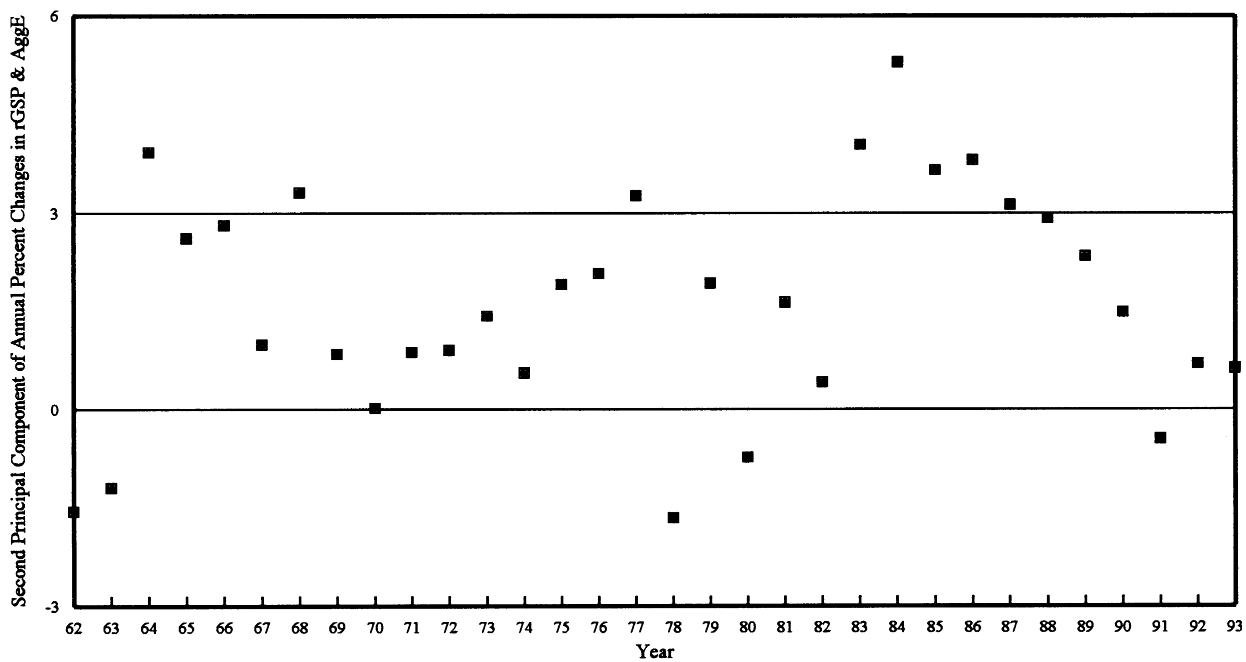


Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at most current report level as of 4/22/97 - see Appendix G, Part 1.

EXHIBIT 1

PART 15

SECOND PRINCIPAL COMPONENT OF ANNUAL PERCENT CHANGES IN rGSP & AggE

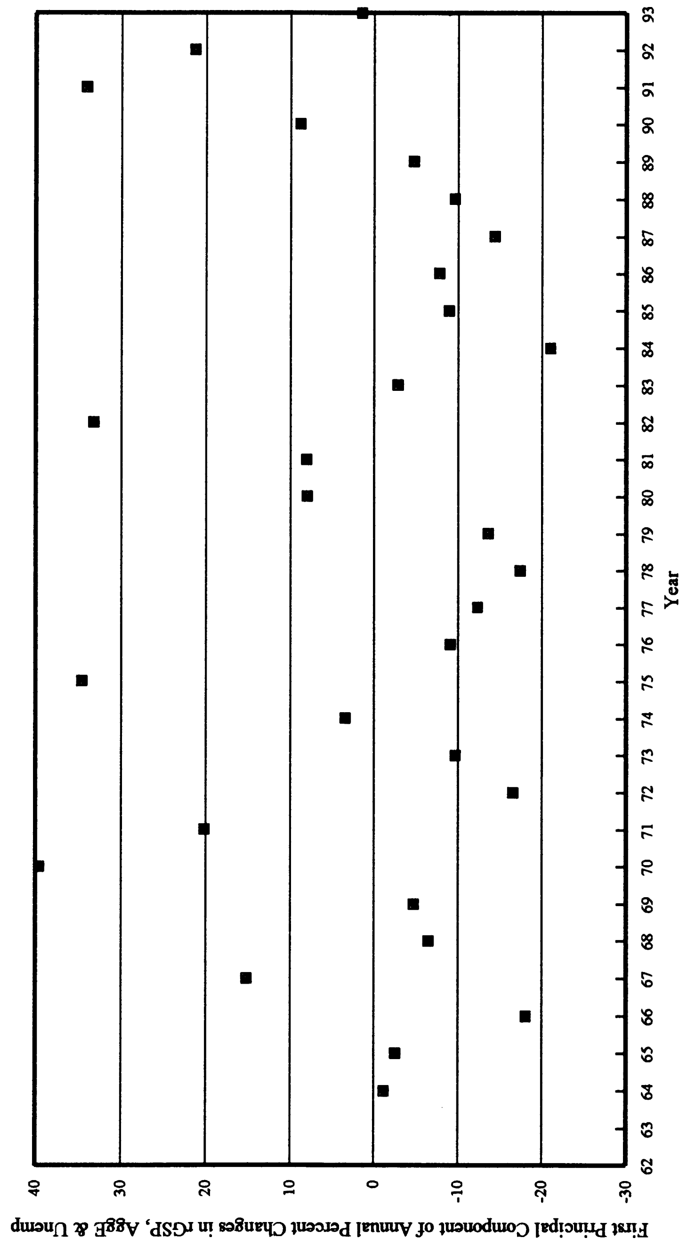


Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at most current report level as of 4/22/97 - see Appendix G, Part 1.

EXHIBIT 1

PART 16

FIRST PRINCIPAL COMPONENT OF ANNUAL PERCENT CHANGES IN rGSP, AggE & Unemp

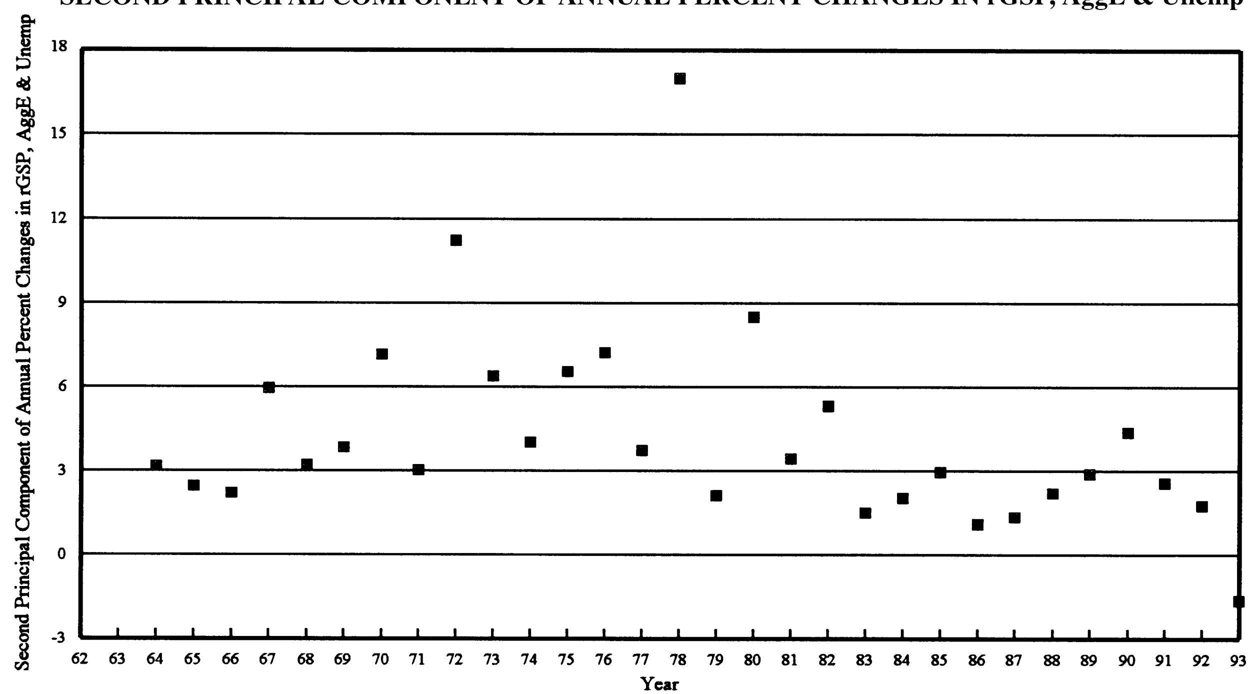


Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at most current report level as of 4/22/97 - see Appendix G, Part 2.

EXHIBIT 1

PART 17

SECOND PRINCIPAL COMPONENT OF ANNUAL PERCENT CHANGES IN rGSP, AggE & Unemp

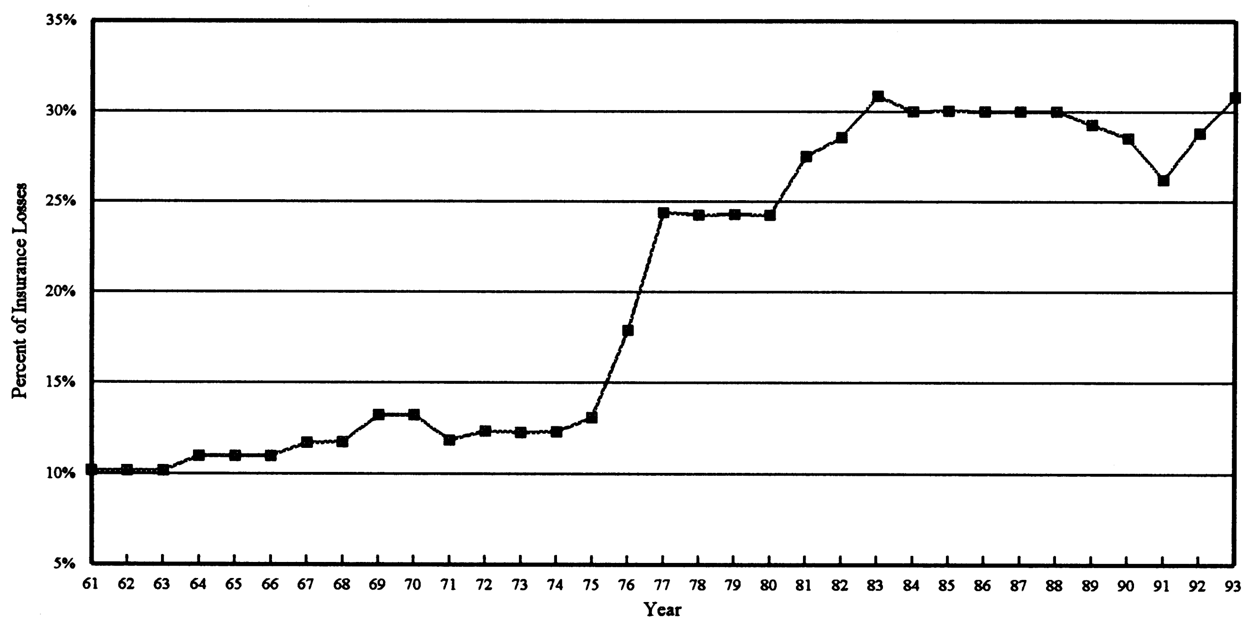


Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at most current report level as of 4/22/97 - see Appendix G, Part 2.

EXHIBIT 1

PART 18—PAGE 1

INSURANCE LOSSES PAID BY SELF-INSUREDS



Source: Social Security Bulletin - Estimates of Workers' Compensation Payments, by State and Type of Insurance.
Note: Series change in Self-Insurance Share Index in 1976. See Appendix H. Series change does not affect annual percent change presented on the following page.

EXHIBIT 1

PART 18—PAGE 2

ANNUAL PERCENT CHANGE IN SELF-INSURANCE SHARE INDEX

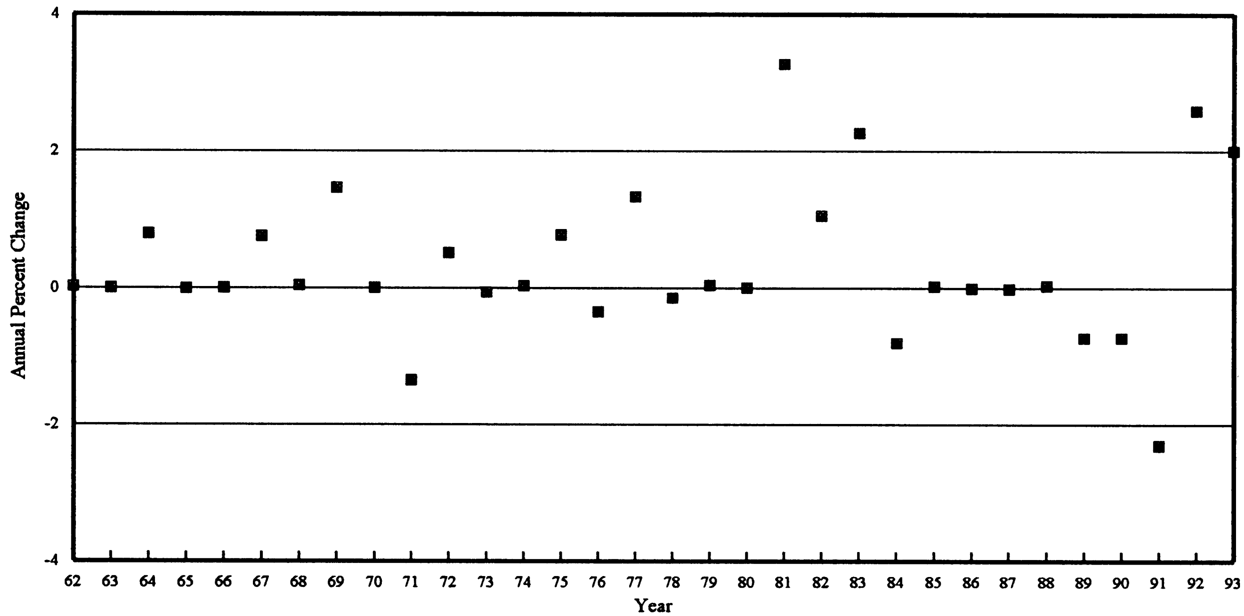


EXHIBIT 2

PART 1

CANDIDATE VARIABLES—TABULAR PRESENTATION ORIGINAL VARIABLES

	Claim Frequency per $1M Payroll (1987)			Cumulative Benefit Level			California Aggregate Emplmt	Real California GSP
Year	Indemnity	Med-Only	Total	Indemnity	Medical	Total		
1961	0.914	6.012	6.926	1.000	1.000	1.001	3,891,683	—
1962	0.934	6.013	6.947	1.001	1.020	1.004	4,071,877	—
1963	0.952	5.952	6.904	1.001	1.080	1.005	4,216,436	210,153
1964	0.956	5.889	6.845	1.001	1.080	1.005	4,346,448	220,848
1965	0.932	5.624	6.557	1.001	1.080	1.005	4,464,625	229,125
1966	0.938	5.583	6.521	1.001	1.158	1.022	4,707,406	240,495
1967	0.966	5.664	6.629	1.001	1.391	1.073	4,840,158	245,762
1968	0.991	5.470	6.461	1.001	1.391	1.073	5,041,894	257,843
1969	1.048	5.579	6.626	1.042	1.429	1.105	5,272,325	264,621
1970	1.014	5.236	6.251	1.042	1.542	1.117	5,240,190	263,933
1971	1.028	5.228	6.255	1.042	1.542	1.117	5,189,637	265,600
1972	1.086	5.636	6.722	1.227	1.581	1.247	5,913,892	281,159
1973	1.237	5.391	6.629	1.283	1.695	1.297	6,383,331	293,735
1974	1.476	4.763	6.240	1.355	1.771	1.362	6,588,356	298,408
1975	1.546	4.682	6.228	1.428	1.995	1.450	6,564,524	304,518
1976	1.630	4.816	6.446	1.433	2.527	1.530	7,130,103	320,160
1977	1.710	4.877	6.588	1.519	2.721	1.626	7,543,268	403,192
1978	1.746	4.904	6.650	1.519	2.721	1.626	9,036,931	424,809
1979	1.736	4.790	6.527	1.519	2.882	1.641	9,448,087	439,868
1980	1.654	4.548	6.203	1.519	3.040	1.655	10,083,911	447,341
1981	1.566	4.249	5.815	1.564	3.256	1.719	10,256,167	457,877
1982	1.520	3.944	5.464	1.564	3.927	1.785	10,131,806	458,036
1983	1.675	4.195	5.870	2.171	4.363	2.200	10,312,305	480,484
1984	1.820	4.206	6.025	2.332	4.738	2.321	10,900,212	517,192
1985	1.770	4.004	5.774	2.332	5.093	2.328	11,378,074	545,612
1986	1.705	3.766	5.470	2.332	5.093	2.328	11,644,237	572,257
1987	1.751	3.751	5.502	2.332	5.278	2.331	12,094,751	599,088
1988	1.752	3.605	5.357	2.332	5.460	2.333	12,556,920	626,079
1989	1.815	3.566	5.380	2.332	5.460	2.333	13,005,986	649,583
1990	1.966	3.495	5.461	2.385	5.684	2.356	13,328,057	665,298
1991	1.867	3.355	5.222	2.502	6.224	2.401	12,796,072	653,197
1992	1.511	3.108	4.619	2.522	6.162	2.407	12,490,570	652,328
1993	1.243	2.763	4.007	2.522	6.162	2.413	12,253,883	—
1994	1.165	2.456	3.621	2.334	6.202	2.186	12,500,754	—
1995	—	—	—	2.425	6.473	2.232	—	—
1996	—	—	—	2.495	6.618	2.268	—	—

Notes: The Principal Components variables are linear combinations of annual percentage changes. Series change in Self-Insurance Share Index in 1976. See Appendix H. Value given for 1976 is average of self-insured share under both series.

California Unemplmt Rate	Indemnity Frequency Haz'ness	Indemnity Pure Premium Haz'ness	Litigation Rate	Cumulative ÷ Indemnity Claims	Principal Components				Self-Insurance Share Index
					PCGA_1	PCGA_2	PCUGA_1	PCUGA_2	
6.9	1.000	1.000	—	—					0.1018
5.8	0.990	0.988	—	—					0.1019
6.0	0.989	0.985	—	—					0.1019
6.0	0.983	0.986	—	—					0.1097
5.9	0.977	0.975	—	—					0.1095
4.9	0.961	0.955	—	—					0.1095
5.7	0.949	0.918	—	—					0.1171
5.4	0.947	0.926	—	—					0.1174
5.2	0.934	0.915	—	—					0.1319
7.3	0.925	0.900	—	—					0.1319
8.8	0.925	0.896	—	—					0.1184
7.6	0.926	0.893	32.81	—					0.1235
7.0	0.925	0.886	28.93	—					0.1228
7.3	0.912	0.866	27.17	—					0.1230
9.9	0.897	0.829	30.59	—					0.1307
9.2	0.902	0.825	33.22	—					0.1788
8.2	0.910	0.821	31.15	2.6375					0.2437
7.1	0.922	0.824	27.37	2.5254		**NOT**			0.2422
6.2	0.920	0.822	28.09	2.3588		**APPLICABLE**			0.2426
6.8	0.909	0.799	25.12	2.2147					0.2425
7.4	0.892	0.774	25.26	2.3447					0.2752
9.9	0.874	0.737	28.76	2.9650					0.2857
9.7	0.872	0.733	27.33	3.2780					0.3083
7.8	0.870	0.737	27.81	3.8679					0.3002
7.2	0.864	0.731	29.03	4.2713					0.3003
6.7	0.855	0.723	35.94	4.4500					0.3001
5.8	0.853	0.721	34.87	4.6127					0.2999
5.3	0.854	0.721	34.86	5.2685					0.3002
5.1	0.852	0.719	34.70	6.4725					0.2927
5.6	0.840	0.706	—	8.4853					0.2853
7.5	0.826	0.692	38.20	9.5761					0.2621
9.1	0.816	0.682	42.85	6.3594					0.2880
9.2	0.811	0.678	—	4.5128					0.3080
—	—	—	—	4.2813					—
—	—	—	—	—					—
—	—	—	—	—					—

EXHIBIT 2

PART 2

CANDIDATE VARIABLES—TABULAR PRESENTATION ANNUAL PERCENT CHANGES

	Claim Frequency per $1M Payroll (1987)			Benefit Level			California Aggregate Emplmt	Real California GSP
Year	Indemnity	Med-Only	Total	Indemnity	Medical	Total		
1961	—	—	—	0.032	0.000	0.092	—	—
1962	2.185	0.015	0.301	0.075	2.012	0.272	4.630	—
1963	1.956	−1.013	−0.614	0.000	5.852	0.158	3.550	—
1964	0.426	−1.063	−0.857	0.000	0.000	0.000	3.083	5.275
1965	−2.484	−4.491	−4.211	0.000	0.000	0.000	2.719	3.759
1966	0.623	−0.731	−0.538	0.000	7.259	1.689	5.438	4.942
1967	2.920	1.441	1.654	0.000	20.083	4.928	2.820	2.069
1968	2.675	−3.423	−2.535	0.000	0.000	0.000	4.168	5.011
1969	5.693	1.992	2.559	4.100	2.747	3.062	4.570	2.533
1970	−3.193	−6.135	−5.670	0.000	7.935	1.043	−0.610	−0.194
1971	1.321	−0.169	0.073	0.000	0.000	0.000	−0.965	0.590
1972	5.659	7.811	7.458	17.732	2.489	11.635	13.956	5.957
1973	13.951	−4.342	−1.387	4.553	7.232	4.049	7.938	4.358
1974	19.315	−11.647	−5.867	5.623	4.461	4.974	3.212	1.747
1975	4.700	−1.707	−0.191	5.418	12.673	6.438	−0.362	1.897
1976	5.430	2.870	3.505	0.300	26.650	5.560	8.616	5.283
1977	4.951	1.263	2.196	6.000	7.702	6.268	5.795	5.545
1978	2.089	0.545	0.946	0.000	0.000	0.000	19.801	5.322
1979	−0.566	−2.309	−1.851	0.000	5.898	0.907	4.550	3.672
1980	−4.716	−5.051	−4.962	0.000	5.479	0.885	6.730	1.633
1981	−5.362	−6.577	−6.253	3.000	7.119	3.842	1.708	2.354
1982	−2.918	−7.188	−6.038	0.000	20.599	3.812	−1.213	0.013
1983	10.225	6.365	7.439	38.800	11.100	23.300	1.782	4.931
1984	8.611	0.257	2.641	7.400	8.600	5.500	5.701	7.665
1985	−2.723	−4.789	−4.165	0.000	7.500	0.300	4.384	5.453
1986	−3.706	−5.957	−5.267	0.000	0.000	0.000	2.339	4.887
1987	2.739	−0.395	0.582	0.000	3.630	0.101	3.869	4.711
1988	0.024	−3.883	−2.639	0.000	3.445	0.099	3.821	4.476
1989	3.597	−1.098	0.438	0.000	0.000	0.000	3.576	3.770
1990	8.353	−1.995	1.495	2.300	4.100	1.000	2.476	2.475
1991	−5.055	−4.003	−4.382	4.900	9.500	1.900	−3.991	−1.900
1992	−19.058	−7.352	−11.537	0.800	−1.000	0.251	−2.387	−0.106
1993	−17.720	−11.090	−13.259	0.000	0.000	0.248	−1.895	—
1994	−6.330	−11.124	−9.637	−7.469	0.646	−9.428	2.015	—
1995	—	—	—	3.919	4.374	2.141	—	—
1996	—	—	—	2.894	2.242	1.596	—	—

Notes: The Principal Components variables are linear combinations of annual percentage changes.
PCGA_1(2) = First (second) principal component of CA Real GSP and Aggregate Employment
PCUGA_1(2) = First (second) principal component of CA Real GSP, Unemployment Rate, and Aggregate Employment
Series change in Self-Insurance Share Index in 1976. See Appendix H. Series change does not affect annual percent change.

California Unemplmt Rate	Indemnity Frequency Haz'ness	Indemnity Pure Premium Haz'ness	Litigation Rate	Cumulative ÷ Indemnity Claims	Principal Components PCGA_1	PCGA_2	PCUGA_1	PCUGA_2	Self-Insurance Share Index
—	—	—	—	—	—	—	—	—	0.002
−15.942	−0.974	−1.233	—	—	4.360	−1.560	—	—	0.009
3.448	−0.128	−0.318	—	—	3.343	−1.196	—	—	−0.000
0.000	−0.565	0.107	—	—	4.680	3.928	−1.188	3.164	0.787
−1.667	−0.629	−1.040	—	—	3.826	2.624	−2.571	2.446	−0.020
−16.949	−1.642	−2.062	—	—	6.785	2.822	−18.124	2.203	−0.000
16.327	−1.237	−3.883	—	—	3.352	0.998	15.155	5.962	0.755
−5.263	−0.249	0.875	—	—	5.612	3.314	−6.495	3.208	0.032
−3.704	−1.361	−1.188	—	—	5.157	0.846	−4.765	3.837	1.455
40.385	−1.003	−1.650	—	—	−0.639	0.023	39.526	7.167	−0.000
20.548	0.082	−0.456	—	—	−0.709	0.881	20.155	3.023	−1.356
−13.636	0.051	−0.287	—	—	15.147	0.907	−16.613	11.223	0.513
−7.895	−0.137	−0.772	−11.834	—	8.942	1.429	−9.696	6.383	−0.071
4.286	−1.398	−2.363	−6.085	—	3.613	0.563	3.374	4.021	0.019
35.616	−1.589	−4.173	12.615	—	0.299	1.908	34.588	6.548	0.769
−7.071	0.500	−0.474	8.598	—	9.892	2.072	−9.127	7.231	−0.358
−10.870	0.949	−0.513	−6.250	—	7.324	3.268	−12.331	3.739	1.331
−13.415	1.249	0.302	−12.133	−4.248	20.437	−1.660	−17.424	16.983	−0.151
−12.676	−0.216	−0.182	2.621	−6.598	5.521	1.925	−13.643	2.121	0.034
9.677	−1.116	−2.861	−10.557	−6.108	6.886	−0.730	7.984	8.507	−0.003
8.824	−1.866	−3.130	0.537	5.871	2.401	1.641	8.013	3.436	3.271
33.784	−2.097	−4.711	13.867	26.453	−1.137	0.421	33.177	5.311	1.047
−2.020	−0.201	−0.534	−4.964	10.556	3.338	4.042	−2.874	1.489	2.261
−19.588	−0.173	0.549	1.773	17.998	7.950	5.297	−21.061	2.026	−0.816
−7.692	−0.748	−0.882	4.382	10.430	5.965	3.658	−8.956	2.964	0.015
−6.944	−1.033	−1.126	23.803	4.183	3.849	3.814	−7.777	1.087	−0.017
−13.433	−0.225	−0.248	−2.974	3.655	5.230	3.132	−14.373	1.335	−0.021
−8.621	0.057	0.090	−0.032	14.219	5.106	2.927	−9.643	2.208	0.026
−3.774	−0.179	−0.384	−0.456	22.853	4.637	2.344	−4.788	2.881	−0.746
9.804	−1.382	−1.756	—	31.096	3.165	1.496	8.811	4.381	−0.746
33.929	−1.718	−1.977	—	12.856	−4.398	−0.444	34.062	2.563	−2.312
21.333	−1.185	−1.526	12.183	−33.591	−2.284	0.704	21.269	1.760	2.590
1.099	−0.572	−0.527	—	−29.038	−1.784	0.638	1.428	−1.647	1.999
—	—	—	—	−5.130	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—

EXHIBIT 3

PART 1

CORRELATIONS AMONG VARIABLES
SAMPLE PERIOD: 1964–1992
PEARSON PRODUCT MOMENT CORRELATION AT LAG = 0

	Claim Frequency per $1M Payroll (1987)			Benefit Level			California Aggregate Emplmt	Real California GSP
	Indemnity	Med-Only	Total	Indemnity	Medical	Total		
Indemnity Claim Frequency	1.000	0.298	0.615	0.385	0.158	0.437	0.343	0.392
Med-Only Claim Frequency	0.298	1.000	0.928	0.521	0.155	0.544	0.445	0.490
Total Claim Frequency	0.615	0.928	1.000	0.552	0.195	0.588	0.484	0.559
Indemnity Benefit Level	0.385	0.521	0.552	1.000	0.110	0.945	0.060	0.204
Medical Benefit Level	0.158	0.155	0.195	0.110	1.000	0.384	−0.102	−0.113
Total Benefit Level	0.437	0.544	0.588	0.945	0.384	1.000	0.082	0.199
California Aggregate Employment	0.343	0.445	0.484	0.060	−0.102	0.082	1.000	0.655
Real California Gross State Product	0.392	0.490	0.559	0.204	−0.113	0.199	0.655	1.000
California Unemployment Rate	−0.347	−0.389	−0.448	−0.110	0.267	−0.059	−0.677	−0.892
Indemnity Frequency Haz'ness	0.260	0.510	0.502	0.127	−0.176	0.093	0.643	0.617
Indemnity Pure Premium Haz'ness	0.169	0.370	0.356	0.105	−0.493	−0.067	0.431	0.638
Litigation Rates	−0.390	−0.155	−0.239	−0.197	0.325	−0.109	−0.543	−0.122
Cumulative ÷ Indemnity Claims	0.690	0.219	0.483	0.112	0.466	0.153	−0.110	0.153
1st PC (rGSP, AggE)	0.367	0.472	0.518	0.085	−0.108	0.104	0.993	0.739
2nd PC (rGSP, AggE)	0.179	0.210	0.262	0.210	−0.049	0.182	−0.116	0.674
1st PC (rGSP, AggE, Unemp)	−0.353	−0.400	−0.459	−0.110	0.261	−0.063	−0.705	−0.897
2nd PC (rGSP, AggE, Unemp)	0.136	0.234	0.231	−0.022	0.119	0.058	0.710	0.040
Self-Insurance Share Index	−0.210	0.014	−0.099	0.317	0.061	0.388	−0.063	0.024

Note: Pearson Product Moment Correlation assumes the variables to be normally distributed.

California Unemplmt Rate	Indemnity Frequency Haz'ness	Indemnity Pure Premium Haz'ness	Litigation Rate	Cumulative ÷ Indemnity Claims	Principal Components				Self-Insurance Share Index
					PCGA_1	PCGA_2	PCUGA_1	PCUGA_2	
−0.347	0.260	0.169	−0.390	0.690	0.367	0.179	−0.353	0.136	−0.210
−0.389	0.510	0.370	−0.155	0.219	0.472	0.210	−0.400	0.234	0.014
−0.448	0.502	0.356	−0.239	0.483	0.518	0.262	−0.459	0.231	−0.099
−0.110	0.127	0.105	−0.197	0.112	0.085	0.210	−0.110	−0.022	0.317
0.267	−0.176	−0.493	0.325	0.466	−0.108	−0.049	0.261	0.119	0.061
−0.059	0.093	−0.067	−0.109	0.153	0.104	0.182	−0.063	0.058	0.388
−0.677	0.643	0.431	−0.543	−0.110	0.993	−0.116	−0.705	0.710	−0.063
−0.892	0.617	0.638	−0.122	0.153	0.739	0.674	−0.897	0.040	0.024
1.000	−0.587	−0.683	0.353	0.025	−0.741	−0.511	0.999	0.038	0.027
−0.587	1.000	0.781	−0.448	−0.156	0.668	0.183	−0.600	0.312	−0.182
−0.683	0.781	1.000	−0.327	−0.116	0.483	0.417	−0.681	−0.068	−0.265
0.353	−0.448	−0.327	1.000	0.374	−0.522	0.357	0.369	−0.368	−0.002
0.025	−0.156	−0.116	0.374	1.000	−0.076	0.286	0.028	−0.119	−0.407
−0.741	0.668	0.483	−0.522	−0.076	1.000	0.000	−0.767	0.639	−0.052
−0.511	0.183	0.417	0.357	0.286	0.000	1.000	−0.490	−0.642	0.094
0.999	−0.600	−0.681	0.369	0.028	−0.767	−0.490	1.000	−0.000	0.029
0.038	0.312	−0.068	−0.368	−0.119	0.639	−0.642	−0.000	1.000	−0.058
0.027	−0.182	−0.265	−0.002	−0.407	−0.052	0.094	0.029	−0.058	1.000

EXHIBIT 3

PART 2

CORRELATIONS AMONG VARIABLES
SAMPLE PERIOD: 1964–1992
SIGNIFICANCE OF CORRELATION AT LAG = 0

	Claim Frequency per $1M Payroll (1987)			Benefit Level			California Aggregate Emplmt	Real California GSP
	Indemnity	Med-Only	Total	Indemnity	Medical	Total		
Indemnity Claim Frequency		0.116	0.000	0.039	0.413	0.018	0.068	0.036
Med-Only Claim Frequency	0.116		0.000	0.004	0.422	0.002	0.016	0.007
Total Claim Frequency	0.000	0.000		0.002	0.310	0.001	0.008	0.002
Indemnity Benefit Level	0.039	0.004	0.002		0.569	0.000	0.758	0.288
Medical Benefit Level	0.413	0.422	0.310	0.569		0.040	0.599	0.560
Total Benefit Level	0.018	0.002	0.001	0.000	0.040		0.673	0.300
California Aggregate Employment	0.068	0.016	0.008	0.758	0.599	0.673		0.000
Real California Gross State Product	0.036	0.007	0.002	0.288	0.560	0.300	0.000	
California Unemployment Rate	0.065	0.037	0.015	0.572	0.162	0.763	0.000	0.000
Indemnity Frequency Haz'ness	0.174	0.005	0.006	0.513	0.361	0.632	0.000	0.000
Indemnity Pure Premium Haz'ness	0.382	0.048	0.058	0.588	0.007	0.732	0.020	0.000
Litigation Rates	0.121	0.553	0.355	0.448	0.203	0.676	0.024	0.640
Cumulative ÷ Indemnity Claims	0.004	0.432	0.068	0.690	0.080	0.587	0.696	0.587
1st PC (rGSP, AggE)	0.050	0.010	0.004	0.662	0.576	0.592	0.000	0.000
2nd PC (rGSP, AggE)	0.353	0.275	0.169	0.274	0.803	0.344	0.549	0.000
1st PC (rGSP, AggE, Unemp)	0.060	0.032	0.012	0.569	0.172	0.746	0.000	0.000
2nd PC (rGSP, AggE, Unemp)	0.483	0.222	0.229	0.911	0.538	0.766	0.000	0.838
Self-Insurance Share Index	0.275	0.942	0.610	0.094	0.753	0.038	0.746	0.900

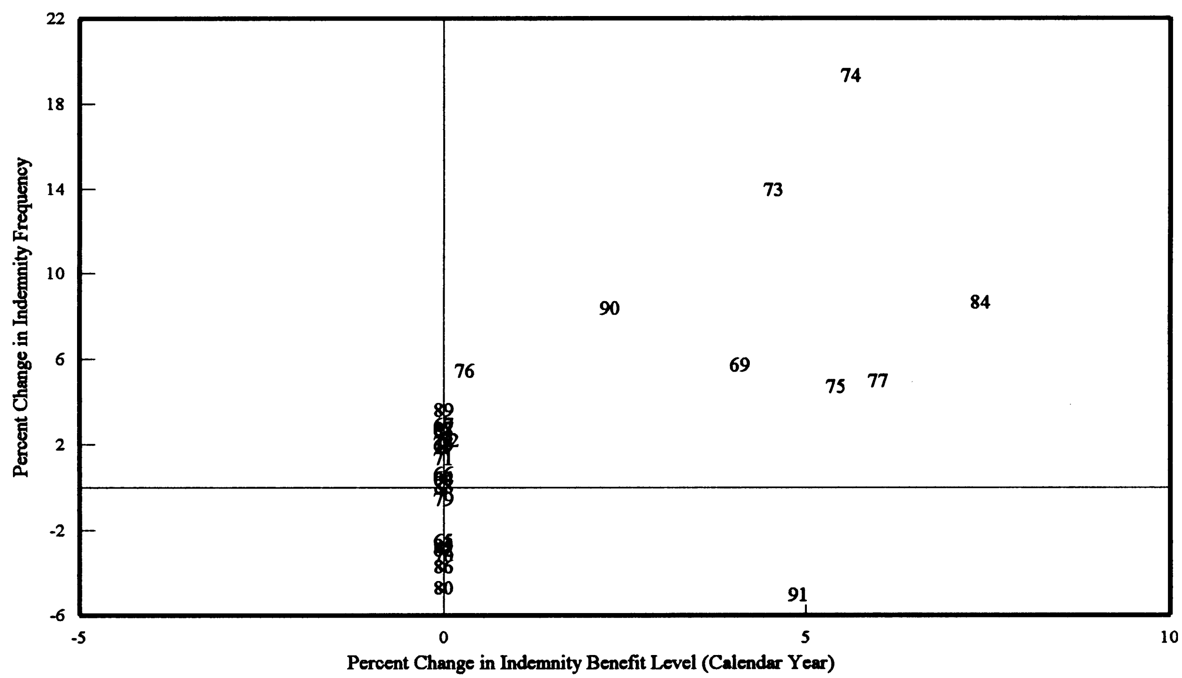
Note: P Value is the probability of observing the indicated SAMPLE correlation coefficient if the True correlation coefficient was actually zero.

California Unemplmt Rate	Indemnity Frequency Haz'ness	Indemnity Pure Premium Haz'ness	Litigation Rate	Cumulative ÷ Indemnity Claims	Principal Components				Self-Insurance Share Index
					PCGA_1	PCGA_2	PCUGA_1	PCUGA_2	
0.065	0.174	0.382	0.121	0.004	0.050	0.353	0.060	0.483	0.275
0.037	0.005	0.048	0.553	0.432	0.010	0.275	0.032	0.222	0.942
0.015	0.006	0.058	0.355	0.068	0.004	0.169	0.012	0.229	0.610
0.572	0.513	0.588	0.448	0.690	0.662	0.274	0.569	0.911	0.094
0.162	0.361	0.007	0.203	0.080	0.576	0.803	0.172	0.538	0.753
0.763	0.632	0.732	0.676	0.587	0.592	0.344	0.746	0.766	0.038
0.000	0.000	0.020	0.024	0.696	0.000	0.549	0.000	0.000	0.746
0.000	0.000	0.000	0.640	0.587	0.000	0.000	0.000	0.838	0.900
	0.001	0.000	0.165	0.928	0.000	0.005	0.000	0.844	0.888
0.001		0.000	0.071	0.578	0.000	0.342	0.001	0.099	0.345
0.000	0.000		0.201	0.680	0.008	0.024	0.000	0.726	0.164
0.165	0.071	0.201		0.232	0.031	0.160	0.145	0.146	0.994
0.928	0.578	0.680	0.232		0.788	0.301	0.921	0.672	0.132
0.000	0.000	0.008	0.031	0.788		1.000	0.000	0.000	0.787
0.005	0.342	0.024	0.160	0.301	1.000		0.007	0.000	0.629
0.000	0.001	0.000	0.145	0.921	0.000	0.007		1.000	0.882
0.844	0.099	0.726	0.146	0.672	0.000	0.000	1.000		0.764
0.888	0.345	0.164	0.994	0.132	0.787	0.629	0.882	0.764	

EXHIBIT 4

PART 1

INDEMNITY FREQUENCY VS INDEMNITY BENEFIT LEVEL



Outliers 1972, 1983 and 1994 used in regression but are not shown in graph

Spearman Rank Correlation Coefficient:	0.58449
Valid Cases	33
Two-tailed Significance	0.00035

Regression Output With Constant		Regression With Dummy Variable and Constant		
Constant	−0.10398	Constant		−1.40102
Std Err of Y Est	7.01511	Std Err of Y Est		6.98836
R Squared	0.15903	R Squared		0.19235
No. of Observations	33	No. of Observations		33
Degrees of Freedom	31	Degrees of Freedom		30
			Ind BL	Dummy
X Coefficient(s)	0.39609	X Coefficient(s)	0.32425	2.91302
Std Err of Coef.	0.16359	Std Err of Coef.	0.17529	2.61825
P-Value	0.02152	P-Value	0.07422	0.27472

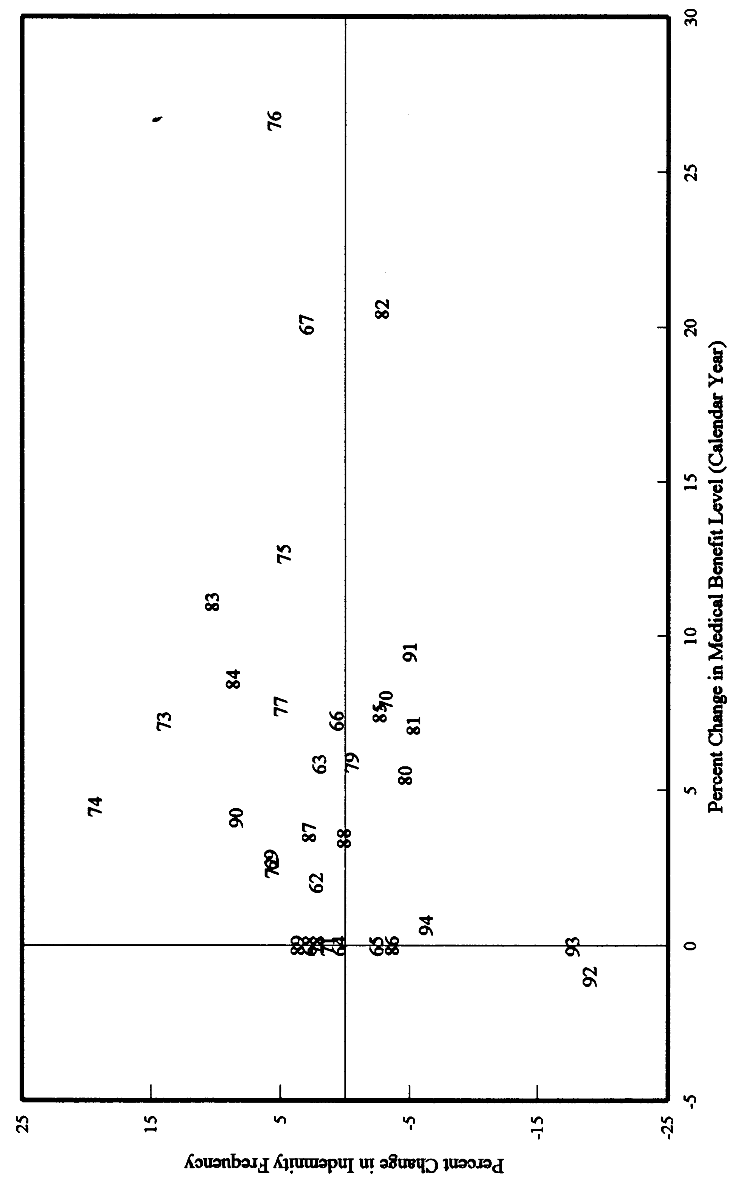
Regression Output Without Constant		Regression With Dummy Variable and No Constant		
Constant	0.00000	Constant		0.00000
Std Err of Y Est	6.90534	Std Err of Y Est		6.94801
R Squared	0.15886	R Squared		0.17504
No. of Observations	33	No. of Observations		33
Degrees of Freedom	32	Degrees of Freedom		31
			Ind BL	Dummy
X Coefficient(s)	0.39146	X Coefficient(s)	0.32425	1.51200
Std Err of Coef.	0.15054	Std Err of Coef.	0.17428	1.93884
P-Value	0.01398	P-Value	0.07233	0.44139

Note: Dummy variable is one for calendar years during which there was a medical benefit change and zero otherwise.

EXHIBIT 4

PART 2

INDEMNITY FREQUENCY VS MEDICAL BENEFIT LEVEL



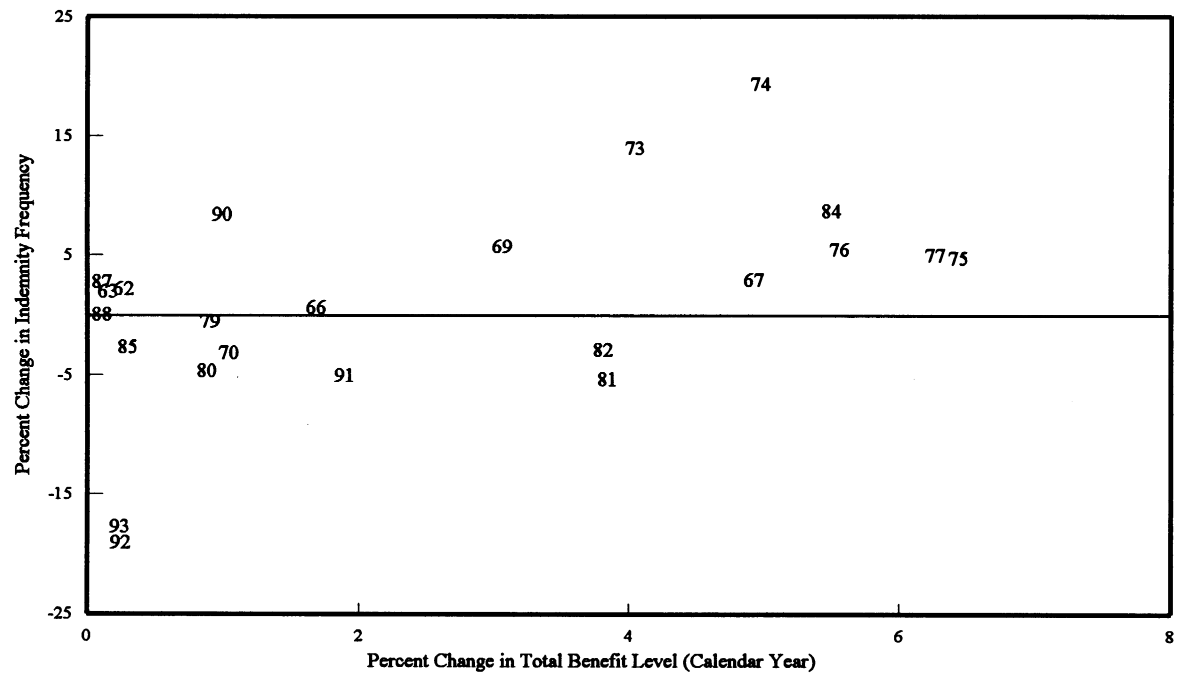
Spearman Rank Correlation Coefficient:	0.26843			
Valid Cases	33			
Two-tailed Significance	0.13092			
Regression Output With Constant:		Regression With Dummy Variable and Constant		
Constant	−0.55947	Constant		−1.72494
Std Err of Y Est	7.43909	Std Err of Y Est		7.51880
R Squared	0.05431	R Squared		0.06509
No. of Observations	33	No. of Observations		33
Degrees of Freedom	31	Degrees of Freedom		30
			Ind BL	Dummy
X Coefficient(s)	0.26884	X Coefficient(s)	0.19654	2.09867
Std Err of Coef.	0.20150	Std Err of Coef.	0.23786	3.56718
P-Value	0.19185	P-Value	0.41518	0.56072
Regression Output Without Constant:		Regression With Dummy Variable and No Constant		
Constant	0.00000	Constant		0.00000
Std Err of Y Est	7.33394	Std Err of Y Est		7.44826
R Squared	0.05120	R Squared		0.05197
No. of Observations	33	No. of Observations		33
Degrees of Freedom	32	Degrees of Freedom		31
			Ind BL	Dummy
X Coefficient(s)	0.22550	X Coefficient(s)	0.19654	0.37373
Std Err of Coef.	0.14668	Std Err of Coef.	0.23563	2.35638
P-Value	0.13403	P-Value	0.41061	0.87501

Note: Dummy variable is one for calendar years during which there was a medical benefit change and zero otherwise.

EXHIBIT 4

PART 3

INDEMNITY FREQUENCY VS TOTAL BENEFIT LEVEL



Outliers 1972, 1983 and 1994 used in regression but are not shown in graph

Spearman Rank Correlation Coefficient:	0.50840
Valid Cases	33
Two-tailed Significance	0.00252

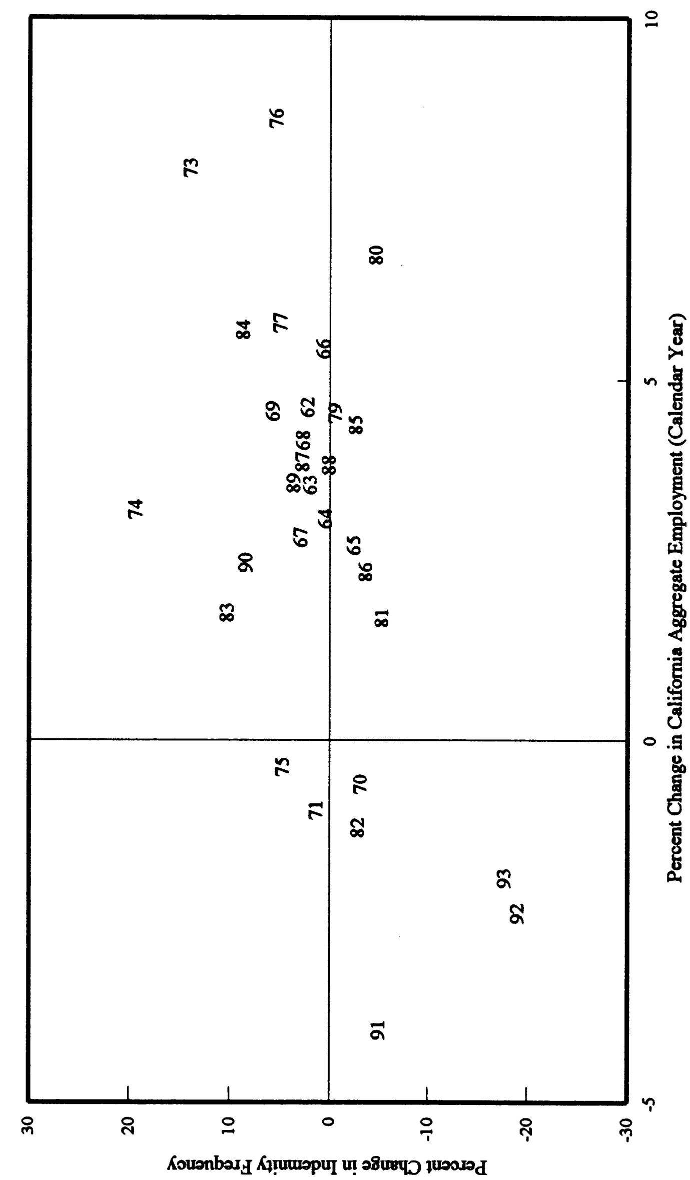
Regression Output With Constant:		Regression With Dummy Variable and Constant		
Constant	−0.68463	Constant		0.56004
Std Err of Y Est	6.81082	Std Err of Y Est		6.88774
R Squared	0.20730	R Squared		0.21544
No. of Observations	33	No. of Observations		33
Degrees of Freedom	31	Degrees of Freedom		30
			Ind BL	Dummy
X Coefficient(s)	0.67889	X Coefficient(s)	0.71540	−1.69602
Std Err of Coef.	0.23844	Std Err of Coef.	0.24985	3.03890
P-Value	0.00776	P-Value	0.00758	0.58092
Regression Output Without Constant:		**Regression With Dummy Variable and No Constant**		
Constant	0.00000	Constant		0.00000
Std Err of Y Est	6.73223	Std Err of Y Est		6.78096
R Squared	0.20050	R Squared		0.21423
No. of Observations	33	No. of Observations		33
Degrees of Freedom	32	Degrees of Freedom		31
			Ind BL	Dummy
X Coefficient(s)	0.62352	X Coefficient(s)	0.71540	−1.13598
Std Err of Coef.	0.21042	Std Err of Coef.	0.24597	1.54338
P-Value	0.00570	P-Value	0.00666	0.46724

Note: Dummy variable is one for calendar years during which there was an indemnity or medical benefit change and zero otherwise.

EXHIBIT 4

PART 4

INDEMNITY FREQUENCY VS CALIFORNIA AGGREGATE EMPLOYMENT



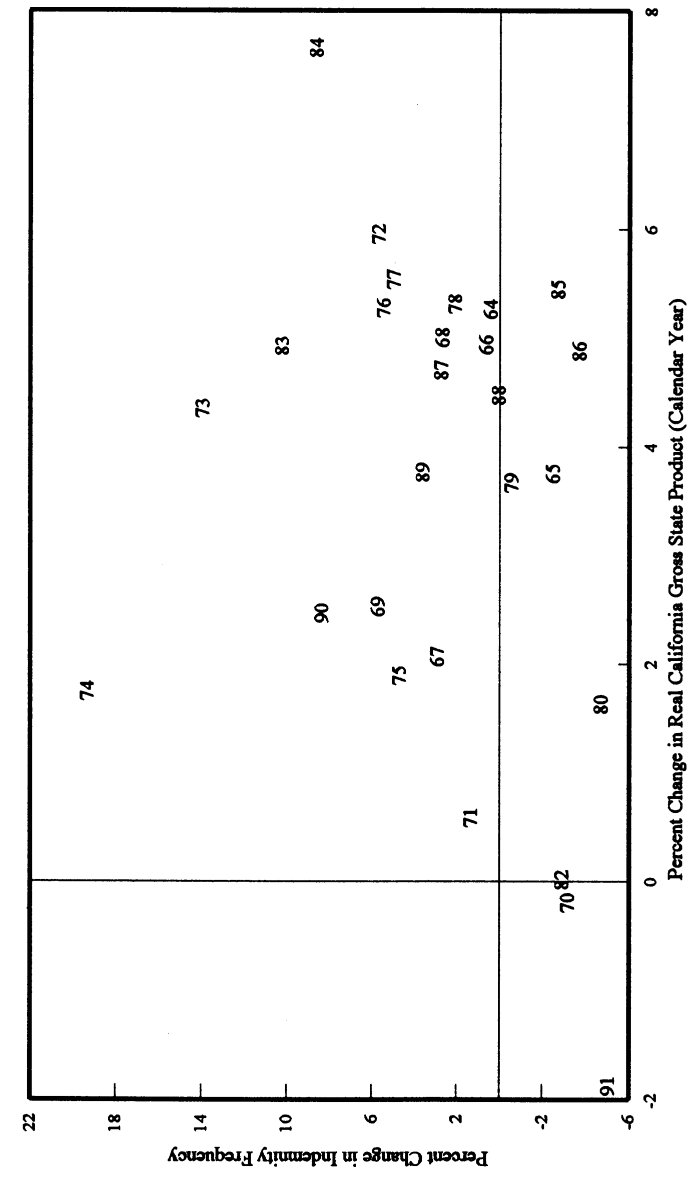
Outliers 1972 and 1978 are used in the regressions but are not shown in the graph

Spearman Rank Correlation Coefficient:	0.47947
Valid Cases	32
Two-tailed Significance	0.00548
Regression Output With Constant:	
Constant	−1.21341
Std Err of Y Est	7.01544
R Squared	0.16030
No. of Observations	32
Degrees of Freedom	30
X Coefficient(s)	0.65749
Std Err of Coef.	0.27474
P-Value	0.02320
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	6.96629
R Squared	0.14443
No. of Observations	32
Degrees of Freedom	31
X Coefficient(s)	0.52540
Std Err of Coef.	0.20998
P-Value	0.01783

EXHIBIT 4

PART 5

INDEMNITY FREQUENCY VS REAL CALIFORNIA GROSS STATE PRODUCT



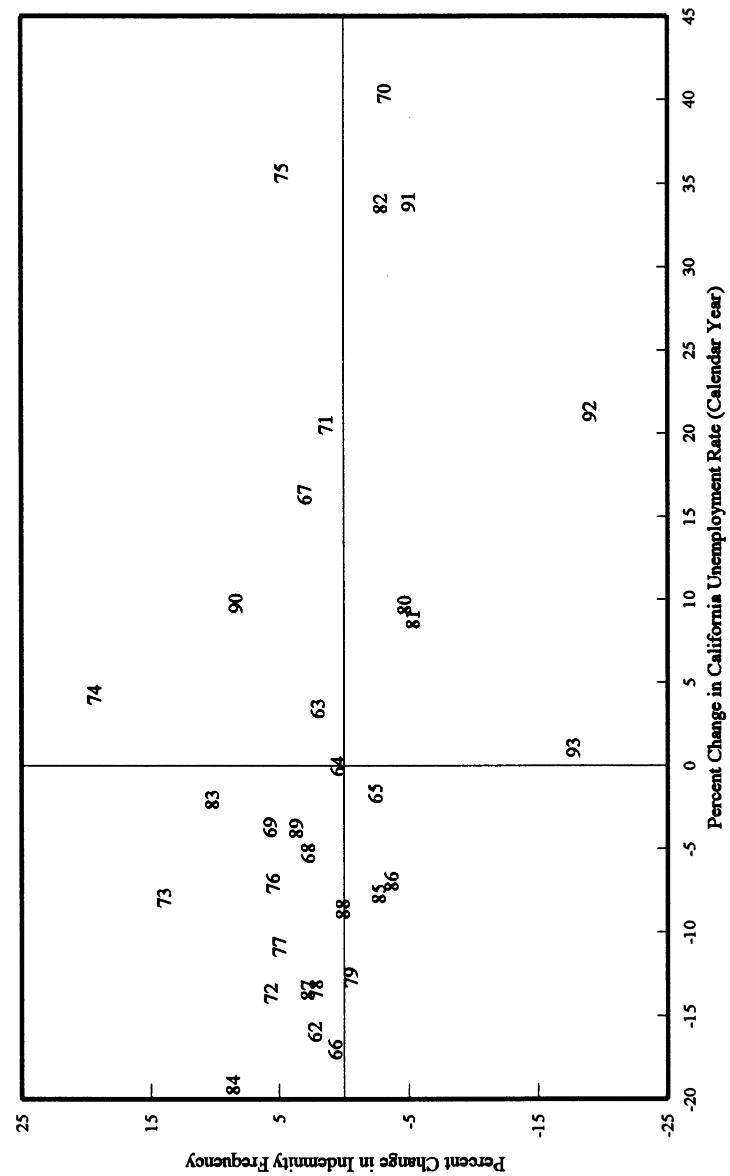
Outlier 1992 is used in the regressions but is not shown in the graph

Spearman Rank Correlation Coefficient:	0.40246
Valid Cases	29
Two-tailed Significance	0.03042
Regression Output With Constant:	
Constant	−2.27752
Std Err of Y Est	6.59466
R Squared	0.15332
No. of Observations	29
Degrees of Freedom	27
X Coefficient(s)	1.21851
Std Err of Coef.	0.55107
P-Value	0.03570
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	6.59963
R Squared	0.12064
No. of Observations	29
Degrees of Freedom	28
X Coefficient(s)	0.74826
Std Err of Coef.	0.30273
P-Value	0.01979

EXHIBIT 4

PART 6

INDEMNITY FREQUENCY VS CALIFORNIA UNEMPLOYMENT RATE

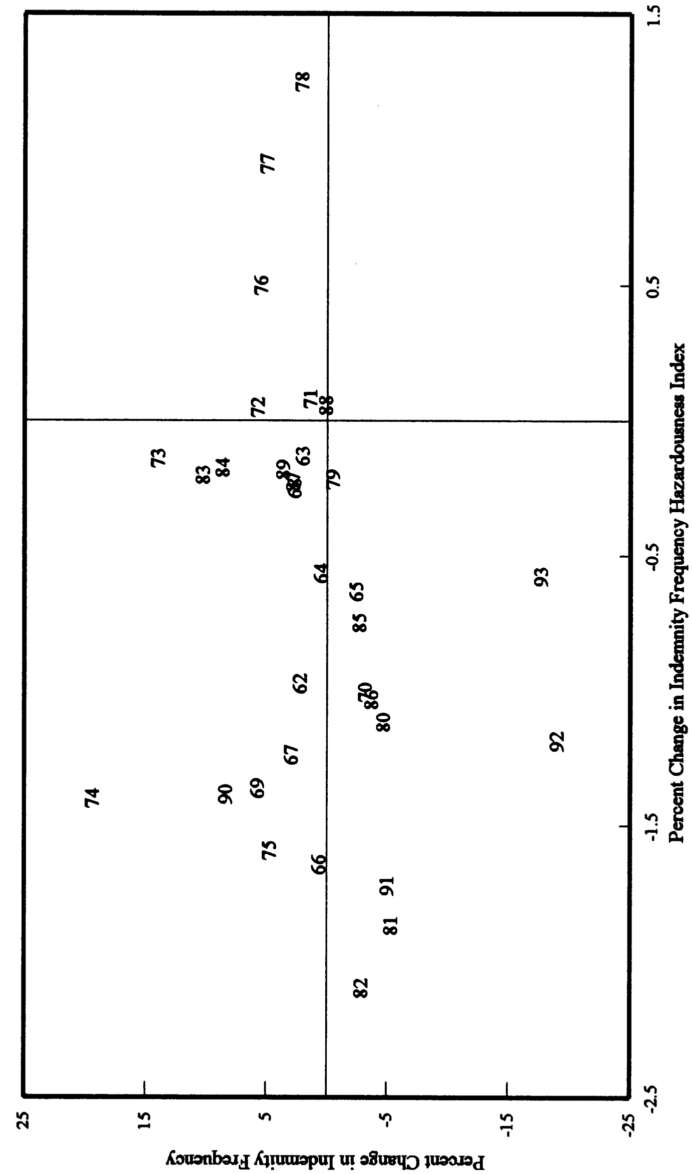


Spearman Rank Correlation Coefficient:	−0.34054
Valid Cases	32
Two-tailed Significance	0.05648
Regression Output With Constant:	
Constant	1.53439
Std Err of Y Est	7.30768
R Squared	0.08889
No. of Observations	32
Degrees of Freedom	30
X Coefficient(s)	−0.13483
Std Err of Coef.	0.07881
P-Value	0.09740
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	7.35321
R Squared	0.04675
No. of Observations	32
Degrees of Freedom	31
X Coefficient(s)	−0.12291
Std Err of Coef.	0.07865
P-Value	0.12824

EXHIBIT 4

PART 7

INDEMNITY FREQUENCY VS INDEMNITY FREQUENCY HAZARDOUSNESS INDEX

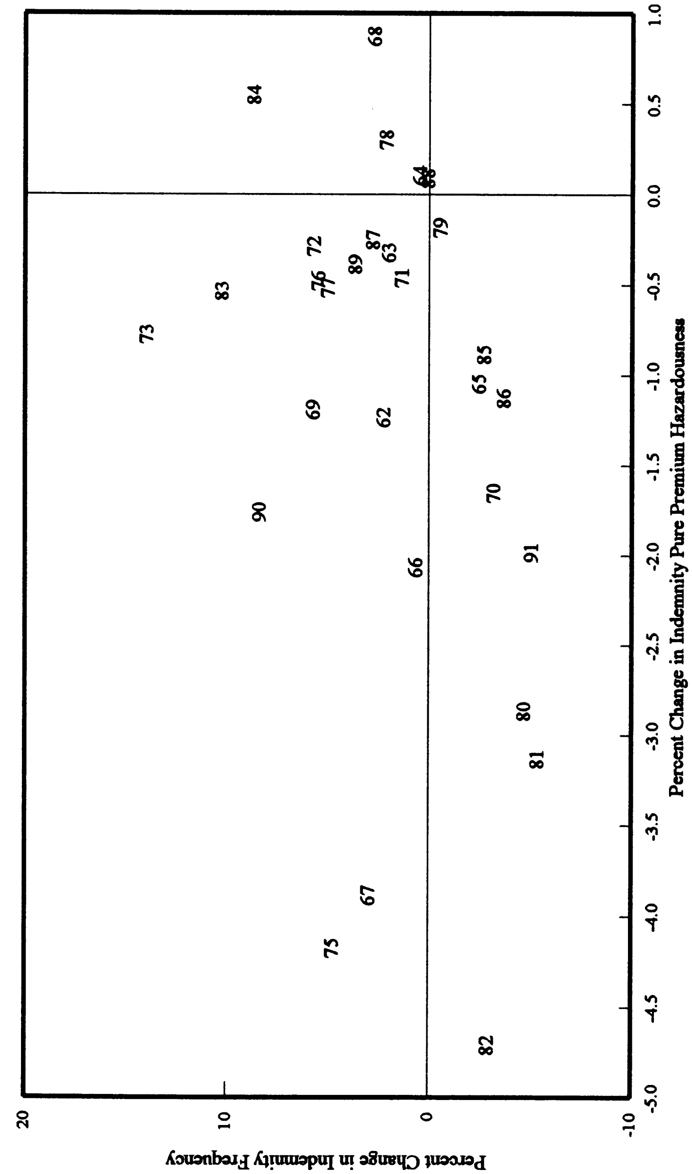


Spearman Rank Correlation Coefficient:	0.32148
Valid Cases	32
Two-tailed Significance	0.07278
Regression Output With Constant:	
Constant	2.58046
Std Err of Y Est	7.46806
R Squared	0.04846
No. of Observations	32
Degrees of Freedom	30
X Coefficient(s)	2.05597
Std Err of Coef.	1.66336
P-Value	0.22600
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	7.62216
R Squared	0.04609
No. of Observations	32
Degrees of Freedom	31
X Coefficient(s)	0.46323
Std Err of Coef.	1.31509
P-Value	0.72703

EXHIBIT 4

PART 8

INDEMNITY FREQUENCY VS INDEMNITY PURE PREMIUM HAZARDOUSNESS INDEX

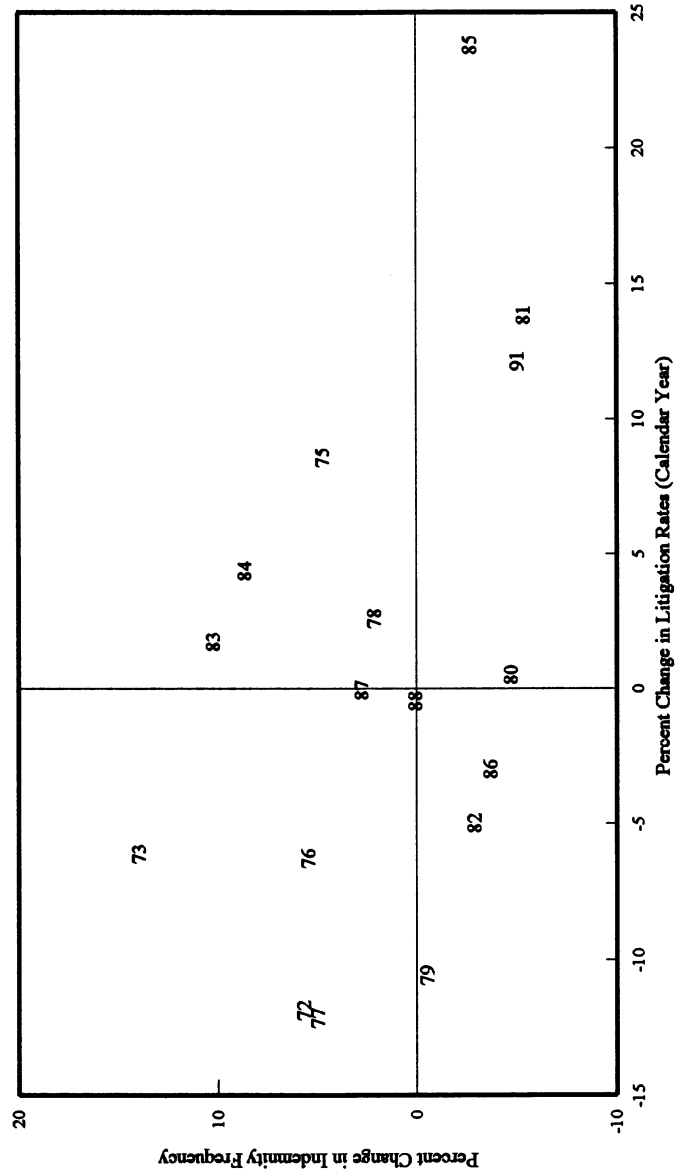


Spearman Rank Correlation Coefficient:	0.20860
Valid Cases	32
Two-tailed Significance	0.24550
Regression Output With Constant:	
Constant	1.95477
Std Err of Y Est	7.61169
R Squared	0.01150
No. of Observations	32
Degrees of Freedom	30
X Coefficient(s)	0.58979
Std Err of Coef.	0.99820
P-Value	0.55904
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	7.63367
R Squared	0.02227
No. of Observations	32
Degrees of Freedom	31
X Coefficient(s)	−0.13025
Std Err of Coef.	0.74832
P-Value	0.86296

EXHIBIT 4

PART 9

INDEMNITY FREQUENCY VS LITIGATION RATE



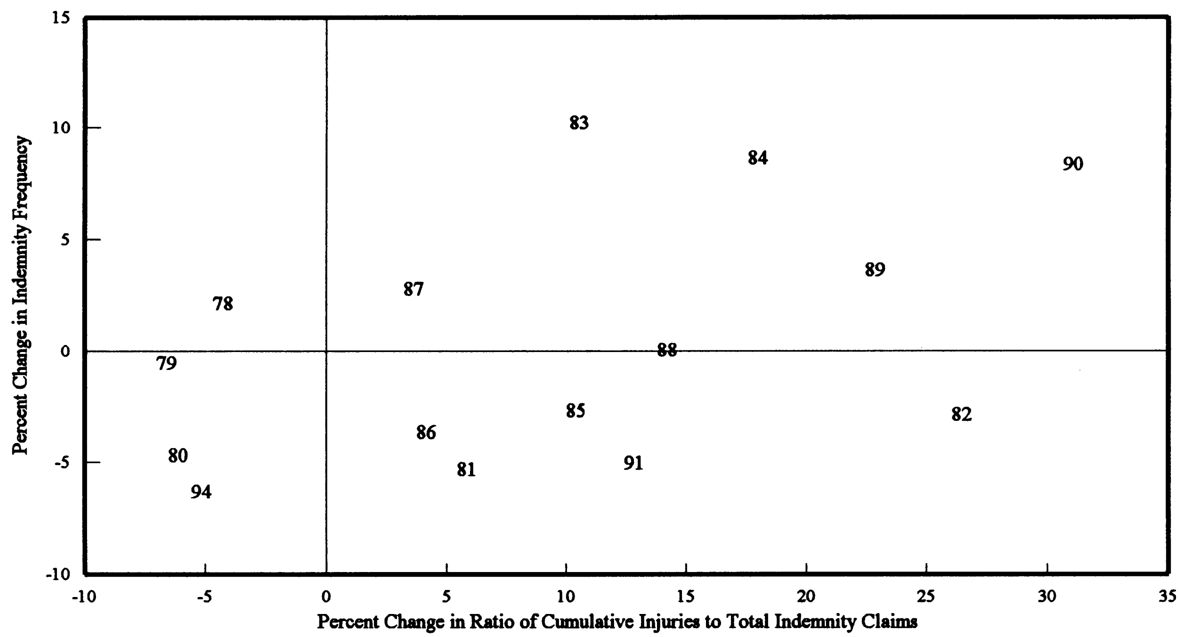
Outliers 1974 and 1985 are NOT used in regressions but are shown on graph

Spearman Rank Correlation Coefficient:	–0.35000
Valid Cases	16
Two-tailed Significance	0.18386
Regression Output With Constant:	
Constant	2.08448
Std Err of Y Est	5.71497
R Squared	0.10549
No. of Observations	16
Degrees of Freedom	14
X Coefficient(s)	–0.23892
Std Err of Coef.	0.18595
P-Value	0.21968
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	7.94003
R Squared	0.26223
No. of Observations	16
Degrees of Freedom	15
X Coefficient(s)	–0.26371
Std Err of Coef.	0.19190
P-Value	0.18956

EXHIBIT 4

PART 10

INDEMNITY FREQUENCY VS RATIO OF CUMULATIVE INJURIES TO TOTAL INDEMNITY CLAIMS



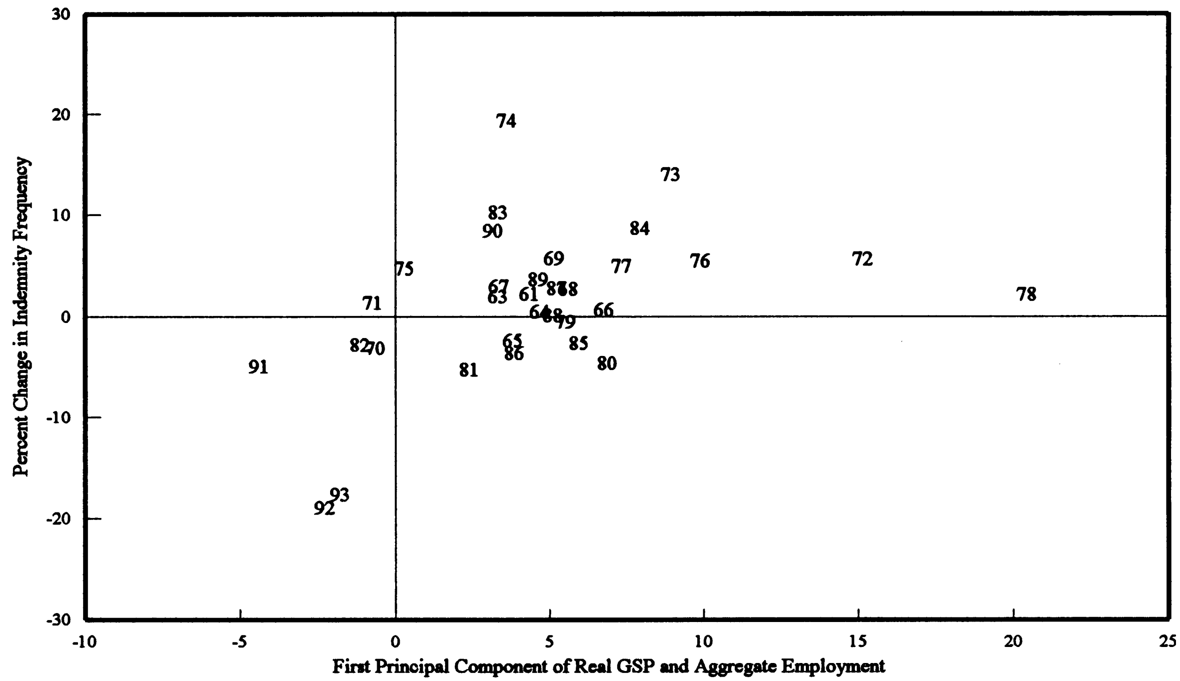
Outliers 1992 and 1993 are used in the regressions but not shown in graph

Spearman Rank Correlation Coefficient:	0.61029
Valid Cases	17
Two-tailed Significance	0.00926
Regression Output With Constant:	
Constant	−3.49152
Std Err of Y Est	5.23706
R Squared	0.60247
No. of Observations	17
Degrees of Freedom	15
X Coefficient(s)	0.35571
Std Err of Coef.	0.07460
P-Value	0.00025
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	6.15150
R Squared	0.41496
No. of Observations	17
Degrees of Freedom	16
X Coefficient(s)	0.30564
Std Err of Coef.	0.08480
P-Value	0.00238

EXHIBIT 4

PART 11

INDEMNITY FREQUENCY VS FIRST PRINCIPAL COMPONENT OF CA REAL GSP AND CA AGGREGATE EMPLOYMENT

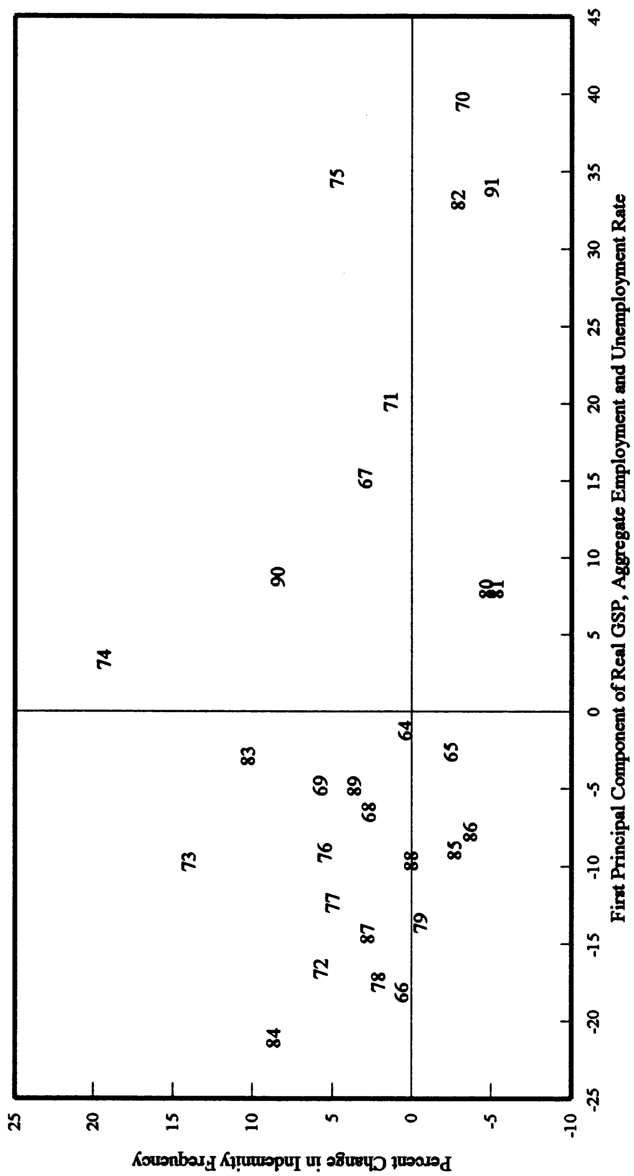


Spearman Rank Correlation Coefficient:	0.43181
Valid Cases	32
Two-tailed Significance	0.01358
Regression Output With Constant:	
Constant	−1.75361
Std Err of Y Est	6.93017
R Squared	0.18059
No. of Observations	32
Degrees of Freedom	30
X Coefficient(s)	0.65856
Std Err of Coef.	0.25612
P-Value	0.01533
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	6.93842
R Squared	0.15126
No. of Observations	32
Degrees of Freedom	31
X Coefficient(s)	0.47547
Std Err of Coef.	0.18563
P-Value	0.01551

EXHIBIT 4

PART 12

INDEMNITY FREQUENCY VS FIRST PRINCIPAL COMPONENT OF CA REAL GSP, CA AGGREGATE EMPLOYMENT AND UNEMPLOYMENT RATE



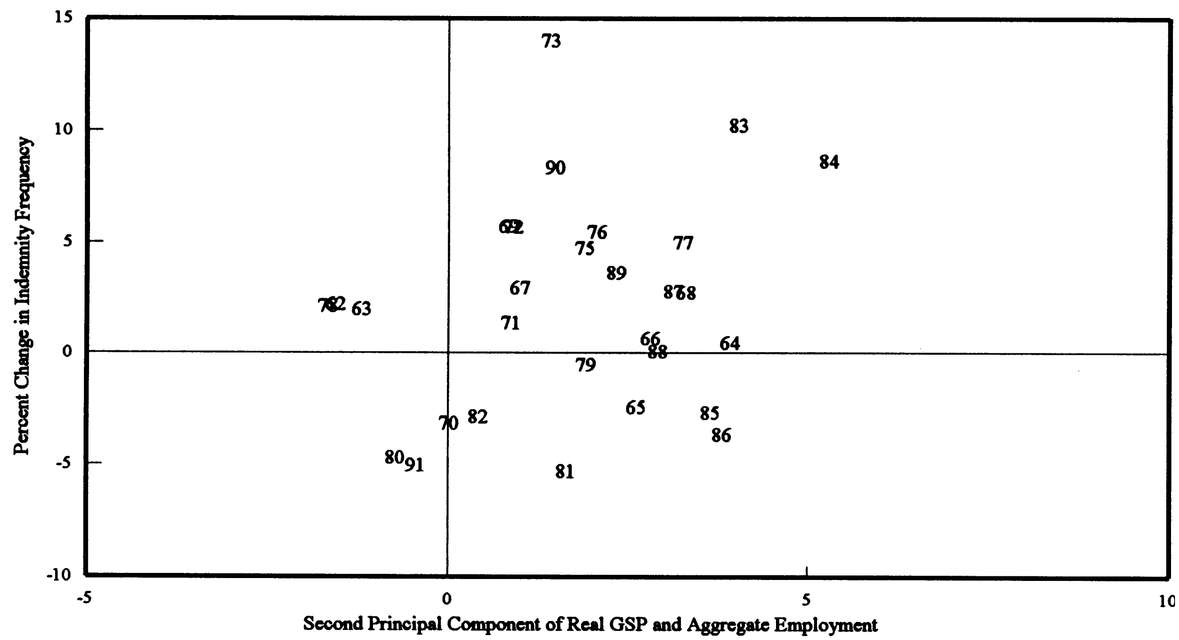
Outliers 1992 and 1993 used in regressions but are not shown in the graph

Spearman Rank Correlation Coefficient:	−0.36418
Valid Cases	30
Two-tailed Significance	0.04786
Regression Output With Constant:	
Constant	1.41012
Std Err of Y Est	7.52272
R Squared	0.09809
No. of Observations	30
Degrees of Freedom	28
X Coefficient(s)	−0.14101
Std Err of Coef.	0.08080
P-Value	0.09121
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	7.52863
R Squared	0.06441
No. of Observations	30
Degrees of Freedom	29
X Coefficient(s)	−0.13357
Std Err of Coef.	0.08054
P-Value	0.10731

EXHIBIT 4

PART 13

INDEMNITY FREQUENCY VS SECOND PRINCIPAL COMPONENT OF CA REAL GSP AND CA AGGREGATE EMPLOYMENT



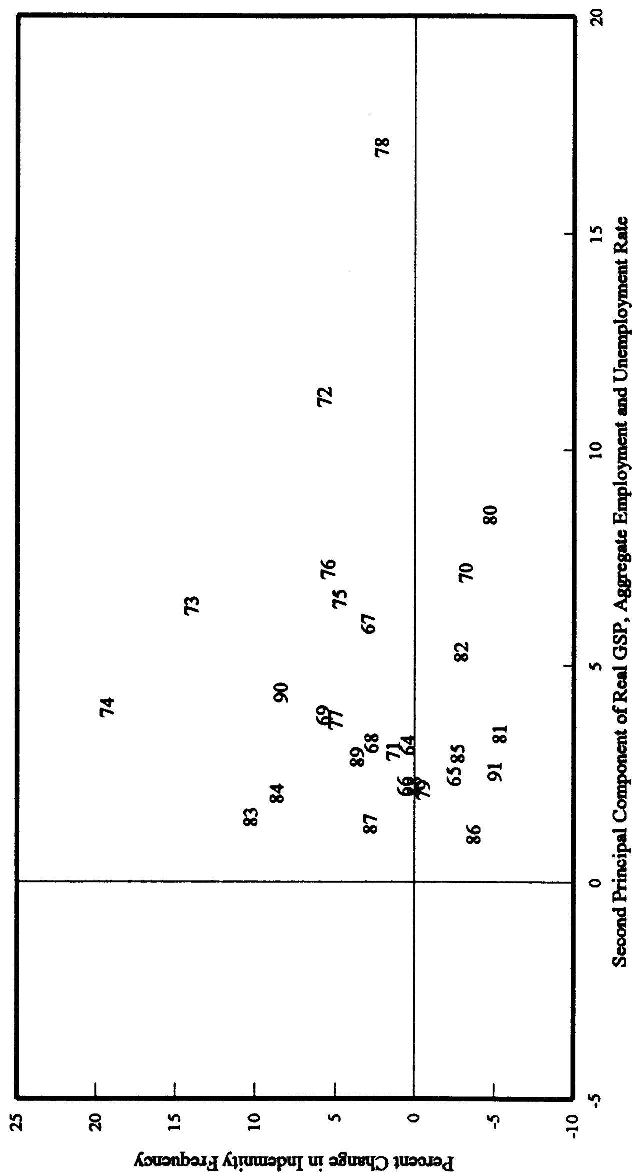
Outliers 1974, 1992 and 1993 used in regression but are not shown in graph

Spearman Rank Correlation Coefficient:	0.22360
Valid Cases	32
Two-tailed Significance	0.21860
Regression Output With Constant:	
Constant	–0.04754
Std Err of Y Est	7.52449
R Squared	0.03402
No. of Observations	32
Degrees of Freedom	30
X Coefficient(s)	0.79696
Std Err of Coef.	0.77531
P-Value	0.31220
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	7.40222
R Squared	0.03400
No. of Observations	32
Degrees of Freedom	31
X Coefficient(s)	0.78313
Std Err of Coef.	0.55359
P-Value	0.16714

EXHIBIT 4

PART 14

INDEMNITY FREQUENCY VS SECOND PRINCIPAL COMPONENT OF CA REAL GSP, CA AGGREGATE EMPLOYMENT AND UNEMPLOYMENT RATE



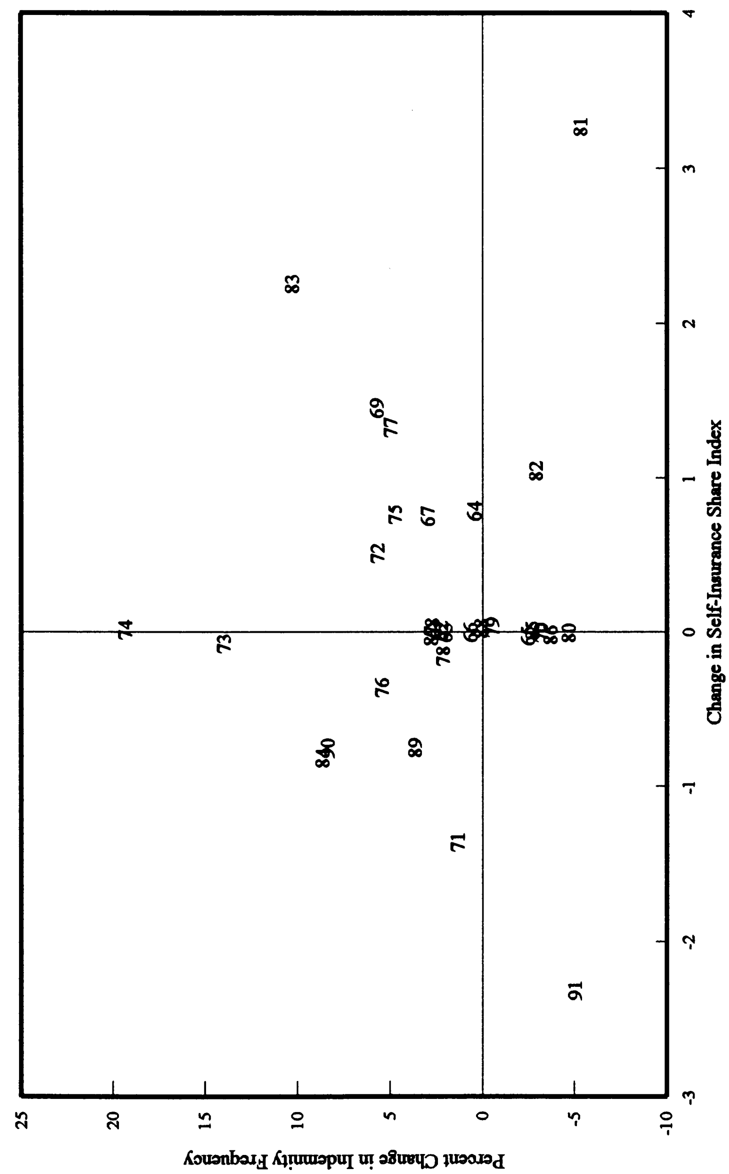
Outliers 1992 and 1993 used in regressions but are not shown in the graph

Spearman Rank Correlation Coefficient:	0.28275
Valid Cases	30
Two-tailed Significance	0.13002
Regression Output With Constant:	
Constant	−1.24561
Std Err of Y Est	7.64980
R Squared	0.06736
No. of Observations	30
Degrees of Freedom	28
X Coefficient(s)	0.57363
Std Err of Coef.	0.40336
P-Value	0.16530
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	7.55920
R Squared	0.05680
No. of Observations	30
Degrees of Freedom	29
X Coefficient(s)	0.39749
Std Err of Coef.	0.25168
P-Value	0.12442

EXHIBIT 4

PART 15

INDEMNITY FREQUENCY VS SELF-INSURANCE SHARE INDEX



Spearman Rank Correlation Coefficient:	−0.10320
Valid Cases	32
Two-tailed Significance	0.56550
Regression Output With Constant:	
Constant	1.55043
Std Err of Y Est	7.54764
R Squared	0.02807
No. of Observations	30
Degrees of Freedom	28
X Coefficient(s)	−1.16260
Std Err of Coef.	1.24904
P-Value	0.35939
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	0.00000
R Squared	0.02807
No. of Observations	30
Degrees of Freedom	29
X Coefficient(s)	−0.82922
Std Err of Coef.	1.21901
P-Value	0.50140

EXHIBIT 5

PART 1

SUMMARY OF SELECTED REGRESSION RESULTS INDEMNITY BENEFIT LEVEL

Independent Variables			Coefficient of Indemnity Benefit Level	Adjusted R^2 ($\times$100)	Mean Residual Error	P-Values for Tests of Normality in Residuals: K-S Test	Shapiro-Wilks	Skewness	Kurtosis
Constant	PCUGA_1 & PCUGA_2	Cum Inj Index	0.286573	87.9086	−0.000001	0.98557	0.726754	0.881409	0.704214
Constant	PCUGA_1	Cum Inj Index	0.261897	85.6142	−0.000001	0.94783	0.748869	0.886143	0.343503
Constant	PCGA_1 & PCGA_2	Cum Inj Index	0.272918	83.4069	−0.000000	0.78244	0.423384	0.548331	0.813895
Constant	rGSP & AggE	Cum Inj Index	0.272919	83.4069	−0.000001	0.78245	0.423383	0.548330	0.813897
Constant	PCGA_1	Cum Inj Index	0.309052	82.2782	0.000000	0.50538	0.105146	0.386600	0.849528
Constant	AggE	Cum Inj Index	0.321087	79.8728	0.000000	0.66497	0.257688	0.398509	0.826385
Constant	rGSP	Cum Inj Index	0.220530	78.2211	−0.000000	0.90252	0.335510	0.997070	0.148195
Origin	PCUGA_1 & PCUGA_2	Cum Inj Index	0.174378	66.0923	−1.821382	0.92753	0.330178	0.930584	0.411106
Origin	PCUGA_1	Cum Inj Index	0.164625	64.4638	−2.660440	0.70927	0.205739	0.925496	0.625992
Origin	AggE	Cum Inj Index	0.168924	32.8023	−3.249618	0.76309	0.425381	0.632718	0.418103
Origin	PCGA_1	Cum Inj Index	0.166104	32.1664	−3.202778	0.74136	0.339507	0.610755	0.411451
Origin	PCGA_1 & PCGA_2	Cum Inj Index	0.236292	31.8970	−2.675669	0.99048	0.736900	0.858949	0.404446
Origin	rGSP & AggE	Cum Inj Index	0.236293	31.8970	−2.675668	0.99048	0.736900	0.858948	0.404446
Origin	rGSP	Cum Inj Index	0.181181	30.8825	−2.692286	0.71617	0.417545	0.613256	0.349558
Origin	PCUGA_1 & PCUGA_2	None	0.310626	26.1143	−0.129588	0.60045	0.113746	0.727394	0.018431
Origin	PCGA_1	None	0.285431	26.0075	−0.745032	0.62419	0.060441	0.890201	0.011669
Origin	AggE	None	0.296739	25.9621	−0.631326	0.74362	0.095821	0.945521	0.014194
Origin	PCUGA_1	None	0.363330	25.5120	0.765343	0.79254	0.116257	0.880094	0.018651
Origin	rGSP	None	0.272067	23.3272	−0.776528	0.52080	0.025223	0.889271	0.007354
Origin	rGSP & AggE	None	0.288721	23.1740	−0.712893	0.64107	0.068267	0.905894	0.012398
Origin	PCGA_1 & PCGA_2	None	0.288721	23.1740	−0.712893	0.64107	0.068266	0.905893	0.012398
Constant	PCGA_1	None	0.321818	20.3379	−0.000000	0.70653	0.169706	0.712973	0.020484
Constant	AggE	None	0.330217	19.3274	0.000001	0.78895	0.207190	0.845635	0.023023
Constant	rGSP	None	0.287312	19.2424	−0.000001	0.71529	0.032419	0.400785	0.008683
Constant	PCUGA_1	None	0.316254	18.7551	0.000001	0.75395	0.091303	0.844291	0.016314
Constant	PCGA_1 & PCGA_2	None	0.300944	18.4262	0.000001	0.68396	0.067943	0.443847	0.013336
Constant	rGSP & AggE	None	0.300945	18.4262	0.000001	0.68396	0.067943	0.443848	0.013336
Constant	PCUGA_1 & PCUGA_2	None	0.319097	17.8047	0.000002	0.56930	0.137531	0.718970	0.020284

PCGA_1(2) = First (second) principal component of California Real GSP and Aggregate Employment.
PCUGA_1(2) = First (second) principal component of California Unemployment Rate, California Real GSP, and Aggregate Employment.

EXHIBIT 5

PART 2

SUMMARY OF SELECTED REGRESSION RESULTS TOTAL BENEFIT LEVEL

Independent Variables			Coefficient of Indemnity Benefit Level	Adjusted R^2 ($\times 100$)	Mean Residual Error	P-Values for Tests of Normality in Residuals: K-S Test	Shapiro-Wilks	Skewness	Kurtosis
Constant	PCUGA_1 & PCUGA_2	Cum Inj Index	0.456894	85.9722	−0.000000	0.92907	0.874813	0.927566	0.611238
Constant	PCUGA_1	Cum Inj Index	0.422782	84.3168	−0.000001	0.95431	0.848256	0.968421	0.361423
Constant	PCGA_1 & PCGA_2	Cum Inj Index	0.420679	80.7484	−0.000000	0.89854	0.322648	0.817466	0.496837
Constant	rGSP & AggE	Cum Inj Index	0.420679	80.7484	0.000000	0.89854	0.322650	0.817465	0.496839
Constant	PCGA_1	Cum Inj Index	0.484488	79.2394	−0.000000	0.92763	0.747651	0.795915	0.775950
Constant	AggE	Cum Inj Index	0.502991	76.5814	−0.000000	0.93771	0.901157	0.842296	0.895562
Constant	rGSP	Cum Inj Index	0.337664	76.3457	0.000001	0.95764	0.381211	0.919454	0.085781
Origin	PCUGA_1 & PCUGA_2	Cum Inj Index	0.256019	64.4102	−1.800296	0.82172	0.359736	0.848332	0.372861
Origin	PCUGA_1	Cum Inj Index	0.225583	62.4887	−2.652812	0.79838	0.328497	0.879385	0.642470
Origin	PCUGA_1	None	0.621759	31.9340	0.083426	0.63017	0.092248	0.700709	0.018230
Origin	PCUGA_1 & PCUGA_2	None	0.564547	30.0771	−0.279990	0.69593	0.082883	0.626873	0.016654
Origin	AggE	Cum Inj Index	0.203652	29.6727	−3.189150	0.87257	0.536635	0.741083	0.302869
Origin	PCGA_1	Cum Inj Index	0.198961	29.1148	−3.152980	0.88414	0.475527	0.715886	0.295250
Origin	rGSP	Cum Inj Index	0.221670	27.7018	−2.714230	0.86316	0.354814	0.659300	0.238581
Origin	AggE	None	0.487566	27.4589	−0.900121	0.46532	0.059950	0.901925	0.012502
Origin	PCGA_1	None	0.472383	27.3324	−0.973372	0.44110	0.047488	0.865926	0.010582
Origin	rGSP & AggE	Cum Inj Index	0.304403	27.2065	−2.715112	0.98855	0.628836	0.823965	0.282299
Origin	PCGA_1 & PCGA_2	Cum Inj Index	0.304402	27.2065	−2.715111	0.98855	0.628837	0.823966	0.282298
Origin	rGSP	None	0.472065	24.9179	−0.931511	0.42116	0.022933	0.909797	0.007748
Origin	rGSP & AggE	None	0.492587	24.6757	−0.875043	0.47991	0.064638	0.914406	0.013135
Origin	PCGA_1 & PCGA_2	None	0.492587	24.6756	−0.875044	0.47991	0.064638	0.914405	0.013135
Constant	PCUGA_1	None	0.609784	24.3041	−0.000003	0.64948	0.086616	0.698150	0.017646
Constant	PCGA_1	None	0.590481	24.0807	−0.000001	0.54070	0.143540	0.593814	0.023128
Constant	rGSP	None	0.547191	23.2327	−0.000001	0.40620	0.024810	0.326243	0.009149
Constant	AggE	None	0.602700	23.0988	−0.000001	0.70428	0.184220	0.713578	0.026753
Constant	PCUGA_1 & PCUGA_2	None	0.600278	22.6779	0.000002	0.81060	0.128743	0.589805	0.020823
Constant	rGSP & AggE	None	0.560971	22.3380	0.000000	0.66443	0.058335	0.356940	0.013897
Constant	PCGA_1 & PCGA_2	None	0.560971	22.3380	0.000000	0.66443	0.058335	0.356940	0.013897

PCGA_1(2) = First (second) principal component of California Real GSP and Aggregate Employment.
PCUGA_1(2) = First (second) principal component of California Unemployment Rate, California Real GSP, and Aggregate Employment.

EXHIBIT 5

PART 3

SUMMARY OF SELECTED REGRESSION RESULTS INDEMNITY AND MEDICAL BENEFIT LEVELS SEPARATELY

Independent Variables			Coefficient of Indemnity Benefit Level	Adjusted R^2 (×100)	Mean Residual Error	P-Values for Tests of Normality in Residuals			
						K-S Test	Shapiro-Wilks	Skewness	Kurtosis
Constant	PCUGA_1 & PCUGA_2	Cum Inj Index	0.292514	86.6292	−0.000000	0.93713	0.469462	0.860224	0.579064
Constant	PCUGA_1	Cum Inj Index	0.265798	84.2016	−0.000001	0.91036	0.648858	0.865121	0.261407
Constant	PCGA_1 & PCGA_2	Cum Inj Index	0.295324	82.4230	−0.000000	0.92179	0.421744	0.377549	0.966732
Constant	rGSP & AggE	Cum Inj Index	0.295324	82.4230	−0.000001	0.92179	0.421743	0.377548	0.966733
Constant	PCGA_1	Cum Inj Index	0.333570	82.2312	−0.000000	0.42860	0.014267	0.168730	0.659239
Constant	AggE	Cum Inj Index	0.349063	80.3324	−0.000001	0.26057	0.005202	0.138342	0.608047
Constant	rGSP	Cum Inj Index	0.248528	77.1472	−0.000001	0.87537	0.434008	0.966419	0.149610
Origin	PCUGA_1	Cum Inj Index	0.272917	72.6089	−1.466101	0.27758	0.024815	0.631453	0.018983
Origin	PCUGA_1 & PCUGA_2	Cum Inj Index	0.264671	70.3213	−1.313504	0.37866	0.025171	0.626510	0.014782
Origin	AggE	Cum Inj Index	0.333620	61.7762	−1.517823	0.98017	0.747490	0.787201	0.379075
Origin	PCGA_1	Cum Inj Index	0.328093	61.1363	−1.520994	0.97121	0.815370	0.876925	0.332588
Origin	PCGA_1 & PCGA_2	Cum Inj Index	0.364976	59.6827	−1.263185	0.93943	0.230026	0.437066	0.726277
Origin	rGSP & AggE	Cum Inj Index	0.364976	59.6827	−1.263183	0.93942	0.230023	0.437065	0.726277
Origin	rGSP	Cum Inj Index	0.319801	57.7186	−1.243197	0.97554	0.703445	0.965119	0.209469
Origin	PCUGA_1	None	0.279791	28.4637	−0.190637	0.68344	0.083596	0.404802	0.017932
Origin	PCUGA_1 & PCUGA_2	None	0.272875	26.0173	−0.388706	0.79310	0.069573	0.414467	0.016002
Origin	AggE	None	0.284992	23.3459	−0.774986	0.62034	0.076152	0.885340	0.013112
Origin	PCGA_1	None	0.277164	23.2979	−0.846570	0.67076	0.056379	0.848956	0.010987
Constant	PCUGA_1	None	0.288158	20.4783	0.000001	0.78526	0.086808	0.367933	0.020017
Origin	rGSP	None	0.267653	20.4679	−0.835981	0.55850	0.025379	0.870139	0.007349
Origin	PCGA_1 & PCGA_2	None	0.282775	20.2872	−0.795989	0.63350	0.068830	0.874672	0.012498
Origin	rGSP & AggE	None	0.282775	20.2872	−0.795988	0.63350	0.068830	0.874672	0.012498
Constant	PCGA_1	None	0.304491	19.9140	−0.000001	0.73658	0.155005	0.392062	0.022554
Constant	rGSP	None	0.266582	18.9563	0.000001	0.26368	0.014179	0.174932	0.007352
Constant	PCUGA_1 & PCUGA_2	None	0.292522	18.7697	0.000003	0.72397	0.074620	0.310532	0.021152
Constant	AggE	None	0.313943	18.6927	−0.000002	0.67057	0.231371	0.513812	0.027657
Constant	rGSP & AggE	None	0.280121	18.2308	−0.000000	0.67307	0.038561	0.189872	0.011297
Constant	PCGA_1 & PCGA_2	None	0.280120	18.2308	0.000004	0.67306	0.038559	0.189866	0.011297

PCGA_1(2) = First (second) principal component of California Real GSP and Aggregate Employment.
PCUGA_1(2) = First (second) principal component of California Unemployment Rate, California Real GSP, and Aggregate Employment.

EXHIBIT 6

PART 1

STATGRAPHICS PLUS REGRESSION RESULTS-MODEL #1

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−4.911830	1.070660	−4.58767	0.0010
CYIndBL	0.286573	0.069859	4.10215	0.0021
PCUGA_1	−0.209370	0.038628	−5.42019	0.0003
PCUGA_2	0.299701	0.170568	1.75708	0.1094
CumInjNDX	0.308297	0.042620	7.23363	0.0000

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	674.4880	4	168.6220	26.4462	0.0000
Residual	63.7604	10	6.3760		
Total (Corr.)	738.2484	14			

R-squared = 91.3633 percent
R-squared (adjusted for d.f.) = 87.9086 percent
Standard Error of Est. = 2.52508
Mean absolute error = 1.65922
Durbin–Watson statistic = 2.14752

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 4 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -4.91183 + 0.286573 * \text{CYIndBL} - 0.20937 * \text{PCUGA\_1}$$
$$+ 0.299701 * \text{PCUGA\_2} + 0.308297 * \text{CumInjNDX}.$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 91.3633% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 87.9086%. The standard error of the estimate shows the standard deviation of the residuals to be 2.52508. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 1.65922 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is greater than 1.4, there is probably not any serious autocorrelation in the residuals.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.1094, belonging to PCUGA_2. Since the P-value is greater or equal to 0.10, that term is not statistically significant at the 90% or higher confidence level. Consequently, you should consider removing PCUGA_2 from the model.

EXHIBIT 6

PART 2

STATGRAPHICS PLUS REGRESSION RESULTS-MODEL #2

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−3.580310	0.824978	−4.33988	0.0012
CYIndBL	0.261897	0.074644	3.50862	0.0049
PCUGA_1	−0.214998	0.041989	−5.12040	0.0003
CumInjNDX	0.301076	0.046272	6.50673	0.0000

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	654.8030	3	218.2677	28.77271	0.0000
Residual	83.4452	11	7.5859		
Total (Corr.)	738.2480	14			

R-squared = 88.6969 percent
R-squared (adjusted for d.f.) = 85.6142 percent
Standard Error of Est. = 2.75426
Mean absolute error = 1.95774
Durbin–Watson statistic = 1.71858

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 3 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -3.58031 + 0.261897 * \text{CYIndBL} - 0.214998 * \text{PCUGA\_1} + 0.301076 * \text{CumInjNDX}$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 88.6969% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 85.6142%. The standard error of the estimate shows the standard deviation of the residuals to be 2.75426. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 1.95774 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is greater than 1.4, there is probably not any serious autocorrelation in the residuals.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0049, belonging to CYIndBL. Since the P-value is less than 0.01, the highest order term is statistically significant at the 99% confidence level. Consequently, you probably don't want to remove any variables from the model.

EXHIBIT 6

PART 3

STATGRAPHICS PLUS REGRESSION RESULTS-MODEL #3

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−7.726190	1.297840	−5.95310	0.0001
CYIndBL	0.272918	0.084933	3.21332	0.0093
PCGA_1	0.649210	0.141971	4.57282	0.0010
PCGA_2	0.584624	0.442156	1.32221	0.2155
CumInjNDX	0.290403	0.051592	5.62879	0.0002

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	650.7490	4	162.6873	18.5931	0.0001
Residual	87.4989	10	8.7499		
Total (Corr.)	738.2479	14			

R-squared = 88.1478 percent
R-squared (adjusted for d.f.) = 83.4069 percent
Standard Error of Est. = 2.95802
Mean absolute error = 1.9507
Durbin–Watson statistic = 2.07557

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 4 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -7.72619 + 0.272918 * \text{CYIndBL} + 0.64921 * \text{PCGA\_1} + 0.584624 * \text{PCGA\_2} + 0.290403 * \text{CumInjNDX}$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 88.1478% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 83.4069%. The standard error of the estimate shows the standard deviation of the residuals to be 2.95802. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 1.9507 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is greater than 1.4, there is probably not any serious autocorrelation in the residuals.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.2155, belonging to PCGA_2. Since the P-value is greater or equal to 0.10, that term is not statistically significant at the 90% or higher confidence level. Consequently, you should consider removing PCGA_2 from the model.

EXHIBIT 6

PART 4

STATGRAPHICS PLUS REGRESSION RESULTS-MODEL #4

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−7.726180	1.297840	−5.95310	0.0001
CYIndBL	0.272919	0.084933	3.21332	0.0093
CYrGSP	0.769158	0.420688	1.82834	0.0974
CYAggE	0.414309	0.196672	2.10660	0.0614
CumInjNDX	0.290403	0.051592	5.62879	0.0002

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	650.7490	4	162.6873	18.5931	0.0001
Residual	87.4989	10	8.7499		
Total (Corr.)	738.2479	14			

R-squared = 88.1478 percent
R-squared (adjusted for d.f.) = 83.4069 percent
Standard Error of Est. = 2.95802
Mean absolute error = 1.9507
Durbin–Watson statistic = 2.07557

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 4 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -7.72618 + 0.272919 * \text{CYIndBL} + 0.769158 * \text{CYrGSP}$$
$$+ 0.414309 * \text{CYAggE} + 0.290403 * \text{CumInjNDX}.$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 88.1478% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 83.4069%. The standard error of the estimate shows the standard deviation of the residuals to be 2.95802. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 1.9507 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is greater than 1.4, there is probably not any serious autocorrelation in the residuals.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0974, belonging to CYrGSP. Since the P-value is less than 0.10, that term is statistically significant at the 90% confidence level. Depending on the confidence level at which you wish to work, you may or may not decide to remove CYrGSP from the model.

EXHIBIT 6

PART 5

STATGRAPHICS PLUS REGRESSION RESULTS-MODEL #5

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−6.852850	1.154560	−5.93544	0.0001
CYIndBL	0.309052	0.083107	3.71872	0.0034
PCGA_1	0.642720	0.146633	4.38319	0.0011
CumInjNDX	0.308337	0.051443	5.99380	0.0001

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	635.4530	3	211.8177	22.6662	0.0001
Residual	102.7960	11	9.3451		
Total (Corr.)	738.2490	14			

R-squared = 86.0757 percent
R-squared (adjusted for d.f.) = 82.2782 percent
Standard Error of Est. = 3.05697
Mean absolute error = 2.09204
Durbin–Watson statistic = 2.27098

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 3 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -6.85285 + 0.309052 * \text{CYIndBL} + 0.64272 * \text{PCGA\_1} + 0.308337 * \text{CumInjNDX}$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 86.0757% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 82.2782%. The standard error of the estimate shows the standard deviation of the residuals to be 3.05697. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 2.09204 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is greater than 1.4, there is probably not any serious autocorrelation in the residuals.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0034, belonging to CYIndBL. Since the P-value is less than 0.01, the highest order term is statistically significant at the 99% confidence level. Consequently, you probably don't want to remove any variables from the model.

EXHIBIT 6

PART 6

STATGRAPHICS PLUS REGRESSION RESULTS-MODEL #6

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−6.384760	1.179070	−5.41509	0.0002
CYIndBL	0.321087	0.088928	3.61065	0.0041
CYAggE	0.648742	0.164242	3.94990	0.0023
CumInjNDX	0.314359	0.054959	5.71994	0.0001

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	621.5000	3	207.1667	19.5192	0.0001
Residual	116.7480	11	10.6135		
Total (Corr.)	738.2480	14			

R-squared = 84.1858 percent
R-squared (adjusted for d.f.) = 79.8728 percent
Standard Error of Est. = 3.25783
Mean absolute error = 2.23537
Durbin–Watson statistic = 2.22488

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 3 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -6.38476 + 0.321087 * \text{CYIndBL} + 0.648742 * \text{CYAggE} + 0.314359 * \text{CumInjNDX}$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 84.1858% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 79.8728%. The standard error of the estimate shows the standard deviation of the residuals to be 3.25783. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 2.23537 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is greater than 1.4, there is probably not any serious autocorrelation in the residuals.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0041, belonging to CYIndBL. Since the P-value is less than 0.01, the highest order term is statistically significant at the 99% confidence level. Consequently, you probably don't want to remove any variables from the model.

EXHIBIT 6

PART 7

STATGRAPHICS PLUS REGRESSION RESULTS-MODEL #7

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−7.771980	1.486670	−5.22778	0.0003
CYIndBL	0.220530	0.093040	2.37028	0.0371
CYAggE	1.346940	0.365450	3.68569	0.0036
CumInjNDX	0.264735	0.057435	4.60930	0.0008

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	611.9200	3	203.9733	17.7608	0.0002
Residual	126.3290	11	11.4845		
Total (Corr.)	738.2490	14			

R-squared = 82.888 percent
R-squared (adjusted for d.f.) = 78.2211 percent
Standard Error of Est. = 3.38887
Mean absolute error = 2.47812
Durbin–Watson statistic = 1.39268

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 3 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -7.77198 + 0.22053*\text{CYIndBL} + 1.34694*\text{CYrGSP} + 0.264735*\text{CumInjNDX}$$

Since the P-value in the ANOVA table is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Squared statistic indicates that the model as fitted explains 82.888% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 78.2211%. The standard error of the estimate shows the standard deviation of the residuals to be 3.38887. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 2.47812 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is less than 1.4, there may be some indication of serial correlation. Plot the residuals versus row order to see if there is any pattern which can be seen.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0371, belonging to CYIndBL. Since the P-value is less than 0.05, that term is statistically significant at the 95% confidence level. Consequently, you probably don't want to remove any variables from the model.

EXHIBIT 7

PART 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #2
ACTUAL VS FITTED CHANGES

PERCENT CHANGE INDEMNITY FREQUENCY = f(Ind BL, PCUGA_1, Cum Inj Index)

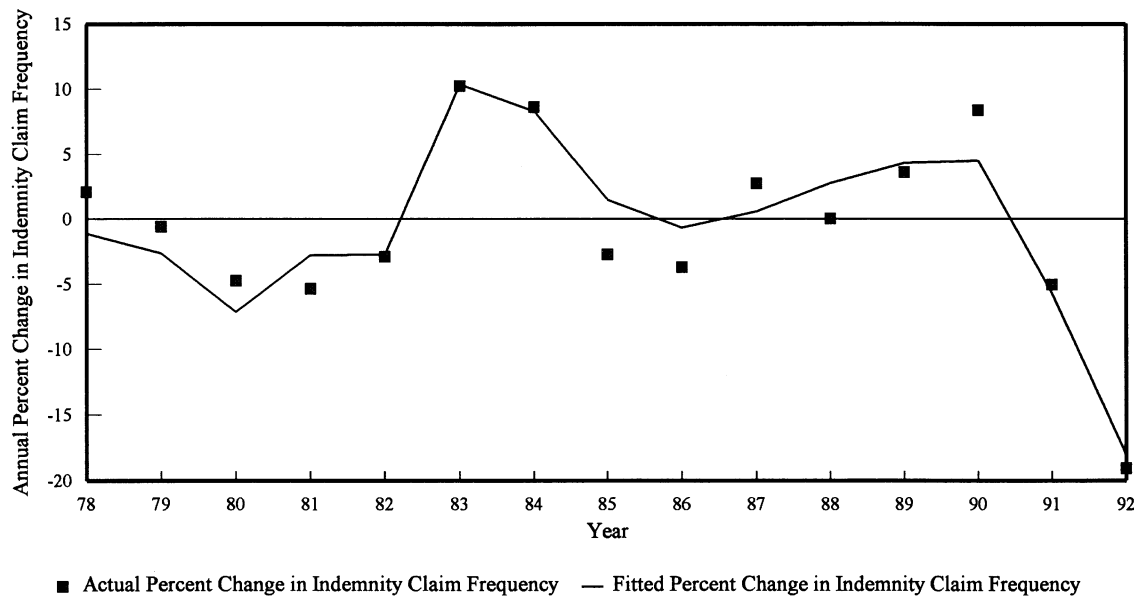


EXHIBIT 7

PART 2—PAGE 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #2
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

PROJECTION INTERVAL = 1 YEAR

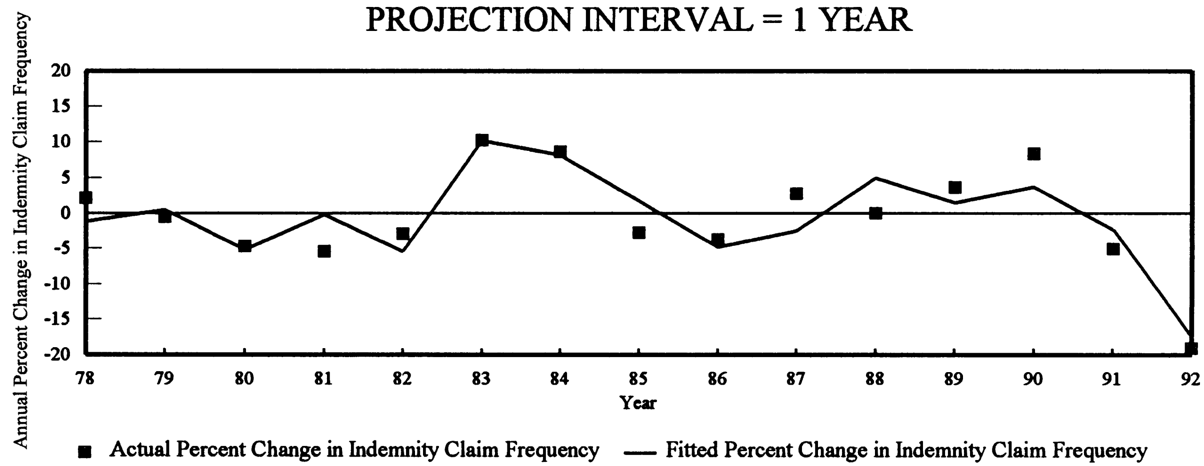


EXHIBIT 7

PART 2—PAGE 2

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #2
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

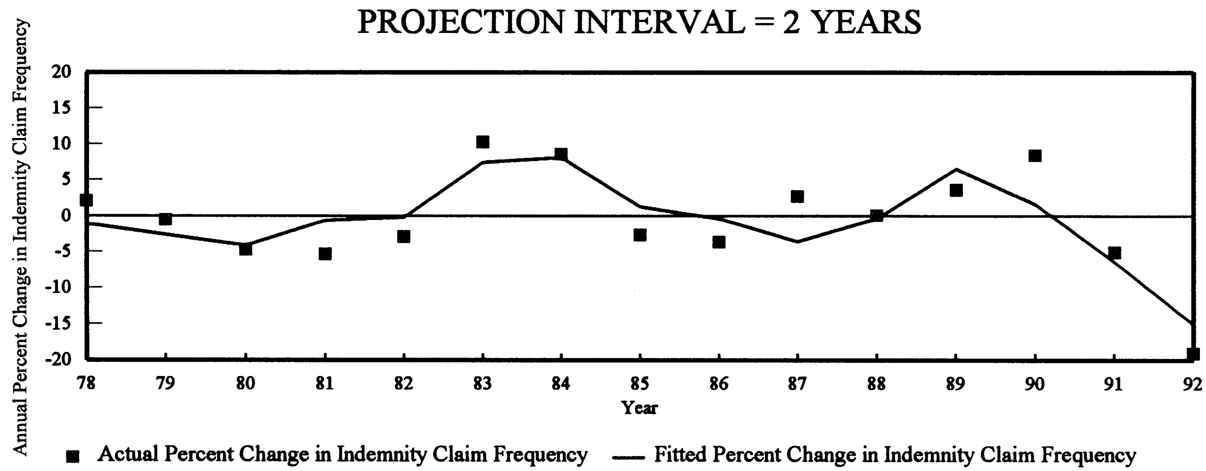


EXHIBIT 7

PART 2—PAGE 3

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #2
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

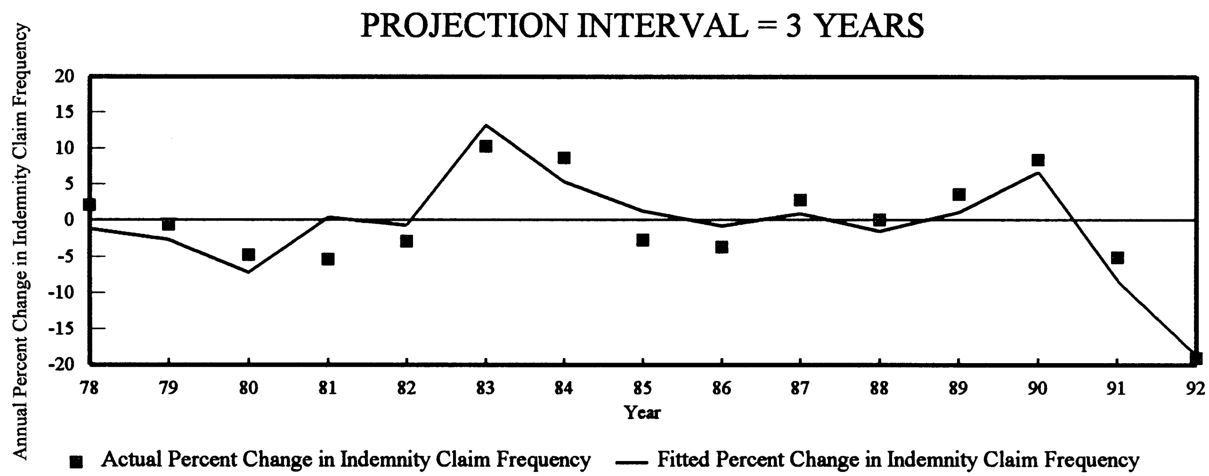


EXHIBIT 7

PART 3—PAGE 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #2
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

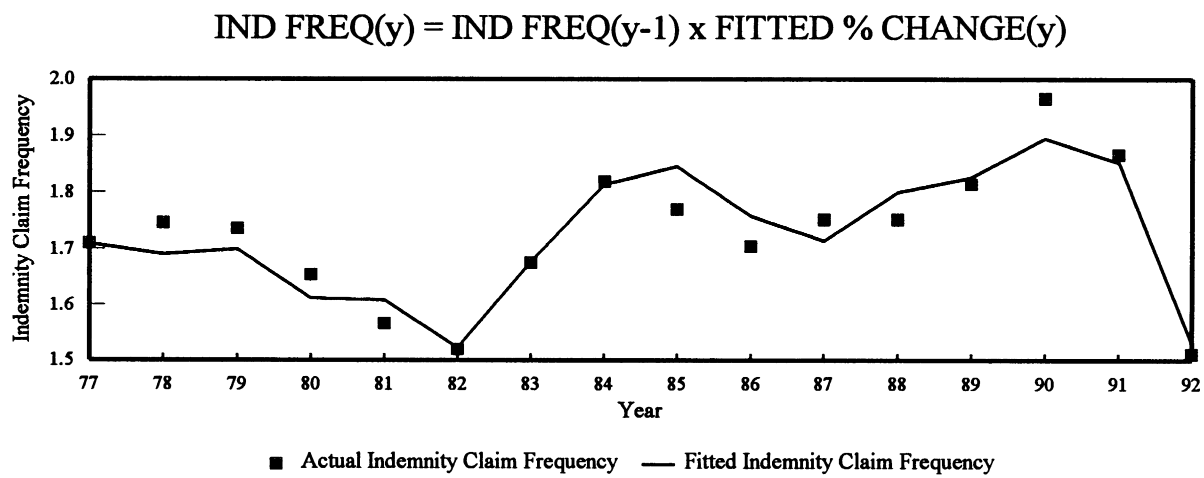


EXHIBIT 7

PART 3—PAGE 2

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #2
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

IND FREQ(y) = IND FREQ(y-2) x FITTED {% CHG(y-1) x % CHG(y)}

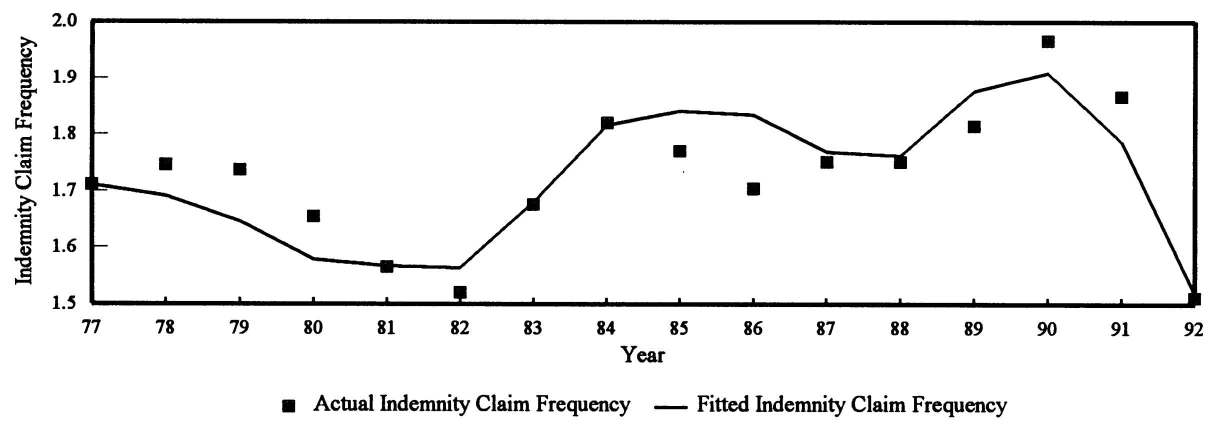


EXHIBIT 7

PART 3—PAGE 3

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #2
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

IND FREQ(y) = IND FREQ(y-3) x FITTED {% CHG(y-2) x % CHG(y-1) x % CHG(y)}

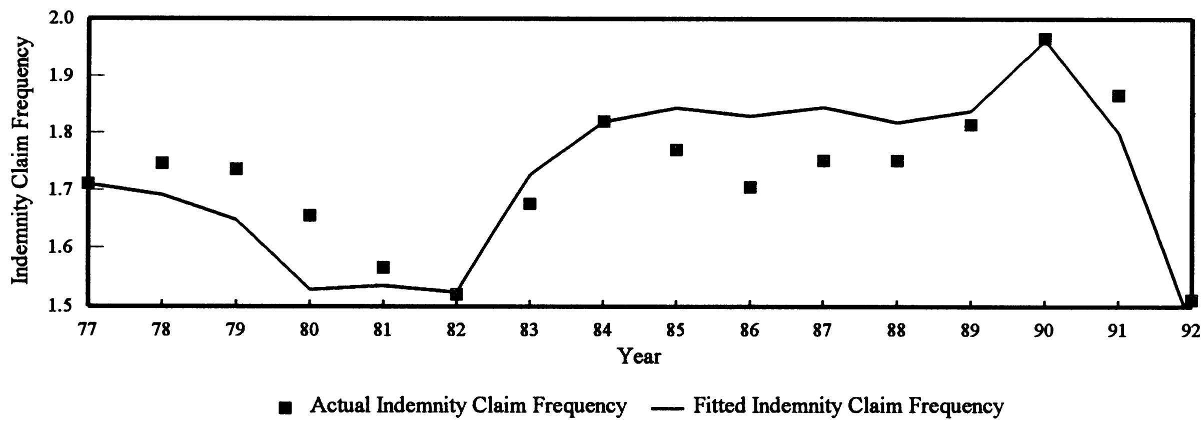


EXHIBIT 8

PART 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #5
ACTUAL VS FITTED CHANGES

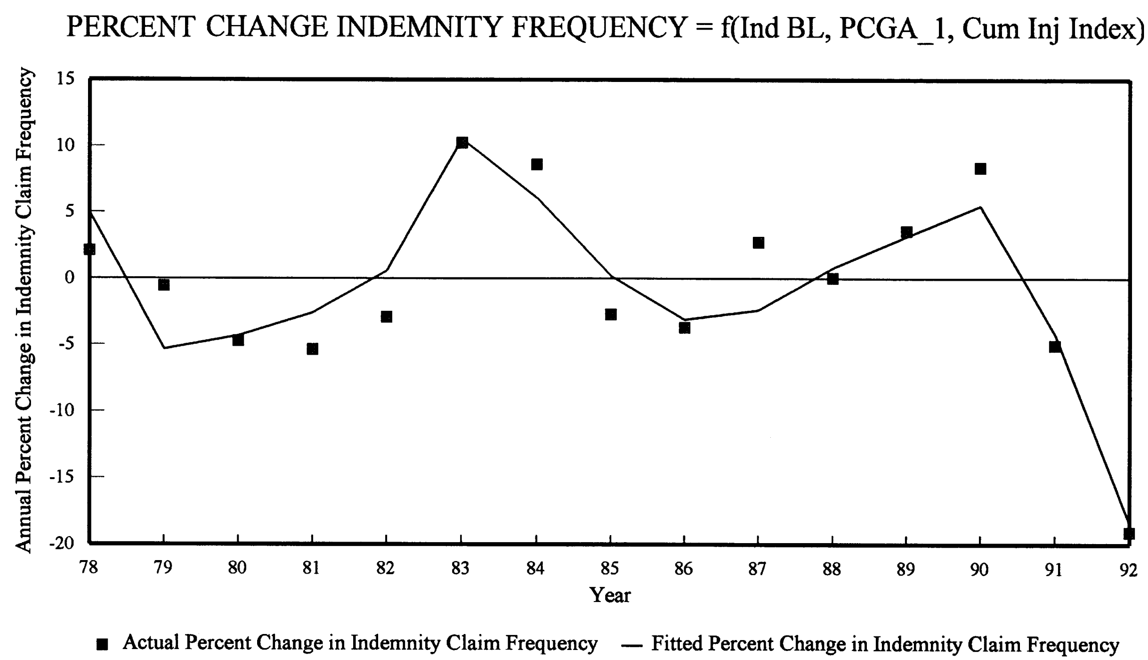


EXHIBIT 8

PART 2—PAGE 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #5
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

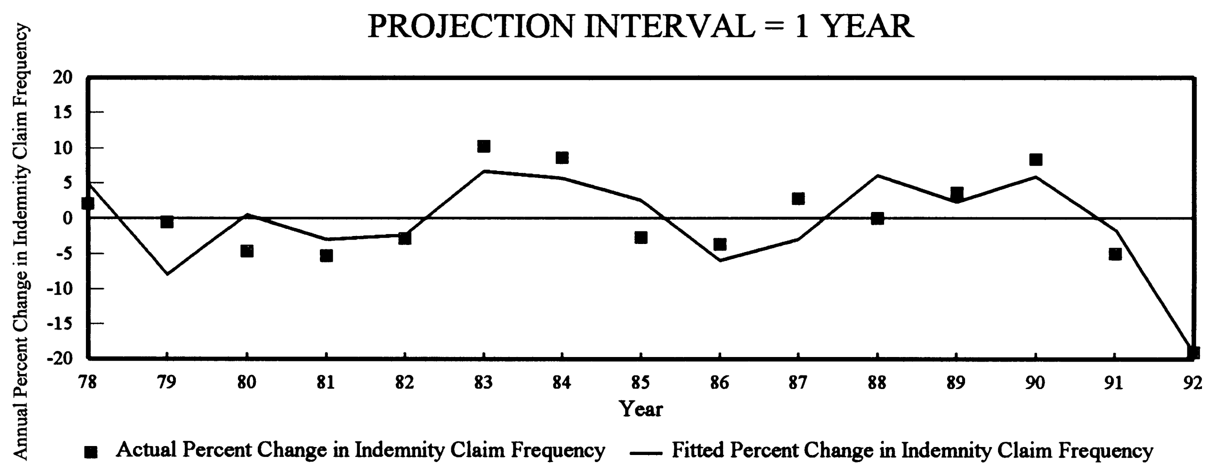


EXHIBIT 8

PART 2—PAGE 2

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #5
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

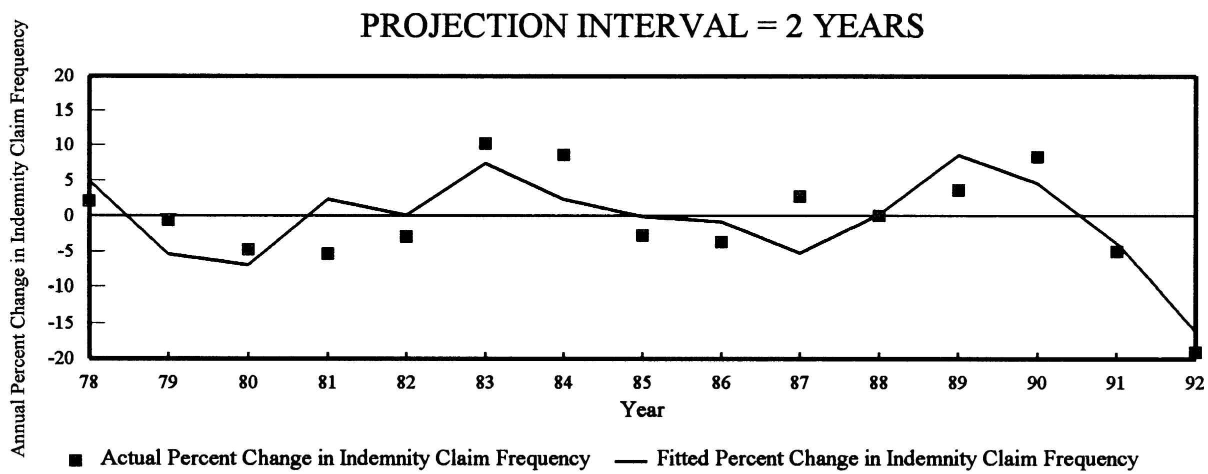


EXHIBIT 8

PART 2—PAGE 3

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #5
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

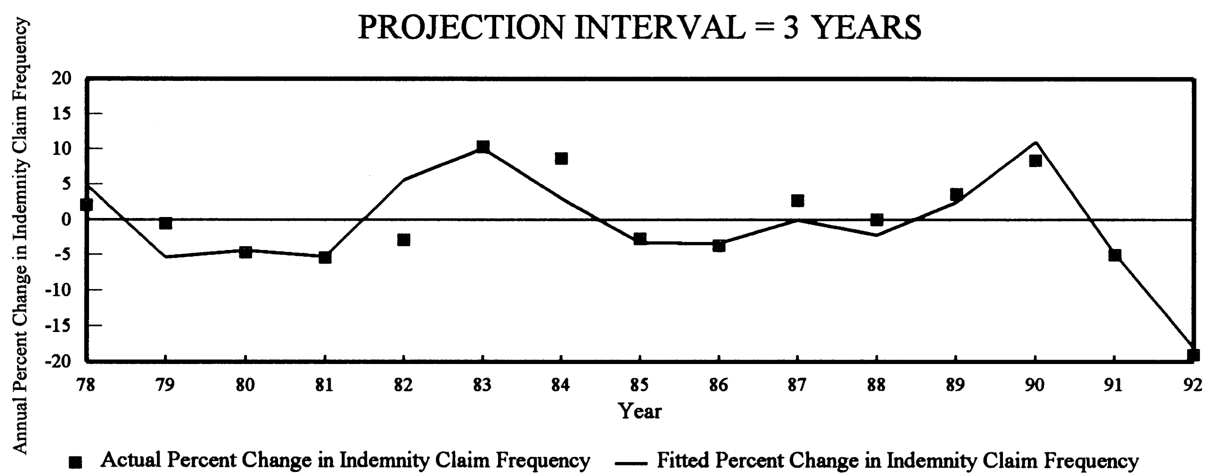


EXHIBIT 8

PART 3—PAGE 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #5
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

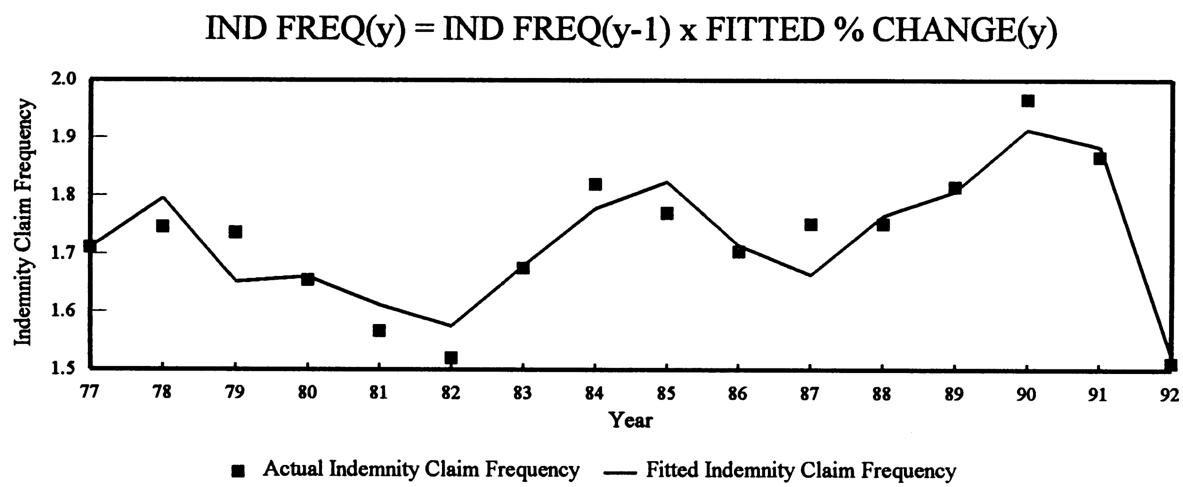


EXHIBIT 8

PART 3—PAGE 2

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #5
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

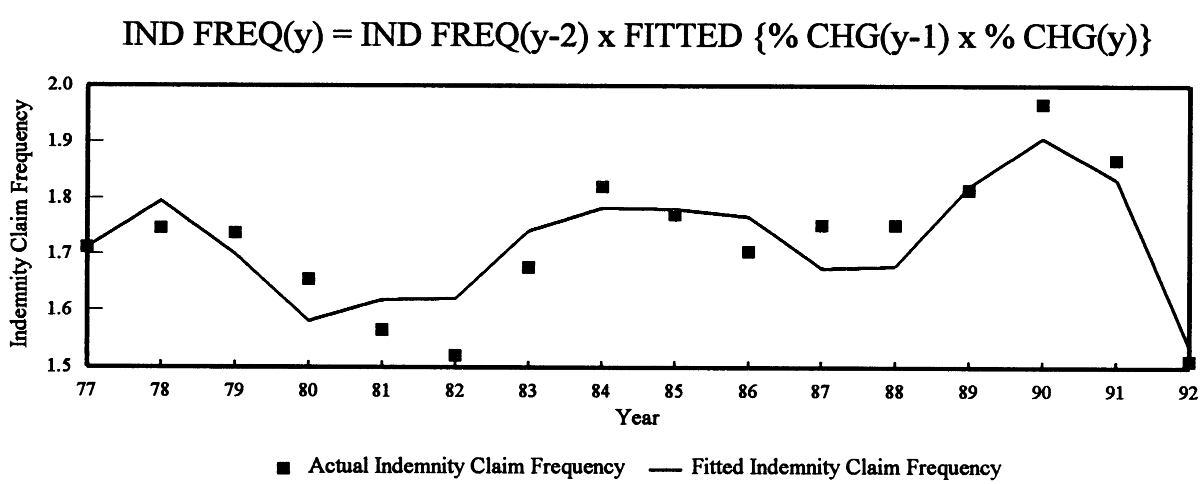


EXHIBIT 8

PART 3—PAGE 3

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #5
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

IND FREQ(y) = IND FREQ(y-3) x FITTED {% CHG(y-2) x % CHG(y-1) x % CHG(y)}

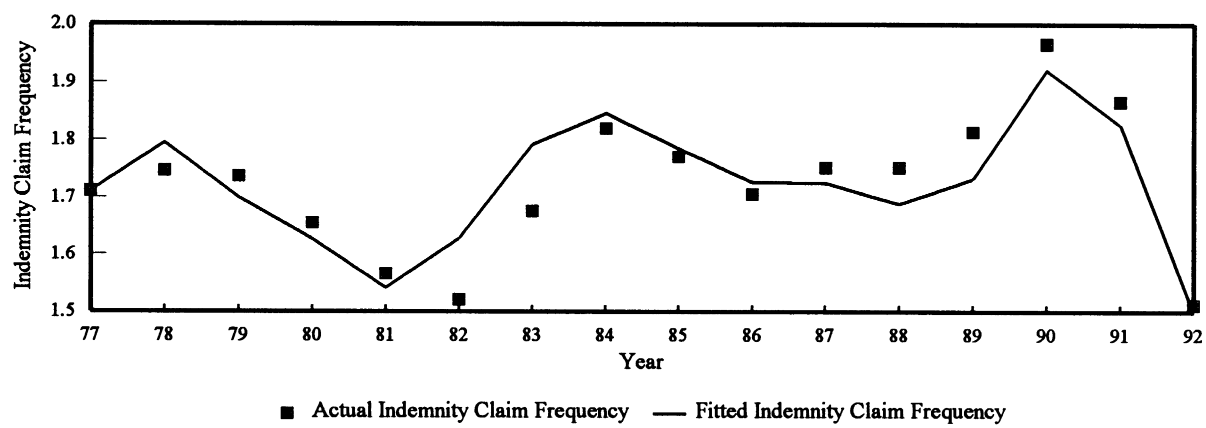


EXHIBIT 9

PART 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #6

ACTUAL VS FITTED CHANGES

PERCENT CHANGE INDEMNITY FREQUENCY = f(Ind BL, AggE, Cum Inj Index)

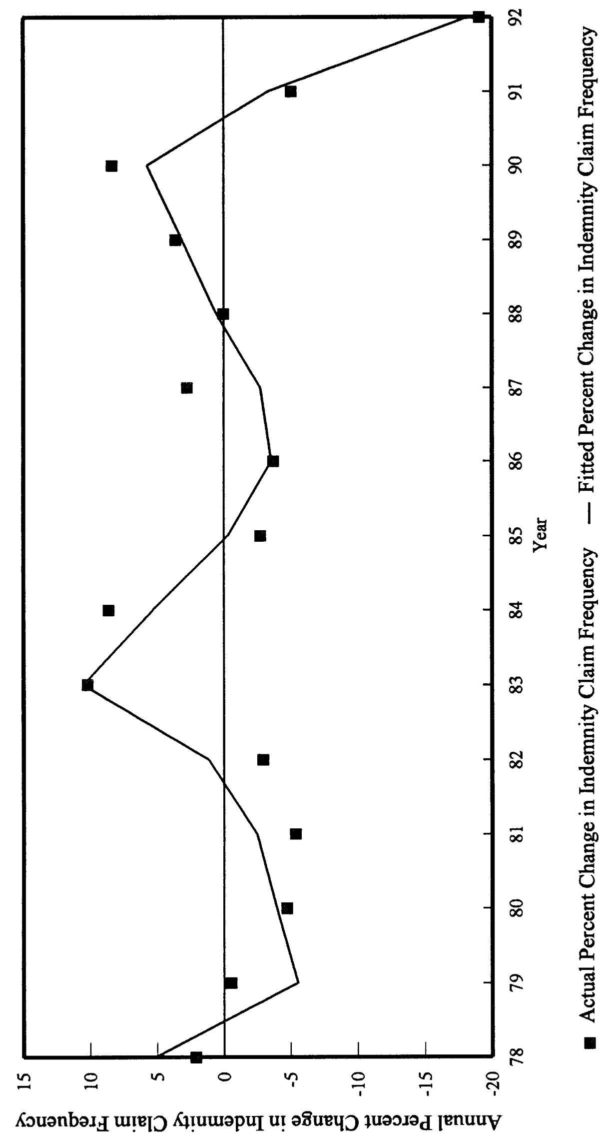


EXHIBIT 9

PART 2—PAGE 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #6
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

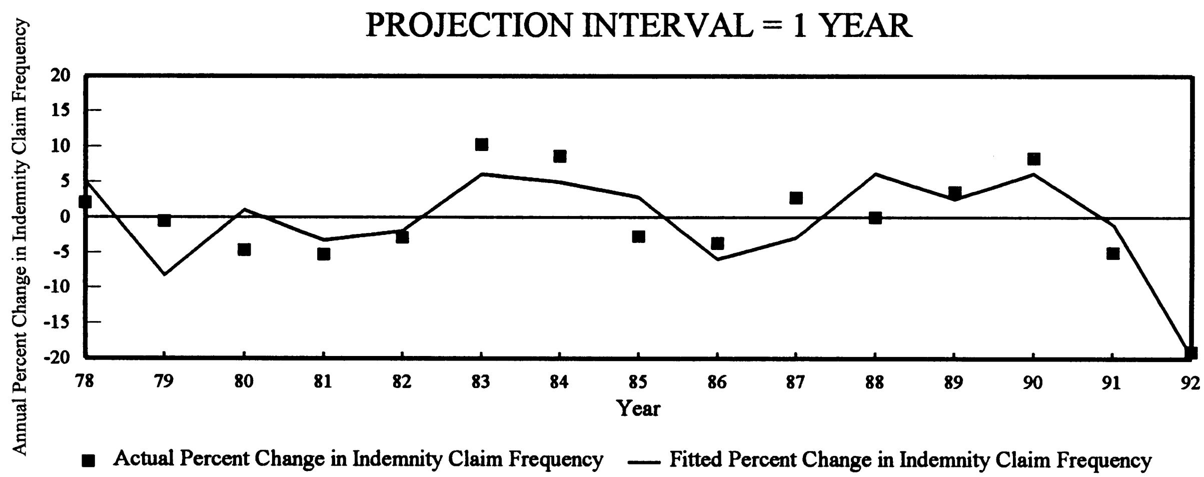


EXHIBIT 9

PART 2—PAGE 2

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #6
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

PROJECTION INTERVAL = 2 YEARS

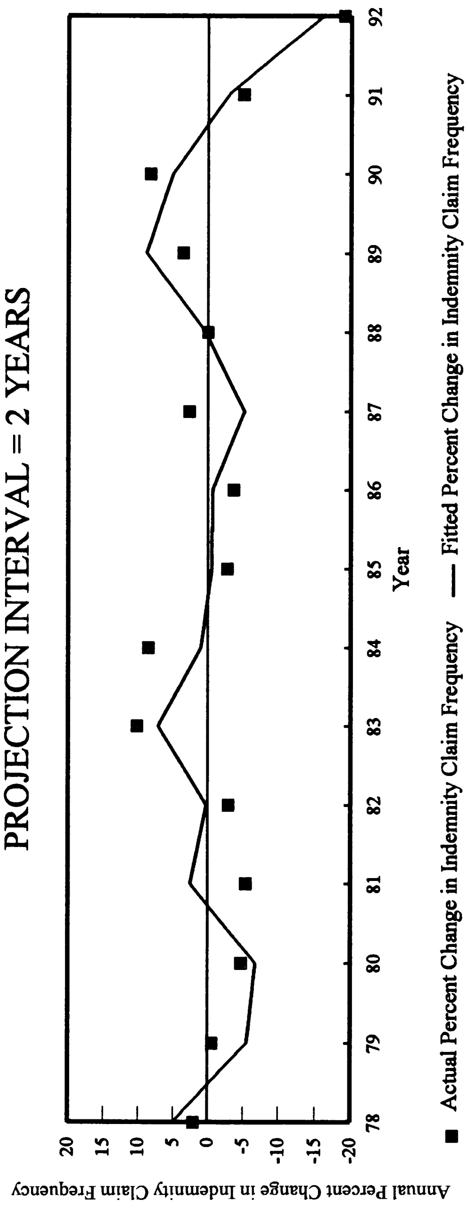


EXHIBIT 9

PART 2—PAGE 3

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #6
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

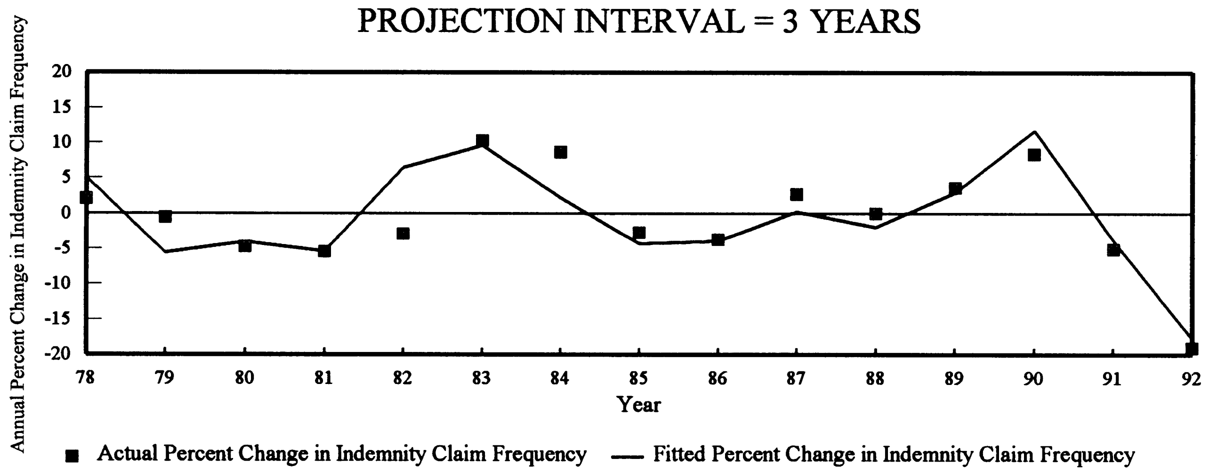


EXHIBIT 9

Part 3—Page 1

Graphical Analysis of Fitted Models: Model #6
Actual vs Fitted Indemnity Claim Frequency at Selected Projection Intervals

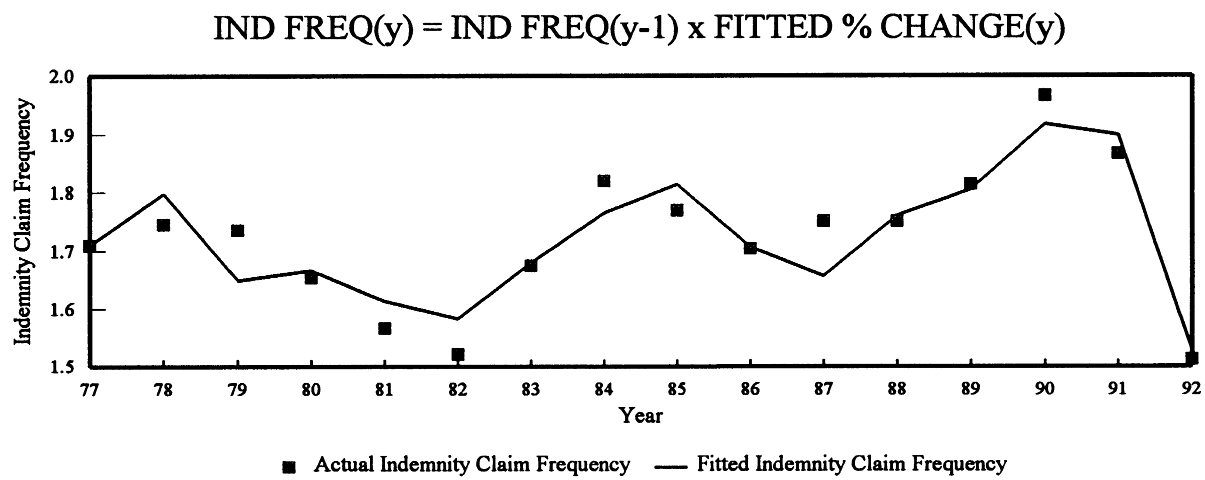


EXHIBIT 9

PART 3—PAGE 2

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #6
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

IND FREQ(y) = IND FREQ(y-2) x FITTED {% CHG(y-1) x % CHG(y)}

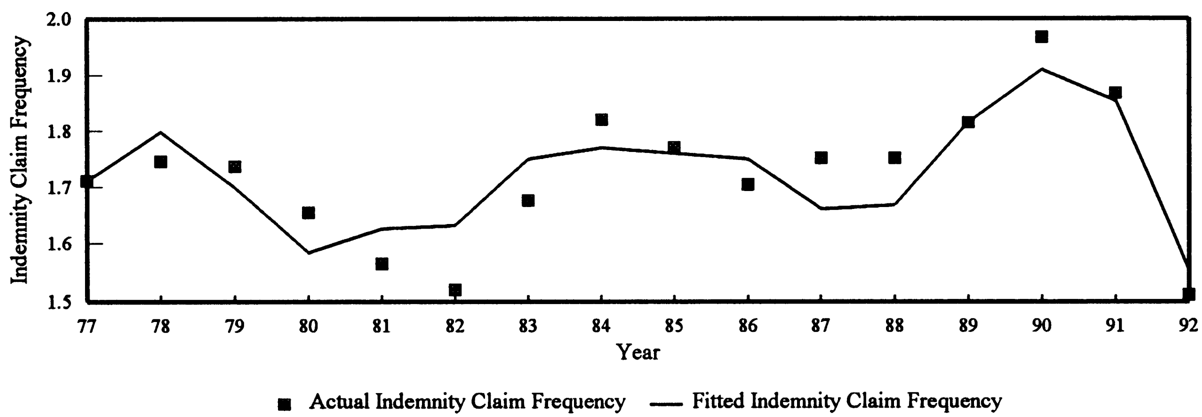


EXHIBIT 9

PART 3—PAGE 3

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #6
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

IND FREQ(y) = IND FREQ(y-3) x FITTED {% CHG(y-2) x % CHG(y-1) x % CHG(y)}

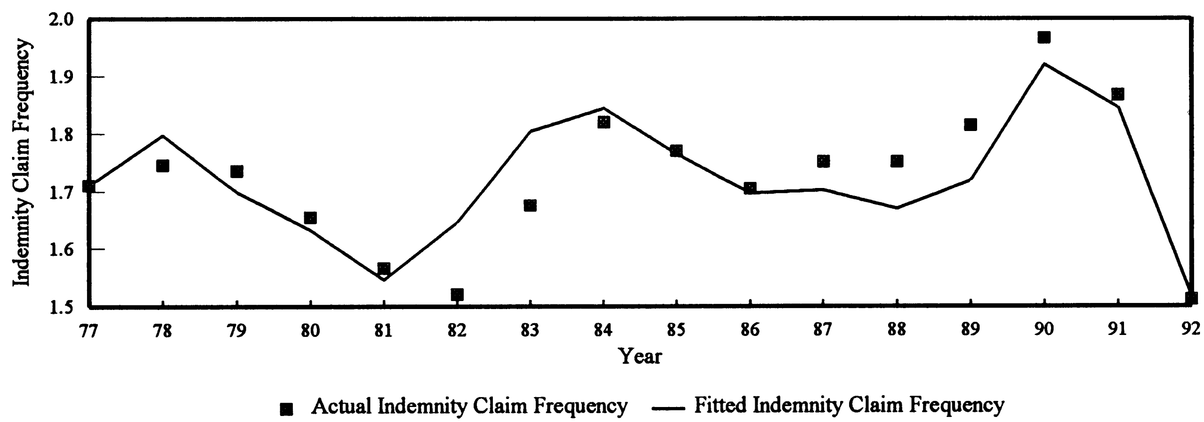


EXHIBIT 10

PART 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #7
ACTUAL VS FITTED CHANGES

PERCENT CHANGE INDEMNITY FREQUENCY = f(Ind BL, rGSP, Cum Inj Index)

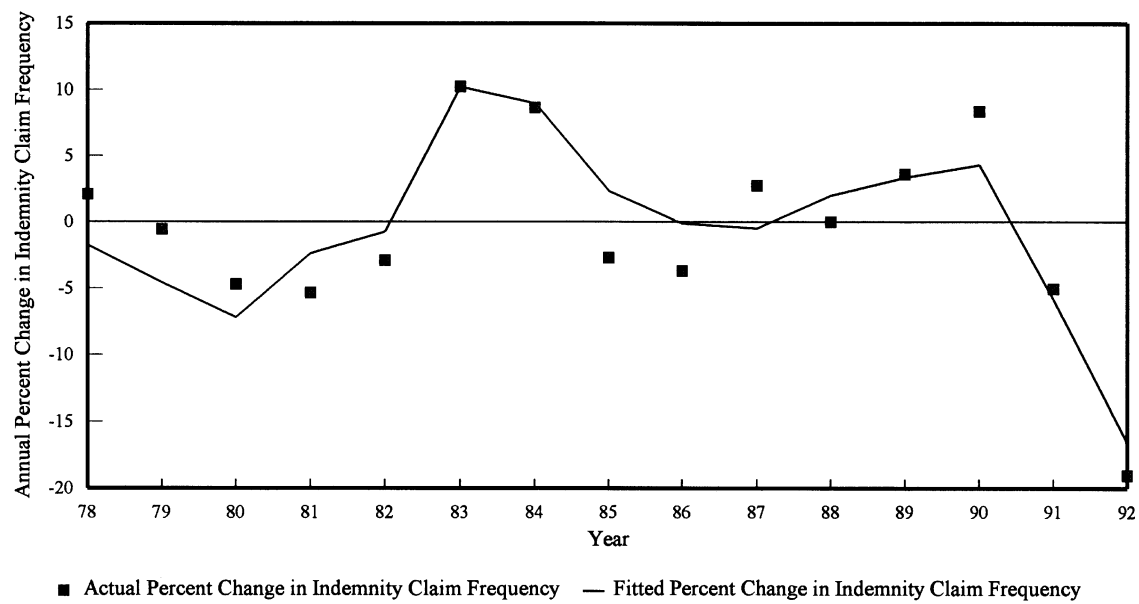


EXHIBIT 10

PART 2—PAGE 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #7
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

PROJECTION PERIOD = 1 YEAR

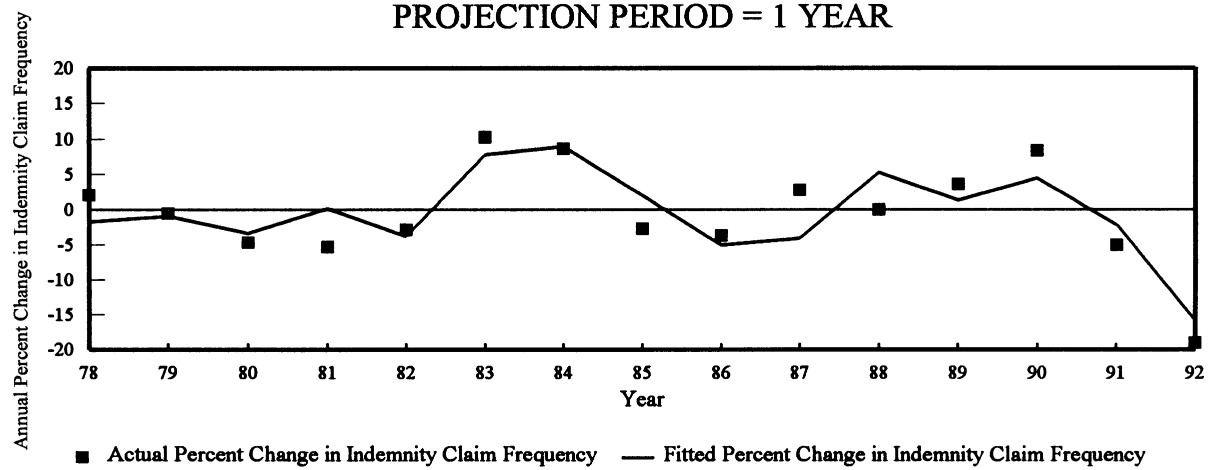


EXHIBIT 10

PART 2—PAGE 2

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #7
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

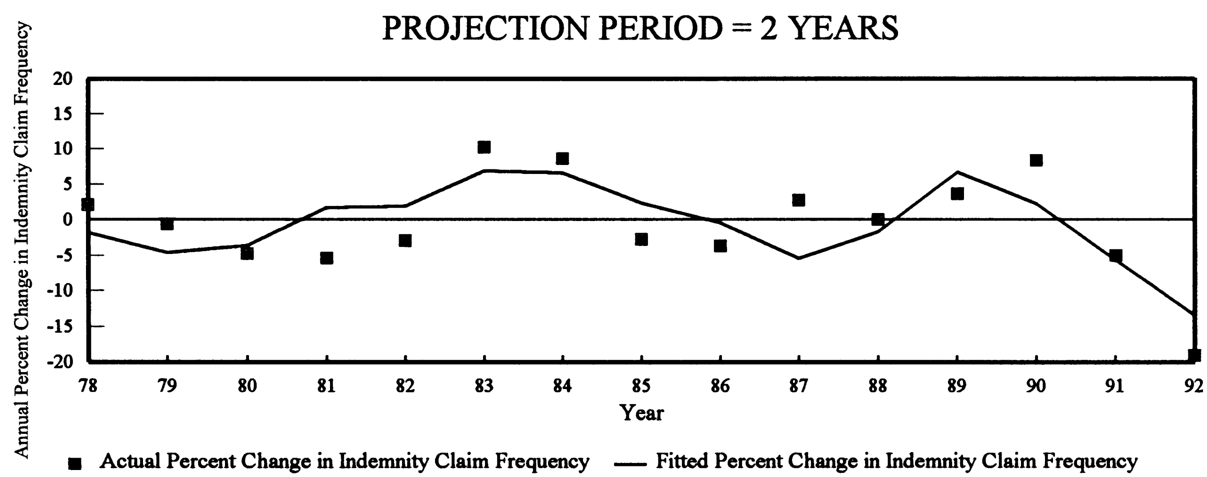


EXHIBIT 10

PART 2—PAGE 3

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #7
ACTUAL VS FITTED ANNUAL CHANGES AT SELECTED PROJECTION INTERVALS

PROJECTION PERIOD = 3 YEARS

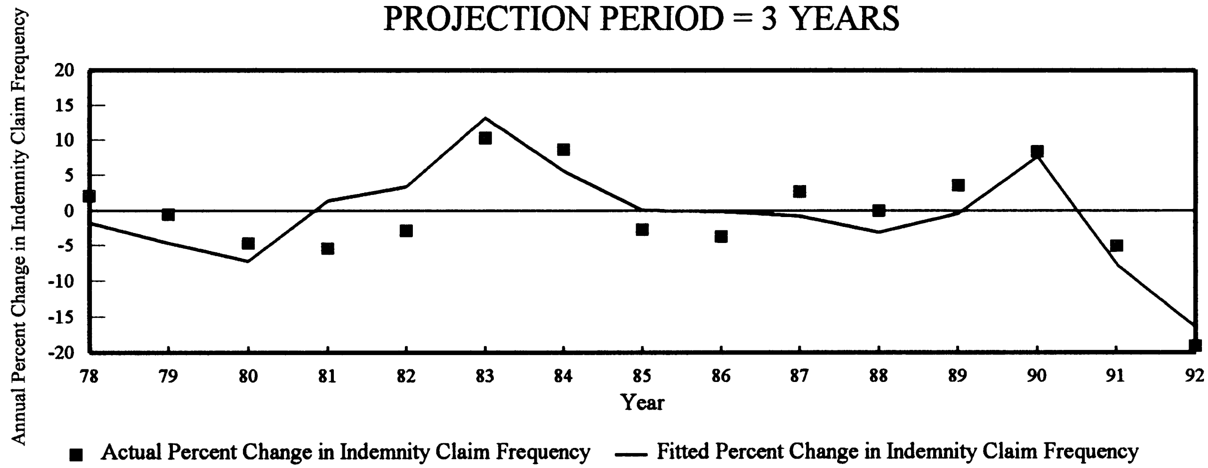


EXHIBIT 10

PART 3—PAGE 1

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #7
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

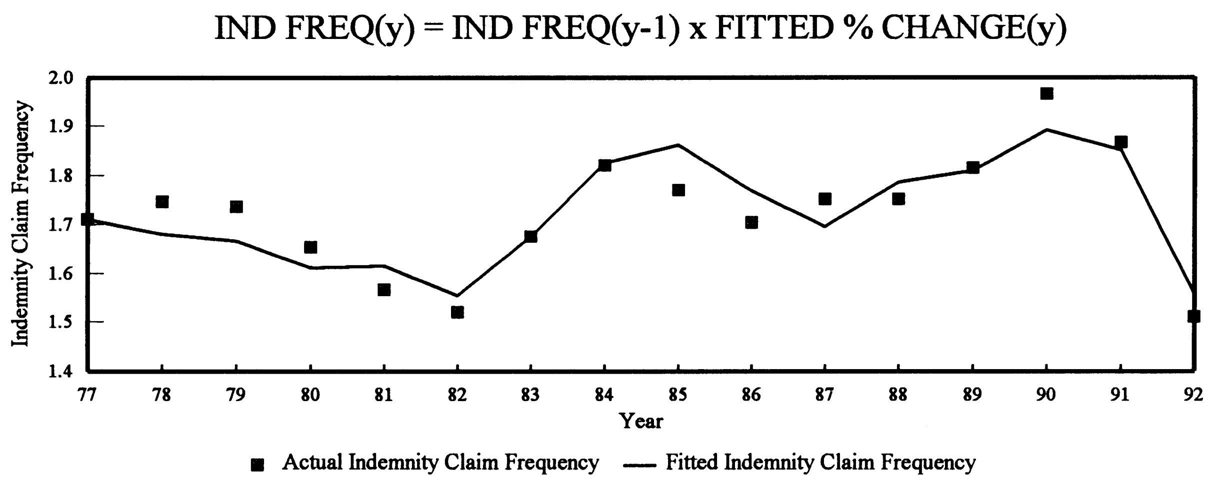


EXHIBIT 10

PART 3—PAGE 2

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #7
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

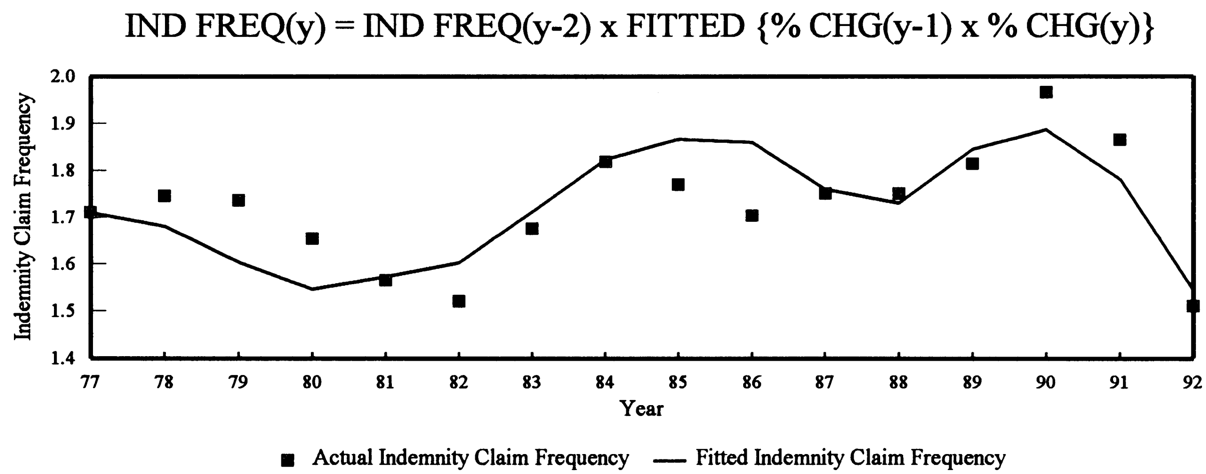


EXHIBIT 10

PART 3—PAGE 3

GRAPHICAL ANALYSIS OF FITTED MODELS: MODEL #7
ACTUAL VS FITTED INDEMNITY CLAIM FREQUENCY AT SELECTED PROJECTION INTERVALS

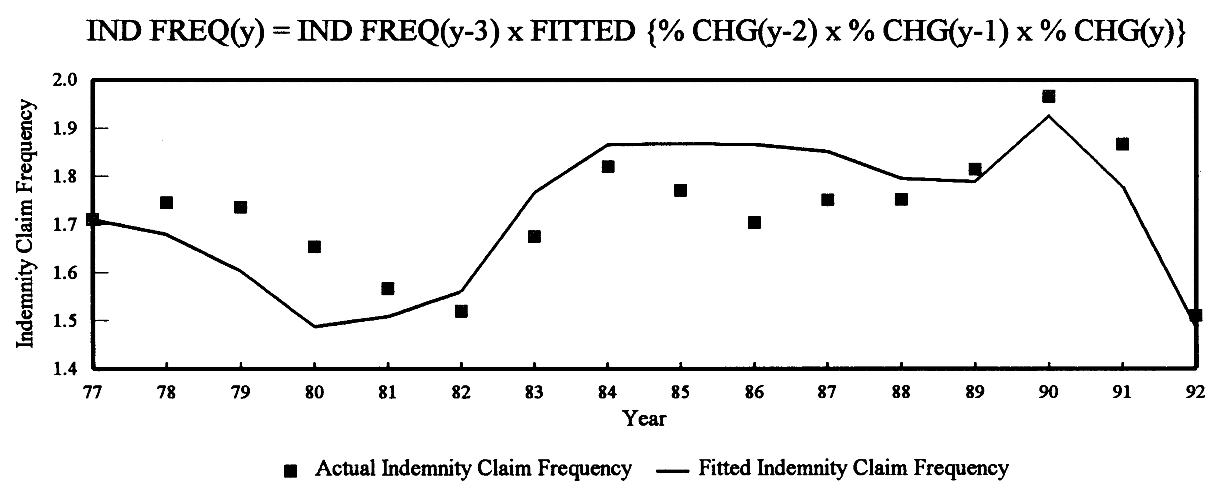


EXHIBIT 11

PART 1

STATGRAPHICS PLUS REGRESSION RESULTS

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−1.579230	1.698230	−0.92993	0.3610
CYIndBL	0.321818	0.153038	2.10287	0.0453
PCGA_1	0.477622	0.240282	1.98775	0.0575

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	360.9690	2	180.4845	4.574216	0.0199
Residual	1025.8800	26	39.4569		
Total (Corr.)	1386.8490	28			

R-squared = 26.028 percent
R-squared (adjusted for d.f.) = 20.3379 percent
Standard Error of Est. = 6.28147
Mean absolute error = 4.3111
Durbin–Watson statistic = 1.19885

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 2 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -1.57923 + 0.321818 * \text{CYIndBL} + 0.477622 * \text{PCGA\_1}$$

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

The R-Squared statistic indicates that the model as fitted explains 26.028% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 20.3379%. The standard error of the estimate shows the standard deviation of the residuals to be 6.28147. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 4.3111 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is less than 1.4, there may be some indication of serial correlation. Plot the residuals versus row order to see if there is any pattern which can be seen.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0575, belonging to PCGA_1. Since the P-value is less than 0.10, that term is statistically significant at the 90% confidence level. Depending on the confidence level at which you wish to work, you may or may not decide to remove PCGA_1 from the model.

EXHIBIT 11

PART 2

STATGRAPHICS PLUS REGRESSION RESULTS

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−1.188410	1.610480	−0.73792	0.4672
CYIndBL	0.330217	0.153726	2.14809	0.0412
CYAggE	0.481486	0.254616	1.89103	0.0698

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	347.9560	2	173.9780	4.354097	0.0234
Residual	1038.8900	26	39.9573		
Total (Corr.)	1386.8460	28			

R-squared = 25.0897 percent
R-squared (adjusted for d.f.) = 19.3274 percent
Standard Error of Est. = 6.32119
Mean absolute error = 4.36516
Durbin–Watson statistic = 1.19294

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 2 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -1.18841 + 0.330217 * \text{CYIndBL} + 0.481486 * \text{CYAggE}.$$

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

The R-Squared statistic indicates that the model as fitted explains 25.0897% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 19.3274%. The standard error of the estimate shows the standard deviation of the residuals to be 6.32119. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 4.36516 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is less than 1.4, there may be some indication of serial correlation. Plot the residuals versus row order to see if there is any pattern which can be seen.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0698, belonging to CYAggE. Since the P-value is less than 0.10, that term is statistically significant at the 90% confidence level. Depending on the confidence level at which you wish to work, you may or may not decide to remove CYAggE from the model.

EXHIBIT 11

PART 3

STATGRAPHICS PLUS REGRESSION RESULTS

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	−2.593800	2.146440	−1.20842	0.2378
CYIndBL	0.287312	0.156838	1.83191	0.0784
CYrGSP	1.016480	0.539883	1.88279	0.0710

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	346.8620	2	173.4310	4.3358	0.0237
Residual	1039.9850	26	39.9994		
Total (Corr.)	1386.8470	28			

R-squared = 25.0108 percent
R-squared (adjusted for d.f.) = 19.2424 percent
Standard Error of Est. = 6.32451
Mean absolute error = 4.08829
Durbin–Watson statistic = 1.09492

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 2 independent variables. The equation of the fitted model is

$$\text{IndFrq} = -2.5938 + 0.287312 * \text{CYIndBL} + 1.01648 * \text{CYrGSP}.$$

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

The R-Squared statistic indicates that the model as fitted explains 25.0108% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 19.2424%. The standard error of the estimate shows the standard deviation of the residuals to be 6.32451. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 4.08829 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is less than 1.4, there may be some indication of serial correlation. Plot the residuals versus row order to see if there is any pattern which can be seen.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0784, belonging to CYIndBL. Since the P-value is less than 0.10, that term is statistically significant at the 90% confidence level. Depending on the confidence level at which you wish to work, you may or may not decide to remove CYIndBL from the model.

EXHIBIT 11

PART 4

STATGRAPHICS PLUS REGRESSION RESULTS

Multiple Regression Analysis

Dependent variable: IndFrq

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	0.938789	1.304640	0.71958	0.4782
CYIndBL	0.316254	0.154940	2.04114	0.0515
PCUGA_1	−0.125809	0.068556	−1.83512	0.0780

Analysis of Variance

Source	Sum of Squares	Degrees of Freedom	Mean Square Error	F-Ratio	P-Value
Model	340.5860	2	170.2930	4.2319	0.0256
Residual	1046.2600	26	40.2408		
Total (Corr.)	1386.8460	28			

R-squared = 24.5583 percent
R-squared (adjusted for d.f.) = 18.7551 percent
Standard Error of Est. = 6.34357
Mean absolute error = 4.26624
Durbin–Watson statistic = 0.989726

The StatAdvisor

The output shows the results of fitting a multiple linear regression model to describe the relationship between IndFrq and 2 independent variables. The equation of the fitted model is

$$IndFrq = 0.938789 + 0.316254 * CYIndBL - 0.125809 * PCUGA\_1.$$

Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95% confidence level.

The R-Squared statistic indicates that the model as fitted explains 24.5583% of the variability in IndFrq. The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 18.7551%. The standard error of the estimate shows the standard deviation of the residuals to be 6.34357. This value can be used to construct prediction limits for new observations by selecting the Reports option from the text menu. The mean absolute error (MAE) of 4.26624 is the average value of the residuals. The Durbin–Watson (DW) statistic tests the residuals to determine if there is any significant correlation based on the order in which they occur in your data file. Since the DW value is less than 1.4, there may be some indication of serial correlation. Plot the residuals versus row order to see if there is any pattern which can be seen.

In determining whether the model can be simplified, notice that the highest P-value on the independent variables is 0.0780, belonging to PCUGA_1. Since the P-value is less than 0.10, that term is statistically significant at the 90% confidence level. Depending on the confidence level at which you wish to work, you may or may not decide to remove PCUGA_1 from the model.

EXHIBIT 12

Summary of Performance Measures for Selected Models Indemnity Benefit Level

Independent Variables			Performance Measures		Projection Period = 1 Year		Projection Period = 2 Years		Projection Period = 3 Years	
			Average Absolute Error	Adjusted R^2 ($\times 100$)	Average Absolute Error	R^2 ($\times 100$)	Average Absolute Error	R^2 ($\times 100$)	Average Absolute Error	R^2 ($\times 100$)
Constant	PCUGA_1 & PCUGA_2	Cum Inj Index	1.6592	87.9086	2.5386	81.6704	2.8823	76.1449	1.9625	90.1574
Constant	PCUGA_1	Cum Inj Index	1.9577	85.6142	2.6385	79.3143	3.0434	74.6150	2.6673	83.3440
Constant	PCGA_1 & PCGA_2	Cum Inj Index	1.9507	83.4069	2.9959	75.5453	3.4739	66.8418	2.3567	83.9297
Constant	rGSP & AggE	Cum Inj Index	1.9507	83.4069	2.9959	75.5453	3.4739	66.8418	2.3567	83.9297
Constant	PCGA_1	Cum Inj Index	2.0920	82.2782	3.4225	69.2466	3.7406	62.9808	2.2048	79.8725
Constant	AggE	Cum Inj Index	2.2354	79.8728	3.6540	66.4066	3.9143	59.7829	2.5036	76.1922
Constant	rGSP	Cum Inj Index	2.4781	78.2211	2.9955	74.2734	3.9902	58.6875	3.4846	72.4205

PCGA_1(2) = First (second) principal component of California Real GSP and Aggregate Employment.
PCUGA_1(2) = First (second) principal component of California Unemployment Rate, California Real GSP, and Aggregate Employment.
R^2 = Square of sample correlation coefficient between the actual and fitted annual percent changes in indemnity claim frequency.

EXHIBIT 13

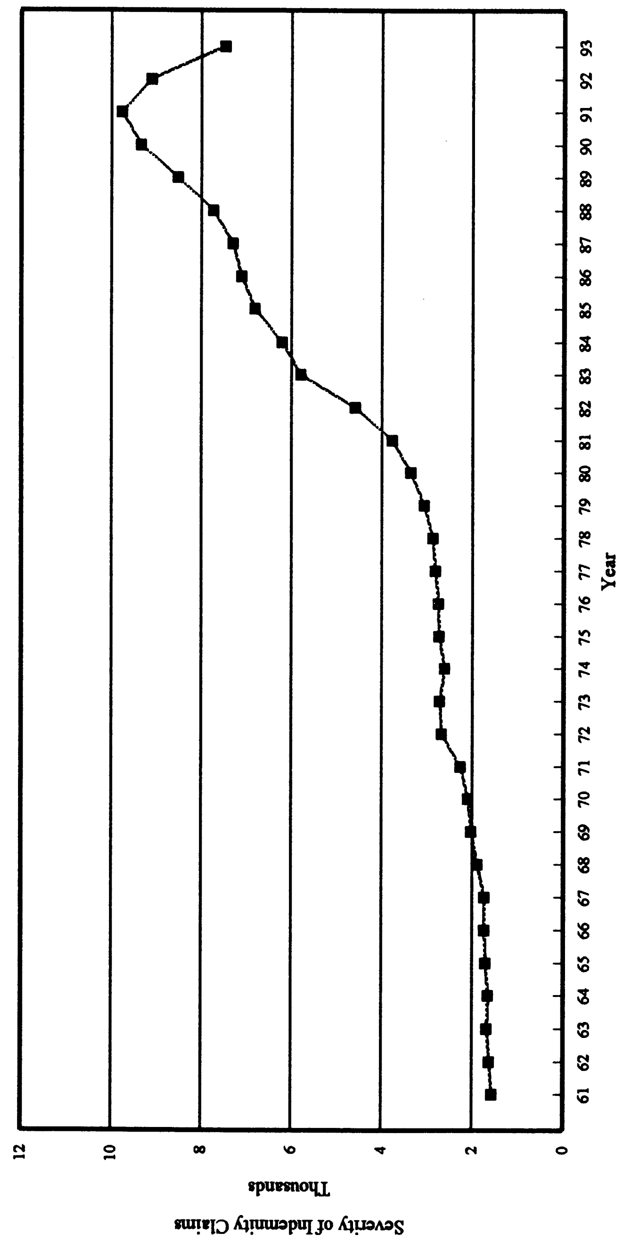
Utilization Point Estimates and Confidence Intervals for Selected Models

Model #2 (Exhibit 7) **Indemnity Benefit Level** **Ratio of Cumulative Injuries to Total Indemnity Claims** **First Principal Component of rGSP, AggE, and Unemp**	**Model #5 (Exhibit 8)** **Indemnity Benefit Level** **Ratio of Cumulative Injuries to Total Indemnity Claims** **First Principal Component of rGSP and AggE**
Point Estimate for percent change in indemnity frequency due to change in indemnity benefit level = 0.2619 × Indemnity Benefit Level Change (CY) **95% Prediction Interval:** [0.2619 ± (2.2010 × 0.0746)] × Indemnity Benefit Level Change (CY) (0.0977, 0.4261) × Indemnity Benefit Level Change (CY) **90% Prediction Interval:** [0.2619 ± (1.7959 × 0.0746)] × Indemnity Benefit Level Change (CY) (0.1279, 0.3959) × Indemnity Benefit Level Change (CY)	**Point Estimate for percent change in indemnity frequency due to change in indemnity benefit level =** 0.3091 × Indemnity Benefit Level Change (CY) **95% Prediction Interval:** [0.3091 ± (2.2010 × 0.0831)] × Indemnity Benefit Level Change (CY) (0.1262, 0.4920) × Indemnity Benefit Level Change (CY) **90% Prediction Interval:** [0.3091 ± (1.7959 × 0.0831)] × Indemnity Benefit Level Change (CY) (0.1599, 0.4583) × Indemnity Benefit Level Change (CY)
Model #6 (Exhibit 9) **Indemnity Benefit Level** **Ratio of Cumulative Injuries to Total Indemnity Claims** **California Aggregate Employment**	**Model #7 (Exhibit 10)** **Indemnity Benefit Level** **Ratio of Cumulative Injuries to Total Indemnity Claims** **California Real Gross State Product**
Point Estimate for percent change in indemnity frequency due to change in indemnity benefit level = 0.3211 × Indemnity Benefit Level Change (CY) **95% Prediction Interval:** [0.3211 ± (2.2010 × 0.0889)] × Indemnity Benefit Level Change (CY) (0.1254, 0.5168) × Indemnity Benefit Level Change (CY) **90% Prediction Interval:** [0.3211 ± (1.7959 × 0.0889)] × Indemnity Benefit Level Change (CY) (0.1614, 0.4808) × Indemnity Benefit Level Change (CY)	**Point Estimate for percent change in indemnity frequency due to change in indemnity benefit level =** 0.2205 × Indemnity Benefit Level Change (CY) **95% Prediction Interval:** [0.2205 ± (2.2010 × 0.0930)] × Indemnity Benefit Level Change (CY) (0.0158, 0.4252) × Indemnity Benefit Level Change (CY) **90% Prediction Interval:** [0.2205 ± (1.7959 × 0.0930)] × Indemnity Benefit Level Change (CY) (0.0535, 0.3875) × Indemnity Benefit Level Change (CY)

EXHIBIT 14

Part 1—Page 1

Indemnity Severity



Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 at 2nd and 1993 at 1st report. See Appendix I.

EXHIBIT 14

PART 1—PAGE 2

ANNUAL PERCENT CHANGE—INDEMNITY SEVERITY

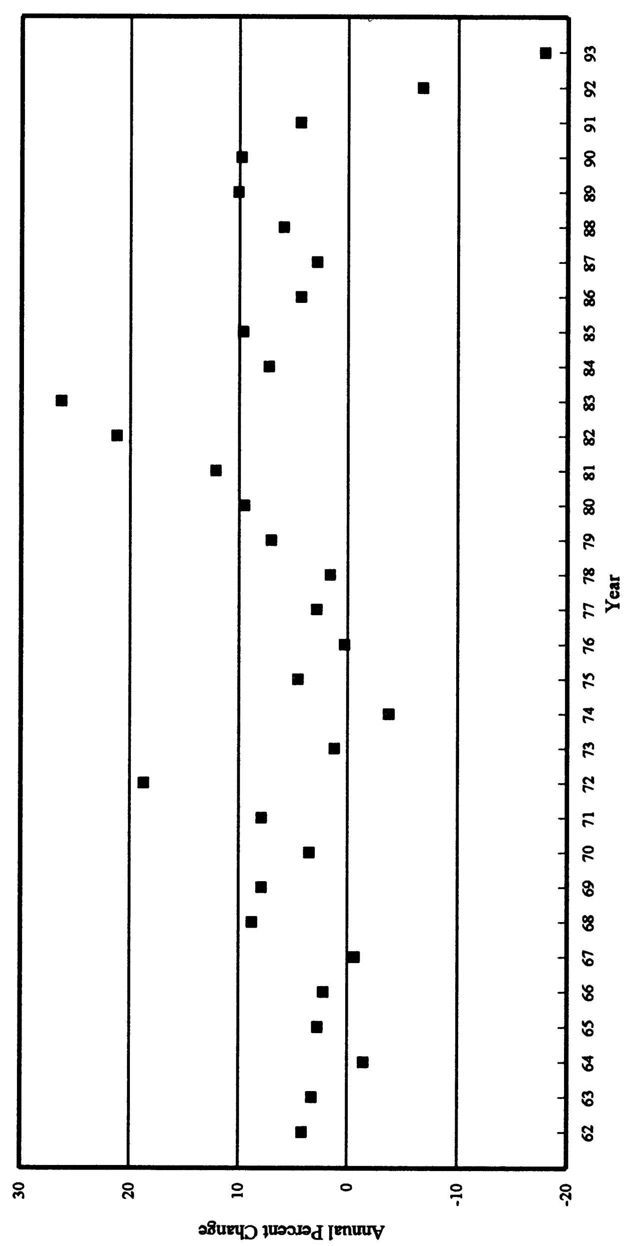
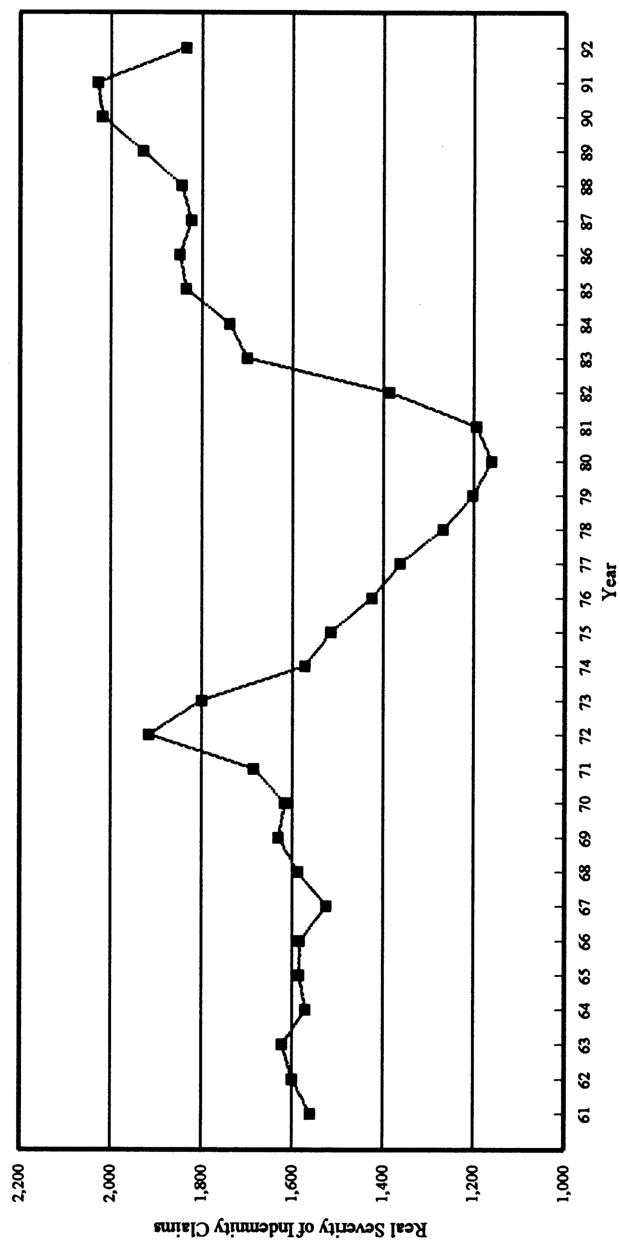


EXHIBIT 14

PART 2—PAGE 1

REAL INDEMNITY SEVERITY



Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 at 2nd and 1993 at 1st report.
Severity brought to a 1982-1984 level using the California CPI. See Appendix I.

EXHIBIT 14

PART 2—PAGE 2

ANNUAL PERCENT CHANGE—REAL INDEMNITY SEVERITY

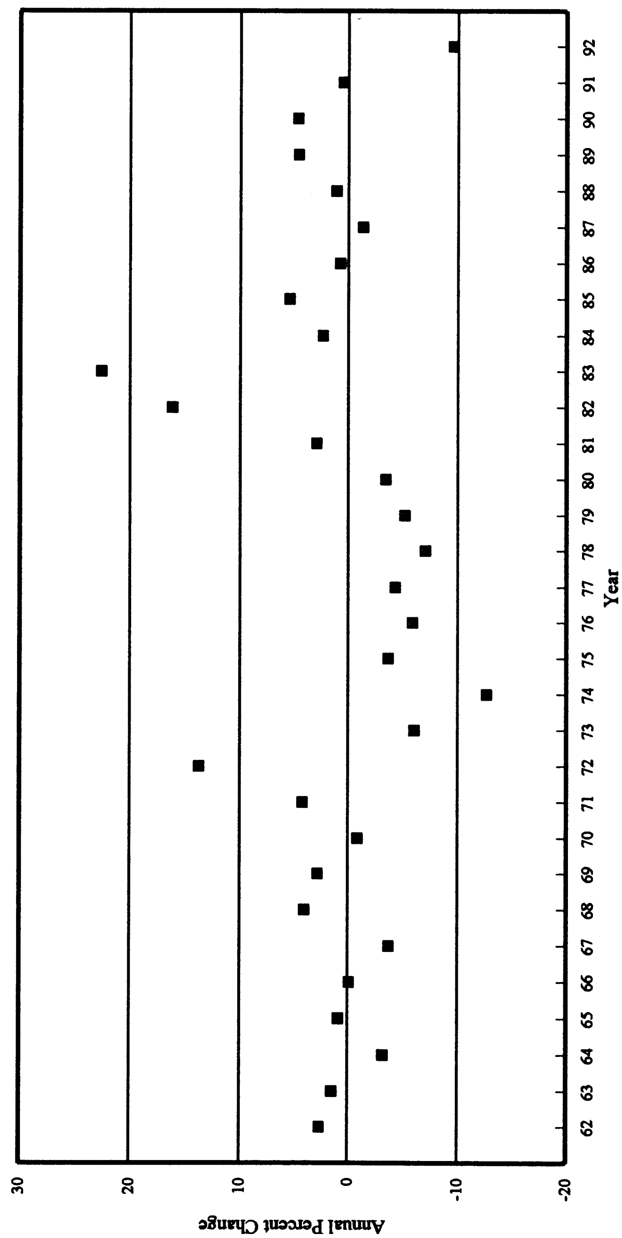


EXHIBIT 15

PART 1

CANDIDATE VARIABLES—TABULAR PRESENTATION ORIGINAL VARIABLES

		Real Severity			Cumulative Benefit Level (Calendar Year)				
		Indemnity Claims							
Year	Indemnity Severity	Indemnity	Medical	All Claims	Indemnity	Medical	Total	California Aggregate Emplmt	Real California GSP
1961	1,559.4	1,559.4	547.0	302.9	1.000	1.000	1.001	3,891,683	–
1962	1,623.8	1,600.0	560.6	315.3	1.001	1.020	1.004	4,071,877	–
1963	1,676.9	1,623.1	584.3	329.4	1.001	1.080	1.005	4,216,436	210,153
1964	1,652.8	1,571.1	581.0	325.8	1.001	1.080	1.005	4,346,448	220,848
1965	1,698.5	1,585.0	638.6	341.6	1.001	1.080	1.005	4,464,625	229,125
1966	1,736.8	1,583.8	700.5	358.5	1.001	1.158	1.022	4,707,406	240,495
1967	1,726.1	1,524.8	695.8	355.8	1.001	1.391	1.073	4,840,158	245,762
1968	1,877.6	1,586.6	725.7	386.7	1.001	1.391	1.073	5,041,894	257,843
1969	2,025.7	1,630.8	739.1	406.9	1.042	1.429	1.105	5,272,325	264,621
1970	2,097.3	1,616.8	751.5	417.3	1.042	1.542	1.117	5,240,190	263,933
1971	2,263.0	1,685.0	806.9	442.2	1.042	1.542	1.117	5,189,637	265,600
1972	2,687.2	1,916.3	830.5	475.6	1.227	1.581	1.247	5,913,892	281,159
1973	2,720.7	1,800.3	808.5	519.3	1.283	1.695	1.297	6,383,331	293,735
1974	2,619.5	1,571.8	732.8	576.8	1.355	1.771	1.362	6,588,356	298,408
1975	2,739.9	1,514.4	765.9	600.0	1.428	1.995	1.450	6,564,524	304,518
1976	2,749.0	1,425.1	783.1	595.7	1.433	2.527	1.530	7,130,103	320,160
1977	2,828.8	1,363.6	782.8	594.2	1.519	2.721	1.626	7,543,268	403,192
1978	2,874.8	1,267.7	787.2	576.1	1.519	2.721	1.626	9,036,931	424,809
1979	3,075.9	1,201.8	792.7	566.9	1.519	2.882	1.641	9,448,087	439,868
1980	3,369.4	1,160.3	824.8	565.8	1.519	3.040	1.655	10,083,911	447,341
1981	3,779.0	1,194.4	910.8	606.7	1.564	3.256	1.719	10,256,167	457,877
1982	4,581.9	1,387.5	1,023.6	714.6	1.564	3.927	1.785	10,131,806	458,036
1983	5,788.3	1,701.2	1,107.8	847.7	2.171	4.363	2.200	10,312,305	480,484
1984	6,207.9	1,740.7	1,104.5	905.7	2.332	4.738	2.321	10,900,212	517,192
1985	6,806.4	1,835.4	1,215.2	983.2	2.332	5.093	2.328	11,378,074	545,612
1986	7,100.4	1,849.1	1,308.9	1,033.2	2.332	5.093	2.328	11,644,237	572,257
1987	7,305.5	1,824.2	1,387.0	1,073.0	2.332	5.278	2.331	12,094,751	599,088
1988	7,737.6	1,844.6	1,457.0	1,131.5	2.332	5.460	2.333	12,556,920	626,079
1989	8,517.8	1,930.2	1,571.0	1,244.0	2.332	5.460	2.333	13,005,986	649,583
1990	9,352.3	2,020.3	1,641.5	1,387.8	2.385	5.684	2.356	13,328,057	665,298
1991	9,760.8	2,029.4	1,593.0	1,368.1	2.502	6.224	2.401	12,796,072	653,197
1992	9,100.9	1,834.4	1,496.2	1,158.2	2.522	6.162	2.407	12,490,570	652,328
1993	7,473.5	–	–	–	2.522	6.162	2.413	12,253,883	–
1994	–	–	–	–	2.334	6.202	2.186	12,500,754	–
1995	–	–	–	–	2.425	6.473	2.232	–	–
1996	–	–	–	–	2.495	6.618	2.268	–	–

Notes: Severity brought to 1982–1984 level using the California CPI.
Series change in Self-Insurance Share Index in 1976. See Appendix H. Value given for 1976 is average of self-insured share under both series.

California Unemplmt Rate	Indemnity Frequency Haz'ness	Indemnity Pure Premium Haz'ness	Litigation Rate	Cumulative Indemnity Claims	Principal Components				Self-Insurance Share Index
					PCGA_1	PCGA_2	PCUGA_1	PCUGA_2	
6.9	1.000	1.000	–	–	NOT APPLICABLE				0.1018
5.8	0.990	0.988	–	–					0.1019
6.0	0.989	0.985	–	–					0.1019
6.0	0.983	0.986	–	–					0.1097
5.9	0.977	0.975	–	–					0.1095
4.9	0.961	0.955	–	–					0.1095
5.7	0.949	0.918	–	–					0.1171
5.4	0.947	0.926	–	–					0.1174
5.2	0.934	0.915	–	–					0.1319
7.3	0.925	0.900	–	–					0.1319
8.8	0.925	0.896	–	–					0.1184
7.6	0.926	0.893	32.81	–					0.1235
7.0	0.925	0.886	28.93	–					0.1228
7.3	0.912	0.866	27.17	–					0.1230
9.9	0.897	0.829	30.59	–					0.1307
9.2	0.902	0.825	33.22	–					0.1788
8.2	0.910	0.821	31.15	2.637					0.2437
7.1	0.922	0.824	27.37	2.525					0.2422
6.2	0.920	0.822	28.09	2.359					0.2426
6.8	0.909	0.799	25.12	2.215					0.2425
7.4	0.892	0.774	25.26	2.345					0.2752
9.9	0.874	0.737	28.76	2.965					0.2857
9.7	0.872	0.733	27.33	3.278					0.3083
7.8	0.870	0.737	27.81	3.868					0.3002
7.2	0.864	0.731	29.03	4.271					0.3003
6.7	0.855	0.723	35.94	4.450					0.3001
5.8	0.853	0.721	34.87	4.613					0.2999
5.3	0.854	0.721	34.86	5.269					0.3002
5.1	0.852	0.719	34.70	6.473					0.2927
5.6	0.840	0.706	–	8.485					0.2853
7.5	0.826	0.692	38.20	9.576					0.2621
9.1	0.816	0.682	42.85	6.359					0.2880
9.2	0.811	0.678	–	4.513					0.3080
–	–	–	–	4.281					–
–	–	–	–	–					–
–	–	–	–	–					–

EXHIBIT 15

PART 2

CANDIDATE VARIABLES—TABULAR PRESENTATION ANNUAL PERCENT CHANGES

		Real Severity			Cumulative Benefit Level (Calendar Year)				
		Indemnity Claims							
Year	Indemnity Severity	Indemnity	Medical	All Claims	Indemnity	Medical	Total	California Aggregate Emplmt	Real California GSP
1961	—	—	—	—	0.032	0.000	0.092	—	—
1962	4.131	2.604	2.484	4.097	0.075	2.012	0.272	4.630	—
1963	3.271	1.446	4.226	4.479	0.000	5.852	0.158	3.550	—
1964	−1.440	−3.204	−0.556	−1.105	0.000	0.000	0.000	3.083	5.275
1965	2.768	0.882	9.916	4.866	0.000	0.000	0.000	2.719	3.759
1966	2.255	−0.076	9.693	4.934	0.000	7.259	1.689	5.438	4.942
1967	−0.616	−3.723	−0.679	−0.762	0.000	20.083	4.928	2.820	2.069
1968	8.778	4.053	4.298	8.688	0.000	0.000	0.000	4.168	5.011
1969	7.888	2.788	1.848	5.241	4.100	2.747	3.062	4.570	2.533
1970	3.533	−0.861	1.672	2.550	0.000	7.935	1.043	−0.610	−0.194
1971	7.899	4.220	7.379	5.956	0.000	0.000	0.000	−0.965	0.590
1972	18.746	13.726	2.919	7.562	17.732	2.489	11.635	13.956	5.957
1973	1.246	−6.056	−2.649	9.186	4.553	7.232	4.049	7.938	4.358
1974	−3.718	−12.692	−9.364	11.072	5.623	4.461	4.974	3.212	1.747
1975	4.595	−3.654	4.521	4.015	5.418	12.673	6.438	−0.362	1.897
1976	0.332	−5.893	2.246	−0.716	0.300	26.650	5.560	8.616	5.283
1977	2.902	−4.316	−0.032	−0.251	6.000	7.702	6.268	5.795	5.545
1978	1.629	−7.033	0.555	−3.039	0.000	0.000	0.000	19.801	5.322
1979	6.993	−5.197	0.698	−1.593	0.000	5.898	0.907	4.550	3.672
1980	9.543	−3.458	4.048	−0.205	0.000	5.479	0.885	6.730	1.633
1981	12.156	2.945	10.432	7.244	3.000	7.119	3.842	1.708	2.354
1982	21.249	16.165	12.384	17.772	0.000	20.599	3.812	−1.213	0.013
1983	26.328	22.604	8.224	18.629	38.800	11.100	23.300	1.782	4.931
1984	7.250	2.325	−0.293	6.839	7.400	8.600	5.500	5.701	7.665
1985	9.641	5.439	10.021	8.557	0.000	7.500	0.300	4.384	5.453
1986	4.319	0.747	7.708	5.084	0.000	0.000	0.000	2.339	4.887
1987	2.889	−1.348	5.969	3.851	0.000	3.630	0.101	3.869	4.711
1988	5.915	1.120	5.049	5.458	0.000	3.445	0.099	3.821	4.476
1989	10.084	4.639	7.823	9.940	0.000	0.000	0.000	3.576	3.770
1990	9.796	4.666	4.486	11.560	2.300	4.100	1.000	2.476	2.475
1991	4.368	0.455	−2.956	−1.420	4.900	9.500	1.900	−3.991	−1.900
1992	−6.761	−9.611	−6.079	−15.340	0.800	−1.000	0.251	−2.387	−0.106
1993	−17.882	—	—	—	0.000	0.000	0.248	−1.895	—
1994	—	—	—	—	−7.469	0.646	−9.428	2.015	—
1995	—	—	—	—	3.919	4.374	2.141	—	—
1996	—	—	—	—	2.894	2.242	1.596	—	—

Notes: PCGA_1(2) = First (second) principal component of CA Real GSP and Aggregate Employment

PCUGA_1(2) = First (second) principal component of CA Real GSP, Unemployment Rate, and Aggregate Employment

Series change in Self-Insurance Share Index in 1976. See Appendix H. Value given for 1976 is average of self-insured share under both series.

California Unemplmt Rate	Indemnity Frequency Haz'ness	Indemnity Pure Premium Haz'ness	Litigation Rate	Cumulative ÷ Indemnity Claims	Principal Components PCGA_1	PCGA_2	PCUGA_1	PCUGA_2	Self-Insurance Share Index
—	—	—	—	—	—	—	—	—	0.002
−15.942	−0.974	−1.233	—	—	4.360	−1.560	—	—	0.009
3.448	−0.128	−0.318	—	—	3.343	−1.196	—	—	−0.000
0.000	−0.565	0.107	—	—	4.680	3.928	−1.188	3.164	0.787
−1.667	−0.629	−1.040	—	—	3.826	2.624	−2.571	2.446	−0.020
−16.949	−1.642	−2.062	—	—	6.785	2.822	−18.124	2.203	−0.000
16.327	−1.237	−3.883	—	—	3.352	0.998	15.155	5.962	0.755
−5.263	−0.249	0.875	—	—	5.612	3.314	−6.495	3.208	0.032
−3.704	−1.361	−1.188	—	—	5.157	0.846	−4.765	3.837	1.455
40.385	−1.003	−1.650	—	—	−0.639	0.023	39.526	7.167	−0.000
20.548	0.082	−0.456	—	—	−0.709	0.881	20.155	3.023	−1.356
−13.636	0.051	−0.287	—	—	15.147	0.907	−16.613	11.223	0.513
−7.895	−0.137	−0.772	−11.834	—	8.942	1.429	−9.696	6.383	−0.071
4.286	−1.398	−2.363	−6.085	—	3.613	0.563	3.374	4.021	0.019
35.616	−1.589	−4.173	12.615	—	0.299	1.908	34.588	6.548	0.769
−7.071	0.500	−0.474	8.598	—	9.892	2.072	−9.127	7.231	−0.358
−10.870	0.949	−0.513	−6.250	—	7.324	3.268	−12.331	3.739	1.331
−13.415	1.249	0.302	−12.133	−4.248	20.437	−1.660	−17.424	16.983	−0.151
−12.676	−0.216	−0.182	2.621	−6.598	5.521	1.925	−13.643	2.121	0.034
9.677	−1.116	−2.861	−10.557	−6.108	6.886	−0.730	7.984	8.507	−0.003
8.824	−1.866	−3.130	0.537	5.871	2.401	1.641	8.013	3.436	3.271
33.784	−2.097	−4.711	13.867	26.453	−1.137	0.421	33.177	5.311	1.047
−2.020	−0.201	−0.534	−4.964	10.556	3.338	4.042	−2.874	1.489	2.261
−19.588	−0.173	0.549	1.773	17.998	7.950	5.297	−21.061	2.026	−0.816
−7.692	−0.748	−0.882	4.382	10.430	5.965	3.658	−8.956	2.964	0.015
−6.944	−1.033	−1.126	23.803	4.183	3.849	3.814	−7.777	1.087	−0.017
−13.433	−0.225	−0.248	−2.974	3.655	5.230	3.132	−14.373	1.335	−0.021
−8.621	0.057	0.090	−0.032	14.219	5.106	2.927	−9.643	2.208	0.026
−3.774	−0.179	−0.384	−0.456	22.853	4.637	2.344	−4.788	2.881	−0.746
9.804	−1.382	−1.756	—	31.096	3.165	1.496	8.811	4.381	−0.746
33.929	−1.718	−1.977	—	12.856	−4.398	−0.444	34.062	2.563	−2.312
21.333	−1.185	−1.526	12.183	−33.591	−2.284	0.704	21.269	1.760	2.590
1.099	−0.572	−0.527	—	−29.038	−1.784	0.638	1.428	−1.647	1.999
—	—	—	—	−5.130	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—

EXHIBIT 16

PART 1

CORRELATIONS AMONG VARIABLES
SAMPLE PERIOD: 1964–1992
PEARSON PRODUCT MOMENT CORRELATION AT LAG = 0

		Real Severity			Benefit Level (Calendar Year)			California Aggregate Emplmt
		Indemnity Claims						
	Indemnity Severity	Indemnity	Medical	All Claims	Indemnity	Medical	Total	
Indemnity Severity								
Real Indemnity Severity	0.927	1.000	0.655	0.696	0.585	0.103	0.525	−0.117
Real Medical Severity	0.614	0.655	1.000	0.504	0.012	0.060	0.011	−0.088
Real Total Severity	0.711	0.696	0.504	1.000	0.423	0.142	0.417	−0.056
Indemnity Benefit Level	0.591	0.585	0.012	0.423	1.000	0.110	0.945	0.060
Medical Benefit Level	0.144	0.103	0.060	0.142	0.110	1.000	0.384	−0.102
Total Benefit Level	0.555	0.525	0.011	0.417	0.945	0.384	1.000	0.082
California Aggregate Employment	0.004	−0.117	−0.088	−0.056	0.060	−0.102	0.082	1.000
Real California Gross State Product	0.088	0.103	0.152	0.167	0.204	−0.113	0.199	0.655
California Unemployment Rate	0.003	0.022	−0.054	−0.074	−0.110	0.267	−0.059	−0.677
Indemnity Frequency Haz'ness	−0.061	−0.091	−0.129	−0.154	0.127	−0.176	0.093	0.643
Indemnity Pure Premium Haz'ness	−0.065	−0.001	−0.126	−0.097	0.105	−0.493	−0.067	0.431
Litigation Rates	0.146	0.255	0.476	0.143	−0.197	0.325	−0.109	−0.543
Cumulative ÷ Indemnity Claims	0.596	0.650	0.557	0.836	0.112	0.466	0.153	−0.110
1st PC (rGSP, AggE)	0.017	−0.088	−0.055	−0.024	0.085	−0.108	0.104	0.993
2nd PC (rGSP, AggE)	0.111	0.249	0.286	0.274	0.210	−0.049	0.182	−0.116
1st PC (rGSP, AggE, Unemp)	0.002	0.025	−0.049	−0.070	−0.110	0.261	−0.063	−0.705
2nd PC (rGSP, AggE, Unemp)	0.010	−0.135	−0.169	−0.145	−0.022	0.119	0.058	0.710
Self-Insurance Share Index	0.167	0.136	0.092	−0.032	0.317	0.061	0.388	−0.063

Note: Pearson Product Moment Correlation assumes the variables to be normally distributed.

Real California GSP	California Unemplmt Rate	Indemnity Frequency Haz`ness	Indemnity Pure Premium Haz`ness	Litigation Rate	Cumulative ÷ Indemnity Claims	Principal Components				Self-Insurance Share Index
						PCGA_1	PCGA_2	PCUGA_1	PCUGA_2	
0.103	0.022	−0.091	−0.001	0.255	0.650	−0.088	0.249	0.025	−0.135	0.136
0.152	−0.054	−0.129	−0.126	0.476	0.557	−0.055	0.286	−0.049	−0.169	0.092
0.167	−0.074	−0.154	−0.097	0.143	0.836	−0.024	0.274	−0.070	−0.145	−0.032
0.204	−0.110	0.127	0.105	−0.197	0.112	0.085	0.210	−0.110	−0.022	0.317
−0.113	0.267	−0.176	−0.493	0.325	0.466	−0.108	−0.049	0.261	0.119	0.061
0.199	−0.059	0.093	−0.067	−0.109	0.153	0.104	0.182	−0.063	0.058	0.388
0.655	−0.677	0.643	0.431	−0.543	−0.110	0.993	−0.116	−0.705	0.710	−0.063
1.000	−0.892	0.617	0.638	−0.122	0.153	0.739	0.674	−0.897	0.040	0.024
−0.892	1.000	−0.587	−0.683	0.353	0.025	−0.741	−0.511	0.999	0.038	0.027
0.617	−0.587	1.000	0.781	−0.448	−0.156	0.668	0.183	−0.600	0.312	−0.182
0.638	−0.683	0.781	1.000	−0.327	−0.116	0.483	0.417	−0.681	−0.068	−0.265
−0.122	0.353	−0.448	−0.327	1.000	0.374	−0.522	0.357	0.369	−0.368	−0.002
0.153	0.025	−0.156	−0.116	0.374	1.000	−0.076	0.286	0.028	−0.119	−0.407
0.739	−0.741	0.668	0.483	−0.522	−0.076	1.000	0.000	−0.767	0.639	−0.052
0.674	−0.511	0.183	0.417	0.357	0.286	0.000	1.000	−0.490	−0.642	0.094
−0.897	0.999	−0.600	−0.681	0.369	0.028	−0.767	−0.490	1.000	−0.000	0.029
0.040	0.038	0.312	−0.068	−0.368	−0.119	0.639	−0.642	−0.000	1.000	−0.058
0.024	0.027	−0.182	−0.265	−0.002	−0.407	−0.052	0.094	0.029	−0.058	1.000

EXHIBIT 16

PART 2

CORRELATIONS AMONG VARIABLES
SAMPLE PERIOD: 1964–1992
SIGNIFICANCE OF CORRELATION AT LAG = 0

		Real Severity			Benefit Level (Calendar Year)			
		Indemnity Claims						
	Indemnity Severity	Indemnity	Medical	All Claims	Indemnity	Medical	Total	California Aggregate Emplmt
Indemnity Severity		0.000	0.000	0.000	0.001	0.456	0.002	0.983
Real Indemnity Severity	0.000		0.000	0.000	0.001	0.595	0.003	0.547
Real Medical Severity	0.000	0.000		0.005	0.952	0.756	0.953	0.649
Real Total Severity	0.000	0.000	0.005		0.022	0.461	0.024	0.774
Indemnity Benefit Level	0.001	0.001	0.952	0.022		0.569	0.000	0.758
Medical Benefit Level	0.456	0.595	0.756	0.461	0.569		0.040	0.599
Total Benefit Level	0.002	0.003	0.953	0.024	0.000	0.040		0.673
California Aggregate Employment	0.983	0.547	0.649	0.774	0.758	0.599	0.673	
Real California Gross State Product	0.650	0.595	0.432	0.387	0.288	0.560	0.300	0.000
California Unemployment Rate	0.987	0.911	0.781	0.702	0.572	0.162	0.763	0.000
Indemnity Frequency Haz'ness	0.754	0.639	0.506	0.425	0.513	0.361	0.632	0.000
Indemnity Pure Premium Haz'ness	0.738	0.997	0.515	0.617	0.588	0.007	0.732	0.020
Litigation Rates	0.577	0.323	0.053	0.584	0.448	0.203	0.676	0.024
Cumulative ÷ Indemnity Claims	0.019	0.009	0.031	0.000	0.690	0.080	0.587	0.696
1st PC (rGSP, AggE)	0.929	0.649	0.776	0.902	0.662	0.576	0.592	0.000
2nd PC (rGSP, AggE)	0.565	0.192	0.133	0.151	0.274	0.803	0.344	0.549
1st PC (rGSP, AggE, Unemp)	0.993	0.897	0.800	0.717	0.569	0.172	0.746	0.000
2nd PC (rGSP, AggE, Unemp)	0.957	0.484	0.380	0.454	0.911	0.538	0.766	0.000
Self-Insurance Share Index	0.385	0.483	0.636	0.868	0.094	0.753	0.038	0.746

Note: P Value is the probability of observing the indicated SAMPLE correlation coefficient if the TRUE correlation coefficient was actually zero.

Real California GSP	California Unemplmt Rate	Indemnity Frequency Haz`ness	Indemnity Pure Premium Haz`ness	Litigation Rate	Cumulative ÷ Indemnity Claims	Principal Components PCGA_1	PCGA_2	PCUGA_1	PCUGA_2	Self-Insurance Share Index
0.650	0.987	0.754	0.738	0.577	0.019	0.929	0.565	0.993	0.957	0.385
0.595	0.911	0.639	0.997	0.323	0.009	0.649	0.192	0.897	0.484	0.483
0.432	0.781	0.506	0.515	0.053	0.031	0.776	0.133	0.800	0.380	0.636
0.387	0.702	0.425	0.617	0.584	0.000	0.902	0.151	0.717	0.454	0.868
0.288	0.572	0.513	0.588	0.448	0.690	0.662	0.274	0.569	0.911	0.094
0.560	0.162	0.361	0.007	0.203	0.080	0.576	0.803	0.172	0.538	0.753
0.300	0.763	0.632	0.732	0.676	0.587	0.592	0.344	0.746	0.766	0.038
0.000	0.000	0.000	0.020	0.024	0.696	0.000	0.549	0.000	0.000	0.746
	0.000	0.000	0.000	0.640	0.587	0.000	0.000	0.000	0.838	0.900
0.000		0.001	0.000	0.165	0.928	0.000	0.005	0.000	0.844	0.888
0.000	0.001		0.000	0.071	0.578	0.000	0.342	0.001	0.099	0.345
0.000	0.000	0.000		0.201	0.680	0.008	0.024	0.000	0.726	0.164
0.640	0.165	0.071	0.201		0.232	0.031	0.160	0.145	0.146	0.994
0.587	0.928	0.578	0.680	0.232		0.788	0.301	0.921	0.672	0.132
0.000	0.000	0.000	0.008	0.031	0.788		1.000	0.000	0.000	0.787
0.000	0.005	0.342	0.024	0.160	0.301	1.000		0.007	0.000	0.629
0.000	0.000	0.001	0.000	0.145	0.921	0.000	0.007		1.000	0.882
0.838	0.844	0.099	0.726	0.146	0.672	0.000	0.000	1.000		0.764
0.900	0.888	0.345	0.164	0.994	0.132	0.787	0.629	0.882	0.764	

EXHIBIT 17

PART 1—PAGE 1

INDEMNITY SEVERITY VS INDEMNITY BENEFIT LEVEL

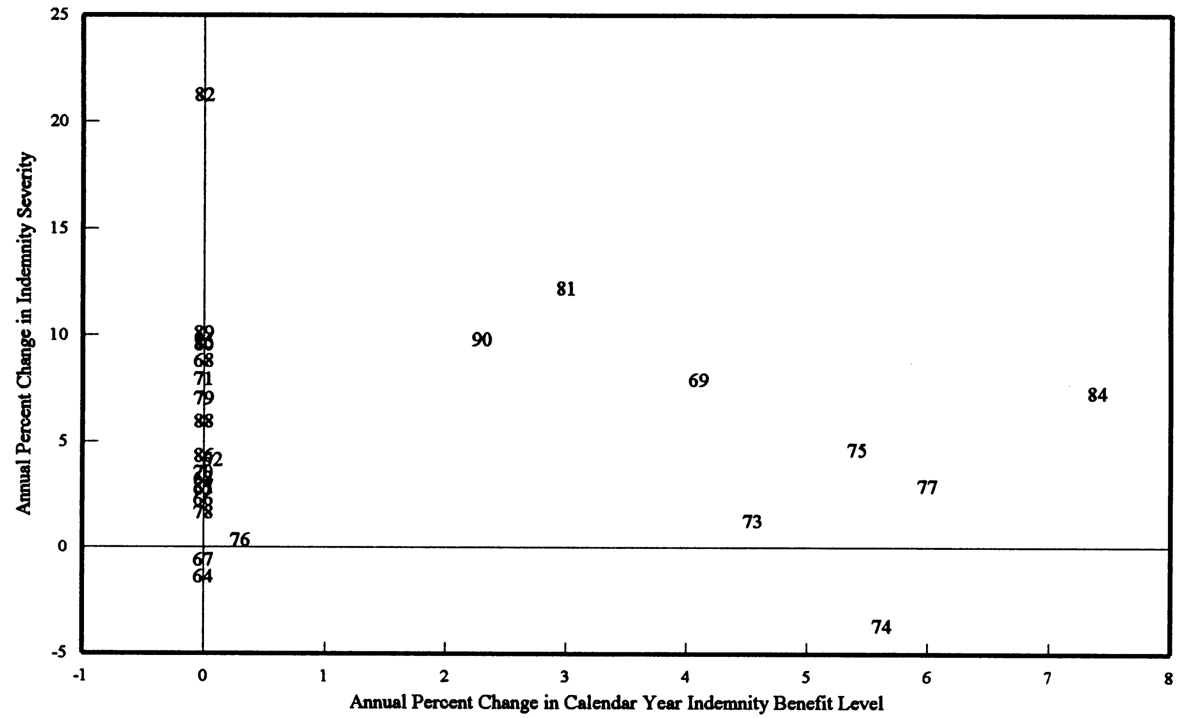


EXHIBIT 17

PART 1—PAGE 2

Outliers 1972 and 1983 used in regression but not shown in graph

Spearman Rank Correlation Coefficient:	0.12728
Valid Cases	31
Two-tailed Significance	0.49500
Regression Output With Constant:	
Constant	4.30094
Std Err of Y Est	5.61927
R Squared	0.35458
No. of Observations	31
Degrees of Freedom	29
X Coefficient(s)	0.54096
Std Err of Coef.	0.13553
P-Value	0.00041
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	6.82402
R Squared	0.01534
No. of Observations	31
Degrees of Freedom	30
X Coefficient(s)	0.75305
Std Err of Coef.	0.15078
P-Value	0.00002

EXHIBIT 17

PART 2—PAGE 1

REAL INDEMNITY SEVERITY VS INDEMNITY BENEFIT LEVEL

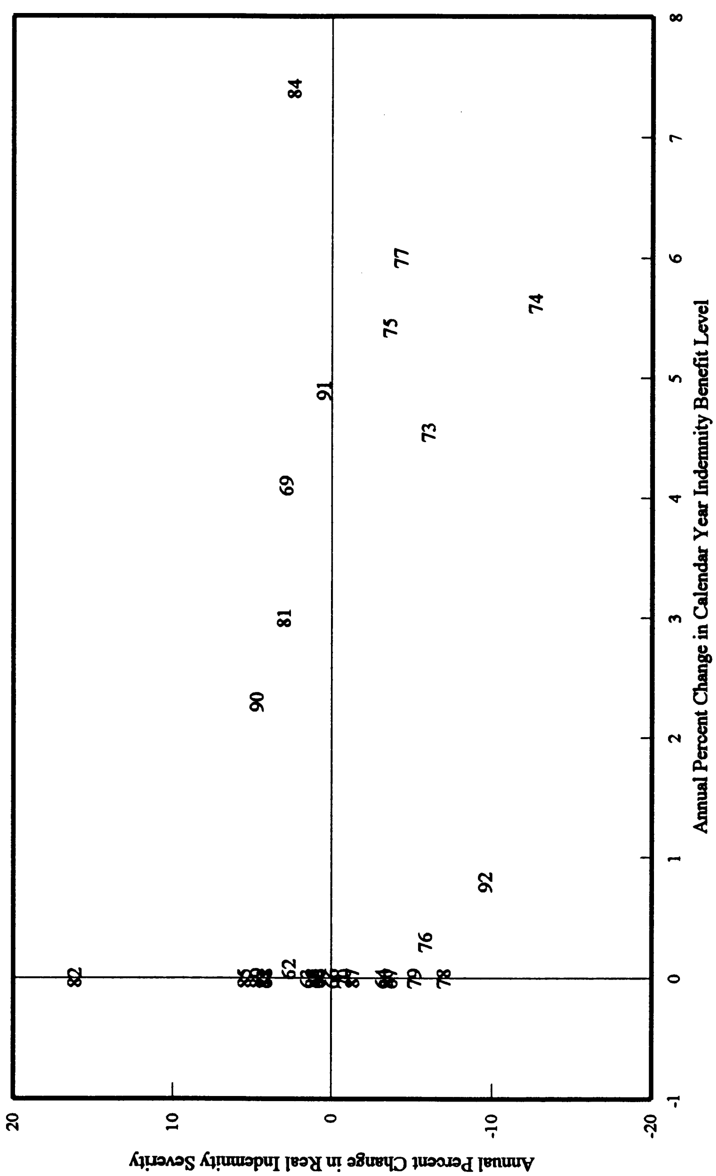


EXHIBIT 17

PART 2—PAGE 2

Outliers 1972 and 1983 used in regression but not shown in graph

Spearman Rank Correlation Coefficient:	0.00661
Valid Cases	31
Two-tailed Significance	0.97180
Regression Output With Constant:	
Constant	-1.01637
Std Err of Y Est	5.99262
R Squared	0.33033
No. of Observations	31
Degrees of Freedom	29
X Coefficient(s)	0.54665
Std Err of Coef.	0.14453
P-Value	0.00072
Regression Output Without Constant:	
Constant	0.00000
Std Err of Y Est	5.96745
R Squared	0.31304
No. of Observations	31
Degrees of Freedom	30
X Coefficient(s)	0.49653
Std Err of Coef.	0.13186
P-Value	0.00072

APPENDIX A

PART 1

DEVELOPMENT OF CANDIDATE VARIABLES CLAIM FREQUENCIES

	Total Exposure (000s)			Incurred Claim Count			Incurred Frequency per $1M (1987)			Annual % Change		
Policy Year	$ Current	Wage Index	$ 1987	Indemnity	Med-Only	Total	Indemnity Claims	Med-Only Claims	Total Claims	Indemnity Claims	Med-Only Claims	Total Claims
1961	21,877,687	22.3	98,107,006	89,647	589,817	679,464	0.9138	6.0120	6.9257	–	–	–
1962	23,612,513	23.2	101,627,186	94,893	611,070	705,963	0.9337	6.0129	6.9466	2.1853	0.0147	0.3011
1963	25,228,415	24.2	104,386,854	99,376	621,304	720,680	0.9520	5.9519	6.9039	1.9557	−1.0132	−0.6141
1964	26,203,849	25.1	104,221,777	99,642	613,730	713,372	0.9561	5.8887	6.8447	0.4265	−1.0626	−0.8573
1965	28,887,463	26.0	111,085,011	103,566	624,766	728,332	0.9323	5.6242	6.5565	−2.4836	−4.4913	−4.2108
1966	31,220,478	27.2	114,696,870	107,600	640,367	747,967	0.9381	5.5831	6.5213	0.6234	−0.7306	−0.5381
1967	33,123,452	29.1	113,908,431	109,981	645,128	755,109	0.9655	5.6636	6.6291	2.9203	1.4408	1.6536
1968	37,504,640	31.4	119,602,193	118,568	654,187	772,755	0.9914	5.4697	6.4610	2.6754	−3.4232	−2.5350
1969	39,913,331	33.7	118,280,474	123,933	659,842	783,775	1.0478	5.5786	6.6264	5.6928	1.9915	2.5594
1970	40,951,049	35.4	115,775,975	117,435	606,247	723,682	1.0143	5.2364	6.2507	−3.1933	−6.1349	−5.6698
1971	43,254,887	36.5	118,637,181	121,927	620,180	742,107	1.0277	5.2275	6.2553	1.3211	−0.1689	0.0729
1972	47,004,364	38.2	122,925,428	133,484	692,791	826,275	1.0859	5.6359	6.7218	5.6595	7.8111	7.4576
1973	50,834,927	40.6	125,208,865	154,932	675,018	829,950	1.2374	5.3911	6.6285	13.9511	−4.3423	−1.3870
1974	54,238,668	42.8	126,869,984	187,310	604,308	791,618	1.4764	4.7632	6.2396	19.3153	−11.6474	−5.8674

1975	57,738,551	46.2	124,895,945	193,063	584,751	777,814	1.5458	4.6819	6.2277	4.7005	−1.7069	−0.1908
1976	62,193,123	49.0	126,958,448	206,908	611,465	818,373	1.6297	4.8163	6.4460	5.4302	2.8697	3.5052
1977	67,671,264	51.8	130,676,822	223,511	637,325	860,836	1.7104	4.8771	6.5875	4.9505	1.2634	2.1956
1978	75,054,494	55.6	134,951,430	235,645	661,759	897,404	1.7461	4.9037	6.6498	2.0893	0.5449	0.9459
1979	82,723,286	60.9	135,929,178	236,010	651,166	887,176	1.7363	4.7905	6.5268	−0.5655	−2.3085	−1.8508
1980	89,813,215	66.6	134,908,287	223,191	613,630	836,821	1.6544	4.5485	6.2029	−4.7159	−5.0513	−4.9621
1981	98,778,141	73.1	135,218,017	211,709	574,590	786,299	1.5657	4.2494	5.8150	−5.3617	−6.5766	−6.2526
1982	103,443,974	77.3	133,843,357	203,441	527,868	731,309	1.5200	3.9439	5.4639	−2.9184	−7.1878	−6.0383
1983	114,266,699	82.0	139,404,129	233,559	584,794	818,353	1.6754	4.1950	5.8704	10.2248	6.3650	7.4387
1984	129,672,576	86.9	149,266,396	271,618	627,773	899,391	1.8197	4.2057	6.0254	8.6114	0.2567	2.6411
1985	140,891,926	91.4	154,095,929	272,771	617,051	889,822	1.7701	4.0043	5.7745	−2.7229	−4.7885	−4.1647
1986	153,916,015	95.3	161,550,696	275,370	608,364	883,734	1.7045	3.7658	5.4703	−3.7057	−5.9574	−5.2671
1987	167,173,336	100.0	167,173,336	292,759	627,052	919,811	1.7512	3.7509	5.5021	2.7390	−0.3948	0.5817
1988	181,245,258	104.9	172,810,122	302,703	623,028	925,731	1.7517	3.6053	5.3569	0.0240	−3.8826	−2.6392
1989	194,374,909	109.2	178,066,971	323,131	634,934	958,065	1.8147	3.5657	5.3804	3.5971	−1.0976	0.4375
1990	197,318,717	112.9	174,807,166	343,711	610,877	954,588	1.9662	3.4946	5.4608	8.3525	−1.9948	1.4951
1991	198,907,627	116.8	170,334,988	317,987	571,418	889,405	1.8668	3.3547	5.2215	− 5.0552	−4.0035	−4.3822
1992	200,370,929	120.5	166,281,823	251,259	516,811	768,070	1.5110	3.1080	4.6191	−19.0585	−7.3518	−11.5373
1993	202,247,504	121.2	166,835,674	207,425	461,029	668,454	1.2433	2.7634	4.0067	−17.7198	−11.0896	−13.2586
1994	210,773,228	122.4	172,186,345	200,526	422,883	623,409	1.1646	2.4560	3.6205	−6.3302	−11.1245	−9.6368

Source: WCIRB of California Class Experience at latest report level as of 11/12/96.
See Appendix A, Part 2 for development of the index to adjust for wage level changes.

APPENDIX A

PART 2

DEVELOPMENT OF INDEX TO ADJUST FOR WAGE LEVEL CHANGES

(A)	(B)	(C)	$\frac{(B) \times 1{,}000}{(C)}$	$\frac{19yy \times 100}{1987}$	Class Experience Exposure (000s)	
Year	Wages (millions)	Employees (thousands)	Avg Wage	Index 1987 = 100	Nominal	Real
1961	30,770	6,036	5,097.75	22.3	21,877,687	98,107,006
1962	33,260	6,262	5,311.40	23.2	23,612,513	101,627,186
1963	35,674	6,457	5,524.86	24.2	25,228,415	104,386,854
1964	38,273	6,659	5,747.56	25.1	26,203,849	104,221,777
1965	40,751	6,855	5,944.71	26.0	28,887,463	111,085,011
1966	44,914	7,218	6,222.50	27.2	31,220,478	114,696,870
1967	48,141	7,242	6,647.47	29.1	33,123,452	113,908,431
1968	52,824	7,369	7,168.41	31.4	37,504,640	119,602,193
1969	57,917	7,508	7,714.04	33.7	39,913,331	118,280,474
1970	61,250	7,575	8,085.81	35.4	40,951,049	115,775,975
1971	63,919	7,669	8,334.72	36.5	43,254,887	118,637,181
1972	69,895	7,996	8,741.25	38.2	47,004,364	122,925,428
1973	76,904	8,286	9,281.20	40.6	50,834,927	125,208,865
1974	84,419	8,638	9,772.98	42.8	54,238,668	126,869,984
1975	90,864	8,598	10,568.04	46.2	57,738,551	124,895,945
1976	100,674	8,990	11,198.44	49.0	62,193,123	126,958,448
1977	112,616	9,513	11,838.12	51.8	67,671,264	130,676,822
1978	128,880	10,137	12,713.82	55.6	75,054,494	134,951,430
1979	146,995	10,566	13,912.08	60.9	82,723,286	135,929,178
1980	164,271	10,794	15,218.73	66.6	89,813,215	134,908,287
1981	182,659	10,938	16,699.49	73.1	98,778,141	135,218,017
1982	193,764	10,967	17,667.91	77.3	103,443,974	133,843,357
1983	207,897	11,095	18,737.90	82.0	114,266,699	139,404,129
1984	230,983	11,631	19,859.26	86.9	129,672,576	149,266,396
1985	251,818	12,048	20,901.23	91.4	140,891,926	154,095,929
1986	270,983	12,442	21,779.70	95.3	153,916,015	161,550,696
1987	295,946	12,946	22,860.03	100.0	167,173,336	167,173,336
1988	320,917	13,385	23,975.87	104.9	181,245,258	172,810,122
1989	343,861	13,780	24,953.63	109.2	193,896,851	177,629,021
1990	368,635	14,286	25,803.93	112.9	197,318,717	174,807,166
1991	373,138	13,978	26,694.66	116.8	198,907,627	170,334,988
1992	383,971	13,939	27,546.52	120.5	200,370,929	166,281,823
1993	384,784	13,885	27,712.21	121.2	202,247,504	166,835,674
1994	395,707	14,141	27,982.96	122.4	210,773,228	172,186,345

Sources: Wages: California Statistical Abstract 1995, "Personal Income in California by Major Source 1969 to 1994"

Employees: California Statistical Abstract, "Employment and Unemployment, California and Metropolitan Areas"

Exposure: WCIRB of California Class Experience (1961—88 3rd Report; 1989–1990 5th Report, 1991 4th Report, 1992 3rd Report, 1993 2nd Report and 1994 1st Report; 1990–1994 Preliminary Summary as of 11/12/96).

APPENDIX B

DEVELOPMENT OF CANDIDATE VARIABLES BENEFIT LEVEL CHANGES

Effective Date	Proportion of Year	Cumulative Benefit Level			Calendar Year Cumulative Benefit Level			Annual Percent Change in CY Benefit Level		
		Indemnity	Medical	Total	Indemnity	Medical	Total	Indemnity	Medical	Total
1-Jan-1961	0.704	1.0000	1.0000	1.0000						
15-Sep-1961	0.296	1.0011	1.0000	1.0031	1.0003	1.0000	1.0009	0.032	0.000	0.092
1-Jan-1962	0.748	1.0011	1.0000	1.0031						
1-Oct-1962	0.252	1.0011	1.0798	1.0052	1.0011	1.0201	1.0036	0.075	2.012	0.272
1-Jan-1963	1.000	1.0011	1.0798	1.0052	1.0011	1.0798	1.0052	0.000	5.852	0.158
1-Jan-1964	1.000	1.0011	1.0798	1.0052	1.0011	1.0798	1.0052	0.000	0.000	0.000
1-Jan-1965	1.000	1.0011	1.0798	1.0052	1.0011	1.0798	1.0052	0.000	0.000	0.000
1-Jan-1966	0.748	1.0011	1.0798	1.0052						
1-Oct-1966	0.252	1.0011	1.3908	1.0726	1.0011	1.1582	1.0222	0.000	7.259	1.689
1-Jan-1967	1.000	1.0011	1.3908	1.0726	1.0011	1.3908	1.0726	0.000	20.083	4.928
1-Jan-1968	1.000	1.0011	1.3908	1.0726	1.0011	1.3908	1.0726	0.000	0.000	0.000
1-Jan-1969	0.748	1.0421	1.3908	1.1015						
1-Oct-1969	0.252	1.0421	1.5424	1.1170	1.0421	1.4290	1.1054	4.100	2.747	3.062
1-Jan-1970	1.000	1.0421	1.5424	1.1170	1.0421	1.5424	1.1170	0.000	7.935	1.043
1-Jan-1971	1.000	1.0421	1.5424	1.1170	1.0421	1.5424	1.1170	0.000	0.000	0.000
1-Jan-1972	0.249	1.0421	1.5424	1.1170						
1-Apr-1972	0.500	1.2881	1.5424	1.2856						
1-Oct-1972	0.251	1.2881	1.6951	1.2985	1.2269	1.5808	1.2469	17.732	2.489	11.635
1-Jan-1973	0.178	1.2881	1.6951	1.2985						
7-Mar-1973	0.822	1.2816	1.6951	1.2972	1.2828	1.6951	1.2974	4.553	7.232	4.049
1-Jan-1974	0.247	1.2944	1.6951	1.3076						
1-Apr-1974	0.501	1.3747	1.6951	1.3638						
1-Oct-1974	0.252	1.3747	1.9951	1.4115	1.3549	1.7707	1.3619	5.623	4.461	4.974
1-Jan-1975	1.000	1.4283	1.9951	1.4496	1.4283	1.9951	1.4496	5.418	12.673	6.438
1-Jan-1976	0.331	1.4326	2.1328	1.4627						
1-May-1976	0.669	1.4326	2.7214	1.5636	1.4326	2.5268	1.5302	0.300	26.650	5.560

APPENDIX B

(Continued)

Effective Date	Proportion of Year	Cumulative Benefit Level			Calendar Year Cumulative Benefit Level			Annual Percent Change in CY Benefit Level		
		Indemnity	Medical	Total	Indemnity	Medical	Total	Indemnity	Medical	Total
1-Jan-1977	1.000	1.5185	2.7214	1.6261	1.5185	2.7214	1.6261	6.000	7.702	6.268
1-Jan-1978	1.000	1.5185	2.7214	1.6261	1.5185	2.7214	1.6261	0.000	0.000	0.000
1-Jan-1979	0.496	1.5185	2.7214	1.6261						
1-Jul-1979	0.504	1.5185	3.0398	1.6554	1.5185	2.8820	1.6409	0.000	5.898	0.907
1-Jan-1980	1.000	1.5185	3.0398	1.6554	1.5185	3.0398	1.6554	0.000	5.479	0.885
1-Jan-1981	0.666	1.5641	3.0398	1.6852						
1-Sep-1981	0.334	1.5641	3.6873	1.7863	1.5641	3.2563	1.7190	3.000	7.119	3.842
1-Jan-1982	1.000	1.5641	3.9270	1.7845	1.5641	3.9270	1.7845	0.000	20.599	3.812
1-Jan-1983	1.000	2.1710	4.3629	2.2003	2.1710	4.3629	2.2003	38.800	11.100	23.300
1-Jan-1984	1.000	2.3316	4.7381	2.3213	2.3316	4.7381	2.3213	7.400	8.600	5.500
1-Jan-1985	1.000	2.3316	5.0935	2.3283	2.3316	5.0935	2.3283	0.000	7.500	0.300
1-Jan-1986	1.000	2.3316	5.0935	2.3283	2.3316	5.0935	2.3283	0.000	0.000	0.000
1-Jan-1987	0.496	2.3316	5.0935	2.3283						
1-Jul-1987	0.504	2.3316	5.4602	2.3330	2.3316	5.2784	2.3307	0.000	3.630	0.101
1-Jan-1988	1.000	2.3316	5.4602	2.3330	2.3316	5.4602	2.3330	0.000	3.445	0.099
1-Jan-1989	1.000	2.3316	5.4602	2.3330	2.3316	5.4602	2.3330	0.000	0.000	0.000
1-Jan-1990	1.000	2.3852	5.6841	2.3563	2.3852	5.6841	2.3563	2.300	4.100	1.000
1-Jan-1991	1.000	2.5021	6.2241	2.4011	2.5021	6.2241	2.4011	4.900	9.500	1.900
1-Jan-1992	0.497	2.5221	6.1618	2.4011						
1-Jul-1992	0.503	2.5221	6.1618	2.4131	2.5221	6.1618	2.4071	0.800	−1.000	0.251
1-Jan-1993	1.000	2.5221	6.1618	2.4131	2.5221	6.1618	2.4131	0.000	0.000	0.248
1-Jan-1994	0.496	2.2775	6.0078	2.1573						
1-Jul-1994	0.504	2.3891	6.3923	2.2134	2.3337	6.2016	2.1856	−7.469	0.646	−9.428
1-Jan-1995	0.496	2.3891	6.3923	2.2134						
1-Jul-1995	0.504	2.4608	6.5521	2.2510	2.4252	6.4728	2.2323	3.919	4.374	2.141
1-Jan-1996	0.497	2.4608	6.5521	2.2510						
1-Jul-1996	0.503	2.5297	6.6831	2.2848	2.4954	6.6180	2.2680	2.894	2.242	1.596

Source: W.C.I.R.B. of California - Analysis of Legislative Benefit Level Changes.

APPENDIX C

PART 1

DEVELOPMENT OF CANDIDATE VARIABLES CALIFORNIA AGGREGATE EMPLOYMENT

Year	Avg Monthly Employees	Annual Percent Change
1961	3,891,683	—
1962	4,071,877	4.6302
1963	4,216,436	3.5502
1964	4,346,448	3.0835
1965	4,464,625	2.7189
1966	4,707,406	5.4379
1967	4,840,158	2.8201
1968	5,041,894	4.1680
1969	5,272,325	4.5703
1970	5,240,190	−0.6095
1971	5,189,637	−0.9647
1972	5,913,892	13.9558
1973	6,383,331	7.9379
1974	6,588,356	3.2119
1975	6,564,524	−0.3617
1976	7,130,103	8.6157
1977	7,543,268	5.7947
1978	9,036,931	19.8013
1979	9,448,087	4.5497
1980	10,083,911	6.7297
1981	10,256,167	1.7082
1982	10,131,806	−1.2125
1983	10,312,305	1.7815
1984	10,900,212	5.7010
1985	11,378,074	4.3840
1986	11,644,237	2.3393
1987	12,094,751	3.8690
1988	12,556,920	3.8212
1989	13,005,986	3.5762
1990	13,328,057	2.4763
1991	12,796,072	−3.9915
1992	12,490,570	−2.3875
1993	12,253,883	−1.8949
1994	12,500,754	2.0146

Source: CA Statistical Abstract—Average Monthly Employment Covered by Unemployment Insurance—All Industries (1970 for 1961–1969; 1995 for 1970–1994).

APPENDIX C

PART 2

DEVELOPMENT OF CANDIDATE VARIABLES CALIFORNIA REAL GROSS STATE PRODUCT

Year	CA GSP $ Millions	Deflator 1982 = 100	Annual Change CA GSP	Annual Change Deflator	Pct. Change CA Real GSP
1961	—	—	—	—	—
1962	—	—	—	—	—
1963	65,905	31.4	—	—	—
1964	70,928	32.1	1.0762	1.0223	5.2747
1965	75,887	33.1	1.0699	1.0312	3.7592
1966	83,006	34.5	1.0938	1.0423	4.9424
1967	88,653	36.1	1.0680	1.0464	2.0695
1968	97,995	38.0	1.1054	1.0526	5.0108
1969	105,766	40.0	1.0793	1.0526	2.5335
1970	111,631	42.3	1.0555	1.0575	−0.1936
1971	119,192	44.9	1.0677	1.0615	0.5903
1972	132,199	47.0	1.1091	1.0468	5.9570
1973	146,473	49.9	1.1080	1.0617	4.3582
1974	160,979	53.9	1.0990	1.0802	1.7474
1975	179,858	59.1	1.1173	1.0965	1.8971
1976	201,536	62.9	1.1205	1.0643	5.2834
1977	227,590	67.3	1.1293	1.0700	5.5446

Series After Department of Commerce Methodology Revised

Year	Current Dollars	Deflator 1987 = 100	Annual Change CA GSP	Annual Change Deflator	Pct. Change CA Real GSP
1977	224,501	55.7	—	—	—
1978	255,552	60.2	1.1383	1.0808	5.3221
1979	287,821	65.4	1.1263	1.0864	3.6721
1980	319,804	71.5	1.1111	1.0933	1.6326
1981	358,920	78.4	1.1223	1.0965	2.3537
1982	382,317	83.5	1.0652	1.0651	0.0128
1983	416,061	86.6	1.0883	1.0371	4.9306
1984	468,127	90.5	1.1251	1.0450	7.6654
1985	511,110	93.7	1.0918	1.0354	5.4532
1986	552,110	96.5	1.0802	1.0299	4.8874
1987	599,088	100.0	1.0851	1.0363	4.7110
1988	650,313	103.9	1.0855	1.0390	4.4759
1989	702,755	108.2	1.0806	1.0414	3.7695
1990	752,761	113.1	1.0712	1.0453	2.4750
1991	767,189	117.5	1.0192	1.0389	−1.8998
1992	787,896	120.8	1.0270	1.0281	−0.1064
1993	—	—	—	—	—
1994	—	—	—	—	—

Source: U.S. Dept of Commerce, Bureau of Economic Analysis (1995 California Statistical Abstract).

APPENDIX C

PART 3

DEVELOPMENT OF CANDIDATE VARIABLES
CALIFORNIA UNEMPLOYMENT RATE

Year	Unemployment Rate	Annual Percent Change
1961	6.9	—
1962	5.8	−15.9420
1963	6.0	3.4483
1964	6.0	0.0000
1965	5.9	−1.6667
1966	4.9	−16.9492
1967	5.7	16.3265
1968	5.4	−5.2632
1969	5.2	−3.7037
1970	7.3	40.3846
1971	8.8	20.5479
1972	7.6	−13.6364
1973	7.0	−7.8947
1974	7.3	4.2857
1975	9.9	35.6164
1976	9.2	−7.0707
1977	8.2	−10.8696
1978	7.1	−13.4146
1979	6.2	−12.6761
1980	6.8	9.6774
1981	7.4	8.8235
1982	9.9	33.7838
1983	9.7	−2.0202
1984	7.8	−19.5876
1985	7.2	−7.6923
1986	6.7	−6.9444
1987	5.8	−13.4328
1988	5.3	−8.6207
1989	5.1	−3.7736
1990	5.6	9.8039
1991	7.5	33.9286
1992	9.1	21.3333
1993	9.2	1.0989

Source: CA Statistical Abstract (1970 for 1961–1967; 1974 for 1967–1969; 1995 for 1970–1994).

APPENDIX D

HAZARDOUSNESS INDICES

Indemnity Frequency Hazardousness Index

To measure the change in hazardousness from policy year to policy year, each classification was first assigned to one of fifteen groups of similar hazardousness of both frequency and severity. The fifteen groups were developed from California's nine retrospective rating hazard groups. Each of the fifteen groups is a subset of one retrospective rating hazard group. That is, all members of a group share the same retrospective rating hazard group or severity profile. Several hazard groups were not subdivided because their classifications' frequency profiles were reasonably homogenous. In all calculations, a class used the frequencies of its respective group.

The change in hazardousness for year t was then calculated in two ways. First, the exposures for year $t+1$ were extended by the indemnity frequencies for year t and this sum divided by the exposures for year t extended by the indemnity frequency for year t. This is the Laspeyres method. Second, the exposures for year $t+1$ were extended by the indemnity frequency for year $t+1$ and this sum divided by the exposures for year t extended by the indemnity frequency for year $t+1$. This is the Paasche method. The geometric mean was then taken of the Laspeyres and Paasche indices. This geometric mean is a Fisher index and the index selected to measure the change in hazardousness for year t.

Indemnity Pure Premium Hazardousness Index

The same procedure was performed to develop the indemnity pure premium hazardousness index except that, instead of using frequencies, indemnity pure premiums were used.

APPENDIX D

PART 1

INDEMNITY FREQUENCY HAZARDOUSNESS INDEX

	Change in Frequency Hazardousness					Changes in Frequency not Accounted for by Changes in Exposure Distribution		
Year	Method 1	Method 2	Geo Mean	Frequency Hazardousness Index	Annual Percent Change	Method 1	Method 2	Geo Mean
1961	1.0000	1.0000	1.0000	1.0000	—	1.0000	1.0000	1.0000
1962	0.9910	0.9895	0.9903	0.9903	−0.9738	1.0455	1.0439	1.0447
1963	0.9988	0.9986	0.9987	0.9890	−0.1282	1.0119	1.0117	1.0118
1964	1.0006	0.9882	0.9944	0.9834	−0.5649	1.0154	1.0029	1.0091
1965	0.9993	0.9882	0.9937	0.9772	−0.6289	0.9875	0.9765	0.9820
1966	0.9841	0.9830	0.9836	0.9612	−1.6416	1.0238	1.0226	1.0232
1967	0.9892	0.9861	0.9876	0.9493	−1.2374	1.0437	1.0404	1.0421
1968	0.9983	0.9967	0.9975	0.9469	−0.2488	1.0302	1.0285	1.0294
1969	0.9868	0.9859	0.9864	0.9340	−1.3610	1.0731	1.0721	1.0726
1970	0.9900	0.9899	0.9900	0.9247	−1.0034	0.9769	0.9768	0.9768
1971	1.0010	1.0007	1.0008	0.9254	0.0824	1.0150	1.0147	1.0149
1972	1.0005	1.0005	1.0005	0.9259	0.0505	1.0534	1.0534	1.0534
1973	0.9989	0.9984	0.9986	0.9246	−0.1369	1.1415	1.1408	1.1411
1974	0.9855	0.9865	0.9860	0.9117	−1.3983	1.2042	1.2054	1.2048
1975	0.9845	0.9838	0.9841	0.8972	−1.5890	1.0698	1.0691	1.0695
1976	1.0053	1.0047	1.0050	0.9017	0.4996	1.0485	1.0479	1.0482
1977	1.0102	1.0088	1.0095	0.9102	0.9490	1.0420	1.0405	1.0413
1978	1.0134	1.0116	1.0125	0.9216	1.2494	1.0376	1.0357	1.0366
1979	0.9982	0.9974	0.9978	0.9196	−0.2159	0.9718	0.9710	0.9714
1980	0.9921	0.9856	0.9888	0.9094	−1.1157	0.9631	0.9567	0.9599

APPENDIX D

PART 1

(Continued)

Year	Change in Frequency Hazardousness			Frequency Hazardousness Index	Annual Percent Change	Changes in Frequency not Accounted for by Changes in Exposure Distribution		
	Method 1	Method 2	Geo Mean			Method 1	Method 2	Geo Mean
1981	0.9818	0.9809	0.9813	0.8924	−1.8664	0.9648	0.9640	0.9644
1982	0.9797	0.9784	0.9790	0.8737	−2.0967	0.9923	0.9909	0.9916
1983	0.9977	0.9982	0.9980	0.8719	−0.2011	1.1042	1.1048	1.1045
1984	0.9984	0.9982	0.9983	0.8704	−0.1728	1.0881	1.0879	1.0880
1985	0.9929	0.9921	0.9925	0.8639	−0.7484	0.9805	0.9797	0.9801
1986	0.9888	0.9906	0.9897	0.8550	−1.0333	0.9721	0.9739	0.9730
1987	0.9981	0.9974	0.9978	0.8531	−0.2249	1.0283	1.0275	1.0279
1988	1.0002	1.0009	1.0006	0.8535	0.0574	1.0010	1.0018	1.0014
1989	0.9983	0.9981	0.9982	0.8520	−0.1788	1.0428	1.0426	1.0427
1990	0.9856	0.9867	0.9862	0.8403	−1.3816	1.0920	1.0932	1.0926
1991	0.9830	0.9826	0.9828	0.8258	−1.7178	0.9668	0.9664	0.9666
1992	0.9889	0.9874	0.9881	0.8160	−1.1852	0.8241	0.8228	0.8235
1993	0.9942	0.9944	0.9943	0.8114	−0.5717	0.8244	0.8246	0.8245

Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 at 2nd and 1993 at 1st report.
Formulas:
Change Due to Exposure Change:
Method 1 |@SUMPRODUCT(New Exposure Dist'n, Old Claims Freq)|/ |@SUMPRODUCT(Old Exposure Dist'n, Old Claims Freq)|
Method 2 |@SUMPRODUCT(New Exposure Dist'n, New Claims Freq)|/ |@SUMPRODUCT(Old Exposure Dist'n, New Claims Freq)|
Change Due to Frequency Change:
Method 1 |@SUMPRODUCT(New Claims Freq, Old Exposure)|/|@SUMPRODUCT(Old Claims Freq, Old Exposure)|
Method 2 |@SUMPRODUCT(New Claims Freq, New Exposure)|/|@SUMPRODUCT(Old Claims Freq, New Exposure)|

APPENDIX D

PART 2

INDEMNITY PURE FREQUENCY HAZARDOUSNESS INDEX

	Change in Pure Premium Hazardousness			Pure Premium Hazardousness	Annual Percent	Changes in Pure Premium not Accounted for by Changes in Exposure Distribution		
Year	Method 1	Method 2	Geo Mean	Index	Change	Method 1	Method 2	Geo Mean
1961	1.0000	1.0000	1.0000	1.0000	—	1.0000	1.0000	1.0000
1962	0.9890	0.9864	0.9877	0.9877	−1.2327	1.0761	1.0733	1.0747
1963	0.9986	0.9950	0.9968	0.9845	−0.3180	1.0129	1.0093	1.0111
1964	1.0076	0.9946	1.0011	0.9856	0.1069	1.0003	0.9873	0.9938
1965	0.9950	0.9842	0.9896	0.9753	−1.0402	1.0374	1.0261	1.0317
1966	0.9800	0.9788	0.9794	0.9552	−2.0622	1.0944	1.0931	1.0938
1967	0.9670	0.9554	0.9612	0.9181	−3.8829	1.0870	1.0740	1.0805
1968	1.0110	1.0065	1.0088	0.9262	0.8754	1.0975	1.0926	1.0951
1969	0.9897	0.9865	0.9881	0.9152	−1.1882	1.1573	1.1536	1.1554
1970	0.9852	0.9818	0.9835	0.9001	−1.6498	1.0258	1.0222	1.0240
1971	0.9963	0.9946	0.9954	0.8960	−0.4560	1.1042	1.1023	1.1033
1972	0.9976	0.9966	0.9971	0.8934	−0.2865	1.2135	1.2123	1.2129
1973	0.9933	0.9912	0.9923	0.8865	−0.7718	1.1754	1.1728	1.1741
1974	0.9771	0.9756	0.9764	0.8655	−2.3630	1.1788	1.1770	1.1779
1975	0.9640	0.9526	0.9583	0.8294	−4.1725	1.1859	1.1718	1.1788
1976	0.9964	0.9941	0.9953	0.8255	−0.4744	1.1032	1.1007	1.1019
1977	0.9973	0.9924	0.9949	0.8213	−0.5134	1.1051	1.0997	1.1024
1978	1.0073	0.9988	1.0030	0.8237	0.3024	1.0771	1.0680	1.0726
1979	1.0009	0.9955	0.9982	0.8222	−0.1822	1.0972	1.0913	1.0942
1980	0.9818	0.9611	0.9714	0.7987	−2.8614	1.1201	1.0965	1.1082

APPENDIX D

PART 2

(Continued)

	Change in Pure Premium Hazardousness			Pure Premium Hazardousness	Annual Percent	Changes in Pure Premium not Accounted for by Changes in Exposure Distribution		
Year	Method 1	Method 2	Geo Mean	Index	Change	Method 1	Method 2	Geo Mean
1981	0.9720	0.9654	0.9687	0.7737	−3.1298	1.1370	1.1293	1.1331
1982	0.9561	0.9497	0.9529	0.7373	−4.7112	1.2180	1.2099	1.2139
1983	0.9967	0.9926	0.9947	0.7333	−0.5343	1.3230	1.3176	1.3203
1984	1.0065	1.0045	1.0055	0.7374	0.5488	1.1430	1.1407	1.1419
1985	0.9893	0.9931	0.9912	0.7308	−0.8818	1.1258	1.1302	1.1280
1986	0.9907	0.9868	0.9887	0.7226	−1.1257	1.0446	1.0405	1.0425
1987	0.9974	0.9977	0.9975	0.7208	−0.2477	1.0901	1.0904	1.0902
1988	1.0007	1.0011	1.0009	0.7215	0.0905	1.0761	1.0765	1.0763
1989	0.9986	0.9938	0.9962	0.7187	−0.3836	1.1310	1.1256	1.1283
1990	0.9815	0.9834	0.9824	0.7061	−1.7561	1.2360	1.2384	1.2372
1991	0.9806	0.9798	0.9802	0.6921	−1.9773	0.9563	0.9555	0.9559
1992	0.9856	0.9839	0.9847	0.6816	−1.5257	0.7895	0.7881	0.7888
1993	0.9943	0.9952	0.9947	0.6780	−0.5266	0.7188	0.7195	0.7191

Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 at 2nd and 1993 at 1st report.
Formulas:
Change Due to Exposure Change:
Method 1 |@SUMPRODUCT(New Exposure Dist'n, Old Claim Severity)|/ |@SUMPRODUCT(Old Exposure Dist'n, Old Claim Severity)|
Method 2 |@SUMPRODUCT(New Exposure Dist'n, New Claim Severity)|/ |@SUMPRODUCT(Old Exposure Dist'n, New Claim Severity)|
Change Due to Frequency Change:
Method 1 |@SUMPRODUCT(New Claim Severity, Old Exposure)|/|@SUMPRODUCT(Old Claim Severity, Old Exposure)|
Method 2 |@SUMPRODUCT(New Claim Severity, New Exposure)|/|@SUMPRODUCT(Old Claim Severity, New Exposure)|

APPENDIX E

PART 1

LITIGATION RATES

A	B	C	D	E	F	G	H		I	J
		CWCI Pre-Reform		CWCI Post-Reform			Converted CWCI Pre-Reform			
Year	Quarter	2nd Quarter Litigation Rate	Policy Year Litigation Rate	2nd Quarter Litigation Rate	Policy Year Litigation Rate	Factor to Adjust Rate to Indemnity Claim Only Basis	2nd Quarter Litigation Rate	Percent Change	Policy Year Litigation Rate	Percent Change
1972	1st	5.9				6.1905	36.52			
1973	1st	5.6				5.3569	30.00			
1972	3rd	4.8				6.1905	29.71			
1972	4th	5.4				6.1905	33.43			
1972	2nd	5.3	5.36			6.1905	32.81		33.20	
1973	2nd	5.4	6.03			5.3569	28.93	−11.8335	32.28	−2.78
1974	2nd	6.4	7.15			4.2448	27.17	−6.0854	30.35	−5.96
1975	2nd	7.6	8.10			4.0255	30.59	12.6145	32.61	7.43
1976	2nd	8.4	8.21			3.9553	33.22	8.5978	32.48	−0.38
1977	2nd	8.1	7.66			3.8454	31.15	−6.2503	29.47	−9.29
1978	2nd	7.4	7.46			3.6984	27.37	−12.1332	27.60	−6.33
1979	2nd	7.5	7.00			3.7448	28.09	2.6212	26.21	−5.02
1980	2nd	6.7	6.76			3.7493	25.12	−10.5571	25.35	−3.27
1981	2nd	6.8	7.55			3.7140	25.26	0.5367	28.04	10.59
1982	2nd	8.0	7.88			3.5947	28.76	13.8667	28.31	0.95
1983	2nd	7.8	8.18			3.5038	27.33	−4.9643	28.64	1.19
1984	2nd	8.4	8.71			3.3112	27.81	1.7725	28.85	0.72
1985	2nd	8.9	10.34			3.2621	29.03	4.3817	33.72	16.89
1986	2nd	11.2	11.14			3.2093	35.94	23.8025	35.74	5.99
1987	2nd	11.1	11.29			3.1419	34.87	−2.9739	35.46	−0.78
1988	2nd	11.4	11.65			3.0582	34.86	−0.0319	35.63	0.46
1989	2nd	11.8	NA			2.9411	34.70	−0.4564	NA	NA
1990	2nd	Suspended	NA			2.7521	Suspended	NA	NA	NA
1991	2nd	13.8	13.99			2.7681	38.20	NA	38.72	NA
1992	2nd	14.1	NA			3.0393	42.85	12.1828	NA	NA

APPENDIX E

PART 1

(Continued)

A	B	C	D	E	F	G	H		I	J
		CWCI Pre-Reform		CWCI Post-Reform		Factor to Adjust Rate to Indemnity Claim Only Basis	Converted CWCI Pre-Reform			
Year	Quarter	2nd Quarter Litigation Rate	Policy Year Litigation Rate	2nd Quarter Litigation Rate	Policy Year Litigation Rate		2nd Quarter Litigation Rate	Percent Change	Policy Year Litigation Rate	Percent Change
1993	2nd			56.0	57.88	3.1553				
1994	2nd			59.0						
1995	2nd									
1996	2nd									

Notes:
C—CWCI Pre-Reform 2nd Quarter Litigation Rate
Total Number of Original Applications for Adjudication filed with the Workers Compensation Appeals Board/Total Number of Claims.
"Second Calendar Year Quarter. Definition Source: CWCI Bulletin, February 6, 1973. Total claims includes med-only claims.
D—Policy Year CWCI Pre-Reform Litigation Rate
P/Y T = 0.375 × C/Y T + 0.625 C/Y T + 1. Based on parallelogram method with no reporting lag. Note that a calendar quarter—not a calendar year—is being used to estimate a policy year.
E—CWCI Post-Reform 2nd Quarter Litigation Rate
Based on a sample of indemnity claims open during a second quarter. Determination of whether an attorney was involved is made for each claim. Med-only claims are excluded.
F—Policy Year CWCI Post-Reform Litigation Rate
P/Y T = 0.375 × C/Y T + 0.625 C/Y T + 1. Based on parallelogram method with no reporting lag.
G—Factor to Adjust Rate to Indemnity Claim Only Basis
The CWCI's Litigation Rates are the number of Applications for Adjudication ÷ Total Number of Claims. Since this is overwhelmed by medical-only claims (which are non-litigated), this factor is used to convert the CWCI's rate from an "All Claims" basis to an "Indemnity Claims" basis. This is the ratio of med-only and indemnity claims to indemnity claims only. The data is from the Bureau's class experience database covering policy years 1961–1991 at 3rd report, 1992 at 2nd report and 1993 at 1st report.
H—Converted CWCI Pre-Reform 2nd Quarter Litigation Rate
CWCI Pre-Reform 2nd Quarter Litigation Rate converted to an indemnity claim only basis: Col C × Col G
I—Converted Policy Year CWCI Pre-Reform Litigation Rate
Policy Year CWCI Pre-Reform Litigation Rate converted to an indemnity claim only basis: Col D × Col G

APPENDIX E

PART 2

FACTOR TO ADJUST LITIGATION RATE TO INDEMNITY CLAIMS ONLY BASIS

	A	B	A + B	(A + B)/B
	Incurred Claims			Correction
Year	Med-Only	Indemnity	Total	Factor
1961	590,123	89,341	679,464	7.6053
1962	610,218	95,745	705,963	7.3734
1963	621,304	99,376	720,680	7.2521
1964	613,813	99,559	713,372	7.1653
1965	624,771	103,549	728,320	7.0336
1966	640,366	107,601	747,967	6.9513
1967	645,128	109,981	755,109	6.8658
1968	654,184	118,573	772,757	6.5171
1969	659,713	124,065	783,778	6.3175
1970	606,247	117,435	723,682	6.1624
1971	619,880	122,227	742,107	6.0715
1972	692,801	133,475	826,276	6.1905
1973	675,018	154,932	829,950	5.3569
1974	605,127	186,491	791,618	4.2448
1975	584,591	193,222	777,813	4.0255
1976	611,465	206,908	818,373	3.9553
1977	636,973	223,863	860,836	3.8454
1978	654,758	242,645	897,403	3.6984
1979	650,266	236,912	887,178	3.7448
1980	613,630	223,191	836,821	3.7493
1981	574,589	211,710	786,299	3.7140
1982	527,867	203,441	731,308	3.5947
1983	584,794	233,559	818,353	3.5038
1984	627,773	271,618	899,391	3.3112
1985	617,048	272,771	889,819	3.2621
1986	608,364	275,370	883,734	3.2093
1987	627,052	292,759	919,811	3.1419
1988	623,028	302,703	925,731	3.0582
1989	630,176	324,655	954,831	2.9411
1990	602,945	344,132	947,077	2.7521
1991	562,022	317,859	879,881	2.7681
1992	514,609	252,344	766,953	3.0393
1993	447,016	207,407	654,423	3.1553

Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at 3rd report except 1992 at 2nd and 1993 at 1st report.

APPENDIX F

RATIO OF CUMULATIVE INJURIES TO TOTAL INDEMNITY CLAIMS

Year	Total Cumulative Injuries	Cumulative Indemnity Injuries	Total Indemnity Claims	Cumulative ÷ Total Indemnity Claims (%)	Percent Change in Ratio
1977	6,665	5,895	223,511	2.6375	—
1978	6,811	5,951	235,645	2.5254	−4.2482
1979	6,347	5,567	236,012	2.3588	−6.5982
1980	5,862	4,943	223,191	2.2147	−6.1084
1981	5,510	4,964	211,709	2.3447	5.8714
1982	6,717	6,032	203,441	2.9650	26.4534
1983	11,122	7,656	233,559	3.2780	10.5560
1984	14,041	10,506	271,618	3.8679	17.9977
1985	16,096	11,651	272,771	4.2713	10.4298
1986	16,195	12,254	275,370	4.4500	4.1829
1987	17,648	13,504	292,759	4.6127	3.6552
1988	21,103	15,948	302,703	5.2685	14.2187
1989	29,190	20,971	324,000	6.4725	22.8527
1990	41,568	29,318	345,517	8.4853	31.0964
1991	45,805	30,437	317,842	9.5761	12.8563
1992	27,075	15,977	251,233	6.3594	−33.5908
1993	17,561	9,360	207,412	4.5128	−29.0384
1994	16,365	8,590	200,642	4.2813	−5.1299

Source: W.C.I.R.B. of California Unit Statistical Reporting Plan.
California Class Experience at most current report level as of 4/22/97.

APPENDIX G

PART 1

DEVELOPMENT OF PRINCIPAL COMPONENTS OF ECONOMIC VARIABLES

STATGRAPHICS PLUS RESULTS

CALIFORNIA AGGREGATE EMPLOYMENT AND REAL GROSS STATE PRODUCT

Analysis Summary

Data variables:
  Annual Percent Change in CY AggE
  Annual Percent Change in CY rGSP

Data input: observations
Number of complete cases: 29
Missing value treatment: listwise
Standardized: no

Number of components extracted: 2

PRINCIPAL COMPONENTS ANALYSIS

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	24.5843	90.363	90.363
2	2.6220	9.637	100.000

This procedure performs a principal components analysis. The purpose of the analysis is to obtain a small number of linear combinations of the two variables which account for most of the variability in the data.

TABLE OF COMPONENT WEIGHTS

Variable (Annual Percent Change)	Component 1	Component 2
CYAggE	0.941545	−0.336888
CYrGSP	0.336888	0.941545

For example, the first principal component has the equation:

$$(0.941545 \times \text{Annual Percent Change CYAggE}) + (0.336888 \times \text{Annual Percent Change CYrGSP})$$

APPENDIX G

PART 2

DEVELOPMENT OF PRINCIPAL COMPONENTS OF ECONOMIC VARIABLES

STATGRAPHICS PLUS RESULTS

CALIFORNIA AGGREGATE EMPLOYMENT, REAL GROSS STATE PRODUCT AND UNEMPLOYMENT RATE

Analysis Summary

Data variables:
 Annual Percent Change in CY AggE
 Annual Percent Change in CY rGSP
 Annual Percent Change in CY UnEmp

Data input: observations
Number of complete cases: 29
Missing value treatment: listwise
Standardized: no

Number of components extracted: 3

PRINCIPAL COMPONENTS ANALYSIS

Component Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	309.5550	96.098	96.098
2	11.5599	3.589	99.687
3	1.0082	0.313	100.000

This procedure performs a principal components analysis. The purpose of the analysis is to obtain a small number of linear combinations of the three variables which account for most of the variability in the data.

TABLE OF COMPONENT WEIGHTS

Variable (Annual Percent Change)	Component 1	Component 2	Component 3
CYAggE	−0.188211	0.980960	−0.047901
CYrGSP	−0.115259	0.026374	0.992985
CYUnEmp	0.975342	0.192412	0.108101

For example, the first principal component has the equation:

$$(-0.188211 \times \text{Annual Percent Change CYAggE})$$
$$- (0.115259 \times \text{Annual Percent Change CYrGSP})$$
$$+ (0.975342 \times \text{Annual Percent Change CYUnEmp})$$

APPENDIX H

DEVELOPMENT OF SELF-INSURANCE SHARE INDEX

	Estimates of Workers Compensation Costs by Type of Insurance (000s)				Share of Total Workers Compensation Costs			
Year	Insurance Losses Incurred by Private Insurance	State and Federal Fund Disburse-ments	Self-Insurance Costs	Total	Insurance Losses Incurred by Private Insurance	State and Federal Fund Disburse-ments	Self-Insurance Costs	Annual Change in Self-Insurance Share Index
1960	100,894	37,124	15,635	153,653	0.6566	0.2416	0.1018	—
1961	115,756	43,813	18,080	177,649	0.6516	0.2466	0.1018	0.0000
1962	130,313	50,971	20,560	201,844	0.6456	0.2525	0.1019	0.0001
1963	147,035	56,566	23,090	226,691	0.6486	0.2495	0.1019	–0.0000
1964	163,720	63,890	28,052	255,662	0.6404	0.2499	0.1097	0.0079
1965	174,367	64,012	29,320	267,699	0.6514	0.2391	0.1095	–0.0002
1966	185,708	68,448	31,260	285,416	0.6507	0.2398	0.1095	–0.0000
1967	215,195	72,071	38,090	325,356	0.6614	0.2215	0.1171	0.0075
1968	223,513	70,615	39,120	333,248	0.6707	0.2119	0.1174	0.0003
1969	245,448	79,090	49,330	373,868	0.6565	0.2115	0.1319	0.0146
1970	278,215	87,677	55,615	421,507	0.6600	0.2080	0.1319	–0.0000
1971	286,177	92,862	50,895	429,934	0.6656	0.2160	0.1184	–0.0136
1972	306,032	105,351	57,970	469,353	0.6520	0.2245	0.1235	0.0051
1973	357,995	123,231	67,370	548,596	0.6526	0.2246	0.1228	–0.0007
1974	402,542	139,348	76,000	617,890	0.6515	0.2255	0.1230	0.0002
1975	472,406	156,161	94,500	723,067	0.6533	0.2160	0.1307	0.0077
1976	557,880	176,918	107,000	841,798	0.6627	0.2102	0.1271	–0.0036
Series After Change in Reporting of Self-Insurance Costs								
1976	557,880	176,918	220,000	954,798	0.5843	0.1853	0.2304	—
1977	658,426	194,901	275,000	1,128,327	0.5835	0.1727	0.2437	0.0133

APPENDIX H

(Continued)

	Estimates of Workers Compensation Costs by Type of Insurance (000s)				Share of Total Workers Compensation Costs			
Year	Insurance Losses Incurred by Private Insurance	State and Federal Fund Disburse-ments	Self-Insurance Costs	Total	Insurance Losses Incurred by Private Insurance	State and Federal Fund Disburse-ments	Self-Insurance Costs	Annual Change in Self-Insurance Share Index
1978	736,873	207,940	302,000	1,246,813	0.5910	0.1668	0.2422	−0.0015
1979	845,126	232,217	345,000	1,422,343	0.5942	0.1633	0.2426	0.0003
1980	950,288	233,427	379,000	1,562,715	0.6081	0.1494	0.2425	−0.0000
1981	1,068,512	242,811	498,000	1,809,323	0.5906	0.1342	0.2752	0.0327
1982	1,192,510	259,317	580,731	2,032,558	0.5867	0.1276	0.2857	0.0105
1983	1,290,575	273,063	697,000	2,260,638	0.5709	0.1208	0.3083	0.0226
1984	1,538,604	319,663	797,000	2,655,267	0.5795	0.1204	0.3002	−0.0082
1985	1,866,429	402,878	974,000	3,243,307	0.5755	0.1242	0.3003	0.0002
1986	2,096,934	523,916	1,124,000	3,744,850	0.5600	0.1399	0.3001	−0.0002
1987	2,328,020	647,921	1,275,000	4,250,941	0.5476	0.1524	0.2999	−0.0002
1988	2,548,616	817,689	1,444,000	4,810,305	0.5298	0.1700	0.3002	0.0003
1989	Data Not Available				*0.5341*	*0.1732*	*0.2927*	*– 0.0075*
1990	3,265,136	1,069,415	1,730,000	6,064,551	0.5384	0.1763	0.2853	−0.0075
1991	4,031,640	1,316,256	1,900,000	7,247,896	0.5562	0.1816	0.2621	−0.0231
1992	4,280,764	1,348,998	2,277,689	7,907,451	0.5414	0.1706	0.2880	0.0259
1993	4,074,854	1,201,452	2,348,756	7,625,062	0.5344	0.1576	0.3080	0.0200
1994	—	—	—	—	—	—	—	—
1995	—	—	—	—	—	—	—	—
1996	—	—	—	—	—	—	—	—

Source: Social Security Bulletin - Estimates of Workers' Compensation Payments, by State and Type of Insurance
All costs are incurred losses except for Self-Insurance costs prior to 1976, which are paid losses.

APPENDIX I

DEVELOPMENT OF CANDIDATE VARIABLES
REAL SEVERITY

	Nominal Claim Severity			California CPI		Real Claim Severity		
	Indemnity Claims					Indemnity Claims		
Year	Medical	Indemnity	Total	Calendar Year	Policy Year	Medical	Indemnity	Total
1961	547.00	1,559.38	302.90	29.5	29.7	547.00	1,559.38	302.90
1962	568.93	1,623.81	320.00	29.9	30.1	560.58	1,599.99	315.31
1963	603.64	1,676.92	340.35	30.4	30.7	584.27	1,623.12	329.43
1964	611.22	1,652.77	342.72	31.0	31.2	581.03	1,571.12	325.79
1965	684.39	1,698.51	366.12	31.5	31.8	638.64	1,584.98	341.64
1966	768.24	1,736.82	393.14	32.2	32.5	700.54	1,583.78	358.50
1967	787.65	1,726.13	402.74	33.0	33.6	695.79	1,524.81	355.77
1968	858.81	1,877.64	457.61	34.4	35.1	725.69	1,586.61	386.68
1969	918.08	2,025.74	505.48	36.1	36.9	739.11	1,630.85	406.94
1970	974.80	2,097.31	541.35	37.9	38.5	751.46	1,616.80	417.32
1971	1,083.67	2,262.98	593.84	39.3	39.8	806.91	1,685.04	442.18
1972	1,164.52	2,687.19	666.94	40.6	41.6	830.46	1,916.34	475.62
1973	1,221.79	2,720.69	784.80	43.0	44.8	808.46	1,800.29	519.31
1974	1,221.20	2,619.53	961.29	47.4	49.4	732.75	1,571.79	576.80
1975	1,385.69	2,739.90	1,085.50	52.3	53.7	765.88	1,514.36	599.96
1976	1,510.52	2,748.98	1,149.00	55.6	57.2	783.08	1,425.12	595.66
1977	1,623.95	2,828.77	1,232.57	59.5	61.5	782.83	1,363.62	594.16
1978	1,785.11	2,874.84	1,306.45	64.4	67.3	787.18	1,267.72	576.11
1979	2,028.72	3,075.86	1,450.95	71.3	75.9	792.68	1,201.83	566.93

APPENDIX I

(Continued)

	Nominal Claim Severity			California CPI		Real Claim Severity		
	Indemnity Claims					Indemnity Claims		
Year	Medical	Indemnity	Total	Calendar Year	Policy Year	Medical	Indemnity	Total
1980	2,395.10	3,369.39	1,642.96	82.4	86.2	824.77	1,160.27	565.76
1981	2,881.61	3,778.96	1,919.62	91.4	93.9	910.81	1,194.44	606.75
1982	3,380.19	4,581.95	2,359.72	97.3	98.0	1,023.60	1,387.52	714.58
1983	3,769.27	5,788.27	2,884.34	98.9	100.9	1,107.78	1,701.16	847.70
1984	3,939.12	6,207.89	3,229.91	103.8	105.8	1,104.54	1,740.71	905.67
1985	4,506.58	6,806.40	3,646.03	108.6	110.0	1,215.22	1,835.38	983.17
1986	5,026.02	7,100.35	3,967.23	112.0	113.9	1,308.89	1,849.10	1,033.16
1987	5,554.70	7,305.46	4,296.94	116.6	118.8	1,387.01	1,824.18	1,072.95
1988	6,111.86	7,737.58	4,746.32	121.9	124.4	1,457.05	1,844.62	1,131.51
1989	6,932.90	8,517.84	5,489.64	128.0	130.9	1,571.04	1,930.19	1,243.98
1990	7,598.95	9,352.25	6,424.39	135.0	137.3	1,641.52	2,020.26	1,387.79
1991	7,661.59	9,760.76	6,579.85	140.6	142.7	1,592.99	2,029.45	1,368.08
1992	7,422.81	9,100.86	5,746.17	145.6	147.2	1,496.16	1,834.39	1,158.21
1993	6,889.82	7,473.47	4,834.73	149.4	—	—	—	—
1994	—	—	—	—	—	—	—	—
1995	—	—	—	—	—	—	—	—
1996	—	—	—	—	—	—	—	—

Nominal Claim Severity from WCIRB of California's Unit Statistical Reports.
Calendar Year California CPI from the California Statistical Abstract—1995 (1982 – 1984 = 100).
Policy Year California CPI = $[(0.5832 \times CPI(t)) + (0.4168 \times CPI(t+1))]$.
Real Severity (t) = Real Severity $(t-1) \times [($Nominal Severity (t)/Nominal Severity$(t-1)) \div ($PY CPI (t)/PY CPI$(t-1))]$.

WORKERS COMPENSATION RESERVE UNCERTAINTY

DOUGLAS M. HODES, SHOLOM FELDBLUM,
AND GARY BLUMSOHN

Abstract

The increased emphasis on solvency monitoring of insurance companies, along with the American Academy of Actuaries' vision of an expanded role for the Appointed Actuary, have stimulated reserving specialists to quantify the uncertainty in their estimates. This paper measures the uncertainty in workers compensation loss reserve indications, compares it to the "implicit interest margin" in statutory (undiscounted) reserves, and examines the implications for capital requirements.

The paper uses a stochastic simulation analysis to model the loss reserving process, with separate but interlinked components for the process risk of loss development, the parameter risk of estimating future age-to-age link ratios, and autocorrelated future interest rates. In addition, the past monetary inflation implicit in paid loss development link ratios is replaced with stochastically generated future inflation rates that are linked to both the concurrent interest rates and the previous year's differential between the inflation rate and the interest rate. Separate simulations are performed for each accident year, and loss development tail factors are generated by an inverse power curve fit to extend the development from 23 years to ultimate.

An "expected policyholder deficit ratio" procedure is used to calibrate the capital needed to guard against reserve uncertainty. Because of the statutory benefits in workers compensation, the steady payment patterns, and the long average duration of compensation reserves, the implicit interest margin in statutory reserves exceeds the *capital required to guard against the variability in the reserve estimates at a 1% expected policyholder deficit level.*

The appendices to the paper contain descriptions of the simulation procedures, as well as a comparison of the paper's conclusions with those of the NAIC's risk-based capital formula.

ACKNOWLEDGEMENT

The authors are indebted to Aaron Halpert, Stephen Lowe, Roger Bovard, Roy Morell, Richard Fein, Louise Francis, and Joseph Boor for reviews of earlier versions of this paper. Any remaining errors, of course, should be attributed to the authors alone.

1. INTRODUCTION

Actuaries have developed a host of techniques for producing point estimates of indicated reserves. Current regulatory concerns, as reflected in the NAIC's risk-based capital requirements, and developing actuarial practice, as reflected in the American Academy of Actuaries' (AAA) vision of the future role of the Appointed Actuary, now stress the uncertainty in the reserve estimates in addition to their expected values. This paper demonstrates how the uncertainty in property/casualty loss reserves may be analyzed, and it draws forth the implications for capital requirements and actuarial opinions.

Genesis of this Paper

This paper was stimulated by the NAIC's risk-based capital efforts and by the AAA's vision of the valuation actuary:

- The reserving risk charge, which measures the potential for unanticipated adverse loss development by line of business, is the centerpiece of the NAIC property/casualty risk-based capital formula, accounting for about 40% of total capital requirements before the covariance adjustment and about 50% after the covariance adjustment (see Feldblum [23]). Because good actuarial analyses of loss reserve uncertainty are still lacking, the reserving risk charges were based on simple extrapolations from past experience, with a large dose of subjective judgment to keep the results reasonable.

- The Appointed Actuary presently opines on the reasonableness of the Annual Statement's point estimates of loss and loss adjustment expense reserves. The American Academy of Actuaries envisions an expanded role, in which the actuary opines on the financial strength of a company under a variety of future conditions (see [1]). The greater the uncertainty in the reserves, the greater the range of reasonable financial conditions that the actuary must consider.

Issues Addressed

This paper focuses on the uncertainty in workers compensation loss reserves. Specifically, it addresses the following issues:

- How should the uncertainty in loss reserves be measured? In other words: How might the variability in the loss reserve estimates best be quantified?
- What insurance characteristics, such as payment patterns and contract obligations, affect reserve uncertainty?
- How does the measure of variability that underlies risk-based capital requirements differ from the measure of variability that underlies the actuarial opinion? More specifically, how does the variability of the discounted, "net" reserves (i.e., loss obligations after consideration of return premiums and additional premiums on retrospectively rated policies, valued on an economic basis) differ from the variability of the undiscounted, "gross" reserves?

The Mixing of Lines

Why concentrate on workers compensation? Why not discuss property/casualty loss reserves in general, of which workers compensation is but one instance?

This is one of the primary errors that have hampered past analyses of loss reserve variability. Many observers have contrasted short-tailed lines like homeowners and commercial property with long-tailed lines like general liability and automobile liability, and they have noted the greater reserve uncertainty associated with the latter lines of business. Consequently, they have reasoned that reserve uncertainty is associated with the length of the average payment lag (i.e., reserves with longer average payment lags have greater uncertainty).

To see the error in this reasoning, let us extend the comparison to life insurance reserves. Single premium traditional life annuities have the longest reserve duration of the major life insurance products. Yet these products have low reserving risk, since the benefits are fixed at policy inception and mortality fluctuations are low.[1]

The bulk of workers compensation loss reserves that persist more than two or three years after the accident date are lifetime pension cases. The indemnity portions of these claims are disabled life annuities, with long duration and low reserve fluctuation for large compensation carriers. For the major insurance companies, the longest workers compensation reserves often have relatively low risk.[2]

[1]These products have significant interest rate risk, which is indeed affected by the average payment lag of the liabilities. For the quantification of interest rate risk for property/casualty insurance companies and the implications for risk-based capital requirements, see Hodes and Feldblum [35].

[2]See Feldblum [19], which compares reserve uncertainty among four property/casualty lines of business: workers compensation, automobile liability, products liability, and property. Compare also Meyers [47], who deals with same issue: "The purpose of this paper is to continue the debate on risk loading and discounting of loss reserves."

Meyers deals with workers compensation pension reserves, which have the highest ratio of implicit interest discount to reserve uncertainty, particularly for a large portfolio

The Peculiarities of Compensation Reserves

The quantification of reserve uncertainty must begin with the characteristics of the line of business. Four aspects of workers compensation reserves that affect the level of uncertainty are dealt with in this paper:

- *Payment Lag and Discount:* The previous section noted that most compensation reserves that persist more than two or three years after the accident date are lifetime pension cases. We compared these to life annuities, which are low risk reserves for large insurance companies. But the analogy is incomplete, since the statutory accounting treatment differs for these two types of business. Life annuities are discounted at rates close to current corporate bond rates. (The statutory discount rate for single premium immediate annuities—the life insurance product most comparable to workers compensation pension cases—issued in the first half of the 1990s is about 7% per annum.)

  Most property-casualty companies discount the indemnity portion of workers compensation lifetime pension cases at 3.5% or 4% per annum, which is well below their actual investment earnings. All other claims, as well as the medical portion of life pension cases, must be shown at undiscounted values in the statutory statements. The analysis in this paper indicates that the low fluctuations in these reserves, combined

---

of reserves. We look at the distribution of age-to-age link ratios, using a lognormal assumption and a Bayesian analysis of parameter risk; Meyers looks at the distribution of ultimate pension costs, using Makeham's mortality curve, again with a Bayesian analysis of the parameter risk. We use an expected policyholder deficit analysis, using a 1% EPD ratio, to calculate capital requirements; Meyers uses a utility function analysis to calculate the needed risk load. The two methods differ, though the results are similar: the implicit interest discount overwhelms the needed risk load or capital requirement. See especially Meyers' Tables 6.1 and 6.2 on page 182. The needed risk load in Meyers' illustration is about $400,000, with some variance depending on the parameters chosen in the utility function. The implicit risk load is $34.5 million assuming no tabular discounts and $9.3 million assuming tabular discounts at a 3.5% annual interest rate.

Hayne [32] shows a method of calibrating the uncertainty in the loss reserves based on loss frequency and loss severity assumptions. Hayne demonstrates his method, but does not provide numerical illustrations based on insurance data.

with the large implicit interest margin, create enormous hidden equity in statutory balance sheets.

- *Statutory Benefits:* What about non-pension cases? Do non-pension compensation reserves have the same uncertainty as many commercial liability reserves have? After all, industry studies have found similarly strong underwriting cycles and reserve adequacy cycles in several of these lines of business.[3]

Yes, underwriting results are driven by industry cycles, and so underwriting results vary greatly from year to year, whether in workers compensation, general liability, or automobile liability. But underwriting cycles reflect primarily the movement of premium levels, not fluctuations in loss experience. Reserve adequacy cycles are a secondary effect, driven by management desires to smooth calendar year operating results. They reflect the accounting treatment of company results, not the uncertainty inherent in the reserves themselves.[4]

When a products liability or medical malpractice accident occurs, the claim may not be reported for some time. Even after the claim is reported, the case may not be settled until years later, and the amount of the loss liability depends on the vagaries of court decisions, societal opinion, and jury awards. This is a major source of reserve uncertainty in the liability lines of business.

In workers compensation, most claims are reported rapidly. (It is hard for the employer to be unaware that a worker has been injured on the job and is on disability leave.) Benefits are mandated by statute, and disputes are resolved relatively quickly by administrative judges. For the major countrywide insurers with broad mixes of business, the paid-loss link ratios, or "age-to-age" factors, are stable in workers compensa-

[3]On the relationship between underwriting cycles and workers compensation reserve fluctuations, see Ryan and Fein [51] and Butsic [14].

[4]On the loss and premium effects of underwriting cycles, see Daykin, Pentikainen, and Pesonen [17], Cummins, Harrington, and Klein [16], and Feldblum [24].

tion, both for pension and for non-pension cases, unlike the comparable factors for the liability lines of business.[5]

- *Tail Development:* But don't workers compensation reserve estimates need large "tail factors," just as liability reserve estimates need? And aren't these tail factors highly uncertain, even as the liability tail factors are?

  Volatile commercial liability tail factors often reflect the emergence or the settlement of claims decades after the occurrence of the accident, such as toxic tort and environmental liability claims. This is true reserve uncertainty.

  Much of the volatility of workers compensation tail factors stems from two causes:

  1. First, mortality among permanently disabled workers, particularly for insurers with smaller blocks of business, is uncertain. For insurers with larger volumes of business, mortality fluctuations are less significant for annuity reserves.

  2. Second, workers compensation tail factors are affected by monetary inflation, both for cost of living adjustments to indemnity benefits and for all aspects of medical benefits. Inflation levels, especially for 30 or 40 years into the future, are extremely uncertain. This is parameter risk, not process risk, so it affects both large and small insurers.

---

[5]This paper emphasizes reserve estimates drawn from paid loss development methods. To avoid issues of company case reserving philosophy, we use loss payments only, not case reserves or reported losses, to quantify the uncertainty in the loss reserve estimates.

Reserving procedures based on case incurred loss development methods depend on company case reserving philosophy and stability. Some of the fluctuations in case reserves stem from different causes than the fluctuations in paid amounts. For instance, many temporary total disability claims are subsequently reclassified as permanent disability claims, causing an immediate change in the case reserve.

We do not have independent information about the reserve uncertainty inherent in case incurred reserving methods. The procedure used in this paper to quantify reserve uncertainty is not directly applicable to "incurred" methods. Analysis of the uncertainty in "paid" methods versus "incurred" methods would be a worthwhile subject for future studies.

This creates great uncertainty in the undiscounted reserve, and the actuary opining on reserve adequacy for statutory statements should consider a wide range of "reasonable" estimates. But the economic value of the reserve is less affected by long-term inflation rates for two reasons: (a) Much of the effect of high long-term inflation rate scenarios appears after 10 or 15 years, when the present value of these payments is much reduced. (b) The effect of high long-term inflation rates is often partially offset by high long-term interest rates.

- *Policy Type:* The type of insurance contract—such as "large dollar deductible" policy versus retrospectively rated policy—affects the degree of reserve uncertainty. A high percentage of the workers compensation contracts covering large employers are retrospectively rated. That is, the premium paid by the employer (the insured) is a function of the incurred losses. If loss reserves develop adversely, the insurer will collect additional retrospective premiums from the employer.

For loss-sensitive contracts, estimates of reserve uncertainty must be distinguished from their implications for capital requirements and actuarial opinions. Risk-based capital requirements reflect the equity needs of the insurer. Similarly, the envisioned future role of the appointed actuary is to opine on the financial strength of the insurer under various future conditions. To the extent that adverse loss development on a book of business is offset by favorable premium development, the financial condition of the insurer is unaffected, and there is less need for additional equity.

Other types of new insurance products have the opposite characteristics. Large dollar deductible policies and excess layers of coverage have higher reserve uncertainty per dollar of "net" reserves (i.e., reserves for the excess layer). A workers compensation reinsurer covering loss layers above high retentions may experience reserve variability unlike that experienced by a primary insurance carrier.

We may summarize the previous discussion in this section as follows. The novice actuary sees an insurer's large book of compensation reserves, notes the long payment lags and the strong underwriting cycles, and concludes: "There must be great uncertainty here. Moreover, unexpected development may severely affect the insurer's financial condition, so much capital is needed to guard against this risk." The experienced actuary replies: "No, because of the steady compensation payment patterns and the long payment lag of these claims, the reserving risk is low enough that it is outweighed by the implicit interest margin in the reserves."

2. MEASURES OF UNCERTAINTY

We have differentiated between the inherent uncertainty in reserve estimates and the accounting illusions caused by discretionary adjustments of reported reserves. Similarly, we may differentiate between actuarial measures of reserve uncertainty and regulatory measures of reserve uncertainty.

The Solvency Regulator and the Actuary

Suppose that the solvency regulator sees wide fluctuation in reported reserve levels and concludes that there is great uncertainty in the reserve estimates. The company actuary responds that the actual reserve indications have been stable. The shift in reported reserve levels from year to year stems simply from a desire to smooth calendar year earnings (see Ryan and Fein [51]).

> "*What difference does that make?"* replies the solvency regulator. *"We are concerned that the reported reserves may not be sufficient to cover the loss obligations of the company. What difference does it make whether the insufficiency stems from an inherent uncertainty in the reserve indications or from discretionary adjustment of the reported reserves?"*

We must differentiate between two types of reserve fluctuations:

- The valuation actuary tells the company's management how much capital it needs to guard against unexpected adverse events. Suppose the actuary's reserve analysis yields a point estimate of $800 million with a range of $650 million to $950 million, and the company is reporting $700 million on its statutory statements. The actuary's recommendation might be that the company needs $250 million of capital: $100 million for reserve "deficiencies" (the difference between the point estimate and the held reserves) and $150 million for reserve uncertainties.[6]
- The solvency regulator can not easily distinguish between adverse loss development stemming from unanticipated random occurrences and adverse loss development stemming from reserve inadequacies. The regulator estimates the variability of reported reserves and applies this figure to some base number. The base number might be the company's reported reserves (if the regulator believes that they are adequate) or an independent estimate of the company's reserve needs (if the regulator lacks confidence in the company's financial statements).

Regulators concerned with reserve uncertainty take the second viewpoint. Our primary interest in this paper is with the uncertainty inherent in the reserve indications themselves, the first viewpoint.

The difference is not in the *magnitude* of the uncertainty, but in the *method* of quantifying the uncertainty.

- The solvency regulator begins with the reserves reported by companies. How the companies determined these reserves, and

[6]In practice, the implicit interest margin in statutory reserves should be included in the valuation actuary's recommendation. To complete the illustration in the text, the actuary might add that there is $200 million of implicit interest margin in the statutory reserves, so only $50 million of capital is needed on an economic basis.

whether the reported reserves accurately reflect the actuary's indications, is irrelevant.

- The actuary examines the factors used to quantify reserve needs, such as age-to-age "link ratios," to determine the uncertainty in the reserve indications. How the company deviates from the reserve indications in its financial statements is not relevant to measuring the uncertainty inherent in the reserves.

Statistical and Financial Measures

We use several measures of reserve uncertainty in this paper: standard deviations, percentiles, and "expected policyholder deficits." The "expected policyholder deficit" (EPD) concept developed by Robert Butsic [13] is used here as a yardstick for the uncertainty in the reserve estimates. The EPD ratio allows us to translate "reserve uncertainty" into a "capital charge," thereby transforming an abstruse actuarial concept into concrete business terms. In Appendix A of this paper, we also discuss the "worst case year" concept used to measure reserve uncertainty for the reserving risk charge in the NAIC risk-based capital formula.[7]

Some readers will rightfully ask: "The NAIC worst case year concept is a simple but arbitrary accounting yardstick that is not supported by financial or actuarial theory. Why include it even in the appendix of an actuarial paper?"

The answer is important. This paper demonstrates that the implicit interest margin in full-value workers compensation reserves exceeds the capital needed to guard against unexpected reserve volatility. Some readers, aware of the 11% workers compensation reserving risk charge in the NAIC's risk-based capital formula, may mistakenly conclude that the "regulatory" and "actuarial" approaches to this problem yield different answers.

This is not so. The NAIC "regulatory" approach yields a similar result to that arrived at here. However, the workers com-

[7] For the NAIC worst case year concept, see Kaufman and Liebers [41] or Feldblum [23].

pensation charges were subjectively modified to produce capital requirements that seemed more reasonable to some regulators.[8] In fact, the apparent "unreasonableness" of the NAIC formula indications to these regulators stemmed from a misunderstanding of statutory accounting and of the risks of workers compensation business, not from any artifacts in the risk-based capital formula. A full discussion of the NAIC approach to reserve uncertainty embodied in the risk-based capital formula is presented in Appendix A.

3. THE QUANTIFICATION OF UNCERTAINTY

Attempts to measure reserve "uncertainty" often dissolve for failure to make clear (i) what exactly we seek to measure and (ii) how we ought to measure it.

This paper combines three elements to analyze the uncertainty of loss reserve estimates:

- A statistical *procedure* to quantify the uncertainty, relying on a stochastic simulation of the loss reserve estimation process.
- A *yardstick* to measure the uncertainty, relying on the expected policyholder deficit ratio.
- The *intuition* that explains the source of the reserve uncertainty, focusing on payment patterns, interest rates, and inflation rates.

Actuarial Procedures

Loss reserve estimates stem from empirical data, such as reported loss amounts or paid loss amounts, combined with actuar-

[8]For example, upon re-examining the workers compensation reserving risk charge in November 1996, using the NAIC formula but with more accurate figures and no subjective adjustments, the American Academy of Actuaries task force on risk-based capital found that the appropriate charge should be −12%, not the +11% in the NAIC risk-based capital *Instructions*. However, the AAA task force noted that any worker's compensation charge less than +10% would be politically infeasible to implement, so no effort was made to change the formula.

ial procedures, such as chain ladder development methods. Loss reserve uncertainty stems from both of these components.

- Random loss fluctuations may cause past experience to give misleading estimates of future loss obligations, and systemic changes (such as managed care) create uncertainty about future patterns.
- Imperfect actuarial analysis of the data may lead to invalid reserve estimates.

The two causes are intertwined. The ideal reserving actuary is ever watchful of data anomalies and will adjust the reserving procedures to avoid the most likely distortions (see, for instance, Berquist and Sherman [5]).

In this paper we do not measure the uncertainty stemming from imperfect actuarial practice. Rather, we assume a standard reserving technique that is often used for workers compensation; namely, a paid loss chain ladder development method.[9]

In practice, reserving actuaries use a variety of techniques. Even when employing a paid loss chain ladder development method, rarely does the reserving actuary follow the method by rote, with no analysis of unusual patterns. To the extent that actuarial judgment improves the reserve estimate, this paper overestimates the reserve uncertainty. To the extent that actuarial judgment masks the true reserve indications, this paper might underestimate the reserve uncertainty.

This paper measures the uncertainty inherent in the empirical data used to produce actuarial reserve estimates. It does not attempt to measure the uncertainty added or subtracted by the quality of actuarial analysis.

[9] We chose this technique, rather than a reported loss chain ladder development technique or Bornhuetter–Ferguson (expected loss) techniques, because it is dependent on claim payment patterns, and not on individual company case reserving practices. Thus, we are measuring the uncertainty caused by fluctuations in actual claim patterns, and not by changes in company case reserving practices.

Empirical Data

How should we measure the uncertainty inherent in the empirical data? The two extremes are described below, neither of which is sufficient by itself.

- We may simulate experience data, develop reserve indications, then continue the simulation to see how accurately the indications forecast the final outcomes.[10] This method is entirely theoretical. The amount of "uncertainty" depends on the simulation procedure. If the simulation procedure is firmly grounded in actual experience, the method works well. If the simulation procedure is chosen more for its mathematical tractability than for its empirical accuracy, the results may not mirror reality.
- We may look at actual experience, develop reserve indications at intermediate points in time, and then compare the indications with the final outcomes.[11] This method is "practical"—so practical, in fact, that the uncertainty measurements are often distorted by historical happenstance.[12]

A good actuarial procedure charts a middle course. We use stochastic simulation of the experience data to ensure statistically valid results. But the simulation parameters are firmly grounded in 25 years of actual paid loss histories from the country's largest workers compensation carrier.[13]

---

[10]See, for instance, Stanard and the Robertson discussion [56].

[11]This is the procedure used by the NAIC risk-based capital formula to estimate reserve uncertainty by line of business.

[12]See the report of the American Academy of Actuaries Task Force on Risk-Based Capital [44].

[13]Some reviewers of earlier drafts of this paper have questioned: Perhaps this insurer has more stable paid loss triangles than other insurers have, because of its size, because of its claim settlement practices, or because of its diversified mix of business. This is a valid comment. Small regional insurers may have different degrees of volatility in their reserve estimates. In particular, smaller insurers have greater process variance in the occurrence of lifetime pension cases, many of which have large total costs, both indemnity and medical. Expansion into new classifications or new states may similarly increase the uncertainty in the reserve estimates. See also the following footnote.

We describe the three elements of the analysis: (i) the stochastic simulation, (ii) the expected policyholder deficit ratio "yardstick," and (iii) the explanatory factors.

The Stochastic Simulation

We begin with 25 years of countrywide paid loss workers compensation experience, separately for indemnity and medical benefits, for accident years 1970 through 1994. From these data we develop 20 columns of paid loss "age-to-age" link ratios, as shown in Exhibits C-1 and C-2.[14]

We fit each column of "age-to-age" link ratios to lognormal curves, determining "mu" (μ) and "sigma" (σ) parameters for each. We perform 10,000 sets of simulations to generate the age-to-age factors that drive the simulated loss payments.

Twenty-five accident years yields 24 columns of "age-to-age" factors. The last four columns contain too few historical factors,

[14]Analysis of the uncertainty inherent in workers compensation loss reserve estimates must be grounded in actual workers compensation experience. The empirical data is the experience of the country's largest workers compensation carrier, with about 10% of the nation's experience during the historical period. To ensure confidentiality of the data, the dollar figures are normalized to a $100 million indicated undiscounted reserve.

Upon reviewing an earlier version of this paper, Stephen Lowe pointed out that "Because of its large market share, [your company's] experience probably does not respond to changes in mix of business by hazard group or state... . For smaller companies, changes in mix of business may add uncertainty beyond what is captured in your model." Similarly, the *Proceedings* referees for this paper write "For many companies, especially those with changes in mix of the type of business they write (different classes, different states) or changes in claims administration practices, the factors are not so stable."

This view is consistent with Allan Kaufman's recommendation that a "small company charge" be added to the risk-based capital formula because small companies experience greater fluctuation in underwriting results and in adverse reserve development. For political reasons, the small company charge was never added to the risk-based capital formula (see Feldblum [23]). In a review of the 1994 risk-based capital results, Barth [2], a senior research associate in the NAIC's research department, similarly concludes that "the R4 RBC i.e., (reserving risk) for companies with large reserves may be higher than necessary, relative to smaller companies."

Lowe, Kaufman, Barth and the *Proceedings* referees are correct. Small companies, or companies entering new markets or developing new products, may experience greater reserve uncertainty than implied here. This paper shows a method for quantifying reserve uncertainty, and it applies the method to the historical data of one particular insurer. To estimate the uncertainty of their own reserve estimates, readers should apply the methods described here to their own company's data. The numerical results in this paper can not necessarily be applied indiscriminately to other insurers.

so instead of fitting these columns to lognormal curves we include these development periods in the "inverse power curve" tail.[15] See Appendix C for a full description of the reserve estimation and simulation procedures.

Standard reserving methods, which forecast best-estimate future age-to-age link ratios, assume that the same factor will recur in each subsequent accident year. In actuarial parlance, when one "squares the triangle," the same age-to-age link ratios appear in each column for all subsequent accident years.

The procedure in this paper uses separate simulations for each subsequent accident year. We are simulating *actual* reserve development, where the process risk in each future accident year is independent of that in the other accident years.

Types of Risk

We categorize risk into two types: process risk and parameter risk (Freifelder [29], Miccolis [48]). We illustrate these components of risk with the fitting procedure described above.

Process Risk: Suppose that we *knew* that the observed (historical) link ratios came from a probability distribution with a mean of μ and a variance of σ^2, or "pdf (μ,σ^2)." For the stochastic analysis, we simulate new realizations of pdf (μ,σ^2).

In this case, we know the *expected* value with certainty. The uncertainty in the reserve estimate derives from the randomness of loss occurrences and loss settlements: that is, from the process risk in loss payments.

Parameter Risk: In truth, we do not know with certainty the expected value of the link ratios or the particular distribution from which they are a realization. We make two assumptions: (a) that the actual link ratios realized in the past and which will be realized in the future come from some distribution and (b)

[15] In addition, the Kreps parameter risk estimation procedure used in this paper does not work when there are only a few historical data points.

that this distribution has a particular form (such as lognormal). We estimate the parameters of the distribution from the historical values that we have observed.

This paper uses a parameter risk procedure developed by Kreps [42]. Using a Bayesian analysis, Kreps shows how to simulate from an unknown lognormal distribution based on a limited sample of data points.

The Kreps procedure is complex. To avoid repeating the mathematics of the Kreps paper [42], we simply note our choice of parameters for the Bayesian prior (readers interested in this subject should refer to that paper). Appendix C of this paper shows the equations we used to quantify the parameter risk. Appendix F of this paper provides a lay explanation of the parameter risk method, without attempting to reproduce the mathematics.

To use the Kreps procedure, one must assume a Bayesian prior distribution. Kreps uses a uniform distribution for the "translation" parameter (μ) and a distribution for the "scaling" parameter (σ) that depends on the user's prior assumptions, as reflected in a θ parameter. If the prior is uniform, then $\theta = 0$. The more conventional choice, if one is using a power-law prior, is to have $\theta = 1$. However, as Kreps pointed out to us (and as our own tests showed), "the conventional choice seems to give large values unreasonably often, given the nature of the business." He noted that $\theta = 2$ generally gives more reasonable results.[16]

Our simulations use $\theta = 2$. Even with this assumption, we found that the simulations occasionally yielded "unreasonable" results. By "unreasonable" we mean that workers compensation payments are based on statutory rules and are generally paid over the duration of a disability. Unlike some general liability claims, one rarely finds huge and unexpected lump-sum payments. Consequently, it is unreasonable to find a link ratio of say 3.0 as the factor for 15 years to 16 years of development.

[16]Kreps has also suggested that one might look for another distribution as a prior, based on our actuarial judgments about the business (private communication).

And yet, on rare occasions, that is what the simulations produce. These rare anomalies greatly affect the mean of the distribution, as well as measures of variability, like the standard deviation and the expected policyholder deficit.

Part of using actuarial judgment is to judge when the numbers being produced by mechanical formulas are not reasonable and to adjust the formulas so the results accord with insurance practice. In our case, we set a rule that if any simulated link ratio fell more than 50 standard deviations above the mean, the simulation is eliminated. In other words, we are trying to eliminate only the most extreme of the unreasonable simulations.

One might be concerned that a rule of this type would eliminate the "high" cases and thus would bias the results downwards. In fact, we found that the rule resulted in insignificant difference in the median result, or even in the 95th percentile of the distribution, and in most cases, the change in the mean was less than 1%. However, the change in the standard deviation and the expected policyholder deficit was more significant, and the results after eliminating the "outliers" are more reasonable.[17]

---

[17] An alternative procedure to quantify parameter uncertainty, which we have also tested on our data, is a procedure developed by Dickson and Zehnwirth [18]. The mean of the sample, μ, is an unbiased estimator of the mean of the distribution. If the distribution has a variance σ^2, and the sample has "n" observations, then the mean of the sample, as an estimator of the true mean of the distribution, has a variance of σ^2/n.

We want to use the sample data to simulate future realizations of the link ratios. The distribution from which these link ratios derive has a variance of σ^2. Furthermore, the whole distribution is "moving around" with a variance of σ^2/n. The total variance of the distribution from which we should simulate future realizations therefore has a variance of $\sigma^2 + \sigma^2/n$. The mean of this distribution is the sample mean, μ, which is an unbiased estimator of the true mean, as noted above. In sum, we must simulate from pdf $(\mu, \sigma^2 + \sigma^2/n)$, not from pdf (μ, σ^2).

Hayne [32] suggests a similar procedure: if the estimate of the μ of the lognormal is assumed to be unknown but to have a normal distribution with mean μ and variance σ'^2, then the final distribution is lognormal with parameters $(\mu, \sigma^2 + \sigma'^2)$.

Dickson and Zehnwirth [18] refer to these two distributions as the fitted curve and the predictive curve. The fitted curve is the best estimate of the probability distribution function; it does not include parameter variance. The predictive curve is the distribution function that one must use to simulate future realizations. It includes parameter variance, which reflects the uncertainty in the choice of parameters for the fitted curve. Our results using the Dickson–Zehnwirth procedure were similar to those using the Kreps [42] procedure. Consequently, we do not show the Dickson–Zehnwirth results in the text.

Shifting Distributions: The parameter risk discussed above assumes that there is a true distribution from which the observed link ratios are drawn, though we do not know this distribution. An additional source of variance is a shift in the true distribution, whether during the past historical period or during the future predictive period. For instance, the increasing involvement of attorneys in workers compensation claims during the 1980s may have contributed to the rising paid loss link ratios during this period, thus shifting the mean and perhaps also the variance of the distribution function. The change in the mix of claims from temporary total disability to permanent partial disability would similarly increase the mean and variance of the distribution (see Kaufman [40]). Conversely, the introduction of managed care in the 1990s may lead to a decrease in the mean of the paid loss link ratios and perhaps also their variance during this decade.

Mahler [46] refers to this as "shifting risk parameters." In his analysis of experience rating plan credibilities, Mahler divides the total expected claim variance into "within variance" and "between variance," and he includes the risk stemming from shifting risk parameters in the "within variance." We proceed similarly in our analysis. Following a suggestion by Mahler (private communication), we divide risk into process risk, specification risk, and parameter risk, where specification risk represents the risk of shifting risk parameters. The variance of the historical age-to-age link ratios stems from both process risk and from specification

---

Mathematically sophisticated readers may note some simplifications here, which are dealt with more fully in the Dickson and Zehnwirth paper. In particular, when we used the Dickson–Zehnwirth procedure, we assumed a lognormal prior distribution with known variance for the mean of the lognormal distribution (see Dickson and Zehnwirth [18, section 2.3, p. 4]. Dickson and Zehnwirth use normal distributions in their paper. As Zehnwirth has explained to the authors (private communication), "the predictive equation is lognormal, with a normal prior for the mean (μ) of the corresponding normal. The prior for $\exp(\mu)$, the median of the lognormal, is a lognormal. The prior for the mean of the lognormal, $\exp(\mu + 0.5 \times \sigma^2)$, is also a lognormal (scaled)."

Dickson and Zehnwirth also provide a parallel derivation for the predictive equation when the observed mean of the lognormal distribution comes from a Gamma prior with unknown variance. The predictive distribution is then a t-distribution, as shown in section 2.4 (pp. 4–5) and Appendix 2 (pp. 17–18) of Dickson and Zehnwirth [18]. See also Francis [27] for a similar comparison of normal and "t" distributions.

risk. Similarly, our quantification of future process includes both process risk and specification risk.

Tail Development

The paid loss development for 25 years is based on observed data. Workers compensation paid loss patterns extend well beyond 25 years. For each simulation, we complete the development pattern as follows:

- Given the 20 paid loss "age-to-age" link ratios from the set of stochastic simulations on the fitted lognormal curves, we fit an inverse power curve to provide the remaining "age-to-age" factors (see Sherman [52]). This fit is deterministic.
- The length of the development period is chosen (stochastically) from a uniform distribution of 30 to 70 years. The paid loss development is truncated at the stochastically selected age.

Because the simulated age-to-age link ratios in the first 20 development periods differ by accident year, the tail factors also differ by accident year.

4. INFLATION AND DISCOUNTING

We are primarily concerned with the economic values, or discounted values, of the reserves, not with undiscounted amounts. The exhibits here show results for undiscounted values in addition to discounted values, because statutory accounting requires the reporting of undiscounted reserves, and the Statement of Actuarial Opinion relates to the statutory figures. Butsic [13], however, emphasizes that his expected policyholder deficit (EPD) procedure, which is used here as one method of quantifying reserve uncertainty, is properly used only when balance sheet entries are stated on an economic basis, thereby avoiding "measurement bias." The EPD ratios are shown for the discounted values, not for undiscounted values.

Standard reserving procedures, when used to estimate discounted reserves, assume a fixed discount rate for unpaid losses.

Similarly, these procedures assume a fixed inflation rate for future loss payments during each development period that equals the inflation rate implicit in the historical age-to-age link ratios.

The treatment of inflation in this paper is more complex. Because of the long loss payment patterns, inflation strongly affects ultimate loss amounts. The effects on reserve variability depend on the manner in which inflation affects the loss amounts. For workers compensation, inflation affects medical benefits through the payment date. In about half of the U.S. jurisdictions, indemnity payments that extend beyond two years have cost of living adjustments (COLA's) that depend on inflation, so inflation affects the indemnity reserves as well.[18]

We use two methods for incorporating the effects of inflation into our simulation. One method leaves the effects of inflation implicit in the simulated link ratios. The other method segregates inflation from "real dollar" development and explicitly simulates future inflation rates. The two methods are described below.

- Unadjusted paid loss development patterns combine true development with the effects of inflation. That is to say, inflation is implicit in each paid loss age-to-age link ratio.[19] Were we to choose a single "best-estimate" link ratio for each development period, that would implicitly fix future inflation at the rate implicit in that "best-estimate" link ratio. Since we stochastically

[18]On the effects of inflation through the "payment date" versus through the "accident date," see Butsic [11], and the discussion by Balcarek.

The statutory rules for cost of living adjustments for indemnity benefits vary greatly by state. Some states have no COLA adjustments. Among the states which do have COLA's, most apply them only to disabilities extending beyond a certain time period, such as two years. In addition, many of these states cap the COLA's at specific levels, such as 5% per annum.

Properly quantifying the effect of the COLA adjustments on workers compensation indemnity reserve indications requires extensive work. For this paper we applied the stochastic inflation model to medical benefits only, where a single index can be used countrywide.

[19]For instance, the link ratio from 12 to 24 months equals the cumulative paid losses at 24 months divided by the cumulative paid losses at 12 months. A higher inflation rate during this development period raises the 24 month figure compared to the 12 month figure.

simulate the link ratios for each future accident year, we have a stochastic projection of inflation rates.

The simulated link ratios are independent of the simulated interest rates, so the implicit inflation rates are also independent of the interest rates. Although this is appropriate for link ratios, it may not be reasonable for inflation rates.

- In the second method, we deal with inflation by (a) stripping out past medical inflation from the historical loss triangles, thereby converting the figures to "real dollar terms," (b) determining "age-to-age" link ratios from the deflated loss amounts, and (c) simulating future inflation patterns and building them back into the projected (future) link ratios.

  Future inflation is simulated based on an autoregressive model that links the inflation rate both with the concurrent interest rate in the future scenarios, and with the discrepancy between the previous year's inflation rate and interest rate. The procedures used for doing this are described below.

Interest Rates

A stochastic model operates by first generating either interest rates or inflation rates—generally by some type of autoregressive function—and then generating the other index by a stochastic model with a partial dependence upon the first index.[20] Numerous methods of generating future interest paths have been developed. We used two of the simpler interest rate generators: an adaptation of the Wilkie/Daykin model, which has been used by the British Solvency Working Party, and the Cox, Ingersoll, Ross (CIR) model, which is used by many financial analysts in the United States. The generators produced comparable results. We describe the equations and results for the Cox, Ingersoll, Ross interest rate generator, which we have used for most of

[20]See Wilkie [58], Daykin, Pentikainen, and Pesonen [17], and the summary and discussion by Francis [28]. For an application to workers compensation reinsurance commutations, see Blumsohn [7].

our simulations. The procedures for the Wilkie/Daykin model are described in the previous version of this paper [34].

We begin with interest rates, simulating short rates for the CIR model, and then we simulate medical inflation rates.

The model begins by postulating a continuous process for interest rates. CIR decomposes the change in the short rate over an instantaneous period of time into a mean-reverting deterministic part and a Brownian motion stochastic part that is proportional to the square root of the current interest rate. That is

$$\partial r = a(b - r)\partial t + \sigma\sqrt{r}\partial Z,$$

where a is the mean-reverting parameter, b is the long-term average interest rate, σ is the annual volatility of the interest rates, and ∂Z is a standard Wiener process.[21] For our runs of the interest rate generator, we used parameters of

- $a = 0.2339$,
- $b = 0.050$,
- $\sigma = 0.0854$.

As a continuous time interest rate process, the CIR model has a "self-reflecting barrier" at $r = 0$. Interest rates cannot become negative, since if the interest rate process ever touches the line $r = 0$, the volatility is zero at that point and the interest rate reverts toward $a \times b$. In addition, CIR model provides for greater volatility as the interest rate becomes larger, which accords with our expectations about interest rate movements.

To run the continuous time CIR model in our simulation, we used monthly increments, with $a = 0.2339/12 = 0.0195$ and with $\sigma = 0.0854/(\sqrt{12}) = 0.0249$.

Some investment analysts concerned with short term bond options dislike equilibrium models, like the CIR model or the

[21] For an introduction to the CIR interest rate process, see Hull [38, Chapter 21].

Wilkie/Daykin model, that do not reproduce the current yield curve. Various arbitrage-free models have been proposed for securities trading operations that depend on interest rate expectations. For long-term dynamic financial analyses—like the quantification of uncertainty in loss reserves—equilibrium models seem satisfactory, and their parsimony perhaps make them preferable.

Inflation Rates

As noted above, there are two methods for dealing with inflation. Traditional reserving methods assume a continuation of the inflation rates implicit in the historical age-to-age link ratios. This procedure takes no account (i) of the autocorrelation in inflation rates or (ii) of the partial correlation with interest rates.

For the analysis in this paper, we strip inflation out of the historical age-to-age paid loss link ratios, and we stochastically simulate future inflation rates.

If we desired to simulate future inflation independently of future interest rates, we might use a procedure analogous to the autoregressive interest rate model, such as

$$\begin{aligned}\text{inflation rate} = {} & \text{average inflation rate}\\ & + \beta^{*}(\text{last year's inflation rate} - \text{average inflation rate})\\ & + \text{an error term.}\end{aligned}$$

Similarly, one could use a formula analogous to the CIR model for inflation rates. The parameters in each model would differ, of course, such as the average rates, the β coefficient, the form of the error term, the volatility parameter, and the starting value.

The stochastic inflation rate path would be independent of the stochastic interest rate path, even over the long term. Since interest rates and inflation rates are in fact correlated, the resulting scenario set would have many unrealistic elements.

Instead, we construct the autocorrelated model to include the current interest rate. There are no "standard" models for the dual generation of interest rates and inflation rates. We have used a model developed by Kreps, namely:

$$\text{Inflation}_t = c + d^*(\text{inflation}_{t-1}) - e^*(\text{Interest rate}_{t-1}) + f^*(\text{interest rate}_t) + \text{error}(t).$$

The fitted parameters are:

$$c = 1.33\%, \qquad d = 0.546, \qquad e = 0.264, \qquad f = 0.484.$$

The error term is normal, with a mean of zero and a standard deviation of 1.83%.

Inflation and Loss Development

To separately account for the effects of inflation on reserve development, we make the following adjustments to the data:[22]

- We convert the paid medical losses to real dollar amounts, using the medical component of the CPI. We then determine paid loss age-to-age link ratios from the deflated figures, we fit lognormal curves to each column of historical link ratios, and we run the simulation 10,000 times to determine the future link ratios.
- For each simulation, we stochastically generate a future interest rate path and a future inflation rate path, using the models described above.
- For each set of simulated link ratios and future inflation rates, we determine two required reserve amounts:
    1. The undiscounted (full value) reserves, using the link ratio and the inflation rate scenarios, and

[22]For a similar adjustment to reserving point estimates, see Richards [50, p. 387]: "These steps are designed to factor out the effects of inflation from historical loss data prior to forecasting, forecast the reserve using the current methodology and then replace the effects of inflation including an assumption of future inflation."

TABLE 1

INFLATION IMPLICIT IN LINK RATIOS; UNCORRELATED ACCIDENT YEARS

	Average Reserve Amount	Standard Deviation of Reserve	95th Percentile of Reserve	5th Percentile of Reserve	Capital Needed for 1% EPD Ratio
Undiscounted	100.0	14.5	125.0	80.4	—
Discounted	57.4	6.4	68.4	47.3	6.4

2. The discounted reserves, using the link ratio, inflation rate, and interest rate scenarios.

5. RESULTS

Table 1 shows results when inflation rates are not simulated separately; rather, the effects of future inflation are implicit in the simulated link ratios. Table 2 shows the results when inflation is removed from the historical link ratios and independently generated inflation rate paths are used for future years.

Exhibits 1 and 2 show the shapes of the probability distributions for the discounted and the undiscounted reserves. Exhibit 1, like Table 1, has no separate simulation of future inflation rates. Rather, the inflation implicit in the historical link ratios is presumed to continue into the future. Exhibit 2, like Table 2, uses the separate stochastic model for future inflation rates, as discussed above.

In Table 1, the average full value reserves are normalized to $100 million to facilitate the interpretation of the figures. The average discounted reserves are $57.4 million, with a standard deviation of $6.4 million. The 5th percentile of the distribution of required reserves is $47.3 million, and the 95th percentile is $68.4 million. To achieve a 1% EPD ratio, capital of $6.4 million is needed, above the $57.4 million of assets needed to support the expected (discounted) loss payments.

TABLE 2

INDEPENDENTLY GENERATED INFLATION RATES; UNCORRELATED ACCIDENT YEARS

	Average Reserve Amount	Standard Deviation of Reserve	95th Percentile of Reserve	5th Percentile of Reserve	Capital Needed for 1% EPD Ratio
Undiscounted	84.2	14.4	109.9	64.9	—
Discounted	49.8	4.6	57.6	42.6	4.1

Table 2 shows the corresponding figures when the future inflation rates are stochastically generated. The following items are noteworthy:

- The average discounted reserve decreases to \$49.8 million, with a standard deviation of \$4.6 million. High inflation scenarios, which strongly affect medium and long duration loss payments, have a lesser effect on discounted reserves. Moreover, high long-term inflation rates are often partially offset by high long-term interest rates.
- Nominal losses decrease to \$84.2 million, since we are projecting lower future inflation than is implicit in the historical loss triangle.
- The capital needed to achieve a 1% EPD ratio declines from \$6.4 million to \$4.1 million. The rationale is similar to that mentioned in the preceding paragraph. The high inflation scenarios that increase the capital requirement when inflation is implicit in future link ratios have a dampened effect when future inflation rates are linked to future interest rates.[23]

[23] When reserves are fully discounted, interest rate risk rises. This is particularly true for lines of business that are inflation sensitive, where the ultimate value of the loss payments depends on inflation up to the payment date. When inflation accelerates, nominal loss payments increase and market values of bonds decrease (if interest rates are linked to inflation rates). For further discussion of the capital required for interest rate risk, as well as the interplay with the capital required for reserving risk, see Hodes and Feldblum [35].

TABLE 3

INDEPENDENTLY GENERATED INFLATION RATES;
CORRELATED ACCIDENT YEARS

	Average Reserve Amount	Standard Deviation of Reserve	95th Percentile of Reserve	5th Percentile of Reserve	Capital Needed for 1% EPD Ratio
Undiscounted	84.5	18.4	119.2	62.9	—
Discounted	49.8	5.7	59.9	41.7	6.5

For Tables 1 and 2, the age-to-age link ratios are separately simulated for each future accident year. In other words, from each column of historical link ratios, we fitted a lognormal curve from which to simulate the future link ratios. We did not simulate a single link ratio which would be applied to all accident years that had not yet reached that stage of development. Rather, we separately simulated link ratios for each future accident year.

This assumes that the development in each accident year is independent of the development in other accident years at the same maturity. To test the results if the opposite assumption is made—namely, that the development at any given maturity is the same in all future accident years—we simulated a single age-to-age link ratio for each maturity and used it for all accident years. The results are shown below in Table 3. Since this procedure assumes perfect correlation among accident years, high or low link ratios are repeated in all accident years and the reserve uncertainty increases.

Uncertainty and Discounting

A common view is that discounted reserves are simply smaller than undiscounted reserves, but they exhibit the same degree of variability. This is not correct. As Exhibits 1 and 2 show, the probability distributions for undiscounted reserves are wide, whereas the corresponding probability distributions for discounted reserves are far more compact. The rationale for this is two-fold.

- First, much of the reserve variability comes from uncertainty in distant tail factors, which strongly wag estimates of undiscounted reserves but have less effect on discounted reserve estimates.
- Second, when using stochastic inflation rate paths with strong autocorrelation, much additional reserve variability results from high or low inflation scenarios. For discounted reserves, part of this variability is offset by corresponding high and low interest rate scenarios.

The magnitude of the difference between the two distributions depends on the parameters of the interest rate generator and the stochastic inflation process. The greater the volatility of interest rate and inflation rates, and the stronger the correlation between them, the greater the difference between the nominal and present value distributions.

Because statutory accounting mandates that insurers hold undiscounted reserves, we have shown results both for discounted reserves and for undiscounted (or "nominal") reserves in the exhibits. In particular, the means, standard deviations, and percentiles of the distributions are shown for both nominal and discounted reserves, though the capital requirements based on the expected policyholder deficit of 1% are applicable only to the discounted values. (See the discussion below in the text and in Appendix B regarding the expected policyholder deficit.) Moreover, the difference between the discounted and undiscounted reserve amounts is the "implicit interest margin" in the reserves, which is important for assessing the implications of the reserve uncertainty on the financial position of the insurance company.

Assumptions and Results

It is instructive to consider the relative reserve variability resulting from the different assumptions. Specifically, will the independent generation of future inflation rate paths increase or decrease reserve variability?

We begin with the results for our "base case," and we consider how each change in assumptions affects the estimated uncertainty. The base case assumes that:

- Link ratios are generated stochastically, incorporating both process risk and parameter risk.
- For the discounted reserves, autocorrelated interest rate paths are generated stochastically.
- Future inflation rates are not generated independently. Rather, the inflation embedded in the observed link ratios is assumed (implicitly) to continue into the future.

For *nominal* reserves, the independent generation of stochastic inflation rate paths adds an additional element of variability to the reserves. Accordingly, the standard deviation of the nominal reserve distribution is higher when inflation rates are independently generated. The coefficient of variation for the base case (Table 1) is about 14.5%, whereas it is about 17.1% when inflation rates are independently generated (Table 2).

For *discounted* reserves, the opposite is true. In the base case, the reserve discount rates are generated independently of the link ratios, in which the inflation rates are implicitly embedded, so reserve variability is high. When inflation rates are generated independently of the link ratios, they are correlated with the stochastically generated interest rates, and their effects partially offset each other, thereby dampening the reserve variability. For the discounted reserves, the capital ratio required for a 1% expected policyholder deficit ratio is 11.1% for the base case, while it is 9.2% when inflation rates are independently generated.

The implications of these results are important for the capital structure of a workers compensation insurer. In our illustration, the average undiscounted required reserves developed from a traditional reserve analysis, with no independent generation of future inflation rates, is $100 million. Most companies use tabular discounts for lifetime pension indemnity benefits, and some companies do not fully account for inflation of medical benefits. For most companies, the held statutory reserves would be between \$80 million and \$90 million.[24]

The average discounted required reserve is \$49.8 million. The implicit interest margin in the statutory reserves is about \$25 million to \$40 million.[25]

The capital required to achieve a 1% EPD ratio because of reserve uncertainty is about \$4.1 million, which is less than a fifth of the implicit interest margin in the statutory reserves. In other words, most insurers would need no additional capital to support the uncertainty in their workers compensation reserves.[26]

A common view is that workers compensation reserve estimates are highly uncertain, because of the long payment lags and because of the unlimited nature of the insurance contract form. This uncertainty creates a great need for capital to hedge against unexpected reserve development.

In fact, the risks in workers compensation lie elsewhere. There is great *underwriting* uncertainty in workers compensation, and regulatory constraints on the pricing and marketing of this line of business have disrupted markets and contributed to the financial distress of several carriers. But once the policy term has expired and the accidents have occurred, less uncertainty remains. The difference between the economic value of the reserves and the reported (statutory) reserves, or the implicit interest margin, is generally greater than the capital needed to hedge against reserve uncertainty.[27]

---

[24]The *Proceedings* reviewers have pointed out that some companies do not carry full value reserves, even on the statutory blank. For such companies, the held statutory reserves would be lower.

[25]The size of the implicit interest margin depends on the prevailing interest rates; it is larger in the high interest rate environments of the 1980's and smaller in the low interest rate environments of the 1990's.

[26]As noted earlier, some additional capital would be needed to support the default risks, market risks, and interest rate risks on the assets supporting the reserves.

[27]The implications for capital allocation to lines of business are important; for a full discussion, see Hodes, et al., [36]. For companies that carry adequate statutory reserves,

The EPD Yardstick

Several elements of our analysis may require further explanation. The following sections provide brief qualitative discussions of certain aspects of the analysis. The appendices provide more complete quantitative descriptions, as well as full documentation of our procedures.

As a yardstick to measure reserve uncertainty, we use the "expected policyholder deficit" (EPD) ratio developed by Butsic [13] for solvency applications. The EPD ratio allows us to:

- Compare reserve uncertainty across different lines of business,
- Compare reserve uncertainty with either explicit margins in held reserves or with the "implicit interest margins" in undiscounted reserves,
- Quantify the effects of various factors (such as the presumed variability of future inflation rates or the premium sensitivity on loss sensitive contracts) on reserve uncertainty, and
- Translate actuarial concepts of reserve uncertainty into more established measures of financial solidity.[28]

The Expected Policyholder Deficit

Were there no uncertainty in the future loss payments, the insurer need hold funds just equal to the reserve amount to meet its loss obligations. Since future loss payments are not certain, funds equal to the expected loss amount sometimes will suffice to meet future obligations, and sometimes they will fall short. The "policyholder deficit" is this shortfall.

---

the capital needed to support compensation reserves is negative, though positive capital is needed to support workers compensation underwriting operations. This is in contrast to the statutory accounting procedures used in many surplus allocation procedures in insurance pricing models. See, for instance, Feldblum [22], and particularly the Cummins/NCCI dispute there on the proper funding of the underwriting loss in the internal rate of return model.

[28]For a full discussion of the use of the EPD yardstick for measuring uncertainty, see Appendix B.

When the present value of the future loss obligations is less than the funds held by the insurance company to meet these obligations, the policyholder deficit is zero. When the present value of the future loss obligations is greater than the funds held, the policyholder deficit is the difference between the two. The expected policyholder deficit (EPD) is the average deficit over all scenarios, weighted by the probability of each scenario. In the analysis here, the expected deficit is the average deficit over all simulations, each of which is weighted equally.

Let us illustrate with the workers compensation reserve simulations in this paper. Suppose first that the company holds no capital besides the funds supporting the reserves. For the discounted analysis, the average reserve amount is $49.8 million (see Table 2). About half the simulations give reserve amounts less than $49.8 million. In these cases, the deficit is zero. The remaining simulations give reserve amounts greater than $49.8 million; these give positive deficits. The average deficit over all 10,000 simulations is the EPD. The "EPD ratio" is the ratio of the EPD to the expected losses, which are $49.8 million in this case.

Clearly, if the probability distribution of the needed reserve amounts is "compact," or "tight," then the EPD ratio is relatively low. Conversely, if the probability distribution of the needed reserve amounts is "diffuse" —that is, if there is much uncertainty in the loss reserves—then the EPD ratio is relatively great.

We have two ways of proceeding:

- We could assume that the company holds no assets besides those needed to support the expected loss obligations, and compare EPD ratios for different lines of business or operating environments.
- We may "fix" the EPD ratio at a desired level of financial solidity and determine how much capital is needed to achieve this EPD ratio.

The second approach translates EPD ratios into capital amounts, so we follow this method. We use a 1% EPD ratio as our benchmark, since Butsic notes that the reserving risk charges in the NAIC property-casualty insurance company risk-based capital formula are of similar magnitude as the charges needed for a 1% EPD ratio.[29]

Suppose the desired EPD ratio is 1%. If the reserve distribution were extremely compact, then even if the insurer held no capital beyond that required to fund the expected loss payments, the EPD ratio might be 1% or less. If the reserve distribution is more diffuse, then the insurer must hold additional capital to achieve an EPD ratio of 1%. The greater the reserve uncertainty, the greater the required capital.

Trends and Correlations

Two additional issues are of importance to reserving actuaries: correlations among link ratios and trends in link ratios.

- *Correlations*: The simulation procedure assumes that a particular link ratio is independent of the other link ratios in the same row. If the link ratios are not independent, the results may be overstated or understated.

  For instance, suppose that accident year 1988 shows a high paid loss link ratio from 24 to 36 months. Should one expect a higher than average link ratio or a lower than average link ratio from 36 to 48 months?

  The answer depends on the cause of the high 24 to 36 month link ratio. If it is caused by a speeding up of the payment pattern, but the ultimate loss amount has not changed, then one should expect a lower than average link ratio from 36 to 48 months. If it is caused by higher ultimate loss amounts

---

[29]For private solvency monitoring analyses, Butsic suggests that a higher ratio may be appropriate, such as 0.1%; see Butsic [13].

(e.g., because of lengthening durations of disability for indemnity benefits or because of greater utilization of medical services), then one should expect a higher than average link ratio from 36 to 48 months.[30]

- *Trends*: Our procedure uses unweighted averages of the link ratios in each column. During the 1980s, industry-wide paid loss link ratios showed strong upward trends, though this trend ceased in the early 1990s.[31] How would the recognition of such trends affect the variability of the reserves estimates as discussed here?

These two issues are related. First, the observed correlations among the columns of link ratios in the historical data result from the trends in these link ratios. When the trends are removed, the correlations largely disappear. Second, the trends affect the proper reserve estimate. The reserving actuary must investigate these trends and their causes, and then project their likely effect on future loss payments. That is not our interest in this paper. Rather, we ask: "What is the inherent variability in the reserve estimation process itself?"[32]

---

[30]For further explanation, see the discussion by H. G. White to Bornhuetter and Ferguson [8], as well as Brosius [10]. Compare also Holmberg [37, p. 254]:

> *There are different reasons we might expect development at different stages to be correlated. For instance, if unusually high loss development in one period were the result of accelerated reporting, subsequent development would be lower than average as the losses that would ordinarily be reported in those later periods would have already been reported. In this instance, correlation between one stage and subsequent stages would be negative. Positive correlation would occur if there were a tendency for weaker-than-average initial reserving to be corrected over a period of several years. In that case, an unusually high degree of development in one period would be a warning of more to come.*

Holmberg looks at incurred loss development. (To circumvent the effects of company case reserving practices on the variability of reserve estimates, we use paid loss development in this analysis.) Hayne [33] also discusses the possible correlations in the reserve estimation procedure, though he deals with them in a different fashion.

[31]See Feldblum, [25, section 7, and the references cited therein].

[32]To incorporate trends in this model, one would restate ("detrend") each column of historical link ratios to the current calendar year level before fitting these observed link ratios to a lognormal curve (see Berquist and Sherman [5]).

Let us take each of these issues in turn.

- *Correlations among columns*: Suppose one has two columns of observed link ratios, each from accident years 1971 through 1993, from 12 to 24 months and from 24 to 36 months, and that they are *not* correlated. We then apply a strong upward trend to both columns. That is, we increase the accident year 1972 link ratios by 1.02, the accident year 1973 link ratios by $(1.02)^2$, the accident year 1974 link ratios by $(1.02)^3$, and so forth.

  The resulting link ratio show a strong positive correlation. Indeed, we observe such a correlation in the historical link ratios used in our simulation. But if we remove the trend, the correlation disappears.

  This trend was caused primarily by the increasing liberalization of workers compensation benefit systems between the mid-1970s and the late 1980s. This liberalization, along with its associated effects (increasing paid loss link ratios, statewide rate inadequacies, growth of involuntary markets) ceased by the early 1990s, and has even reversed in many jurisdictions. The advent of managed care, along with workers compensation reforms in several state legislatures, may lead to further reduction in paid loss link ratios.

- *Correlations among years:* The chain ladder reserving technique involves "squaring the triangle." From each column in the observed triangle of age-to-age link ratio, we estimate a future link ratio, which is applied to all cells in that column of the triangle of future link ratios. When determining point estimates of indicated reserves, it is appropriate to use the same projected "best estimate" link ratio for all future accident years (i.e., for all the remaining cells in each column).

  The analysis here is different. We are not simulating a reserve estimate, or a reserve indication. Rather, we are simulating the potential future realization of loss development. In any simulation, the actual development will differ by accident year.

This is particularly important when studying reserve uncertainty. Our concern is not simply to quantify the expected development but to measure the variability of this development. Thus, when performing a stochastic analysis to determine reserve variability, it is proper to separately simulate the projected link ratios for each future accident year.[33]

For instance, suppose we have accident years 1970 through 1994, valued through December 31, 1994, and we are simulating the link ratios for the 48 months to 60 months development period. We need projected link ratios for accident years 1991, 1992, 1993, and 1994. We perform the stochastic simulation using the predictive curve four times, to give independently simulated link ratios for these four accident years. Similarly, once we have the projected link ratios, we fit inverse power curves to each accident year, to generate separate tail factors for each year.

Practicing actuaries may wonder about the materiality of this issue: does the increase in simulations increase or decrease the resultant reserve variability, and how large is this increase or decrease?

Consider the difference between (i) simulating once and using the same projected link ratio for all four accident years and (ii) simulating four times, once for each future accident year. The more separate (independent) pieces there are in each simulation of the total reserve requirements (as in the latter procedure), the tighter will be the distribution of the total reserve requirement. The fewer separate pieces there are in each simulation of the total reserve requirement (as in the former procedure), the greater will be the effect of individual

[33]The statement in the text is true if the variability stems from process risk. For the parameter risk component of the variability, one might argue that it is more proper to simulate once and to use the same factor each future accident year.

"outlying" factors, and the distribution of the total reserve requirement will be more widely spread.

Thus, the use of separate simulations decreases the estimated reserve variability. The effect is small, though, since there are many independent development periods in each simulation. The figures are shown in Table 3.

- *Trends*: Yes, there were trends, at least in the 1980s. Moreover, there are multiple reserving methods. The mark of the skilled actuary is to take the various reserve indications and the manifold causes for discrepancies among them and to project an estimate as close as possible to the true, unfolding loss payments.

  In our analysis, we have used the full column of observed link ratios to fit the lognormal curve, and then we have compared the simulated loss payments with their averages. Had we incorporated the "trends," and had we ignored old link ratios (because they are not relevant for today's environment), we might have produced tighter reserve distributions.

  If one places faith in the skills of reserving actuaries, then the use of a solitary reserving method overstates the uncertainty of the reserving process. Suppose the simulation produces actual loss payments considerably higher than the reserve estimate. Oftentimes, the experienced actuary would have noted signs that the paid loss estimate was underestimating the actual reserve need, and that other methods were giving higher indications. By combining the indications from several methods, the actuary might come closer to the actual reserve need, thereby reducing the uncertainty in the estimates.

Perhaps uncertainty can be reduced by actuarial judgments of trends and by actuarial weighing of various indications. The concern of this paper is more fundamental: even in rote applications of basic reserving techniques, how much uncertainty is produced by the fluctuations in loss data?

Federal Income Taxes

We have ignored income taxes, since their effect is uniform for most scenarios. Federal income taxes reduce the potential profits of the insurance company, but they also reduce the potential losses.

Suppose we determined that if there were no income taxes, an insurer has a 5% chance of exhausting its surplus because of the variability in loss reserves. Then with an income tax rate of 35%, the chance of exhausting its surplus is less than 5% for this insurer.

In effect, the U.S. government acts as a pro-rata reinsurer for all the company's business. It takes 35% of the revenue, and it pays 35% of losses plus expenses.

The risk on any particular insurance contract is not affected by federal income taxes. Rather, the contract is reduced in size: all revenues and expenditures are multiplied by 65%. Similarly, the variability in the loss reserves is not affected by federal income taxes. Rather, the reserves are simply reduced in size by a factor of 65%. Yardsticks such as percentiles or the coefficient of variation are not affected by federal income taxes.

Yardsticks such as the probability of ruin and the expected policyholder deficit ratio, however, relate reserve variability to the company's capital. The capital is on a post-tax basis, so the federal income tax rate is relevant. In addition, since the expected losses are on the company's books, taxes have already been paid on the assumption that these will be the ultimate losses. This means that the company's surplus reflects taxes at the expected level of losses. If one needs a certain amount of capital to pass a given "probability of ruin" test or a given "EPD ratio" test when one does not take into account federal income taxes, then one needs only 65% as much capital to pass the same test if one *does* take into account federal income taxes.[34]

[34] Similarly, Butsic [13] recommended that the charges in the NAIC risk-based capital formula be reduced for the offsetting effects of federal income tax recoupments, though his proposal was never implemented.

Because the potential federal income tax returns are affected by a host of factors, including the amount of taxes paid in the past three years and the amount of taxable income in the insurance enterprise's other operations, we have stated all our results on a pre-tax basis. For comparative analyses, a pre-tax basis is sufficient, such as for comparing reserve uncertainty among lines of business or among different policy forms. Practicing actuaries measuring capital requirements, however, should convert the results to a post-tax basis, using the particular tax situation of their own company or client.

6. STATUTORY BENEFITS

For the insurer from which these data were drawn, workers compensation reserves have about the same average payment lags as general liability GL reserves. There is great uncertainty in this company's GL reserves, as an equivalent analysis to that shown in this paper would show.[35] The causes of the GL reserve uncertainty illuminate the reasons for the compactness of the workers compensation reserve distribution.

- *IBNR Emergence*: Many GL claims are not reported to the insurer until years after the accident. For toxic tort and environmental impairment exposures, claims are still being reported decades after the exposure period (see, for instance, ISO [39] or Simpson, Smith, and Babbitt [53]). In contrast, most workers compensation accidents are known to the em-

[35] A full actuarial study of reserve uncertainty would apply the techniques used in this paper to all lines of business and compare the reserve distributions, EPD ratios, or capital requirements among them. The analysis must take into account the factors specific to each line that affect reserve fluctuations. For instance, just as we examine loss sensitive contracts for workers compensation, we must examine latent injury claims, such as those stemming from asbestos and pollution exposures, for general liability. For lines of business like general liability, results about reserve uncertainty can not always be generalized, since company practices vary so widely: some companies write premises and operations coverage for retail establishments, while other companies insure large manufacturing concerns; some companies are inundated by asbestos claims, while other companies have few of these cases. The extent of such analysis, of course, puts it beyond the scope of this paper.

ployer within days of the accident, and insurance companies are notified soon thereafter.

- *Claim Payment Patterns*: General liability losses depend upon judicial decisions and jury awards. Ultimate costs may not be known until years after the claim has been reported to the insurer. Even cases settled out-of-court are often settled "on the courthouse steps," after pre-trial discovery and litigation efforts have provided good indications of the expected judicial outcome.

Workers compensation benefits, in contrast, are fixed by statute, both in magnitude and in timing. The benefits may be determined either by agreement between the insurer and the injured worker, or by a workers compensation hearing officer. The major uncertainty in indemnity benefits is the duration of disability on non-permanent cases and the mortality rates on permanent cases. For sufficiently large blocks of business, both of these have relatively compact distributions. The major uncertainty for medical benefits is the rate of inflation and the extent of utilization of medical services. Over a large enough block of business, these risks also have relatively compact distributions, particularly when reserves are discounted.[36]

Butsic [12, p. 179], summarizes this view as follows:

> *For example, Workers Compensation reserves should have a lower risk than Other Liability reserves, even though the average payment durations are about the same, because Workers Compensation loss reserves consist partly of fixed, more predictable, life pension benefits.*

[36]Changes in the workers compensation system may either increase or decrease the reserve uncertainty. For instance, the advent of managed care may increase the uncertainty of ultimate loss payments, since the efficacy of managed care is not well known. It is equally possible that managed care will decrease reserve uncertainty, since the medical benefits may become easier to estimate. Our analysis partially incorporates this "specification risk" (to use Mahler's [46] term) in the process risk of the lognormal distribution (see the discussion above).

This paper provides the statistical support for the workers compensation half of this citation from Butsic.

7. CONCLUSIONS

Casualty actuaries have developed numerous methods of estimating required loss reserves. But reserves are uncertain, and actuaries are now being asked to quantify the uncertainty inherent in the reserve estimates.

Many past attempts to address this subject have foundered on one of two shoals. Some attempts are silver vessels of pure theory: loss frequencies are simulated by Poisson functions, loss severity is simulated by lognormal distributions, inflation is simulated by Brownian movements, and the results are much prized by hypothetical companies. Other attempts are steel vessels of actual experience: actual reserve changes, taken from financial statements, reveal how companies have acted in the past, though they offer imperfect clues about the uncertainties inherent in the reserve estimation process itself.

This paper glides between the shoals. Loss reserve uncertainty must be tied to the line of business. The uncertainty in workers compensation reserves is different from the uncertainty in general liability reserves even as it is different from the uncertainty in life insurance or annuity reserves. We begin with extensive data—twenty five years of experience from the nation's premier workers compensation carrier.

These data allow the actuary to develop reserve indications. Our concerns in this paper are different. We fit these data to families of curves to develop probability distributions of required reserves. The power of stochastic simulation techniques enables us to develop thousands of potential outcomes that are solidly rooted in the empirical data.

The analysis shows that workers compensation reserves, when valued on a discounted basis, have a highly compact distribution. To measure uncertainty, we use the "expected policyholder deficit" (EPD) ratio. For workers compensation, the amount of capital needed to achieve a 1% EPD ratio is only a small fraction of the "implicit interest margin" in the reserves themselves.

The vicissitudes of inflation are a major cause of workers compensation reserve fluctuations, and changes in interest rates strongly influence discounted values. This paper uses stochastically generated interest rates and inflation rates to model the reserve uncertainty.

The combination of rigorous actuarial theory with an extensive empirical database enables us to examine the uncertainty in the reserves themselves. Similar analyses should be performed for other lines of business, such as automobile insurance or general liability. Comparisons among the lines, as well as comparisons of reserve uncertainty with underwriting risks and with asset risks, would allow us to exchange preconceived notions with well-supported facts.

REFERENCES

[1] American Academy of Actuaries, "Position Statement on Insurer Solvency," *Actuarial Update*, September 1992.

[2] Barth, Michael, "Risk-Based Capital Results for the Property-Casualty Industry," *NAIC Research Quarterly* II, I, January 1996, pp. 17–31.

[3] Beard, R. E., T. Pentikainen, and E. Pesonen, *Risk Theory: The Stochastic Basis of Insurance*, Third Edition, London: Chapman and Hall, 1984.

[4] Bender, Robert K., "Aggregate Retrospective Premium Ratio as a Function of the Aggregate Incurred Loss Ratio," *PCAS* LXXXI, 1994, pp. 36–74; discussion by Howard C. Mahler, pp. 75–90.

[5] Berquist, James R., and Richard E. Sherman, "Loss Reserve Adequacy Testing: A Comprehensive Systematic Approach," *PCAS* LXIV, 1977, pp. 123–184; discussion by J. O. Thorne, *PCAS* LXV, 1978, pp. 10–34.

[6] Berry, Charles H., "A Method for Setting Retro Reserves," *PCAS* LXVII, 1980, pp. 226–238.

[7] Blumsohn, Gary, "Levels of Determinism in Workers' Compensation Reinsurance Commutations," *PCAS* LXXXVI, 1999.

[8] Bornhuetter, Ronald L., and Ronald E. Ferguson, "The Actuary and IBNR," *PCAS* LX, 1973, pp. 165–168.

[9] Bowers, Newton L., Jr., Hans U. Gerber, James C. Hickman, Donald A. Jones, and Cecil J. Nesbitt, *Actuarial Mathematics*, Itasca, Illinois: Society of Actuaries, 1986.

[10] Brosius, J. Eric, "Loss Development Using Credibility," CAS Part 7 examination study note, December 1992.

[11] Butsic, Robert P., "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. 51–102; discussion by Rafal J. Balcarek, pp. 103–109.

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

[13] Butsic, Robert P., "Solvency Measurement for Property-Liability Risk-Based Capital Applications," *Journal of Risk and Insurance* 61, 4, December 1994, pp. 656–690.

[14] Butsic, Robert P., "The Underwriting Cycle: A Necessary Evil?," *The Actuarial Digest* 8, 2, April/May 1989.

[15] Cook, Charles F., "Trend and Loss Development Factors," *PCAS* LVII, 1970, pp. 1–26.

[16] Cummins, David J., Scott E. Harrington, and Robert W. Klein, *Cycles and Crises in Property/Casualty Insurance: Causes and Implications for Public Policy*, National Association of Insurance Commissioners, 1991.

[17] Daykin, Chris D., Teivo Pentikainen, and M. Pesonen, *Practical Risk Theory for Actuaries*, First Edition, Chapman and Hall, 1994.

[18] Dickson, David C. M., and Ben Zehnwirth, "Predictive Aggregate Claims Distributions," Research Paper No. 27, Centre for Actuarial Studies, Department of Economics, The University of Melbourne, Australia, February 1996.

[19] Feldblum, Sholom, "Author's Reply to Discussion by Stephen Philbrick of Risk Loads for Insurers," *PCAS* LXXX, 1993, pp. 371–373.

[20] Feldblum, Sholom, "Completing and Using Schedule P," *Regulation and the Casualty Actuary*, edited by Sholom Feldblum and Gregory Krohm, NAIC, 1997.

[21] Feldblum, Sholom, Discussion of Teng and Perkins: "Estimating the Premium Asset on Retrospectively Rated Policies," *PCAS* LXXXV, 1998.

[22] Feldblum, Sholom, "Pricing Insurance Policies: The Internal Rate of Return Model," Casualty Actuarial Society Part 10A Examination Study Note, May 1992.

[23] Feldblum, Sholom, "NAIC Property/Casualty Insurance Company Risk-Based Capital Requirements," *PCAS* LXXXIII, 1996, pp. 297–435.

[24] Feldblum, Sholom, "Underwriting Cycles and Business Strategies," Casualty Actuarial Society *Forum*, Spring 1990, pp. 63–132.

[25] Feldblum, Sholom, "Workers' Compensation Ratemaking," Casualty Actuarial Society Part 6 Study Note, September 1993.

[26] Fitzgibbon, Walter J., Jr., "Reserving for Retrospective Returns," *PCAS* LII, 1965, pp. 203–214.

[27] Francis, Louise A., "A Model for Combining Timing, Interest Rate, and Aggregate Loss Risk," *Valuation Issues*, Casualty Actuarial Society Discussion Paper Program, 1989, pp. 155–216.

[28] Francis, Louise A., "Modelling Asset Variability in Assessing Insurer Solvency," *Insurer Financial Solvency*, Casualty Actuarial Society Discussion Paper Program, 1992, II, pp. 585–656.

[29] Freifelder, R. L., *A Decision Theoretic Approach to Insurance Ratemaking*, Homewood, IL: Richard D. Irwin, 1976.

[30] Gillam, William R., and Richard H. Snader, "Fundamentals of Individual Risk Rating," 1992, available from the CAS.

[31] Greene, Howard W., "Retrospectively-Rated Workers Compensation Policies and Bankrupt Insureds," *Journal of Risk and Insurance* 7, 1, September 1988, pp. 52–58.

[32] Hayne, Roger, "Application of Collective Risk Theory to Estimate Variability in Loss Reserves," *PCAS* LXXVI, 1989, pp. 77–110.

[33] Hayne, Roger, "A Method to Estimate Probability Level for Loss Reserves," Casualty Actuarial Society *Forum* I, 1994, pp. 297–356.

[34] Hodes, Douglas M., Sholom Feldblum, and Gary Blumsohn, "Workers Compensation Reserve Uncertainty," Casualty Actuarial Society *Forum*, Summer 1996, pp. 61–149.

[35] Hodes, Douglas M., and Sholom Feldblum, "Interest Rate Risk and Capital Requirements for Property-Casualty Insurance Companies," *PCAS* LXXXIII, 1996.

[36] Hodes, Douglas M., Sholom Feldblum, and Tony Neghaiwi, "The Financial Modeling of Property-Casualty Insurance Companies," *North American Actuarial Journal*, July 1999, pp. 41–69.

[37] Holmberg, Randall D., "Correlation and the Measurement of Loss Reserve Variability," Casualty Actuarial Society *Forum* I, 1994, pp. 247–278.

[38] Hull, John C., *Options, Futures, and Other Derivatives*, Fourth Edition, Englewood Cliffs, NJ: Prentice Hall, 2000.

[39] Insurance Services Office, *Superfund and the Insurance Issues Surrounding Abandoned Hazardous Waste Sites*, December 1995.

[40] Kaufman, Allan M., "Evaluating Workers Compensation Trends Using Data by Type of Disability," *Trends*, Casualty Actuarial Society Discussion Paper Program, 1990, pp. 425–461.

[41] Kaufman, Allan M., and Elise C. Liebers, "NAIC Risk Based Capital Efforts in 1990–91," *Insurer Financial Solvency*, Casualty Actuarial Society Discussion Paper Program, I, 1992, pp. 123–178.

[42] Kreps, Rodney, "Parameter Uncertainty in (Log)Normal Distributions," *PCAS* LXXXIV, 1997, pp. 553–580.

[43] Lee, Yoong-Sin, "The Mathematics of Excess of Loss Coverages and Retrospective Rating—A Graphical Approach," *PCAS* LXXV, 1988, pp. 49–78.

[44] Lowe, Stephen P., "Report on Reserve and Underwriting Risk Factors," Casualty Actuarial Society *Forum*, Summer 1993, pp. 105–171.

[45] Mahler, Howard C., Discussion of Bender "Aggregate Retrospective Premium Ratio as a Function of the Aggregate Incurred Loss Ratio," *PCAS* LXXXI, 1994, pp. 75–90.

[46] Mahler, Howard C., "An Example of Credibility and Shifting Risk Parameters," *PCAS* LXXVII, 1990, pp. 225–282.

[47] Meyers, Glenn G., "Risk Theoretic Issues in Loss Reserving: The Case of Workers Compensation Pension Reserves," *PCAS* LXXVI, 1989, pp. 171–192.

[48] Miccolis, Robert S., "On the Theory of Increased Limits and Excess of Loss Pricing," *PCAS* LXIV, 1977, pp. 27–59.

[49] Philbrick, Stephen, "Accounting for Risk Margins," Casualty Actuarial Society *Forum* I, 1994, pp. 1–90.

[50] Richards, William F., "Evaluating the Impact of Inflation on Loss Reserves," *Inflation Implications for Property/Casualty Insurers*, Casualty Actuarial Society Discussion Paper Program, 1981, pp. 384–400.

[51] Ryan, Kevin M., and Richard I. Fein, "A Forecast for Workers Compensation," *NCCI Digest* III, Issue IV, December 1988, pp. 43–50.

[52] Sherman, Richard, "Extrapolating, Smoothing, and Interpolating Development Factors," *PCAS* LXXI, 1984, pp. 122–192; discussion by Stephen Lowe and David F. Mohrman, *PCAS* LXXII, 1985, p. 182; author's reply to the discussion, p. 190.

[53] Simpson, Eric M., W. Dolson Smith, and Cynthia S. Babbitt, "P/C Industry Begins to Face Environmental and Asbestos Liabilities," *BestWeek*, January 1996.

[54] Simon, LeRoy J., "The 1965 Table M," *PCAS* LII, 1965, pp. 1–45.

[55] Skurnick, David, "The California Table L," *PCAS* LXI, 1974, pp. 117–140.

[56] Stanard, James N., "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques," *PCAS* LXXII, 1985, pp. 124–153.

[57] Teng, Michael T. S., and Miriam Perkins, "Estimating the Premium Asset on Retrospectively Rated Policies," *PCAS* LXXXIII, 1996, pp. 611–647.

[58] Wilkie, A. D., "A Stochastic Investment Model for Actuarial Use," *Transactions of the Faculty of Actuaries*, Vol. 39, 1986, p. 341.

EXHIBIT 1

DISTRIBUTIONS OF RESERVES
WITHOUT INFLATION ADJUSTMENTS

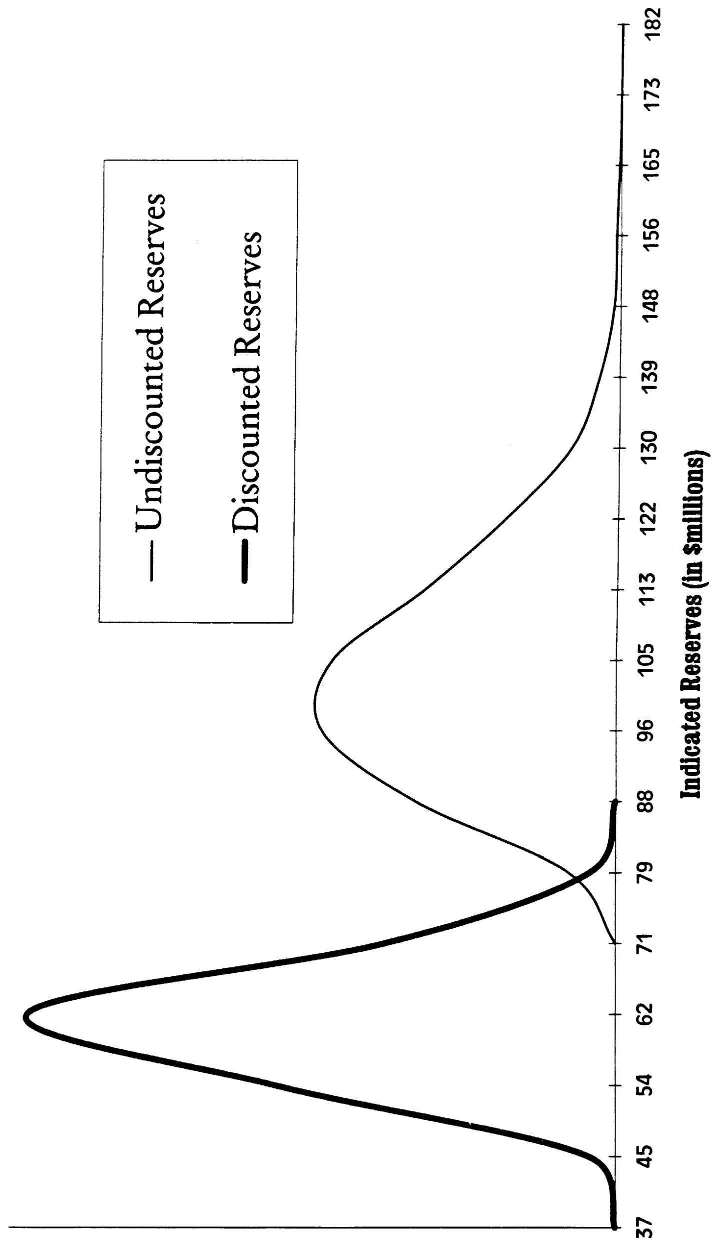
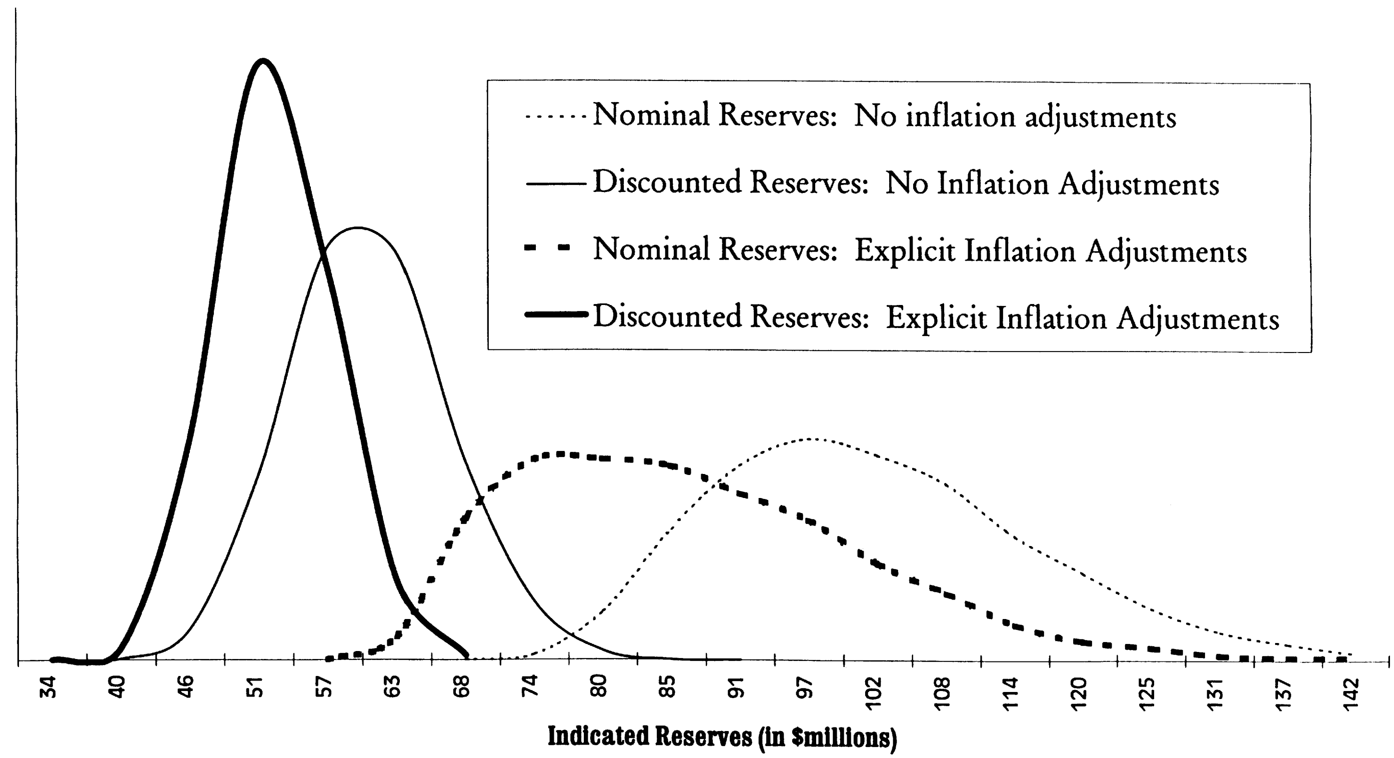


EXHIBIT 2

DISTRIBUTIONS OF RESERVES
WITH AND WITHOUT INFLATION ADJUSTMENTS



APPENDIX A

WORKERS COMPENSATION RESERVES AND RISK-BASED CAPITAL REQUIREMENTS

The text of this paper distinguishes between "regulatory measures" of reserving risk, as used in the NAIC's risk-based capital formula, and "actuarial measures" of reserving risk, as quantified here. The analysis in this paper shows that the volatility inherent in workers compensation reserve estimates is well below the implicit interest margin in statutory (undiscounted) reserves. The NAIC risk-based capital formula, however, has a reserving risk charge of 11% for workers compensation, even after incorporation of the expected investment income on the assets supporting the reserves.

An actuary unfamiliar with the development of the workers compensation reserving risk charge in the risk-based capital formula might conclude that "regulatory measures" of workers compensation reserving risk give high capital charges whereas "actuarial measures" give low charges. This is not correct. The risk-based capital formula gives a low charge for workers compensation reserving risk, even as the actuarial analysis in this paper provides. The final 11% charge in the risk-based capital formula is an *ad hoc* revision intended to provide more "reasonable" capital requirements.

The workers compensation reserving risk charge was one of the most contested aspects of the risk-based capital formula, and the derivation of the final 11% charge was never publicly revealed. This appendix explains the issues relating to the workers compensation reserving risk charge, and it shows the charge resulting from the NAIC "worst-case year" method.

Adverse Development and Loss Reserve Discounting

The reserving risk charge in the risk-based capital formula bases the capital requirements on the historical adverse loss development in each line of business. The "worst-case" industry-wide adverse loss development as a percentage of initial reserves is determined from Schedule P data, and this figure is then reduced by a conservative estimate of expected investment income.

For workers compensation, the original risk-based capital formula produced a charge of 0.4%.[37] The 1992 Best's *Aggregates and Averages* shows a gross "worst-case year" adverse development of 24.2%, as derived in Exhibit A-1.

Two considerations related to loss reserve discounting complicate the estimation of the reserving risk charge for workers compensation.

- Statutory accounting conventions for property/casualty insurers are conservative, particularly with regard to the reporting of loss reserves. The Annual Statement shows undiscounted reserves, leaving a large margin in the reserves themselves, particularly for long-tailed lines of business.

  In other words, property/casualty insurers have two potential margins to ensure adequacy of loss reserves: an implicit interest margin in the reserves themselves, and an explicit capital requirement provided by the reserving risk charge. To avoid "double counting," the risk-based capital formula offsets the implicit interest margin against the explicit reserving risk charge.

- The "double margin" occurs when reserves are reported on an undiscounted basis. But some property/casualty reserves are reported on at least a partially discounted basis. For instance, many carriers use tabular discounts for workers compensation lifetime pension claims. The special statutory treatment of workers compensation lifetime pension cases necessitates adjustments to the reserving risk charge.

---

[37]For a full description of the risk-based capital reserving risk charges, see Feldblum [23].

Both the NAIC Risk-Based Capital Working Group and the American Academy of Actuaries task force on risk-based capital spent months working on these two topics. The issues are complex, and no clear explanation is available for either regulators or for industry personnel. To clarify the issues, this appendix discusses the treatment of the implicit interest margin in statutory reserves and the adjustments needed for tabular loss reserve discounts in workers compensation.

Payment Patterns and Discount Rates

The amount of the implicit interest margin, or the difference between undiscounted (full-value) reserves and discounted (economic) reserves, depends on two items: the payout pattern of the loss reserves and the interest rate used to discount them.

For most lines of business, the NAIC risk-based capital formula uses the IRS loss reserve payment pattern along with a flat 5% discount rate. These choices were made for simplicity. Using the IRS discounting pattern avoids the need to examine loss reserve payout patterns, and using a flat 5% discount rate avoids the need to examine investment yields. For some lines of business, these choices are acceptable proxies for good solvency regulation. For workers compensation, greater complexities arise.

- *Payment Pattern*: The IRS procedure assumes that all losses are paid out within 15 years. Moreover, the pattern is based on the industry data for the first 10 years as reported in Schedule P.

For short-tailed lines of business, this is not unreasonable, since most losses are indeed paid out before the Schedule P triangles end. Workers compensation reserves, however, have a payout schedule of about 50 years, since permanent total disability cases—which are a small percentage of the claim count but a large percentage of the dollar amount—extend for the lifetime of the injured worker.

- *Discount Rate*: For its discount rate, the IRS uses a 60 month rolling average of the federal midterm rate, which is defined as the average yield on outstanding Treasury securities with maturities between 3 and 9 years. Since 1986, the IRS discount rate has ranged between 6% and 8%.

Actual portfolio yields have been about 100 to 200 basis points higher, since insurance companies invest not only in Treasury securities but also in corporate bonds, common stocks, real estate, and mortgages. However, these latter investment vehicles have additional risks, such as default risks, market risks, and liquidity risks. As a loss reserve discounting rate, many casualty actuaries would prefer the 6% to 8% "risk-free" Treasury rate to the 8% to 10% portfolio rate, particularly for statutory financial statements which emphasize solvency.

The NAIC risk-based capital formula uses a flat 5% discount rate. A variety of justifications have been given, such as:

- The 5% rate is simple, obviating any need to examine actual investment yields and cutting off any arguments about the "appropriate" rate.
- The 5% rate adds an additional margin of conservatism, since it is 1 to 3 points lower than the corresponding IRS rate.

For lines of business where the implicit interest margin in the reserves is small, the difference between the 5% NAIC rate and the 6% to 8% IRS rate is not that important in setting capital requirements. For a line of business like workers compensation, however, where the discount factor ranges from 60% to 83%, depending upon the assumptions, the choice of discount rate has a great effect.

We begin the analysis below with the current NAIC risk-based capital assumptions to see the unadjusted charge produced by the formula. We then turn to actual payment patterns and investment yields to address the fundamental questions: "What is the risk associated with workers compensation loss reserves? And how much capital ought insurance companies to hold to guard against this risk?"

The IRS Discount Factor

The IRS determines the loss reserves payout pattern by examining the ratio of paid losses to incurred losses by line of business for each accident year from Part 1 of Schedule P. The data are drawn from Best's *Aggregates and Averages*, and the payout pattern is redetermined every five years.

Schedule P shows only 10 years of data, though several lines of business, such as workers compensation, have payout schedules extending up to 50 years. The IRS allows an extension of the payout pattern beyond the 10 years shown in Schedule P for up to an additional 6 years. The extension of the payout pattern does not rely on either empirical data or financial expectations. Rather, the payout percentage in the tenth year is repeated for each succeeding year until all reserves are paid out.

Accident Years vs. Aggregate Reserves

The IRS determines a discount factor for each accident year. The risk-based capital formula uses a single discount factor for all accident years combined. Thus, one must use a weighted average of the discount factors, based on the expected reserves by accident year.[38]

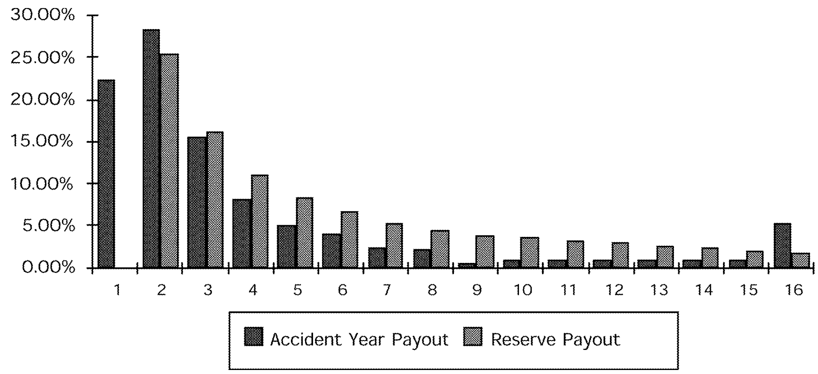
Exhibit A-2 shows the workers compensation payment pattern using the IRS procedures and the Best's *Aggregates and Averages* Schedule P data.

- The left-most column shows the payment year. Because workers compensation reserves are paid out so slowly, the IRS extends the payment schedule for the full 16 years. It is still far too short, particularly for lifetime pension cases.

[38]For simplicity, the calculations in this paper assume that the volume of workers compensation business is remaining steady from year to year. A theoretical refinement would be to use the actual volume of industrywide workers compensation reserves in each of the past ten years, though there is no significant difference in the result.

FIGURE 1

WORKERS COMPENSATION PAYOUT PATTERNS



- The middle column shows the payment schedule for an individual accident year. This payment schedule says that 22.34% of an accident year's incurred losses are paid in the first calendar year, 28.36% in the next calendar year, and so forth.
- The right-most column shows the payment schedule for the aggregate reserves, assuming no change in business volume over the 16 year period. This payment schedule says that 25.42% of the reserves will be paid in the immediately following calendar year, 16.14% in the next calendar year, and so forth.

Figure 1 shows the payout patterns for an individual accident year and for the aggregate reserves. The horizontal axis represents time since the inception of the most recent accident year. The accident year payout pattern begins with the first losses paid on the policy, soon after the inception of the accident year. The valuation date of the reserves in the graph is the conclusion of the most recent accident year, so the payout pattern begins in the second year since inception.

The payout pattern is combined with an annual interest rate to give the discount factor, or the ratio of discounted reserves to undiscounted reserves. With an interest rate of 5% per annum, the discount factor for the reserves is 82.98%. The risk-based capital formula would therefore indicate a reserving risk charge of

$$[1.242 \times 82.98\%] - 1 = 3.06\%.$$

The 3% reserving risk charge depends upon the conservative 5% annual interest rate and the short IRS payment pattern. More realistic interest rates and payment patterns, even when still containing margins for conservatism, lead to a negative charge. We discuss these in conjunction with tabular loss reserve discounts below.

Discounted Reserves

What if an insurer holds discounted reserves, or partially discounted reserves? How should the reserving risk procedure described above be modified to account for the reserve discount?

This question is most relevant for workers compensation. Statutory accounting normally requires that insurers report undiscounted, or full-value, reserves. An exception is made for workers compensation lifetime pension cases, where insurers are allowed to value indemnity (lost income) reserves on a discounted basis. State statutes often mandate conservative discount rates, usually between 3.5% and 5% per annum, with the most common being 4%. These reserve discounts are termed "tabular" discounts, since they are determined from mortality tables, not from aggregate cash flow analyses.

Adverse Development and Interest Unwinding

The combination of three factors—(a) adverse development, (b) the unwinding of interest discounts, and (c) weekly claim payments—produces intricate results that are difficult even for the most technically oriented readers to follow. So let us begin with a simple example, which illustrates the concepts discussed above.

Suppose we have one claim, which will be used for determining both the "worst case" adverse loss development and the interest discount factor. The claim occurred in 1987, and it will be paid in 1997 for $10,000.

Suppose first that the company accurately estimates the ultimate settlement amount and sets up this value at its initial reserve. Adverse loss development in this "worst case year" is 0%. Since there is a substantial implicit interest offset—the claim is paid 10 years after it occurs—the final reserving risk charge would be negative. In practice, there are no negative charges in the NAIC risk-based capital formula, since all charges are bounded below by 0%.

How large is the offset for the implicit interest discount? For a claim paid ten years after it occurs with a 5% per annum discount factor, the offset is $1 \div (1.05)^{10} = 61.39\%$. The final reserving risk charge in this simplified illustration is 38.61%.

What if the company holds the reserve on a discounted basis, using a 4% per annum discount rate? In 1987, the company sets up a reserve of [$\$10,000 \div (1.04)^{10}$], or $6,756. In 1988, the discounted reserve increases to [$\$10,000 \div (1.04)^{9}$], or $7,026. In 1989, the discounted reserve increases to [$\$10,000 \div (1.04)^{8}$], or $7,307.

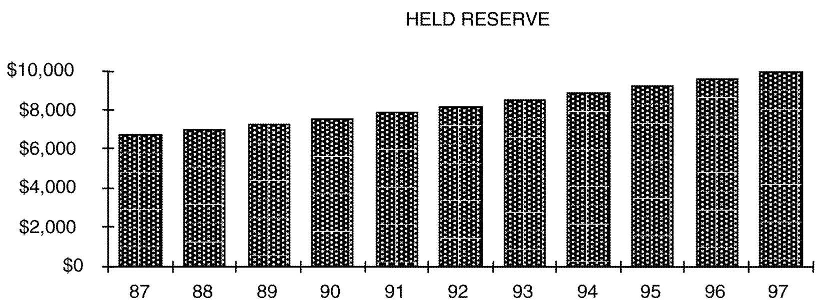
The increases in the held reserve, from $6,756 to $7,026 in 1988, and from $7,026 to $7,307 in 1989, stem from the "unwinding" of the interest discount. However, they show up in Schedule P of the Annual Statement just like any other adverse development.[39]

Figure 2 shows the unwinding of the 4% interest discount over the course of the ten years that the reserve is on the company's

[39]This was true for the *pre-1995* Schedule P, when Part 2 was net of tabular discounts, though it was gross of non-tabular discounts. In 1995 and subsequent Annual Statements, Part 2 of Schedule P is gross of all discounts, so the unwinding of the interest discount no longer shows up as adverse development (see Feldblum [20]). The NAIC risk-based capital reserving risk charges were derived from the 1992 Schedule P.

FIGURE 2

UNWINDING OF INTEREST DISCOUNT



books. Between 1987 and 1992, the held reserve increases from \$6,756 to \$8,219, for observed adverse loss development during this period of 21.67% [= $(8,219 - 6,756) \div 6,756$].

The unwinding of the interest discount during 1987 through 1992 is reflected in the observed adverse development, so it is picked up by the NAIC calculation of the reserving risk charge. That is,

- A valuation basis that uses undiscounted reserves shows no adverse loss development on this claim.
- A valuation basis that uses reserves discounted at a 4% annual rate shows 21.67% of observed loss development.

The higher risk-based capital reserving risk charge generated by the discounted reserves is offset by the lower reserves held by the company.

Future Interest Unwinding

The unwinding of the interest discount continues from 1992 through 1997. Since this future unwinding is not yet reflected in the Schedule P exhibits of historical adverse loss development, a modification of the standard reserving risk charge calculation is needed.

What adjustment is needed? Consider the assumptions underlying the reserving risk charge. The reserving risk charge implicitly says:

> *Let us select the "worst case" adverse loss development that happened between 1983 and 1992, and let us assume that it might happen again.*

This procedure assumes that the 1992 reserves are adequate. That is to say, we should not *expect* either adverse or favorable development of the 1992 reserves.[40]

This is the proper assumption for the risk-based capital formula. The observed adverse loss development is meant to capture unanticipated external factors that cause higher or lower settlement values for insurance claims. A line of business may show adverse loss development even if the initial reserves were properly set on a "best estimate" basis. If a company is indeed holding inadequate reserves, it is the task of the financial examiners of the domiciliary state's insurance department to correct the situation. This is not the role of the generic risk-based capital formula.

If the reserves are valued on a discounted basis, however, they will continue to show (apparent) adverse development until all the claims are settled. In the example above,

- The unwinding of the interest discount between 1987 and 1992 is reflected in the observed adverse loss development, and no further adjustments are needed.
- The unwinding of the interest discount between 1992 and 1997 is not reflected anywhere, so an adjustment to the calculation procedure must be made.

---

[40]We do not *expect* either adverse or favorable development of the 1992 reserves. The risk-based capital requirement guards against *unexpected* adverse development of the reserves.

Alternative Adjustments

There are two ways to make this adjustment: either in the "worst case year" industry adverse loss development or in the offset for the implicit interest discount.

- *Adverse loss development*: One might add the expected future unwinding of the interest discount that will occur after the final valuation date to the "worst case year" observed adverse loss development. In the example above, the observed adverse loss development from 1987 to 1992 is $1,464, giving a factor of +21.7% as a percentage of beginning reserves. We expect further adverse loss development of $1,781 from 1992 to 1997 because of continued unwinding of the interest discount. The total adverse loss development is therefore $3,245, or +48.0% as a percentage of beginning reserves.
- *Implicit interest discount*: The further unwinding of the actual interest discount in the reserves may be used to reduce the offset for the implicit interest discount. In the example above, the observed adverse loss development is offset by ten years of implicit interest discount at a 5% annual rate. However, there are five years of unwinding of the actual 4% interest discount that are still to come (1992 through 1997), and that are not reflected in the observed adverse development.

In our illustration, ten years of implicit interest discount at a 5% annual rate gives a discount factor of 61.4%. Five future years of actual interest unwinding at a 4% annual rate gives a discount factor of 82.2%. The interest margin that should offset the "worst case year" adverse loss development is the *excess* of the implicit interest cushion over the actual interest discount, or 74.7% [= 61.2% ÷ 82.2%].

Diversity and Other Obstacles

In practice, the needed adjustments for tabular discounts are difficult to determine for a variety of reasons.

- *Industry Practice*: There is great disparity among insurance companies in the use of tabular reserve discounts. The prevalent practice is to use tabular discounts on indemnity benefits for lifetime pension cases. But there are companies that do not use tabular reserve discounts at all, and that report aggregate loss reserves on a full-value basis.[41]

- *Pension Identification*: Some companies show tabular discounts only for claims that have been identified as lifetime pension cases. Other companies show tabular discounts for the expected amount of claims that will ultimately be coded as lifetime pension cases.

  The distinction between "identified" and "unidentified" lifetime pension cases is analogous to the distinction between "reported" and "IBNR" claims. A workers compensation claim may be reported to the company soon after it occurs, but it may remain "unidentified" as a lifetime pension case for several years.

- *Indemnity vs. Medical Benefits*: Workers compensation benefits comprise two parts: indemnity benefits, which cover the loss of income, and medical benefits, which cover such expenses as hospital stays and physicians' fees.

  Lifetime pension cases may show continuing payments of both types. For instance, an injured worker who becomes a quadriplegic may receive a weekly indemnity check for loss of income as well as compensation for the medical costs of around-the-clock nursing care.

  Some insurers will discount only the indemnity benefits, since the weekly benefits are fixed by statute.[42] Other insurers will discount the medical benefits as well, since the payments

[41] More precisely, the case reserves generally show the tabular discounts. However, these discounts are "grossed up," or eliminated, by the actuarial "bulk" reserves.

[42] In some states, the indemnity benefit may depend on cost of living adjustments, so the amounts are not entirely "fixed."

are regular and do not vary significantly, even if they are not fixed by statute.

- *Interest Rates*: The interest rate used for the tabular reserve discounts varies by company and by state of domicile. Some companies use a 3.5% annual rate, since this is the interest rate used in the NCCI statistical plan. Several New York and Pennsylvania domiciled companies use a 5% annual rate, since this is the rate permitted by statute in these states. Other companies may use a 4% annual rate, since this is the most common rate in other state statutes.

Pension Discounts

The 3.06% reserving risk charge calculated above uses the conservative 5% interest rate in the risk-based capital formula and the short IRS payment pattern.

As we have discussed above, the NAIC reserving risk charge presumes that loss reserves are reported at undiscounted values. If reserves are valued on a discounted basis—as is true for certain workers compensation cases—then one expects future "adverse development," so the NAIC procedure is incomplete.

What is the expected effect of tabular discounts on the reserving risk charge for workers compensation? Analysts unfamiliar with workers compensation are tempted to say: *It should increase the charge.*

This would indeed be true if lifetime pension cases had the same payment pattern as other workers compensation claims and the only difference between pension cases and other compensation claims were that the pension cases are reported on a discounted basis whereas the other compensation claims are reported on an undiscounted basis. But this is not so. In fact, the very reason that tabular reserve discounts are permitted for lifetime pension cases is that they are paid slowly but steadily over the course of decades.

In other words, to properly incorporate tabular discounts into the workers compensation reserving risk charge, two changes are needed:

- One must increase the "worst case year" adverse development to include the future unwinding of the interest discount on the pension cases. Alternatively, one may adjust the "implicit interest discount" offset to account for the discount already included in the reported reserves.
- One must adjust the payout pattern from the IRS sixteen year pattern to the longer pattern appropriate for lifetime pension cases.

The net effect is to reduce the reserving risk charge. In fact, the indicated charge becomes negative, so it would be capped at 0% by the NAIC formula rules.

This is expected. The NAIC risk-based capital formula imposes a reserving risk charge when the "worst case" adverse development exceeds the implicit interest margin in the reserves. For lines of business like products liability and non-proportional reinsurance, the potential adverse development may far exceed the implicit interest margin, so companies must hold substantial amounts of capital to guard against reserving risk. For workers compensation "non-pension" cases, the mandated statutory benefits reduce the risk of adverse development while the slow payment pattern increases the implicit interest discount, so that the latter almost entirely offsets the former, resulting in the 3% charge calculated above with the RBC formula's exceedingly conservative assumptions. For workers compensation lifetime pension cases, true adverse development practically disappears, since mortality rates do not fluctuate randomly, and only the unwinding of the tabular discount remains. Because of the extremely long payout pattern for lifetime pension cases and the low interest rate allowed for tabular discounts, the implicit interest margin in lifetime pension reserves is well in excess of the "worst case" adverse development.

To calculate the appropriate reserving risk charge for workers compensation, after taking into consideration the tabular discounts on lifetime pension cases, we make the two adjustments discussed above.

- We replace the IRS payment pattern with a 50 year payment pattern derived from the historical experience of the nation's largest compensation carrier. At a 5% per annum interest rate, the present value of the reserves is 65.6% of the ultimate value, as shown in Exhibit A-3.[43]
- We increase the "worst case year" adverse development to incorporate the future interest unwinding on lifetime pension cases. The observed "worst case year" adverse development is 24.2% of initial reserves, from the 1985 statement date to the 1992 statement date. This includes the unwinding of tabular interest discount between 1985 and 1992. The post-1992 unwinding of interest discount on these pension cases adds between 6% and 8% to this figure. To be conservative, we use the 8% endpoint, giving a total adverse development of 34.1%.[44]
- The resulting reserving risk charge is $(1.341 \times 0.656) - 1$, or -14.1%. In other words, industry-wide workers compensation reserves have always been adequate on a discounted basis, even during the worst of years.

[43] Are statistics from a single carrier, no matter how large, a valid proxy for industry-wide figures? For loss ratios, expense ratios, and profit margins they are not appropriate, since each carrier has its own operating strategy. But workers compensation payment patterns are determined by statute; they do not differ significantly among companies, assuming that they have a similar mix of business by state. In November 1996, the American Academy of Actuaries task force on risk-based capital verified the pattern shown in the exhibits here, using data from eight large workers compensation carriers.

[44] For the unwinding of the tabular interest discount, it is no longer appropriate to use a single company's experience as a proxy for the industry. Insurers vary in whether they use tabular discounts at all, what types of benefits they apply the discounts to, and what interest rate they use to discount the reserves. The "6% to 8%" range in the text results from extended observation of reserving practices in workers compensation, along with detailed analysis of one company's own experience. With the reporting of tabular discounts in the 1994 Schedule P, more refined estimates of industry-wide practice may soon be available.

EXHIBIT A-1

NAIC METHOD

Consolidated Industry 1992 Schedule P, Part 2D (Workers Compensation)
Incurred Losses and ALAE

	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
All Prior	18,141,872	18,124,544	18,133,835	18,522,890	18,876,893	19,168,300	19,695,156	20,083,948	20,568,671	21,085,073
1983	10,285,007	10,518,014	10,615,001	10,800,631	10,904,709	11,053,667	11,087,456	11,163,710	11,309,445	11,364,446
1984		11,935,500	12,483,704	12,996,457	13,398,843	13,641,258	13,807,452	13,890,249	14,025,270	14,170,486
1985			13,506,212	14,148,315	14,560,000	15,036,193	15,289,931	15,451,016	15,648,178	15,824,280
1986				15,657,270	16,137,074	16,618,301	16,738,489	16,875,565	17,142,404	17,341,361
1987					18,543,543	18,630,232	18,849,648	18,945,479	19,228,271	19,492,604
1988						21,144,056	21,525,659	21,824,122	22,103,365	22,403,642
1989							23,337,805	23,983,219	24,549,997	24,863,843
1990								25,687,116	26,642,155	26,948,591
1991									27,107,842	27,477,716
1992										25,391,687
Total Incurred	28,426,879	40,578,058	54,738,752	72,125,563	92,421,062	115,292,007	140,331,596	167,904,424	198,325,598	226,363,729
Latest View of Incurreds	32,449,519	46,620,005	62,444,285	79,785,646	99,278,250	121,681,892	146,545,735	173,494,326	200,972,042	226,363,729
Adverse Development	4,022,640	6,041,947	7,705,533	7,660,083	6,857,188	6,389,885	6,214,139	5,589,902	2,646,444	0

Consolidated Industry 1992 Schedule P, Part 3D (Workers Compensation)
Paid Losses and ALAE

	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
All Prior	0	3,644,371	6,120,130	7,945,566	9,435,974	10,681,184	11,778,823	12,725,216	13,559,492	14,283,870
1983	2,595,880	5,414,887	6,989,233	8,016,341	8,714,150	9,206,673	9,565,906	9,842,305	10,035,523	10,182,271
1984		3,098,456	6,475,507	8,592,479	9,951,477	10,858,138	11,474,586	11,929,005	12,268,081	12,519,733
1985			3,307,517	7,223,536	9,609,598	11,175,251	12,188,709	12,884,942	13,409,641	13,785,641
1986				3,399,423	7,693,744	10,430,068	12,205,296	13,319,794	14,100,702	14,642,580
1987					3,823,180	8,916,751	12,037,953	13,992,209	15,236,501	16,062,480
1988						4,517,537	10,522,224	14,272,224	16,542,809	17,976,821
1989							4,923,056	11,851,679	16,021,809	18,519,232
1990								5,283,149	12,856,717	17,435,376
1991									5,481,562	12,644,529
1992										4,795,009

Consolidated Industry 1992 Schedule P, [{Part 2D}–{Part 3D}] (Workers Compensation)
Loss and ALAE Reserves

	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
All Prior	18,141,872	14,480,173	12,013,705	10,577,324	9,440,919	8,487,116	7,916,333	7,358,732	7,009,179	6,801,203
1983	7,689,127	5,103,127	3,625,768	2,784,290	2,190,559	1,846,994	1,521,550	1,321,405	1,273,922	1,182,175
1984		8,837,044	6,008,197	4,403,978	3,447,366	2,783,120	2,332,866	1,961,244	1,757,189	1,650,753
1985			10,198,695	6,924,779	4,950,402	3,860,942	3,101,222	2,566,074	2,238,537	2,038,639
1986				12,257,847	8,443,330	6,188,233	4,533,193	3,555,771	3,041,702	2,698,781
1987					14,720,363	9,713,481	6,811,695	4,953,270	3,991,770	3,430,124
1988						16,626,519	11,003,435	7,551,898	5,560,556	4,426,821
1989							18,414,749	12,131,540	8,528,188	6,344,611
1990								20,403,967	13,785,438	9,513,215
1991									21,626,280	14,833,187
1992										20,596,678
Reserves Held (1)	25,830,999	28,420,344	31,846,365	36,948,218	43,192,939	49,506,405	55,635,043	61,803,901	68,812,761	73,516,187
Adverse Development (2)	4,022,640	6,041,947	7,705,533	7,660,083	6,857,188	6,389,885	6,214,139	5,589,902	2,646,444	0
(2)/(1)	**15.6%**	**21.3%**	**24.2%**	**20.7%**	**15.9%**	**12.9%**	**11.2%**	**9.0%**	**3.8%**	**0.0%**
Worst Year Development:			**24.2%**							

EXHIBIT A-2

IRS PAYMENT PATTERN (1992–1996)

Payment Year	Payment Pattern (Single Accident Year)	Payment Pattern (Stationary Book)
	Accident Year Payout	Reserve Payout
1	22.34%	0.00%
2	28.36%	25.42%
3	15.49%	16.14%
4	8.23%	11.07%
5	5.14%	8.37%
6	4.16%	6.69%
7	2.41%	5.33%
8	2.31%	4.54%
9	0.52%	3.78%
10	0.96%	3.61%
11	0.96%	3.30%
12	0.96%	2.98%
13	0.96%	2.67%
14	0.96%	2.35%
15	0.96%	2.03%
16	5.25%	1.72%

EXHIBIT A-3

WORKERS' COMPENSATION PAYMENT PATTERN

Year	Payment Pattern (Single Accident Year)	Payment Pattern (Stationary Book)
1	0.190	
2	0.213	0.127
3	0.127	0.094
4	0.083	0.074
5	0.057	0.061
6	0.041	0.052
7	0.032	0.045
8	0.025	0.041
9	0.021	0.037
10	0.016	0.033
11	0.014	0.031
12	0.013	0.028
13	0.011	0.026
14	0.010	0.025
15	0.009	0.023
16	0.009	0.022
17	0.009	0.020
18	0.007	0.019
19	0.006	0.018
20	0.006	0.017
21	0.006	0.016
22	0.005	0.015
23	0.006	0.014
24	0.005	0.013
25	0.005	0.013
26	0.004	0.012
27	0.004	0.011
28	0.004	0.010
29	0.004	0.010
30	0.004	0.009
31	0.004	0.009
32	0.003	0.008
33	0.003	0.008
34	0.003	0.007
35	0.003	0.006
36	0.003	0.006
37	0.003	0.006
38	0.003	0.005
39	0.003	0.005
40	0.003	0.004
41	0.003	0.004
42	0.003	0.003
43	0.003	0.003
44	0.002	0.003
45	0.002	0.002
46	0.002	0.002
47	0.002	0.001
48	0.002	0.001
49	0.002	0.001
50	0.002	0.000
Total (Excluding first 12 months)	0.810	1.000
Present Value @ 5%	0.767	0.656

APPENDIX B

THE "EXPECTED POLICYHOLDER DEFICIT" YARDSTICK

Quantifying Reserve Uncertainty

Reserve uncertainty is a slippery concept, difficult to grasp and even more difficult to quantify. The actuary's skill is in forming a "best estimate" that accords with the data and that is appropriate for the particular business environment, such as the insurance marketplace for the premium rates, a statutory financial statement for the reserve requirements, or a merger transaction for the company valuation.

Quantifying reserve uncertainty is complex. A statistician might discuss reserve uncertainty as a probability distribution. One might show the mean of the distribution, its variance, and its higher moments; one might show various percentiles; one might even try to fit the empirical distribution to a mathematical curve. Accordingly, the exhibits in this paper show the mean, the standard deviation, the 95th percentile, and the 5th percentile of each of the distributions.

Capital Requirements

In recent years, state and federal regulators have been setting capital requirements for financial institutions, such as for banks and insurance companies. In theory, "risk-based capital requirements" relate the capital requirements to the uncertainty in various balance sheet items. In practice, most of the risk-based capital formulas that have been implemented in recent years use crude, generic charges that are based more on *ad hoc* considerations of what constitutes a "reasonable" charge than on rigorous actuarial or financial analyses.

Risk-based capital theory, however, is a siren for some actuaries and academicians, who have examined the relationship between uncertainty and capital requirements. In an ideal risk-based capital system, capital requirements should be calibrated among the balance sheet items in proportion to the risk that each poses to the company's solvency. Suppose a company has $100 million of bonds and $100 million of loss reserves, and the theoretically correct risk-based capital system says that the company needs $5 million of capital to guard against the uncertainty in the bond returns and $15 million of capital to guard against the uncertainty in the loss reserve payments. Then we can say that the uncertainty in the loss reserve portfolio is "three times as great" as the uncertainty in the bond portfolio.

Of course, we don't really mean that "uncertainty" is an absolute quantity that can be three times as great as some other figure. Rather, our measuring rod gives us a figure that we use as a proxy for the amount of uncertainty.

Moreover, our interest is not in absolute capital requirements but in the *relative* uncertainty among the company's various components. The regulator must indeed calibrate the absolute capital requirements, deciding between (i) $5 million of capital for bond risk and $15 million of capital for reserve risk versus (ii) $10 million of capital for bond risk and $30 million of capital for reserve risk. For the measurement of uncertainty, however, we are most interested in relative figures, such as the relative amount of capital needed to guard against reserve risk versus the amount needed to guard against bond risk, or the percentage reduction in capital for business written on loss sensitive contracts.

Calibrating Capital Requirements

There are two "actuarial" methods of calibrating capital requirements.

- The "probability of ruin" method says: How much capital is needed such that the chance of the company's insolvency during the coming time period is equal to or less than a given percentage?

- The "expected policyholder deficit" method says: How much capital is needed such that the expected loss to policyholders and claimants during the coming time period—as a percentage of the company's obligations to them—is equal to or less than a given amount?[45]

In this paper, we use the "expected policyholder deficit" (EPD) approach. The results would be no different if we used a "probability of ruin" approach.

Computing the Expected Policyholder Deficit

The "expected policyholder deficit" is a relatively new concept, having first been introduced in 1992. This appendix provides a brief explanation of the EPD analysis used in the paper.

Let us repeat the underlying question. The EPD analysis says: "Given a probability distribution for an uncertain balance sheet item, how much capital must the company hold such that the ratio of the expected loss to policyholders to the obligations to policyholders is less than or equal to a desired amount?" The format of the analysis depends on the type of probability distribution.

- For a simple discrete distribution, we can work out by hand the exact capital requirement. The type of simple discrete distribution that we illustrate below never occurs in real life. We use it only as a heuristic example, since the same procedure is used in our simulation analysis.
- If the empirical probability distribution can be modeled by a mathematically tractable curve, a closed-form analytic expression for the EPD can sometimes be found. In his previously cited paper, Butsic [14] does this for the normal and lognor-

[45]The "probability of ruin" method is explained in Daykin, Pentikainen, and Pesonen [17]. Probability of ruin analyses have long been used by European actuaries; see especially Beard, Pentikainen, and Pesonen [3] and Bowers, Gerber, Hickman, Jones, and Nesbitt [9]. The "expected policyholder deficit" method is explained in Butsic [13].

mal distributions, which can serve as reasonable proxies for many balance sheet items.

- The distributions in this paper are derived by means of stochastic simulation. Each distribution results from 10,000 Monte Carlo simulations. We determine the amount of capital needed to achieve a desired EPD ratio, as explained below.

Let us begin with the first case, the simple discrete distribution, to illustrate how the analysis proceeds. The extension to the full stochastic simulation merely requires greater computer power; there is no difference in the structure of the analysis.

Scenarios and Deficits

The distributions used in this paper are based on 10,000 simulations each. Think of this as 10,000 different scenarios. In fact, however, these simulations are *stochastic*. We do not know what these simulations are until after they have been realized. In other words, there are an infinite number of *possible* scenarios, 10,000 of which will be realized in the simulation.

To clarify the meaning of the "expected policyholder deficit," let us assume that an insurer with $250 million of assets faces two possible scenarios:

- In the *favorable* scenario, the company's interpretation of its insurance contracts will be upheld by the courts, and it must pay losses of $200 million.
- In the *adverse* scenario, the company's interpretation will *not* be upheld by the courts, and it must pay losses of $300 million.

Suppose also that there is a 60% chance of the favorable scenario being realized and a 40% chance of the adverse scenario being realized.[46]

[46] In the simulation analysis in this paper, only reserves are uncertain; assets are not uncertain. However, the same type of analysis applies to both assets and liabilities. Indeed, a more complete model would examine the external (economic and financial) factors that lead to variability in ultimate loss reserves, and it would analyze their effects on asset values as well.

What is the expected policyholder deficit? In the favorable scenario, the company has a positive net worth at the end. Since we are concerned only with deficits, a positive outcome of any size is considered a \$0 deficit.

In the adverse scenario, the final deficit is a \$50 million deficit, or −\$50 million. Since there is a 40% chance of an adverse outcome, the *expected* policyholder deficit is

$$\$0 \text{ million} \times 60\% + (-\$50 \text{ million} \times 40\%) = -\$20 \text{ million}.$$

The EPD Ratio

The definition of the EPD ratio is:

$$\text{EPD ratio} = (\text{expected policyholder deficit}) \div (\text{expected loss}).$$

In the example above, there is a 60% chance of a \$200 million payment to claimants and a 40% chance of a \$300 million payment to claimants. Thus, the expected loss is:

$$(\$200 \text{ million} \times 60\%) + (\$300 \text{ million} \times 40\%) = \$240 \text{ million}.$$

The EPD ratio is:

$$\$20 \text{ million} \div \$240 \text{ million} = 8.33\%.$$

Consistency

We use a 1% expected policyholder deficit ratio to determine the capital requirements. We use 1% to be consistent with the charges in the NAIC risk-based capital formula. In memoranda submitted to the American Academy of Actuaries task force on risk-based capital, Butsic estimates that the overall industrywide reserving risk charge in the NAIC risk-based capital formula amounts to approximately a 1% EPD ratio.

This allows us to compare the workers compensation loss reserve uncertainty to other sources of insurance company risk. If one believes that the overall capital requirements in the NAIC risk-based capital formula are reasonable, so a 1% EPD ratio is appropriate, then the degree of workers compensation loss reserve uncertainty measured in this paper can be viewed in light of the other NAIC capital requirements. As Butsic [12] says:

> *The amount of risk-based capital for each source of risk (e.g., underwriting, investment, or credit) must be such that the risk of insolvency (or other applicable impairment) is directly proportional to the amount of risk-based capital for each source of risk.*

Capital Requirements

We illustrate the calculation of capital requirements with the example given above. The capital required depends on the EPD ratio that the company (or the solvency regulator) seeks to maintain. We use a 1% target EPD ratio for this illustration.

If the company holds no capital, then its EPD ratio equals:

$$\text{(expected policyholder deficit)} \div \text{(expected loss)}$$

$$= \$20 \text{ million} \div \$240 \text{ million} = 8.33\%.$$

This exceeds the 1% target EPD ratio. The company must hold sufficient capital such that its revised EPD, or EPD*, satisfies the relationship:

$$\text{EPD}^* \div \text{(expected loss)} = \text{EPD}^* \div \$240 \text{ million} = 1\%,$$

$$\text{or} \qquad \text{EPD}^* = \$2.4 \text{ million}.$$

In the favorable scenario, the company already has sufficient funds to pay the losses. Adding capital will not change the policyholder deficit. In the adverse scenario, the company's assets are not sufficient to pay the losses. Adding capital will reduce the policyholder deficit. To achieve an EPD* of \$2.4 million, we solve:

$$40\% \times (\text{current assets} + \text{additional capital} - \text{liabilities})$$
$$= -\$2.4 \text{ million},$$
$$40\% \times (\$250 \text{ million} + \text{additional capital} - \$300 \text{ million})$$
$$= -\$2.4 \text{ million},$$
$$-\$50 \text{ million} + \text{additional capital}$$
$$= -\$6.0 \text{ million},$$
$$\text{additional capital} = \$44 \text{ million}.$$

Since the current assets are \$250 million, the additional capital required is \$44 million, and the expected losses are \$240 million, the total capital requirement for the company is \$250 million + \$44 million – \$240 million = \$54 million.

Full Simulation

The full analysis in this paper proceeds in the same fashion. The 10,000 simulations are run, each of which produces a "realization" for the loss amount. The average of these 10,000 realizations is the expected loss. The probability of each realization is 0.01%.

We first assume that the asset amount equals the expected loss, and we determine the loss payment and the deficit in each realization.

- If the loss amount is less than the asset amount, then the loss payment equals the loss amount, and the deficit is zero.
- If the loss amount exceeds the asset amount, then the loss payment equals the asset amount, and the deficit is the difference between the loss amount and the asset amount.

We sum the deficits in the 10,000 realizations, and we divide by 10,000. This gives the expected policyholder deficit. We then divide by the expected loss amount to give the EPD ratio.

If the probability distribution for the loss reserves is extremely compact, then the EPD ratio may be less than 1% even if no capital is held. For instance, suppose that the probability distribution is uniform over the range $100 million ± $4 million. Then the expected policyholder deficit is 1% if no capital is held.[47] This makes sense—if the loss payments are practically certain, there would be little need for surplus to support the reserves.

In practice, of course, the loss payments are not certain, and the EPD ratio would be greater than 1% if no capital is held. We proceed iteratively. We add capital and redetermine the loss payment and deficit in each scenario. This gives a new expected policyholder deficit and a new EPD ratio. If the EPD ratio still exceeds 1%, we must add more capital. If the EPD ratio is now less than 1%, we can subtract capital. With sufficient computer power, we quickly converge to a 1% EPD ratio.

[47] If the actual loss is less than $100 million, then the deficit is zero. If the actual loss exceeds $100 million, then the deficit is uniform over [$0, $4 million], for an average of $2 million. The expected deficit over all cases is therefore $1 million, for an EPD ratio of 1%.

APPENDIX C

The Simulation Procedure

Casualty actuaries are accustomed to providing point estimates of indicated reserves. The traditional procedures—such as a chain ladder loss development using 25 accident years of experience, supplemented by an "inverse power curve" tail factor—provide a sound basis for estimating workers compensation reserve needs. The actuary's task is to examine the historical experience for trends, evaluate the effects of internal (operational) changes on case reserving practices and settlement patterns, and forecast the likely influence of future economic and legal developments on the company's loss obligations.

Our perspective in this paper is different. We are not determining a point estimate of the reserve need; rather, we are determining a probability distribution for the reserve need. We use the same procedure and the same data as we would use for the point estimate: a chain ladder loss development based on 25 accident years of experience, along with a tail factor based on an inverse power curve fit. But now each step turns stochastic, and the probability distribution is determined by a Monte Carlo simulation.

The traditional procedures for determining point estimates are documented in various textbooks. This appendix shows the corresponding procedures for determining the probability distribution.

Data

We use a chain ladder *paid* loss development, since payment patterns for workers compensation are relatively stable whereas case reserving practices often differ from company to company and from year to year. This enables readers to replicate our results using their own companies' data.

We begin with accident year triangles with 25 years of cumulative paid losses, separately for indemnity (wage loss) and medical benefits. Indemnity and medical benefits have different loss payment patterns, and they are affected by different factors. For instance, medical benefits are strongly affected by medical inflation and by changes in medical utilization rates.

From the historical data we determine paid loss "age-to-age" factors (or "link ratios"). Exhibit C-1 shows 20 columns of paid loss age-to-age factors for countrywide indemnity plus ALAE benefits. For instance, the column labeled "12–24" shows the ratio of cumulative paid indemnity losses at 24 months to the corresponding cumulative paid indemnity losses at 12 months for each accident year. Similarly, Exhibit C-2 shows the paid loss age-to-age factors for countrywide medical benefits.

Point Estimates versus Realizations

The reserving actuary, when determining a point estimate, would examine these factors for trends. For a point estimate, the reserving actuary might use an average of the most recent five factors, instead of an average of all the factors in the column.

In this paper, our goal is to estimate the uncertainty in the reserve indications. Just as there was an upward trend in the age-to-age factors during the 1980s, there may be subsequent upward or downward trends in the 1990s. We therefore use the entire column of factors in our analysis. An "outlying" factor that is not a good estimator of the expected future value is an important element in measuring the potential variability of the future value.

We want to use the historical factors to simulate future "realizations." We do this by fitting the observed factors to a curve, thereby obtaining a probability distribution for the "12 to 24" age-to-age factors. Note carefully—this is *not* the probability distribution of the loss reserves, which will be the *output* of the simulation and which is *not* modeled by any mathematical function. This is the probability distribution of the age-to-age factors, which is the *input* to the simulation and is modeled by a curve.

Lognormal Curve Fitting

In this analysis, we used lognormal curves, which gave good fits to the data. Exhibit C-3 shows the curve fitting procedure for the first column of "indemnity plus ALAE" age-to-age factors.

For the lognormal curve, the probability distribution function is

$$f(x) = \frac{e^{-.5(\ln(x)-\mu/\sigma)^2}}{x\sigma\sqrt{2}\pi}$$

and the cumulative distribution function is

$$F(x) = \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)$$

We fit the function with the "development" part of the link ratios, or the "age-to-age factor minus one," as shown in Column 2 of Exhibit C-3. Column 3 shows the natural logarithms of the factors in Column 2. We use the method of moments to find the parameters of the fitted curve. The "mu" (μ) parameter is the mean of the figures in Column 3, and the "sigma" (σ) parameter is the standard deviation of the figures in Column 3.

We do the same for each "age-to-age" development column. The fitted parameters shown in the box in Exhibit C-3 are carried back to the final two rows in Exhibit C-1. Thus, each column has its own lognormal probability distribution function. We do this for development through 252 months. There is still paid loss development after 252 months, but there is insufficient historical experience to generate the factors, so we use an inverse power curve to estimate the loss development "tail" (discussed below).

For each run, we use a random number generator (Excel's built-in "RAND" function) to obtain simulated "age-to-age" factors in each column. Column 3 of Exhibit C-4 shows the results of one simulation for indemnity plus ALAE payments. For instance, the simulated age-to-age factor for 12 to 24 months of development is 2.312. The simulations for each of the 20 columns are independent of each other. For instance, the simulated 1.401 factor for "24 to 36" months in Column 3 of Exhibit C-4 is independent of the simulated 2.312 factor for "12 to 24" months.[48]

Parameter Variance

Two types of variance affect the simulation of future age-to-age link ratios: process variance and parameter variance.

- Process variance is the variance caused by the random nature of insurance losses. Even if the expected link ratios were known with certainty, the observed link ratios would differ from them because more losses than expected or less losses than expected might be paid in any given period.
- Parameter variance reflects the actuary's uncertainty about the expected losses. We estimate the probability distribution of the age-to-age link ratios from historical data. Our estimate may not be perfectly accurate; that is, we may have misestimated the parameters of the fitted probability distribution.

Quantifying Parameter Variance

To quantify parameter variance, we use a model developed by Kreps [42]. We assume that the observed age-to-age link ratios in a given development period come from the same lognormal probability distribution. We estimate the parameters of the fitted distribution as documented above, giving a lognormal distribution with parameters μ and σ.

[48]Our analysis assumes independence among columns. Dependence among columns may raise or lower the reserve variability, depending on whether the columns are positively or negatively correlated with each other. See the text of this paper for further discussion of trends in "age-to-age" factors on any observed correlations between columns, and see Holmberg [37] for methods of quantifying these correlations.

Think of our problem in the following fashion. We want to use our fitted distribution to simulate new observations. The actual future value may differ from the expected (mean) value because of process variance. In addition, we are uncertain whether we have chosen the proper expected (mean) value.

The Kreps procedure works as follows:

We fit a lognormal distribution to each column of the triangle. We assume that there are n age-to-age factors in the column, and that these are represented by $x_1, x_2 \ldots x_n$.

1. Calculate

$$\mu_0 = (1/n) \times \Sigma \ln(x_i)$$
$$\sigma_0 = \{(1/n) \times \Sigma[\ln(x_i) - \mu_0]^2\}^{0.5}.$$

   These are the maximum likelihood estimators that would typically be used for simulation in the absence of parameter uncertainty.

2. Generate 3 random variables:

   i) z, which has a standard normal distribution,

   ii) w, which has a Chi-squared distribution with parameter $(n + \theta - 1)$,

   iii) $v \times (n + \theta - 2)^{0.5}$, which has a t distribution, with parameter $(n + \theta - 2)$.

   The value of θ depends on the Bayesian prior that is used. If the prior is a uniform distribution, $\theta = 0$. A power-law prior gives $\theta = 1$. The lower the value of θ, the more the effect of the parameter uncertainty. In private correspondence, Kreps pointed out to us that selecting $\theta = 0$ or 1 can give unreasonable results more often than one would like. In experimenting with this modeling, we found the same thing: every so often, the model would generate a gigantic age-to-age factor that would be totally unreasonable, given the nature of the business. Consequently, we used $\theta = 2$. (On rare occasions, even this gave unreasonable results—see the discussion below.)

3. Calculate

$$z_{\text{eff}} = v + z^*\{n^*(1+v^2)/w\}^{0.5}.$$

4. Calculate

$$x = \exp[\mu_0 + \sigma_0^* z_{\text{eff}}].$$

x is then a single simulation from the lognormal, taking parameter uncertainty into account.

As noted above, we occasionally found that an "unreasonably" large age-to-age factor would be generated. These factors were so large that they ended up dominating the simulation results. To eliminate these unreasonable cases, we set a rule that if any of the simulated age-to-age factors was more than 50 standard deviations from the mean, then that whole simulation would be eliminated. We are dealing with 20 years of workers compensation paid losses and are simulating a separate ATA factor for every point in the lower triangle. Also, since we have separate triangles for medical and indemnity, we are simulating 420 age-to-age factors each time. The rule applies to each one individually; in other words, if even 1 of the 420 was outside of 50 standard deviations, we threw them all out, and simulated again. Even so, we ended up throwing out the results in fewer than 3% of the cases. (In private correspondence, Kreps described this rule as "very generous" and suggested that he might have limited factors to within 10 standard deviations.)

Accident Year Correlations

In standard reserve analyses, the actuary derives a "best-estimate" age-to-age link ratio for each development period and uses that estimate for all future accident years. The actuary seeks a best-estimate reserve indication, so the best-estimate link ratios should be used for all years.

Our concern in this paper is with simulating the actual (future) development of the reserves. Each future year will have a distinct age-to-age link ratio in each period. To accurately model the future development of the reserves, we simulate separate future link ratios for each future accident year.

For example, suppose that our most recent accident year is 1994, our current valuation date is December 31, 1994, and we are simulating age-to-age link ratios from 48 to 60 months from accident years 1990 and prior. We use the simulated 48 to 60 month link ratios to develop accident years 1991 through 1994. We do four separate simulations to obtain four different link ratios for these four accident years.

Using separate simulated link ratios for each accident year assumes that the years are uncorrelated with respect to loss development. Using a single simulated link ratio for all accident years assumes that the accident years are perfectly correlated with each other. The independence assumption leads to a lower estimate of reserve uncertainty, since high development in one accident year may be offset by low development in another accident year. The dependence assumption leads to a higher estimate of reserve uncertainty, since high development in one accident year is associated with high development in all accident years.

The practical effect of using separate simulations versus using a single simulation for the link ratio for all future accident years in a given development period depends on the number of independent development periods in the simulation. The model in this paper uses 20 independent development periods plus a tail factor. Since the development periods are independent of each other, high development in one period is generally offset by low or average development in other periods. Therefore, the difference between independence among the accident years and dependence among the accident years is not great.

The tables in the text of this paper show results for both the independence assumption and the dependence assumption. The discussion in the paper uses the results for the independence assumption (i.e., for separate simulations by accident year).

Tail Development

Exhibit C-4 shows the fitting of the inverse power curve for one simulation. To clarify the procedure, let us *contrast* this with fitting an inverse power curve for a "best-estimate" reserve indication. For the "best-estimate" indication, we would use "selected" age-to-age factors in Column 3, such as averages of the factors in each column, or averages of the most recent years, or perhaps averages that exclude high and low factors. For the indemnity plus ALAE "12 to 24" months factor, the overall average is 2.685 and the average of the most recent five factors is 2.887. For a "best estimate," we would probably choose a factor such as 2.500.

In our analysis, the 20 factors in Column 3 are the results of *simulations* from the 20 fitted lognormal curves. For instance, the 2.312 factor is a simulation from the lognormal curve representing the probability distribution for the 12 to 24 month column.

From these *simulated* age-to-age factors, we fit an inverse power curve to estimate the "tail" development.[49] The inverse power curve will vary from simulation to simulation, since we have different "age-to-age" factors in each run. Moreover, the inverse power curve varies from accident year to accident year, since the simulated age-to-age link ratios vary by accident year.

The inverse power curve models the age-to-age ("ATA") factors as

$$\text{ATA} = 1 + at^{-b}$$

where "t" represents the "development year," and "a" and "b" are the parameters that we must fit. In workers compensation,

[49] For the rationale of using an inverse power curve for the tail development, see Sherman [52].

the shape of the loss payment pattern differs greatly between the first several years and subsequent years. In early years, there are many temporary total claims with rapid payment patterns. By the tenth year, most of the remaining reserves are for lifetime pension cases (fatalities and permanent total disability cases) with slow payment patterns. Therefore, we fit the inverse power curve using the simulated factors from the tenth through the 20th columns only.[50]

Columns (4) and (5) of Exhibit C-4 show the fitting procedure. Column (4) is the natural logarithm of the development year in Column (2), and Column (5) is the natural logarithm of the "simulated age-to-age [ATA] factor minus one" in Column (3). The inverse power curve can be written as

$$\ln(\text{ATA} - 1) = \ln(a) - b \times \ln(t).$$

We use a least squares procedure to determine the parameters a and b from the figures in Columns (4) and (5), giving $\ln(a) = -0.722$, or $a = 0.486$, and $b = 1.498$, as shown in the box at the bottom of Exhibit C-4.

The fitted inverse power curve provides age-to-age factors for development years 21 through 70. We don't really know how long paid loss development continues for workers compensation. Moreover, the factors are small. For development years 30 through 39 in this simulation, the age-to-age factors are about 1.002, and for development years 40 through 70, the factors are about 1.001. (The actual factors, of course, differ in the subsequent decimal places.) We therefore choose the length of the tail development stochastically; that is, the length of the total development is chosen randomly from a uniform distribution between 30 and 70 years.

[50]For actual reserve indications, one would probably segment the data between non-pension cases (temporary total and permanent partial cases) and lifetime pension cases (fatalities and permanent total cases).

Parameter Variance in the Tail

We have included both process variance and parameter variance in the simulated age-to-age link ratios for the first 20 development periods. The tail factors are an inverse power curve extension of each set of simulated age-to-age link ratios.

The tail factor selection procedure is a deterministic fit to the simulated age-to-age link ratios.[51] To the extent that process risk and/or parameter risk affect the variability of the age-to-age link ratios, they affect the variability of the tail factors.

One reviewer of an earlier draft of this paper wondered whether parameter variance might be incorporated independently in the tail factors. Specifically, the model currently has the following steps:

- We stochastically simulate age-to-age link ratios separately for each accident year and each development period, incorporating both process variance and parameter variance.
- We stochastically select the length of the development period, between 30 years and 70 years.
- We fit an inverse power curve to the simulated age-to-age link ratios to generate a tail factor.

The revised procedure would expand the third step in the list above as follows:

- Fit an inverse power curve to the simulated age-to-age link ratios. The inverse power curve is a two parameter family of curves. The fitting procedure gives "best estimates" for each of the two parameters.
- The current procedure considers the fitted parameters as the final values for each simulation. In place of this, assume a "structure function" for the distribution of these two param-

[51] The length of the tail development, though, is an independent stochastic choice, unrelated to the set of age-to-age link ratios.

eters. The values derived by fitting the inverse power curve would be the means of the distributions. The variance of the distribution, as well as the type of distribution, would be chosen subjectively.

- Stochastically select values for these two parameters from their assumed probability distribution. Use these simulated values of the two parameters to generate the inverse power curve tail factor.

Although this procedure is complex, it is important to consider all sources of variability, and to incorporate them, when feasible, into an actuarial model. Two factors, however, hampered the implementation of this procedure in our analysis.

- We had no *a priori* expectations about the type of structure function or the variance of the structure function.
- For the parameter risk in the link ratio estimation, we used a mathematically tractable approximation to simplify the simulation. For the parameter risk in the tail factor estimation, we are not aware of any corresponding approximation.

Thus, the procedures used in this paper do not separately incorporate parameter risk into the tail factor estimation.

Selected Factors

In the simulation shown in Exhibit C-5, the stochastic selection produced a development period of 54 years. We therefore have three sets of age-to-age factors:

- For development years 1 through 20, we use the simulated age-to-age factors generated by the lognormal curves for each column. For these development years, the "selected ATA" in Column (4) equals the "simulated ATA" in Column (2), not the "fitted ATA" in Column (3).
- For development years 21 through 53, we use the age-to-age factors from the fitted inverse power curve. For these devel-

opment years, the "selected ATA" in Column (4) equals the "fitted ATA" in Column (3).

- For development years 54 through 70, we use age-to-age factors of unity.

We now have all the age-to-age factors for this simulation. We "square the triangle" in the standard reserving fashion to determine ultimate incurred losses, and we subtract cumulative paid losses to date to obtain the required reserves. Exhibit C-6 shows the determination of the required medical reserves for one simulation. The "ultimate paids" in Exhibit C-6 are the "paid-to-date" times the "age-to-ultimate" factors, and the "indicated reserves" are the "ultimate paids" minus the "paid-to-date." The right-most two columns of Exhibit C-6 show the determination of the present value of the reserves. The "present value factors" are discussed in Appendix D, which has a full explanation of inflation effects.

We perform this simulation 10,000 times, giving a complete probability distribution of the required reserves, and we determine the mean, standard deviation, 95th percentile, and 5th percentile of this distribution. For the manner of determining the "capital required to achieve a 1% expected policyholder deficit ratio" (the right-most column of the exhibits in the text of this paper), see Appendix B.

EXHIBIT C-1

AGE-TO-AGE FACTORS FOR PAID INDEMNITY AND PAID ALAE

	12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	120–132	132–144	144–156	156–168	168–180	180–192	192–204	204–216	216–228	228–240	240–252	252–264	264–276
1970					1.055	1.040	1.028	1.021	1.018	1.016	1.012	1.013	1.009	1.008	1.007	1.021	1.001	1.004	1.006	1.005	1.004	1.004
1971				1.094	1.055	1.041	1.026	1.024	1.016	1.011	1.010	1.011	1.007	1.010	1.012	1.009	1.005	1.006	1.005	1.005	1.004	1.006
1972			1.168	1.093	1.065	1.043	1.032	1.025	1.016	1.018	1.018	1.010	1.012	1.011	1.008	1.008	1.008	1.007	1.007	1.006	1.006	1.008
1973		1.386	1.169	1.096	1.062	1.049	1.033	1.025	1.020	1.017	1.012	1.013	1.012	1.011	1.008	1.008	1.007	1.007	1.007	1.005	1.007	
1974	2.334	1.385	1.164	1.093	1.068	1.044	1.034	1.022	1.019	1.016	1.013	1.010	1.013	1.009	1.008	1.009	1.007	1.006	1.008	1.007		
1975	2.310	1.398	1.190	1.116	1.076	1.051	1.037	1.026	1.021	1.016	1.013	1.014	1.012	1.011	1.010	1.009	1.010	1.010	1.010			
1976	2.262	1.388	1.195	1.117	1.069	1.048	1.031	1.027	1.020	1.017	1.013	1.013	1.012	1.008	1.011	1.009	1.009	1.009				
1977	2.192	1.397	1.191	1.111	1.070	1.048	1.031	1.023	1.019	1.016	1.015	1.013	1.011	1.010	1.009	1.009	1.009					
1978	2.246	1.407	1.193	1.113	1.068	1.048	1.031	1.027	1.022	1.019	1.016	1.014	1.012	1.012	1.011	1.009						
1979	2.199	1.409	1.192	1.109	1.068	1.045	1.036	1.027	1.023	1.020	1.019	1.013	1.011	1.012	1.012							
1980	2.169	1.400	1.209	1.107	1.074	1.050	1.038	1.030	1.023	1.020	1.017	1.017	1.011	1.011								
1981	2.191	1.400	1.185	1.115	1.075	1.055	1.041	1.032	1.025	1.019	1.017	1.016	1.011									
1982	2.179	1.395	1.207	1.131	1.098	1.059	1.046	1.043	1.026	1.024	1.020	1.015										
1983	2.283	1.437	1.227	1.140	1.088	1.064	1.048	1.037	1.025	1.022	1.017											
1984	2.345	1.473	1.228	1.134	1.089	1.064	1.044	1.033	1.027	1.018												
1985	2.422	1.473	1.245	1.140	1.087	1.057	1.041	1.030	1.020													
1986	2.377	1.500	1.237	1.133	1.085	1.055	1.038	1.026														
1987	2.452	1.496	1.234	1.127	1.080	1.053	1.034															
1988	2.496	1.498	1.228	1.126	1.074	1.047																
1989	2.502	1.512	1.231	1.121	1.068																	
1990	2.666	1.520	1.232	1.109																		
1991	2.529	1.507	1.217																			
1992	2.454	1.470																				
1993	2.426																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Avg of ln(ATA-1)	0.30	−0.82	−1.58	−2.16	−2.62	−3.00	−3.33	−3.59	−3.86	−4.03	−4.21	−4.34	−4.52	−4.61	−4.65	−4.64	−5.21	−5.00	−4.95	−5.19	−5.23	−5.12
Var of ln(ATA-1)	0.010	0.013	0.015	0.018	0.024	0.019	0.028	0.033	0.025	0.030	0.042	0.027	0.031	0.026	0.029	0.085	0.972	0.078	0.052	0.036	0.051	0.079
Lognormal Parameters, ignoring parameter risk:																						
mu	0.30	−0.82	−1.58	−2.16	−2.62	−3.00	−3.33	−3.59	−3.86	−4.03	−4.21	−4.34	−4.52	−4.61	−4.65	−4.64	−5.21	−5.00	−4.95	−5.19	−5.23	−5.12
sigma	0.104	0.116	0.127	0.136	0.158	0.142	0.172	0.188	0.162	0.180	0.211	0.169	0.184	0.168	0.178	0.307	1.046	0.299	0.247	0.209	0.253	0.324

EXHIBIT C-2

AGE-TO-AGE FACTORS FOR MEDICAL BENEFITS

	12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	120–132	132–144	144–156	156–168	168–180	180–192	192–204	204–216	216–228	228–240	240–252	252–264	264–276
1970					1.013	1.011	1.010	1.008	1.005	1.007	1.006	1.008	1.007	1.007	1.005	1.006	1.001	1.007	1.005	1.007	1.008	1.009
1971				1.019	1.015	1.014	1.008	1.010	1.010	1.008	1.009	1.011	1.012	1.008	1.008	1.007	1.007	1.008	1.011	1.012	1.012	1.008
1972			1.045	1.026	1.018	1.014	1.013	1.012	1.009	1.013	1.010	1.011	1.018	1.011	1.008	1.010	1.007	1.009	1.010	1.011	1.008	1.010
1973		1.105	1.044	1.030	1.017	1.017	1.012	1.012	1.011	1.010	1.010	1.009	1.010	1.009	1.009	1.008	1.012	1.007	1.007	1.009	1.008	
1974	1.895	1.108	1.050	1.028	1.021	1.013	1.013	1.010	1.014	1.009	1.009	1.011	1.010	1.007	1.014	1.014	1.009	1.009	1.007	1.012		
1975	1.898	1.122	1.055	1.034	1.023	1.019	1.017	1.014	1.011	1.009	1.012	1.012	1.011	1.011	1.014	1.015	1.012	1.010	1.009			
1976	1.893	1.113	1.056	1.035	1.026	1.019	1.016	1.016	1.014	1.014	1.012	1.009	1.010	1.010	1.010	1.011	1.010	1.009				
1977	1.865	1.119	1.055	1.035	1.020	1.021	1.014	1.013	1.011	1.011	1.008	1.011	1.009	1.009	1.010	1.010	1.008					
1978	1.912	1.122	1.057	1.036	1.025	1.018	1.019	1.016	1.014	1.013	1.017	1.015	1.014	1.013	1.012	1.012						
1979	1.869	1.120	1.056	1.036	1.025	1.020	1.016	1.015	1.013	1.011	1.014	1.012	1.012	1.011	1.009							
1980	1.849	1.126	1.063	1.031	1.030	1.023	1.017	1.016	1.014	1.014	1.013	1.010	1.017	1.011								
1981	1.836	1.127	1.054	1.037	1.028	1.021	1.018	1.018	1.015	1.015	1.012	1.010	1.007									
1982	1.808	1.126	1.063	1.040	1.025	1.022	1.020	1.016	1.013	1.013	1.010	1.009										
1983	1.898	1.135	1.071	1.041	1.030	1.028	1.022	1.020	1.018	1.014	1.012											
1984	1.948	1.158	1.072	1.047	1.036	1.027	1.024	1.020	1.013	1.010												
1985	1.949	1.156	1.081	1.048	1.035	1.028	1.022	1.016	1.015													
1986	1.808	1.162	1.082	1.051	1.032	1.026	1.019	1.015														
1987	1.906	1.172	1.084	1.049	1.037	1.025	1.017															
1988	1.871	1.178	1.083	1.051	1.030	1.021																
1989	1.934	1.172	1.079	1.043	1.024																	
1990	1.898	1.171	1.071	1.036																		
1991	1.869	1.153	1.058																			
1992	1.773	1.126																				
1993	1.772																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Avg of ln(ATA-1)	−0.14	−1.99	−2.77	−3.31	−3.71	−3.93	−4.14	−4.25	−4.41	−4.49	−4.53	−4.56	−4.52	−4.66	−4.64	−4.63	−5.06	−4.79	−4.87	−4.61	−4.72	−4.68
Var of ln(ATA-1)	0.004	0.029	0.042	0.061	0.082	0.071	0.086	0.061	0.076	0.048	0.057	0.025	0.085	0.045	0.079	0.095	1.098	0.023	0.098	0.055	0.044	0.014
Lognormal Parameters:																						
mu	−0.14	−1.99	−2.77	−3.31	−3.71	−3.93	−4.14	−4.25	−4.41	−4.49	−4.53	−4.56	−4.52	−4.66	−4.64	−4.63	−5.06	−4.79	−4.87	−4.61	−4.72	−4.68
sigma	0.062	0.175	0.209	0.253	0.293	0.274	0.301	0.254	0.285	0.226	0.247	0.165	0.304	0.223	0.295	0.325	1.111	0.162	0.338	0.256	0.235	0.138

EXHIBIT C-3

Illustration of Fitting Lognormal Distributions to Age-to-Age Factors

	(1) 12–24 Factors for Indemnity & ALAE	(2) Age-to-Age Factor minus 1 (1) −1	(3) Natural Logs of (Age-to-Age Factors minus 1) ln (2)
1974	2.334	1.334	0.288
1975	2.310	1.310	0.270
1976	2.262	1.262	0.232
1977	2.192	1.192	0.175
1978	2.246	1.246	0.220
1979	2.199	1.199	0.181
1980	2.169	1.169	0.156
1981	2.191	1.191	0.175
1982	2.179	1.179	0.165
1983	2.283	1.283	0.249
1984	2.345	1.345	0.297
1985	2.422	1.422	0.352
1986	2.377	1.377	0.320
1987	2.452	1.452	0.373
1988	2.496	1.496	0.403
1989	2.502	1.502	0.407
1990	2.666	1.666	0.510
1991	2.529	1.529	0.425
1992	2.454	1.454	0.375
1993	2.426	1.426	0.355
Average	2.352	1.352	0.296
Variance	0.018	0.018	0.010

Fitted Lognormal	
μ_0 [= mean of the logs of (ATA-1)]	0.296
σ_0 [= standard deviation of logs of (ATA-1)]	0.099
Parameter Risk Procedure	
n (= number of ATA factors)	20
Θ	2
z [= Std normal random variable (simulated)]	−0.509
w [= Chi-square$_{(n+\Theta-1)}$ random variable (simulated)]	14.475
v [= $t_{(n+\Theta-2)}$ random variable (simulated) $\div(n+\Theta-2)^{0.5}$]	0.419
z_{eff} [= $v + z \times \{n \times (1+v^2)/w\}^{0.5}$]	−0.230
Simulated ATA [= $1+\exp(\mu_0 + \sigma_0 \times z_{eff})$]	2.315

The simulated age-to-age factor is a single pick from a lognormal distribution with parameter risk taken into account [Note that the "1+" at the start of the expression for the simulated ATA is needed because we fit the curve to (ATA-1)] For each simulated ATA factor, we need to simulate from 3 random variables, to get z, w, and v This was done in Excel, by inverting the cumulative density functions of the respective distributions.

EXHIBIT C-4

Illustration of Fitting an Inverse Power Curve to the Simulated Age-to-Age Factors

(1) Development Period	(2) Year	(3) Simulated ATA	(4) ln(year) ln(2)	(5) ln(ATA-1) ln[(3)−1]	(6) Fitted ATA $1 + a \times (2)^{[-b]}$
12–24	1	2.312			1.486
24–36	2	1.401			1.172
36–48	3	1.208			1.094
48–60	4	1.106			1.061
60–72	5	1.072			1.044
72–84	6	1.055			1.033
84–96	7	1.040			1.026
96–108	8	1.037			1.022
108–120	9	1.029			1.018
120–132	10	1.015	2.303	−4.211	1.015
132–144	11	1.011	2.398	−4.484	1.013
144–156	12	1.013	2.485	−4.360	1.012
156–168	13	1.012	2.565	−4.439	1.010
168–180	14	1.011	2.639	−4.544	1.009
180–192	15	1.013	2.708	−4.362	1.008
192–204	16	1.008	2.773	−4.807	1.008
204–216	17	1.003	2.833	−5.770	1.007
216–228	18	1.005	2.890	−5.365	1.006
228–240	19	1.008	2.944	−4.856	1.006
240–252	20	1.007	2.996	−4.985	1.005

Fitting a least squares line to columns (4) and (5), with (5) as the dependent variable gives the following fitted paramenters:

$$\text{slope} = -1.498$$
$$\text{Intercept} = -0.722$$

Since the inverse power curve can be written in the form: $\ln(\text{ATA-1}) = \ln(a) - b\ln(t)$, we have the following parameters for the inverse power curve:

$$a = \exp(\text{intercept}) = 0.486$$
$$b = -\text{slope} = 1.498$$

EXHIBIT C-5

Illustration of Selecting Age-to-Age Factors

(1) Year	(2) Simulated ATA	(3) Fitted ATA $a = 0.486$ $b = 1.498$	(4) Selected ATA Cut-off for tail* 54	(1) Year	(3) Fitted ATA $a = 0.486$ $b = 1.498$	(4) Selected ATA Cut-off for tail* 54
1	2.312	2.626	2.312	36	1.008	1.008
2	1.401	1.576	1.401	37	1.007	1.007
3	1.208	1.314	1.208	38	1.007	1.007
4	1.106	1.204	1.106	39	1.007	1.007
5	1.072	1.146	1.072	40	1.006	1.006
6	1.055	1.111	1.055	41	1.006	1.006
7	1.040	1.088	1.040	42	1.006	1.006
8	1.037	1.072	1.037	43	1.006	1.006
9	1.029	1.060	1.029	44	1.006	1.006
10	1.015	1.052	1.015	45	1.005	1.005
11	1.011	1.045	1.011	46	1.005	1.005
12	1.013	1.039	1.013	47	1.005	1.005
13	1.012	1.035	1.012	48	1.005	1.005
14	1.011	1.031	1.011	49	1.005	1.005
15	1.013	1.028	1.013	50	1.005	1.005
16	1.008	1.026	1.008	51	1.004	1.004
17	1.003	1.023	1.003	52	1.004	1.004
18	1.005	1.021	1.005	53	1.004	1.004
19	1.008	1.020	1.008	54	1.004	1.000
20	1.007	1.018	1.007	55	1.004	1.000
21		1.017	1.017	56	1.004	1.000
22		1.016	1.016	57	1.004	1.000
23		1.015	1.015	58	1.004	1.000
24		1.014	1.014	59	1.004	1.000
25		1.013	1.013	60	1.004	1.000
26		1.012	1.012	61	1.003	1.000
27		1.012	1.012	62	1.003	1.000
28		1.011	1.011	63	1.003	1.000
29		1.010	1.010	64	1.003	1.000
30		1.010	1.010	65	1.003	1.000
31		1.009	1.009	66	1.003	1.000
32		1.009	1.009	67	1.003	1.000
33		1.009	1.009	68	1.003	1.000
34		1.008	1.008	69	1.003	1.000
35		1.008	1.008	70	1.003	1.000

*The cut off for the tail models the actuarial uncertainty in when to cut off the development from the inverse power curve The cut-off is based on a uniform distribution from 30 to 70.

EXHIBIT C-6

CALCULATION OF REQUIRED RESERVES FOR A SINGLE SIMULATION (MEDICAL PAYMENTS ONLY)

Year	Paid to Date	Age-to-Ultimate	Ultimate Paids	Indicated Reserves	Present Value Factor	Present Value of Reserves
1994	1,787,601	3.202	5,723,852	3,936,251	0.697	2,744,063
1993	3,324,538	1.778	5,910,348	2,585,810	0.579	1,496,946
1992	4,208,871	1.538	6,474,177	2,265,307	0.514	1,164,422
1991	7,017,997	1.462	10,261,961	3,243,963	0.497	1,612,068
1990	7,547,277	1.393	10,511,828	2,964,552	0.470	1,392,487
1989	7,905,743	1.348	10,655,677	2,749,934	0.452	1,243,164
1988	8,507,321	1.306	11,112,168	2,604,846	0.427	1,112,307
1987	7,629,124	1.284	9,798,726	2,169,602	0.422	915,457
1986	6,621,638	1.270	8,409,386	1,787,748	0.426	761,993
1985	5,398,367	1.250	6,746,697	1,348,331	0.418	563,797
1984	3,997,086	1.234	4,932,840	935,754	0.415	388,306
1983	3,198,587	1.222	3,908,208	709,622	0.417	295,599
1982	2,895,279	1.210	3,504,490	609,210	0.418	254,948
1981	2,929,995	1.200	3,517,101	587,106	0.422	248,033
1980	2,704,128	1.192	3,222,023	517,895	0.429	221,946
1979	2,552,368	1.181	3,013,230	460,862	0.428	197,322
1978	2,375,139	1.173	2,786,341	411,202	0.436	179,325
1977	1,986,508	1.172	2,328,957	342,449	0.463	158,711
1976	1,680,001	1.163	1,954,084	274,083	0.469	128,469
1975	1,321,413	1.159	1,531,944	210,531	0.489	103,028
1974	1,154,614	1.146	1,323,337	168,723	0.483	81,430
1973	1,004,449	1.135	1,140,181	135,733	0.478	64,937
1972	908,372	1.124	1,021,158	112,786	0.470	53,015
1971	782,100	1.118	874,591	92,491	0.478	44,228
1970	776,907	1.113	864,352	87,445	0.487	42,566
Total	90,215,423		121,527,659	31,312,236		15,468,566

APPENDIX D

INFLATION ADJUSTMENTS

For certain long-tailed lines of business, much reserve uncertainty stems from changes in the rate of inflation. For workers compensation medical benefits, as an example, the employer is responsible for physician fees, which are affected by the rate of inflation up through the date that the service is rendered.

Paid loss development analyses may overstate the uncertainty in reserve indications, particularly if one is concerned with the economic value of the reserves and not their nominal value. For instance, suppose that the cumulative paid losses *in real dollar terms* will increase by 30% over the coming year, for a "real dollar" age-to-age factor of 1.300. If inflation is high, the nominal age-to-age factor may be 1.350. If inflation is low, the nominal age-to-age factor may be 1.320.

To some extent, this is "apparent" reserve uncertainty, not real reserve uncertainty. We can get a better estimate of reserve uncertainty by

- Stripping inflation out of the historical paid losses,
- Determining "age-to-age" factors in real dollar terms,
- Using the "real dollar" factors to produce all the simulations, and
- Restoring nominal inflation, based upon a stochastically generated inflation rate path, to determine ultimate losses.[52]

Exhibit D-1 shows the procedure used to put the paid loss experience into real dollar terms (at a 1994 price level). We demon-

[52]These adjustments are equally important for standard "point estimates" of indicated reserves. Nominal dollar paid loss "age-to-age" factors have the historical inflation rate built into them (see Cook [15]). If future inflation is expected to be different from past inflation, a rote application of the paid loss chain ladder technique may give misleading reserve indications.

strate the procedure for medical benefits, which we assume to be fully inflation sensitive. Indemnity benefits, in contrast, are only partially inflation sensitive. About half the states have "cost of living" adjustments for wage loss benefits, but generally these adjustments apply only to certain cases (such as cases that extend for two years or more), and they are often capped (say, at 5% per annum).

We begin with the medical component of the Consumer Price Index, shown on the second row of Exhibit D-1. During the 1980s, the rate of increase in workers compensation medical benefits exceeded the medical CPI. This additional WC medical inflation is related to increases in utilization rates or, perhaps, to the incurral of medical services to justify claims for increased indemnity benefits.

For ratemaking, we would need a "loss cost trend factor" for workers compensation medical benefits, of which the medical CPI is but one component. For our purposes, we are concerned only with medical inflation. Changes in utilization rates remain embedded in the paid loss development factors. If the reserving actuary believes that future changes in utilization rates will differ from past changes in utilization rates, this expected difference must be separately quantified.

We must convert the *incremental* paid losses during each calendar year to their "real dollar" (calendar year 1994) values. For ease of application, the one dimensional index in the second row of Exhibit D-1 is converted to a two-dimensional triangle. For instance, the "0.76" in column (5) for accident year 1990 means that accident year losses paid between 48 and 60 months (i.e., between January 1, 1995, and December 31, 1995) must be multiplied by 0.76 to bring them to accident year 1990 levels. The 0.76 factor is derived from the inflation index: $0.76 = 1/(1.0885 \times 1.0805 \times 1.0667 \times 1.0536)$.

We now redo the entire simulation procedure as documented in Appendix C, using the paid losses that have all been adjusted to a 1994 cost level.

Inflation Rate Generator

The derivation of the stochastic medical inflation rate model is shown in Exhibit D-2. We use the medical CPI as the "monetary" inflation component of workers compensation medical benefits, since this is the index that we used to deflate the medical link ratios in Exhibit D-1.

Workers compensation medical loss cost trends are not necessarily the same as the medical CPI, whether year by year or over a long-term average, since other factors (such as utilization rates) affect medical loss cost trends. The historical link ratios are not deflated for this residual trend, so the residual trend is not added back for future periods. If the reserving actuary believes that future utilization rate trends will differ from the historical utilization rate trends, a further adjustment should be made to the simulation model.[53]

Restoring Inflation

To properly estimate reserves, we must "restore" future inflation at the rates stochastically generated for this scenario. To keep the calculations tractable, we assume (i) annual changes in interest rates and inflation rates, and (ii) mid-year loss payments.[54]

The procedure consists of the following steps:

- Remove inflation from the historical link ratios, fit them to a lognormal curve, accounting for parameter risk, and simulate future link ratios for each accident year, as in Appendix C.
- From the simulated link ratios, determine age-to-ultimate factors and payment patterns for each accident year.

[53]The advent of managed care procedures in the 1990s may warrant such an additional adjustment.

[54]Mid-year loss payments is the common proxy for loss payments spread evenly over the year. For payments after the first year, this is a reasonable approximation.

- Stochastically generate an interest rate path and an inflation rate path.
- Assume all payments are made at mid-year. Inflate the "real dollar" loss payments by the future inflation rates to determine nominal loss payments. The sum of the loss payments is the undiscounted required reserve.
- Discount the nominal loss payments by the future interest rates to determine discounted loss payments. The sum of the discounted loss payments is the discounted required reserve.

For example, suppose that in one simulation we had the following figures:

Year	Simulated Link Ratio	Development Factor	Payment Pattern	Inflation Rate	Interest Rate
1	1.776	2.446	0.409	5.7%	7.5%
2	1.105	1.378	0.317	6.3%	6.6%
3	1.057	1.247	0.076	6.2%	6.4%

The simulated link ratios are for a particular accident year in a particular simulation. The simulated development factors are the backward product of the simulated link ratios. For instance, $2.446 = 1.378 \times 1.776$.

The payment pattern is the percent of losses paid in the calendar year preceding the development factor in the adjoining cell. For instance, the development factor at the end of "year 1" is 2.446. This implies that the percent of losses paid in the first 12 months equals $1 \div 2.446$, or 40.9%. At the end of the second year, the development factor is 1.378. This implies that the percent of losses paid in the first 24 month is $1 \div 1.378$, or 72.6%. Since 40.9% of losses have been paid in the first 12 months, 31.7% of losses are paid between 12 and 24 months.

To simplify the exposition of the inflation and discounting procedures, assume that total "real dollar" losses are $1,000,000 for the most recent accident year (1994 in our example). Of this amount, \$409,000 is paid in the first twelve months, and they are not included in the loss reserves held at the end of the year.

Another \$317,000 is paid on July 1 of the following calendar year (1995 in our example). This amount is in December 31, 1994 dollars. The nominal losses paid are therefore $\$317,000 \times (1.057)^{0.5}$. The discounted dollars in this scenario equal $\$317,000 \times (1.057)^{0.5} \div (1.075)^{0.5}$.

Another \$76,000 is paid on July 1 of the next calendar year (1996 in our example). Again, this amount is in December 31, 1994 dollars. The nominal losses paid are therefore $\$76,000 \times (1.057) \times (1.063)^{0.5}$. The discounted dollars in this scenario equal $\$76,000 \times (1.057) \times (1.063)^{0.5} \div \{(1.075) \times (1.066)^{0.5}\}$.

EXHIBIT D-1

Stripping Medical Inflation from the Losses

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
Medical Inflation	6.65%	6.40%	4.75%	3.65%	6.65%	10.65%	10.75%	9.55%	9.00%	8.80%	10.10%	10.85%	11.15%	10.20%	7.50%	6.25%	6.90%	7.05%	6.55%	7.10%	8.35%	8.85%	8.05%	6.67%	5.36%

Accident Year	Index for use in Calendar Year (Multiplying the corresponding element in the traingle by this factor puts the loss back at the medical price level for the accident year)																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1970	1.00	0.94	0.90	0.87	0.81	0.73	0.66	0.60	0.55	0.51	0.46	0.42	0.38	0.34	0.32	0.30	0.28	0.26	0.24	0.23	0.21	0.19	0.18	0.17	0.16
1971	1.00	0.95	0.92	0.86	0.78	0.70	0.64	0.59	0.54	0.49	0.44	0.40	0.36	0.34	0.32	0.30	0.28	0.26	0.24	0.22	0.21	0.19	0.18	0.17	
1972	1.00	0.96	0.90	0.82	0.74	0.67	0.62	0.57	0.52	0.47	0.42	0.38	0.35	0.33	0.31	0.29	0.27	0.25	0.24	0.22	0.20	0.19	0.18		
1973	1.00	0.94	0.85	0.77	0.70	0.64	0.59	0.53	0.48	0.43	0.39	0.37	0.34	0.32	0.30	0.28	0.26	0.24	0.22	0.21	0.19	0.18			
1974	1.00	0.90	0.82	0.74	0.68	0.63	0.57	0.51	0.46	0.42	0.39	0.37	0.34	0.32	0.30	0.28	0.26	0.24	0.22	0.21	0.20				
1975	1.00	0.90	0.82	0.76	0.70	0.63	0.57	0.51	0.46	0.43	0.41	0.38	0.36	0.33	0.31	0.29	0.26	0.24	0.23	0.22					
1976	1.00	0.91	0.84	0.77	0.70	0.63	0.57	0.51	0.48	0.45	0.42	0.39	0.37	0.35	0.32	0.29	0.27	0.25	0.24						
1977	1.00	0.92	0.84	0.77	0.69	0.62	0.56	0.52	0.49	0.46	0.43	0.41	0.38	0.35	0.32	0.30	0.28	0.26							
1978	1.00	0.92	0.83	0.75	0.68	0.61	0.57	0.54	0.50	0.47	0.44	0.41	0.38	0.35	0.32	0.30	0.29								
1979	1.00	0.91	0.82	0.74	0.67	0.62	0.59	0.55	0.51	0.48	0.45	0.41	0.38	0.35	0.33	0.31									
1980	1.00	0.90	0.81	0.74	0.69	0.64	0.60	0.56	0.53	0.49	0.46	0.42	0.39	0.36	0.34										
1981	1.00	0.90	0.82	0.76	0.71	0.67	0.62	0.59	0.55	0.51	0.46	0.43	0.40	0.38											
1982	1.00	0.91	0.84	0.79	0.74	0.69	0.65	0.61	0.56	0.52	0.48	0.45	0.42												
1983	1.00	0.93	0.88	0.82	0.77	0.72	0.67	0.62	0.57	0.53	0.49	0.47													
1984	1.00	0.94	0.88	0.82	0.77	0.72	0.67	0.61	0.57	0.53	0.50														
1985	1.00	0.94	0.87	0.82	0.77	0.71	0.65	0.60	0.56	0.53															
1986	1.00	0.93	0.88	0.82	0.76	0.69	0.64	0.60	0.57																
1987	1.00	0.94	0.88	0.81	0.74	0.69	0.64	0.61																	
1988	1.00	0.93	0.86	0.79	0.73	0.69	0.65																		
1989	1.00	0.92	0.85	0.78	0.74	0.70																			
1990	1.00	0.92	0.85	0.80	0.76																				
1991	1.00	0.93	0.87	0.82																					
1992	1.00	0.94	0.89																						
1993	1.00	0.95																							
1994	1.00																								

EXHIBIT D-2

Page 1

Fitting of Model for Medical Inflation

Model: Medical inflation$_t$ =
$a\times$(interest rate$_t$) + $\beta\times$[(medical inflation$_{t-1}$) $-\,\alpha\times$(interest rate$_{t-1}$)]
+(1 $-\,\beta$)$\times$[(avg. medical inflation)$-\alpha\times$(avg. interest rate)]+error$_t$
$\alpha = 0.484 \qquad \beta = 0.546$
α and β are chosen to minimize the sum of the squared errors in column 6

Year	(1) Medical CPI at December	(2) Annual % Increase in Medical CPI	(3) Yield on Intermediate Term Govt Bonds*	(4) Least-Squares Fit of Medical Inflation Model**	(5) Error***	(6) Squared Error****
1935	10.2					
1936	10.2	0.0%	1.3%			
1937	10.3	1.0%	1.1%	1.5%	−0.56%	0.00003
1938	10.3	0.0%	1.5%	2.3%	−2.30%	0.00053
1939	10.4	1.0%	1.0%	1.4%	−0.43%	0.00002
1940	10.4	0.0%	0.6%	1.9%	−1.88%	0.00035
1941	10.5	1.0%	0.8%	1.6%	−0.61%	0.00004
1942	10.9	3.8%	0.7%	2.0%	1.82%	0.00033
1943	11.4	4.6%	1.5%	3.9%	0.67%	0.00004
1944	11.7	2.6%	1.4%	4.1%	−1.49%	0.00022
1945	12.0	2.6%	1.0%	2.9%	−0.33%	0.00001
1946	13.0	8.3%	1.1%	3.0%	5.34%	0.00285
1947	13.9	6.9%	1.3%	6.2%	0.69%	0.00005
1948	14.7	5.8%	1.5%	5.5%	0.27%	0.00001
1949	14.9	1.4%	1.2%	4.7%	−3.31%	0.00109
1950	15.4	3.4%	1.6%	2.5%	0.83%	0.00007
1951	16.3	5.8%	2.2%	3.8%	2.06%	0.00043
1952	17.0	4.3%	2.4%	5.1%	−0.79%	0.00006
1953	17.6	3.5%	2.2%	4.1%	−0.58%	0.00003
1954	18.0	2.3%	1.7%	3.5%	−1.24%	0.00015
1955	18.6	3.3%	2.8%	3.5%	−0.14%	0.00000
1956	19.2	3.2%	3.6%	4.2%	−0.94%	0.00009
1957	20.1	4.7%	2.8%	3.5%	1.18%	0.00014
1958	21.0	4.5%	3.8%	5.0%	−0.50%	0.00003
1959	21.8	3.8%	5.0%	5.2%	−1.37%	0.00019
1960	22.5	3.2%	3.3%	3.7%	−0.48%	0.00002
1961	23.2	3.1%	3.8%	4.1%	−0.95%	0.00009
1962	23.7	2.2%	3.5%	3.7%	−1.55%	0.00024
1963	24.3	2.5%	4.0%	3.5%	−1.00%	0.00010
1964	24.8	2.1%	4.0%	3.6%	−1.54%	0.00024
1965	25.5	2.8%	4.9%	3.8%	−0.94%	0.00009
1966	27.2	6.7%	4.8%	3.9%	2.77%	0.00077
1967	28.9	6.3%	5.8%	6.5%	−0.24%	0.00001
1968	30.7	6.2%	6.0%	6.1%	0.13%	0.00000
1969	32.6	6.2%	8.3%	7.2%	−0.98%	0.00010

EXHIBIT D-2

PAGE 2

FITTING OF MODEL FOR MEDICAL INFLATION

Year	(1) Medical CPI at December	(2) Annual % Increase in Medical CPI	(3) Yield on Intermediate Term Govt Bonds*	(4) Least-Squares Fit of Medical Inflation Model**	(5) Error***	(6) Squared Error****
1970	35.0	7.4%	5.9%	5.4%	1.99%	0.00040
1971	36.6	4.6%	5.3%	6.3%	−1.76%	0.00031
1972	37.8	3.3%	5.9%	5.3%	−1.99%	0.00040
1973	39.8	5.3%	6.8%	4.9%	0.43%	0.00002
1974	44.8	126%	7.1%	5.9%	6.69%	0.00448
1975	49.2	9.8%	7.2%	9.8%	0.04%	0.00000
1976	54.1	10.0%	6.0%	7.7%	2.27%	0.00051
1977	58.9	8.9%	7.5%	8.8%	0.06%	0.00000
1978	64.1	8.8%	8.8%	8.5%	0.37%	0.00001
1979	70.6	10.1%	10.3%	8.8%	1.33%	0.00018
1980	77.6	9.9%	12.5%	10.2%	−0.24%	0.00001
1981	87.3	12.5%	14.0%	10.2%	2.29%	0.00053
1982	96.9	11.0%	9.9%	9.3%	1.74%	0.00030
1983	103.1	6.4%	11.4%	10.2%	−3.84%	0.00147
1984	109.4	6.1%	11.0%	7.1%	−1.04%	0.00011
1985	116.8	6.8%	8.6%	5.9%	0.88%	0.00008
1986	125.8	7.7%	6.9%	6.1%	1.63%	0.00027
1987	133.1	5.8%	8.3%	7.8%	−1.95%	0.00038
1988	142.3	6.9%	9.2%	6.7%	0.18%	0.00000
1989	154.4	8.5%	7.9%	6.5%	1.98%	0.00039
1990	169.2	9.6%	7.7%	7.6%	1.99%	0.00040
1991	182.6	7.9%	6.0%	7.4%	0.50%	0.00003
1992	194.7	6.6%	6.1%	7.0%	−0.40%	0.00002
1993	205.2	5.4%	5.2%	5.9%	−0.46%	0.00002
1994	215.3	4.9%	7.8%	6.7%	−1.75%	0.00030
Mean		5.4%	5.0%		0.04%	0.00033
					183% = Std Dev of errors	0.01901 = Sum of square errors

* Source: Ibbotson Associates: Stocks, Bonds, Bills, and Inflation, 1995 Edition
** Column 4 = α[Col. 3 for current year]+β[Col. 2 for previous year−α(Col. 3 for previous year)]+$(1-\beta)$ [Avg. of Col. 2 − α(Avg. of Col. 3)]
*** Column 5 = Column 2−Column 4
**** Column 6 = $\{\text{Column } 5\}^2$
Fitted α and β minimize the sum of column 6.
The error term for the model is a normal distribution, with mean = 0.00% and standard deviation = 1.83%

APPENDIX E

LOSS-SENSITIVE CONTRACTS

In the text of this paper, we examine the uncertainty in the loss reserves. In practice, reserve uncertainty varies with the type of insurance contract. For instance, high-level workers compensation excess-of-loss covers, as well as large dollar deductible policies offered to large employers, have greater reserve uncertainty, particularly in the early policy years when the insurer's estimated liability is subject to great variation.

For business written on loss-sensitive contracts, such as retrospectively rated plans for large workers compensation risks or reinsurance treaties with sliding scale reinsurance commissions, the opposite is true. Companies are concerned with the uncertainty in the net reserves, or the future loss payments after adjustment for retrospective premiums and variable commissions.[55]

Large dollar deductible policies are relatively new, and we do not yet have the requisite data to estimate the reserve uncertainty. In addition, the slow payment patterns of workers compensation excess covers and of large dollar deductible policies will delay the empirical quantification of their reserving risk.

In contrast, we have relatively complete data on loss-sensitive contracts. Moreover, the effects of loss sensitive contracts on reserve uncertainty has become a significant regulatory and actuarial issue in recent years. The NAIC risk-based capital formula contains an offset of 15% to 30% to the reserving risk charge for business written on loss-sensitive contracts (Feldblum [23]). In

[55] The discussion here assumes familiarity with retrospective rating plans and with their parameters, such as loss limits, premium maximums, and premium minimums, as well as with standard reserving techniques for retrospective premiums. More detailed information on the retrospective rating plan pricing parameters may be found in Simon [54], Skurnick [55], Lee [43], Gillam and Snader [30], Bender [4], and Mahler [45]. The retrospective premium reserving techniques that underlie the analysis in this paper are discussed in Fitzgibbon [26], Berry [6], Teng and Perkins [57], and Feldblum [21].

1995, a new Part 7 was added to Schedule P of the Fire and Casualty Annual Statement to quantify the risk-based capital loss-sensitive contract offset and to measure the premium sensitivity to losses on loss-sensitive contracts (Feldblum [20], [21]).

This appendix presents an analysis of reserving risk on retrospectively rated policies. Insurers writing excess layers of coverage or large dollar deductible policies should perform a similar analysis on those policy types.

When the retrospective rating plan contains loss limits or premium maximums and minimums, reserving risk remains, though it is dampened. These plans are more risky in some ways and less risky in other ways than traditional first dollar coverages are. The "pure insurance portion" of the plan is more risky, since

- The consideration paid by the insured is the "insurance charge", and
- The benefits paid by the insurer are the difference between (a) the value of the uncapped and unbounded premium and (b) the value of the capped and bounded premiums.[56]

The "pure insurance portion" is like excess-of-loss reinsurance, where the loss limit provides coverage like that of per-accident excess-of-loss and the premium bounds provide coverage like that of aggregate excess-of-loss. The variability of reserves for excess layers of coverage, per dollar of reserve, is generally greater than the corresponding variability of reserves for first dollar coverage.

If the retrospectively rated policy is considered as a whole—(both the insurance portion and the "pass-through" portion)—the retrospectively rated plan is less risky, per dollar of loss, than

[56] "Caps" refer to the loss limits; "bounds" refer to the premium maximums and minimums. "Ratable losses" are paid by the insurer but reimbursed by the employer, so there is no insurance risk. Acquisition expenses, underwriting expenses, and adjustment expenses are paid by the insurer but reimbursed in the basic premium and in the loss conversion factor, again eliminating much of the risk to the insurer.

traditional first dollar coverage. In fact, if there are no loss limits and no maximum or minimum bounds on the premium, then the insurance contract becomes simply a financing vehicle and the insurance company serves as a claims administrator, not as a risk-taker. There is no underwriting or reserving uncertainty at all, though there is still “credit risk” (see Greene [31]).

Premium Sensitivity

How potent are loss sensitive contracts in reducing “net” loss reserve uncertainty? (By “net” loss reserve uncertainty, we mean the variability in the insurer’s total reserves, or loss reserves minus retrospective premium reserves. The “accrued retrospective premium reserves” are carried as an asset on statutory financial statements, whereas loss reserves are carried as a liability.) The answer depends on the “premium sensitivity” of the plan; that is, the amount of additional premium generated by each additional dollar of loss.

We quantify the net loss reserve uncertainty in the same fashion as we did earlier, by asking: “How does reserve uncertainty affect the financial condition of the insurer?” For instance, if the required reserves turn out to be 15% higher than our current estimates, how much additional funds will the company need to meet its loss obligations?

For business which is not written on loss sensitive contracts, the answer is simple. The additional funds needed equal the additional dollars of loss minus the amount of any implicit interest cushion in the reserves.

For business written on loss sensitive contracts, the answer is more complex, as the following illustration shows. Suppose that the indicated workers compensation reserves are $800 million. As a conservative range to guard against reserve uncertainty, the valuation actuary chooses an upper bound of $1,050 million as the worst case reserve estimate. The actuary estimates that there would be about $200 million of implicit interest margin in this scenario, so the capital needed to guard against reserve uncertainty is $50 million.[57]

Suppose now that half of the company's workers compensation business is written on retrospectively rated policies, of two types:

- Large accounts have plans with wide swings; loss limits and premium maximums are high, so each additional dollar of loss generates about a dollar of premium.
- Small and medium-size accounts have plans with narrower swings. Loss limits and premium maximums are lower and constrain the retro premiums. On average, each additional dollar of loss generates about 65¢ of additional premium.

For the entire book of retrospectively rated contracts, the premium sensitivity is 80%; that is, each additional dollar of loss generates about 80¢ of additional premium.

How much capital should this insurer hold to guard against reserve uncertainty? Suppose the needed reserves increase to the "worst case" scenario of $1,050 million. Half of this business is written on retrospectively rated plans, and the average premium sensitivity is 80%. In other words, of the adverse loss development of $250 million, $125 million occurs on retrospectively rated business. With a premium sensitivity of 80%, adverse loss development of $125 million generates $100 million of additional premium.

We add the $100 million of additional premium to the $200 of implicit interest margin to arrive at a solvency cushion of $300 million. Since the worst case adverse loss development is $250 million, the company already has a $50 million surplus

[57]For the illustration, we assume that the company wishes to hold a margin for reserve uncertainty even greater than the implicit interest margin. The text of this paper shows that for workers compensation, this implies a very low EPD ratio.

solvency cushion in the carried reserves, so no additional capital is needed.[58]

In sum, loss sensitive contracts have potent implications for the quantification of reserve uncertainty. We examine this subject from two perspectives:

- A theoretical perspective, showing the factors affecting the risks in loss sensitive contracts, and
- A simulation perspective, showing the effects of loss sensitive contracts on our measures of reserve uncertainty.

Underwriting Risk and Reserving Risk

Before turning to reserve uncertainty, let us broaden our inquiry and ask: "To what extent do retrospectively rated policies mitigate underwriting uncertainty in general?" We can answer this question empirically, by comparing the variability of standard loss ratios and net loss ratios on a large and mature book of retrospectively rated workers compensation policies.

- *Standard loss ratios* are incurred losses divided by standard earned premium. These loss ratios are influenced by random loss occurrences and premium rate fluctuations, and they vary considerably over time.
- *Net loss ratios* are incurred losses divided by the final earned premiums, as modified by retrospective adjustments. These adjustments counteract both the random loss occurrences and the

[58] An adjustment is needed to bring the accrued retrospective premiums to present value. The magnitude of this adjustment depends on the type of retrospective rating plan. For "paid loss" retro plans, the additional premium is collected when the losses are paid, so the present value of the retro premium is less than $100 million. For "incurred loss" retro plans, the additional premium is billed and collected when the case reserves develop adversely, so a smaller adjustment is needed. In this illustration, the implicit interest margin in the loss reserves is $200 million÷$1,050 million, or 19%. If all the retro plans in this illustration are paid loss retros, and the additional premium is collected when the losses are paid, the present value of the additional premiums is $81 million.

fluctuations in manual rate levels, so the net loss ratios should be more stable over time.

Exhibit E-1 shows these loss ratios for retrospectively rated policies issued by a large workers compensation insurer. Only mature policies are used in this comparison, to ensure that the net loss ratios are not subject to significant additional retrospective adjustments.[59]

As expected, the mean loss ratios are similar for standard and net—77.0% for standard and 78.8% for net. (The net loss ratios are slightly higher, since more retrospective premiums are returned than are collected.) The variances and standard deviations, however, differ greatly. The standard loss ratios show a variance of 46.9% and a standard deviation of 68.5%. Retrospective rating dampens the fluctuations in the loss ratios, leading to a variance of 11.2% and a standard deviation of 33.4%.

Reserve Uncertainty

Exhibit E-1 deals with (prospective) underwriting risk, or the risk that future underwriting returns will be lower than anticipated. Let us return now to reserving risk. We ask "To what extent is adverse development on existing losses mitigated by loss sensitive contracts?"

To resolve this issue, we must know the premium sensitivity of the retrospective rating plans, or the amount of additional premium received for each dollar of additional loss. Let us examine the variables that affect the premium sensitivity: the plan parameters, the current loss ratio, and the maturity of the reserves.[60]

---

[59]The exhibit in this paper, along with the variances and standard deviations, was produced by Miriam Perkins. An earlier exhibit from the same book of business, produced by Dr. J. Eric Brosius, was provided by the authors to the American Academy of Actuaries task force on risk-based capital. It was used by the Tillinghast consulting firm to support the recommendations of the task force regarding the loss-sensitive contract offset to the reserving and underwriting risk charges in the NAIC risk-based capital formula.

[60]Compare Bender [4, p. 36]: "The aggregate premium returned to a group of individual risks that are subject to retrospective rating depends upon the retrospective rating formula, the aggregate loss ratio of the risks, and the distribution of the individual risks' loss ratios around the aggregate."

Plan Parameters

If the retrospective rating plan had no loss limits and no constraints on the final premium, the premium sensitivity would equal the loss conversion factor times the tax multiplier, which is generally equal to or greater than one. In most cases—and particularly for smaller risks—the loss limits and the premium maximums constrain the swing of the plan, and the premium sensitivity is lower than one.

Generally, larger insureds choose retrospective rating plans with wide swings, while smaller insureds choose more constrained plans. To quantify premium sensitivity, therefore, the book of business should be divided into relatively homogeneous groups by size of risk, such as between medium sized risks and "national accounts."[61] (Small risks rarely use retrospective rating plans.)

The differences are dramatic. National accounts in our own book of business, with annual premium of $2 million or more per risk, almost always have wide swing plans, and the average premium sensitivity is close to one. Medium sized risks in our

---

There are several additional items which should also be examined for a complete analysis of the effects of loss-sensitive contracts on reserve uncertainty. As noted earlier, we should look at the effects of "incurred loss" retros versus "paid loss" retros on the implicit interest margin in the accrued retrospective premiums. To be conservative, we assume here that all plans are paid loss retros; since the additional loss payments and the additional premium collections occur at the same time, we simply net them out. Incurred loss retros would show even greater dampening of the loss reserve uncertainty; since the premiums have less implicit interest margin, the effective premium sensitivity is greater than a nominal dollar analysis indicates.

In addition, a complete analysis should look at the effects of the plan parameters on the credit risk of the company and on the size of the implicit interest margin. The accrued retrospective premiums are a receivable, not an investable asset. As is true for losses, they are held on statutory financial statements at ultimate value, not at present value. If loss reserves are backed by accrued retrospective premiums, then either these premium reserves should be reduced to present value or the implicit interest margin in the loss reserves should be reduced.

[61] This subdivision of the data by size of insured or by "underwriting market" is generally available in company files. Of course, if the company keeps data by type of plan (wide swing plans vs. narrow swing plans and so forth), this more accurate subdivision is preferable.

book of business, with more constrained plans, have an average premium sensitivity of about 65%.[62]

Loss Ratio

The premium maximum and the loss limits constrain the swing of the plan. Ideally, we wish to know whether adverse loss development causes the retrospectively rated premium on each policy to hit the premium maximum or the loss to hit the loss limit. However, we do not have information on each individual change in reported losses. Actuaries estimate from aggregates, not from details. We must determine which aggregate statistics are suitable predictors of the average amount of retrospective premium that will be collected.

Given the parameters of any retrospectively rated plan, the loss ratio determines whether the retrospective premium will be capped at the maximum. Given a distribution of loss ratios in a book of business, all of which are written on similar retrospectively rated plans, we can estimate the percent of plans that will hit the maximum premium. If the *shape* of this distribution does not depend significantly upon the average loss ratio of the book of business, and if we know the average loss ratio, then we can determine the percent of plans that will hit the maximum premium.

The general rule is that *premium sensitivity declines as the aggregate loss ratio increases.* During poor underwriting years,

---

[62]These are empirical figures, using actual ratios of retrospective premium collected to historical loss development. Bender [4], using theoretical relationships based on the NCCI's "Table M," estimates premium sensitivity for various risk sizes. Bender's analysis is a useful check on our procedure, but it is not a substitute. His analysis posits that the Table M relationships are correct and that compensation carriers actually use the NCCI Table M insurance charges to price their retrospectively rated policies. In practice, insurers use a variety of plans for their large insureds, and they often negotiate the loss limits, premium maximum, and plan parameters in each case for their national accounts.

As emphasized in Howard Mahler's [45] discussion of Bender's paper, the premium sensitivity is strongly dependent on the size of the risk. Bender analyzes primarily small risks, where the premium sensitivity is weak. The sensitivity rises rapidly with the size of the risk; see especially Bender's [4] Table 5 on page 50, which shows the "slope" of the plan as a function of the "loss group," and Mahler's [45] comments on pages 76–78.

when loss ratios are higher, adverse loss development leads to less additional premium than in good underwriting years, when loss ratios are lower.

Reserve Maturity

In workers compensation, adverse loss development at early maturities stems from delayed reporting of some cases and primarily from the reclassification of less serious cases to more serious cases. For instance, almost all lower back sprains and strains are initially classified as short-term temporary total cases. Significant case reserve development is expected in the first two or three years, as some of these claims develop into permanent partial or permanent total cases. Much of this development is within the "ratable" area of the retrospective rating plan; for instance, a $10,000 claim is reclassified as a $100,000 claim, so premium sensitivity is high.

At later maturities, adverse loss development stems primarily from re-estimation of the costs of permanent cases. For a plan with low or even moderate loss limits, most of the adverse loss reserve development after five or six years occurs in the "non-ratable" portion of the retrospective rating plan. For instance, a $300,000 claim may be re-estimated at $400,000, when it becomes evident that the worker will not soon be returning to work. For plans sold to medium-sized employers, the premium sensitivity for this change is generally low.

Furthermore, many companies "close" their retrospective rating plans after, say, six or seven years, with a final accounting between the company and the insured. Adverse development occurring after this date would not affect the retrospective premiums.[63]

[63]Retrospectively rated plans sold to large accounts are frequently kept open for longer periods. In fact, plans sold to "national accounts" are often kept open indefinitely, or at least until the insurer and the employer agree on a final reckoning.

Effects on the Simulation

For the simulation, we use premium sensitivity factors based on observed long-term patterns by market and by reserve maturity in our countrywide book of business.[64] From the empirical data we produce two curves, each showing premium sensitivity by reserve duration, one for national accounts and one for medium-sized risks. We weight these two curves by the volume of business in these two markets.

In the simulation analysis, we first repeat the steps outlined earlier. Based upon historical experience, we estimate (deterministically) the amount of case reserves associated with each cumulative paid loss amount at each duration. From the change in reported losses, we determine the change in retrospective premiums, and thereby the change in "net reserves."

The effects of loss sensitive contracts vary greatly by type of plan and by company practice. Several reviewers of drafts of this paper have pointed out to us: "Your company writes primarily large accounts and uses highly sensitive, wide swing plans. For this type of business, the net reserve uncertainty is clearly mitigated. What about other companies, which use less sensitive plans, recognize the adverse development later, and close their plans after several years? Would they also show a significant reduction in net reserve uncertainty?"

Accordingly, we made three adjustments, to model the loss sensitive contracts often used for medium-sized risks:

- We assume that the retrospective plans are relatively insensitive. For the most recent accident year, the assumed premium

[64]To avoid undue complexity, we do not consider aggregate loss ratios in the simulation analysis. To incorporate the aggregate loss ratio dimension, we would have to evaluate the effect of each simulated link ratio on the new accident year loss ratio and determine a new premium sensitivity factor for every cell in every simulation. Moreover, since we are using paid loss age-to-age factors, we would have to convert paid loss ratios to incurred loss ratios. The benefits from these refinements are far less than the additional effort.

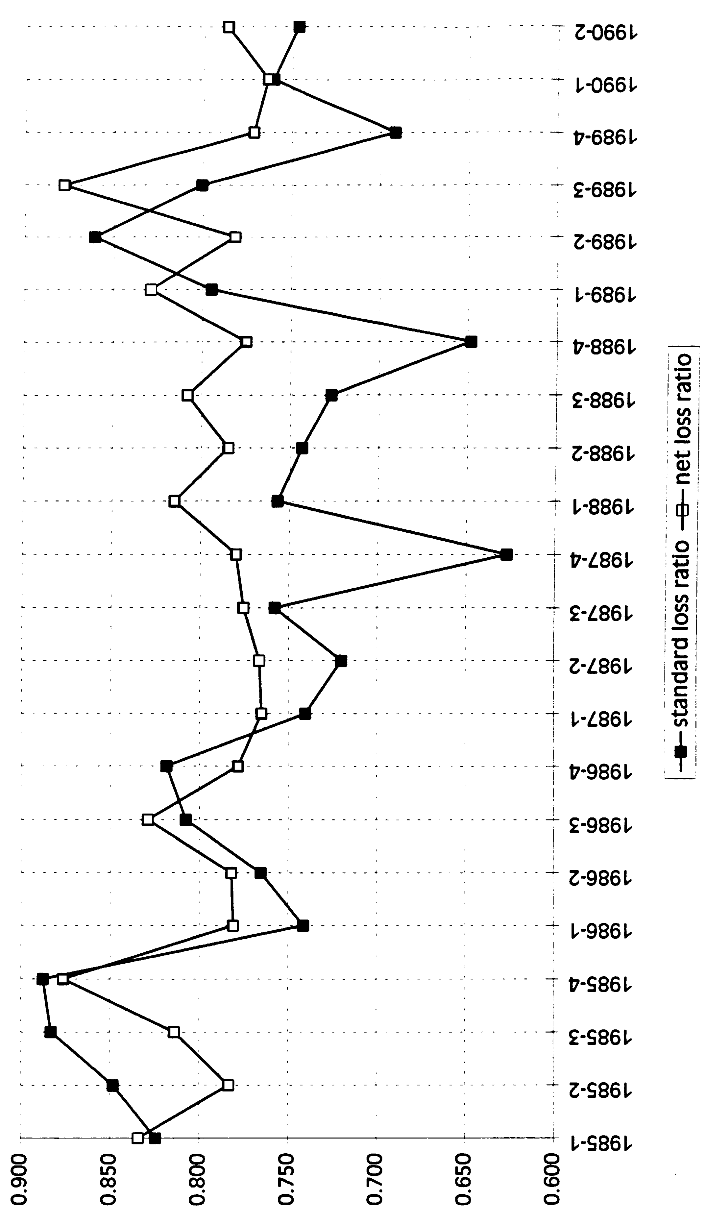
sensitivity is 49%, with the sensitivity factor decreasing for each older accident year.

- We assume that most adverse development is recognized late, when premium sensitivity is lower.
- We assume that the plans are closed, on average, about five to ten years after policy inception. With the late recognition of the adverse development and the relative early closure of the plans, even the limited premium sensitivity is markedly reduced for older accident years.

We ran corresponding stochastic simulations for the loss-sensitive book of business. Even with the assumptions listed above, the projected reserve distribution is more compact, and there is less "reserve uncertainty." Specifically, the use of loss sensitive contracts reduces the standard deviation of the reserve realizations by about 35%, and it reduces the capital needed for a 1% EPD ratio by about 20%.

EXHIBIT E-1

WORKERS COMPENSATION
RETROSPECTIVELY RATED POLICIES



APPENDIX F

PARAMETER UNCERTAINTY IN RESERVE ESTIMATES: THE KREPS PROCEDURE

The analysis in this paper estimates the uncertainty in workers compensation loss reserves. The text and the other appendices explain the method and its rationale, and they provide the simulation equations in sufficient detail that practicing actuaries can replicate our results. Most elements of our procedure are easily visualized, so that the intuition behind each step is clear.

This is less true of the Kreps parameter risk estimation process. The procedure itself is relatively new, having first appeared in the 1997 issue of the *Proceedings of the CAS*. The simulation equations that are shown in Appendix C are taken directly from Kreps [42], which provides the justification for this process. These equations are not self-explanatory, and we have not reproduced the derivations that Kreps provides. Moreover, the magnitude of the parameter risk depends on the choice of the Bayesian prior selected by the analyst, which can be a difficult decision. To aid the reader in understanding our procedure, this appendix provides an intuitive overview of parameter risk and of the Kreps estimation process.

Actuaries generally distinguish between two sources of uncertainty: process risk and parameter risk. Process risk is the risk that actual results will differ from our expected results because of random loss occurrences. Parameter risk is the risk that our expected results are not the true expected results because we have misestimated the parameters of our distributions.

Process risk can generally be estimated directly, as long as one properly identifies all the sources of process risk. In the analysis in the text of this paper, we consider the process risk from age-to-age link ratios, from loss development tail factors, from future interest rates, and from future inflation rates. Parameter risk is more difficult to quantify. Some actuaries would argue that it is impossible to quantify completely, since any estimate of parameter risk relies on assumptions about the nature of the distributions.

In this paper, we use a procedure developed by Kreps [42] to estimate parameter risk. The mathematically adept reader is referred to Kreps's 1997 *Proceedings* paper, which is the basis for the simulations which we use. Appendix C shows the equations that we used in the simulations in incorporate parameter risk. Kreps provides a similar but independent analysis of Homeowners reserve uncertainty, using lognormal distributions of paid loss age-to-age link ratios. Kreps uses fewer data points and a more diffuse Bayesian prior, thus magnifying the parameter risk compared to the process risk. However, workers compensation has much larger paid loss development factors than Homeowners, and the development extends over a much longer period, so the total reserve uncertainty is greater in our analysis than in his.

This appendix does not purport to summarize Kreps's paper, which is already a succinct and clear exposition of a complex topic. Rather, this appendix provides a non-mathematical "intuitive" explanation of what we are doing. It explains where the parameter uncertainty resides in our analysis, what aspects of the parameter uncertainty we purport to measure, how we do so, and what choices we make in the estimation process.

Parameter Risk

Process risk and parameter risk are frequently discussed in relation to policy pricing, particularly for estimating needed profit margins and risk loads. We briefly summarize the pricing distinction between these two sources of risk, and then we extend the distinction to loss reserving.

In traditional ratemaking, the pricing actuary estimates the mean of future loss costs. This mean is based on both historical data, such as two or three years of experience, and various adjustment factors, such as development factors and trend factors.

The traditional procedure gives an expected mean for future loss costs, frequently called a "best-estimate."[65] The traditional procedure does not indicate how much uncertainty is associated with the expected future loss costs.

The uncertainty can arise from two sources: process risk and parameter risk. The pricing actuary is setting a premium rate, which considers only the expected value of the future loss costs. But losses are random occurrences, and actual losses will almost never precisely equal the expected losses. Process risk is the risk that actual losses will differ from the true expected losses.

The total pricing uncertainty, however, is the risk that actual losses will differ from our estimate of future loss costs, not from the true expected loss costs. Parameter risk is the risk that our estimate of future loss costs differs from the true expected loss costs. Parameter risk arises because the components of our pricing procedure are estimates, not known values. This is clear for such items as trend factors, since we can only estimate the effects of monetary inflation and other "social" influences on insurance losses. This is equally true, though, for our historical data. The pricing actuary begins with past experience, which he or she trends to a future policy period. In truth, the pricing actuary wishes to begin with the expected past experience, or the losses that were expected in the historical experience period. Sometimes the actual past losses are the best estimate of the expected past losses. At other times, the pricing actuary makes explicit corrections to actual past experience; the smoothing of catastrophe experience and the credibility weightings of historical loss ratios are two examples of this. Parameter risk includes the risk that

[65] In fact, this estimated mean may not be the true "best estimate"; that is, it may not be the true mean of the estimated distribution. This is because the distributions used to generate the future loss costs, such as the distribution of historical losses, the distribution of development factors, and the distribution of trend factors, are often highly skewed and correlated. For example, the trend factor used in ratemaking is the product of several annual trend rates, and these rates are autocorrelated. The mean of the product of several skewed and correlated distributions is not the same as the product of the means of these distributions.

the historical experience was not the expected experience even in the past.

Parameter Risk: Reserving

Loss reserve estimates show the same two sources of uncertainty. Chain ladder loss development methods derive age-to-age link ratios from past experience and use them to estimate future development. Process risk is the risk that actual loss development link ratios experienced in the future will differ from the true expected link ratios, since the occurrence of IBNR claims, the durations and the extent of disability on known claims, and the decisions of hearing officers and courts on contested claims are all unknown factors that influence the ultimate losses.

Traditional reserve analyses use the average historical link ratios as estimates of future ones, adjusted perhaps for outlying observations, "high" and "low" values, and systematic changes in claims operations or in the insurance environment. In this paper, we do not project "best-estimate" age-to-age link ratios. Instead, we use the historical link ratios to estimate the distribution from which future link ratios may emerge. We assume that the actual link ratios in any given development period are members of a lognormal family. We fit the parameters of the lognormal curve for each development period from the historical observations.

Parameter risk may take several forms. Some types of parameter risk are dealt with in other parts of our simulation procedure. For instance, the traditional reserve analysis is hampered by the possibility that changes in inflation rates will modify the distribution of link ratios. Our simulation procedure makes this risk explicit by stochastically generating future inflation rate paths.

Another type of parameter risk is the risk that the distribution of age-to-age link ratios is better modeled by some other curve, not by a lognormal. Curve families differ in their skewness and in the thickness of their tails, which affect the future (simulated) link ratios.

This risk definitely exists; the distributions of link ratios are presumably not perfectly lognormal. To a large extent, this risk is implicitly incorporated in our parameter risk estimation procedure, since the family of all lognormal distributions probably covers most of the variability in the actual future link ratios.[66] However, the reader should be aware that we have assumed that the distribution of link ratios is lognormal.

The parameter risk that we model here is the risk that we have incorrectly chosen the parameters of the lognormal distribution. If we had an unlimited number of observations from a distribution, we would be fully confident that the fitted distribution was indeed the true distribution. With the small sample of observations in actual reserving practice, the parameters of the fitted distribution may differ from the parameters of the true distribution.

There are other possible reasons for an incorrect choice of parameters. Perhaps we chose parameters which were correct for the historical period, but the distribution has since changed. A workers compensation insurance analogue would be a change in the types of claims over time. For instance, temporary total disability claims have low paid loss link ratios, whereas permanent partial disability claims have higher paid loss link ratios. If the mix of claims has been changing from temporary total to permanent partial, this will cause a change in the overall paid loss link ratios.

In his analysis of experience rating plan credibilities, Mahler [46] divides the total expected variance into two parts: the within variance and the between variance. He further divides the within variance into two parts: the process risk for any individual, and the change of the individual's distribution over time (the fluctuation of risk parameters over time). The standard techniques

[66] As noted by Hayne [32, p. 96], "Estimates of parameter variability may address some of the uncertainty inherent in the choice of a particular distribution for the model."

for estimating within variance usually incorporate both of these types of risk.

We have followed Mahler's approach in our analysis. We have estimated the distribution of link ratios from the full historical experience. To the extent that this distribution has been changing over time, the historical observations exhibit more variance than would otherwise be the case. The process risk estimated in our paper includes both the process risk from a stable distribution as well as the risk stemming from changing distributions over time, which Mahler terms "specification risk" (private communication).

The parameter risk incorporated in our analysis is the risk that the historical sample of observed link ratios does not accurately reflect the parameters of the true distribution. The magnitude of this parameter risk depends on three items: (i) the size of the sample, (ii) the variance of the sample observations, and (iii) our prior knowledge (or our assumed prior knowledge) of the distribution of link ratios. These factors have a strong effect on our results. We explain the intuition by illustration.

Suppose that we are estimating paid loss link ratios for 24 months to 36 months. The historical experience gives us 5 observations, of 1.400, 1.450, 1.600, 1.425, and 1.500. The average of these numbers is 1.475. We presume that these observations come from a distribution with a mean of 1.475.

With only five observations, none of which is exactly 1.475, our estimate of this mean is hardly certain. The true mean is probably close to 1.475, but it could be 1.500, 1.525, or even 2.500. The more observations we have, the more confidence we would have that the true mean is close to the sample mean. In our parameter risk quantification procedure, fewer observations we have, the greater the parameter risk, and the greater the reserve uncertainty.

Similarly, the variation in our observations also affects our confidence in the sample mean. Suppose that instead of the five observations in our illustration, we had five observations of 1.200, 1.150, 1.450, 1.900, and 1.675. The sample mean is still 1.475, but now we have less confidence that the true mean is close to 1.475. We might think now that the true mean is probably between 1.200 and 1.700. Conversely, if our observations were 1.470, 1.475, 1.480, 1.473, and 1.477, we would have greater confidence that the true mean is about 1.475.

This is a simplistic explanation; the mathematically precise version is Bayesian estimation. The chance of obtaining five observations of 1.470, 1.475, 1.480, 1.473, and 1.477 from a distribution with a mean of 1.475 and a small variance is much greater than the chance of obtaining these same five observations from a distribution with a mean of 1.600 and a larger variance. If the five observations are 1.200, 1.150, 1.450, 1.900, and 1.675, the chance of obtaining these observations from a distribution with a mean of 1.475 is still greater than the chance of obtaining them from a distribution with a mean of 1.600, but it is no longer than much greater.

In Bayesian analysis, we are concerned not just with the mean and variance of our observations. Bayesian analysis looks at every individual observation. That is, we examine the likelihood of obtaining each observation from the universe of lognormal distributions.

Our prior expectations of the true mean of the distribution also affects the parameter risk. Suppose that we knew absolutely nothing about link ratios. We have no prior expectations at all. For all that we know, the true mean might lie anywhere from $-\infty$ to $+\infty$. The sample of five observations tells us something about the true mean, but we are not about to rule out any possibilities yet.

Suppose, however, that we are experienced reserving actuaries. We have a good feel for the expected link ratio in this development period for this book of business. Even before seeing any observations, we are certain that the true mean is between 1.000 and 2.000. From our reserving experience, we are fairly confident that the mean is between 1.400 and 1.600. Given the actual observations, we are much more confident that the true mean is about 1.475.

Let us return to lognormal distributions of link ratios. The intuition behind the Kreps estimation procedure for parameter risk does not depend on the type of distributions. However, the mathematics leading to Kreps's quantification equations shown in Appendix C assume a lognormal or a normal distribution of the variable which we are estimating.[67]

With our sample observations (the historical link ratios), we fit a lognormal curve and we determine the fitted parameters μ and σ. Because we have only a limited number of observations for each development period (between 5 and 25), there is significant parameter risk; that is, our fitted μ and σ parameters may not be the parameters of the true distribution. We turn to Bayesian analysis. We take the universe of lognormal distributions, and we say: "For each member of this universe of lognormal distributions, but is the chance that it would produce a sample like the one which we observe?" This is a standard likelihood question, and Kreps uses a negative loglikelihood test. Bayesian analysis allows us to invert this relationship and to say: "For the given sample of observations, what is the chance that the true distribution is any given member of the universe of lognormal distributions?"

Fitted Distribution and Predictive Distribution

To clarify what is happening, we must distinguish between the fitted distribution and the predictive distribution. Suppose that we had an infinite number of observations, so there is no error stemming from small sample size. That is to say, if all the

[67]The equations in Appendix C are for a lognormal distribution. The equations for a normal distribution are similar.

observations come from the same distribution, then the mean of the sample is almost certainly the mean of the distribution.[68]

We use the sample to fit the lognormal distribution. There is no parameter uncertainty here (or, more accurately, the parameter uncertainty is 0%), so we use the fitted distribution to generate additional values for our stochastic simulation. In this case, the fitted distribution is also the predictive distribution.

Suppose instead that we have a finite sample. Once again, we fit a lognormal distribution. Our fitted lognormal may be the exact same distribution that we fit with the infinite sample. With the finite sample, though, there is parameter risk. That is, we are not certain that the parameters of the fitted curve are indeed the parameters of the true distribution.

In this case, we do not generate future realizations from the fitted curve. The fitted curve is the most likely true distribution, but it is not the only possible true distribution. In fact, with continuous parameters, as is true in the illustrations in this paper, there are an infinite number of potential distributions.

Think of our Bayesian analysis as telling us the chance that each possible lognormal distribution is the true distribution. That is, the Bayesian analysis gives us a distribution of lognormal curves. Think of our simulations as a two stage process. First we simulate from this distribution of lognormal curves to get the particular curve that we will use. We then simulate from this lognormal distribution to get a future observation.

The "two stage process" was simply a manner of speaking; we do not actually simulate in two stages. We are simulating

[68]"Almost certainly" means with 100% confidence. This is not the same as "definitely." Statistically, we can be 100% sure that the mean of the sample is the mean of the distribution, yet the two means can certainly be different, even widely different. As a heuristic example, suppose that the distribution is all integers between 1 and 10. The mean of this distribution is 5.5. The probability of an observation being greater than 5 is 50% . It is clearly possible for every observation to be greater than 5, though the probability of an infinite stream of such observations is 50% to the infinite power, or 0%. This is an example where the mean of an infinite sample differs from the mean of the distribution, though the probability of this happening is 0%.

in a single stage, but we are not simulating from a lognormal distribution. We are simulating from another distribution, from a distribution with more parameters than a lognormal has.[69] This is the predictive distribution, which is used to generate future observations.

What is this distribution from which are simulating, this predictive distribution? There is a particular distribution, though it depends not only on the historical observations and the assumption that they are members of a lognormal distribution, but also on the Bayesian prior that we use in the analysis. We could consider this question empirically, as a heuristic exercise; we can't actually do this in practice. That is, we simulate several thousand, or several million, observations, and we examine the new sample to determine what distribution it comes from.

This method is good for thought experiments only; it is not feasible. Instead, Kreps shows the analytic solution: the maximum likelihoods, the Bayesian analysis, the negative loglikelihood procedure, and the formulation of the predictive distribution. One might think: "The result must be awfully complex." Yes, it is complex in the general case. But if we assume that the distribution is a normal or lognormal distribution, and if we make certain assumptions about the Bayesian prior, then the mathematics is tractable, and Kreps obtains simple equations for the simulation. These are the equations shown in Appendix C.

One view sometimes heard on this subject runs as follows: "We know that our observations come from a lognormal distribution; this is the assumption underlying the whole procedure. We are not certain about the parameters of this lognormal distribution because of the small sample size of our historical observations. This is the source of the parameter risk. This parameter risk concerns the values of the parameters of the lognormal dis-

---

[69]The number of "parameters" of this distribution depends on our prior assumptions about the universe of lognormal curves, or our "Bayesian prior"; we get to this in a moment.

tribution; it is not a question of what type of distribution the observations come from. The predictive distribution may not be the same as the fitted distribution, but it still must be a lognormal distribution."

This argument is specious. The predictive distribution is not a lognormal; in fact, it is not even a two parameter distribution. What kind of distribution is it? That depends on the Bayesian prior that we use in the analysis.

Bayesian Priors

We have made several references already to Bayesian priors; it's time that we defined what we're talking about. Suppose that we knew that the link ratios come from a lognormal distributions, but that we have no prior information at all about what type of lognormal distribution it is. That is to say, we know that the link ratios come from a lognormal distribution with parameters μ and σ, but we have no assumptions about what μ and σ might be. Mathematically, we say that our prior assumption about the distribution of the μ parameter is that it is uniform over all numbers. It is just as likely that it equals 1 as that it equals 100 or one million. The σ parameter must be positive, but that is the only assumption that we make, so the prior distribution is uniform over all positive numbers. In statistical jargon, we say that we have a diffuse prior. Think of this as our having no prior assumptions about the universe of lognormal distributions; every one is just as reasonable as another.

Could we use this diffuse universe of lognormal distributions as our predictive distribution? That is, if we have no observations at all, could we use this diffuse universe of lognormals? Of course not. All we know is that the desired numbers come from a lognormal distributions, but this could be any lognormal distribution at all. The predictive distribution is so diffuse that it has infinite variance. The simulations will not converge, no matter how many simulations we use.

The preceding statement warrants further explanation, since this is a problem even for simulations which do converge. Suppose that we have no observations, and we have no prior assumptions, so we simulate from the diffuse universe of lognormals. Think of this in the two stage process: we first pick parameters μ and σ by choosing a real number for μ and a positive number for σ. We have set no bounds for these numbers; they could be anything. We then simulate a realization from this lognormal; this realization is unbounded. No matter how many realization we use, the expected mean of our realizations is unbounded.

If we have some observations, the Bayesian analysis makes our posterior universe of lognormal distributions less diffuse. If our five observations are 1.400, 1.450, 1.600, 1.425, and 1.500, then it is much more likely that the true lognormal distribution has a mean of 1.475 than that it is has a mean of 10 or of 100.

Parameters for the Bayesian Prior

In practice, a completely diffuse Bayesian prior is often unworkable; moreover, it sometimes fails to make sense even in theory. To clarify the procedure used in this paper, we must examine the method of choosing the Bayesian prior in the Kreps procedure. Kreps determines μ_0 and σ_0 from the observations, and he calculates a negative loglikelihood from these values for a lognormal with parameters μ and σ (equation 2.25 on page 558). To simplify the analysis, he rescales the problem by defining normalized variables v and y such that:

$$\mu = \mu_0 + v\sigma_0$$

and

$$\sigma = y\sigma_0.$$

The Bayesian prior for the distribution of μ and σ can be restated as a prior assumption for the distribution of v and y.

Kreps [42, pp. 559–560]:

> We take a Bayesian approach and use diffuse prior distributions for v and y. Since v runs along the full axis from minus infinity to plus infinity, the prior used is just 1. Since y runs along the semi-axis, the suggested prior is proportional to $1/y^{\theta}$ where θ is either 0 or 1, depending on one's preference. The choice $\theta = 1$ emphasizes small values of y and corresponds to the assumption that the prior distribution of $\ln(y)$ is flat; the choice $\theta = 0$ assumes that the prior distribution of y is flat. Venter has emphasized that any choice of prior has strong implications. Ideally, the nature of the data being fitted would give some clues as to proper priors.

The comment by Venter referred to above is that "on a semi-axis a flat prior corresponds to assuming that it is as likely for the variable to lie between a million and a million and one as it is for the variable to lie between zero and one, and that it is infinitely more likely to be excess of any finite amount than to be less than that amount" (Kreps [42, footnote 7]).

Even with a θ of 1, our simulations produced unreasonable results. The text of our paper explains what we mean by "unreasonable." After much discussion with Dr. Kreps, we used a θ of 2. Dr. Kreps sums up the theory as follows (private communication):

> On pages 83–74 of section 3.2.2 of *Statistical Decision Theory and Bayesian Analysis*, second edition, by James O. Berger (Springer, 1980), there is the section "Noninformative Priors for Location and Scale Problems" which outlines the arguments and the problems with the Bayesian priors. The crude result is that for a location parameter, the density is 1 and for a scale parameter it is $1/\theta$. Berger goes on to talk about the Jeffreys results in the next section, which in the case

> of normals reduce to powers of sigma. Which power depends on what you like, but the choice theta = 2 is actually the computational Jeffreys result even if Jeffreys himself prefers theta = 1. So you take your choice; personally I think we always know something about the data and a noninformative prior is something like laziness on our part.

For workers compensation paid loss link ratios, we know a great deal about the data. Simply picking a value of θ is indeed laziness. The problem, however, is two-fold. First, we have great difficulty conceptualizing what any value of θ means for the universe of lognormals as potentials distributions for paid loss link ratios. Yes, we can state the mathematics, but we have difficulty visualizing whether a $\theta = 2$ is more reasonable than a $\theta = 5$ or vice versa. Second, if we use other ways of stating our prior assumptions, we can't work these assumptions into Kreps's equations.

Our final choice is summarized in the text of the paper. We chose a θ of 2, to ensure as diffuse a Bayesian as practicable for our application, and we discarded the extreme realizations with means more than 50 standard deviations away from the overall average. This may not be the ideal procedure, but we do not even know if it is too conservative or too liberal.

The Kreps parameter risk estimation procedure had one additional effect on our method. We noted above that the variance of our predictive distribution depends on both the Bayesian prior and the number of observations ("n"). Kreps discusses this problem in terms of the variance of z_{eff}, where z_{eff} is the effective deviate of ln(x), where x is the variable which we are simulating. Kreps shows that for $n + \theta \leq 4$, the variance of z_{eff} is infinite, and he notes that "this formula also tempts one to choose $\theta = 5$ so that var(z_{eff}) = 1 for all n" (page 564). Similarly, in discussing the standard deviation of the underlying distribution, Kreps says:

"The standard deviation does not exist if $n + \theta \leq 4$, but goes to zero as the sample size increases" (page 561).

This is the problem of convergence discussed earlier. Kreps [42, p. 561] says:

> In simulation situations if the underlying distribution does not have a finite variance then the mean of the simulation will not converge, because the mean of the simulation itself will have an infinite standard deviation. In practice, this shows up as occasional large jumps in the mean, even with millions of simulations (in fact, no matter how many simulations are done).

We choose $\theta = 2$. We deal with the variance problem by using only 20 columns of age-to-age link ratios, so that we always have a sufficient number of observations. For development beyond the 21st year, we use the inverse power curve tail factor approximation.

Conclusion

Neither the Kreps paper nor this paper is the definitive word on parameter risk. Even with the Kreps procedure, the analyst must choose a Bayesian prior based upon his or her own reserving knowledge and prejudices. Nevertheless, the thrust or the Kreps paper is that parameter risk is a significant source of reserve uncertainty. Our analysis illustrates this uncertainty, though we do not even pretend to have authoritatively measured it. However, by choosing a relatively diffuse Bayesian prior, and by discarding only those realizations that were extremely far from the sample mean, we have presumably erred on the side of caution, by overestimating the parameter risk.

A SYSTEMATIC RELATIONSHIP BETWEEN MINIMUM BIAS AND GENERALIZED LINEAR MODELS

STEPHEN MILDENHALL

Abstract

The minimum bias method is a natural tool to use in parameterizing classification ratemaking plans. Such plans build rates for a large, heterogeneous group of insureds using arithmetic operations to combine a small set of parameters in many different ways. Since the arithmetic structure of a class plan is usually not wholly appropriate, rates for some individual classification cells may be biased. Classification ratemaking therefore requires measures of bias, and minimum bias is a natural objective to use when determining rates.

This paper introduces a family of linear bias measures and shows how classification rates with minimum (zero) linear bias for each class are the same as those obtained by solving a related generalized linear model using maximum likelihood. The examples considered include the standard additive and multiplicative models used by the Insurance Services Office (ISO) for private passenger auto ratemaking and general liability ratemaking (see ISO [11] *and Graves and Castillo* [8]*, respectively).*

Knowing how to associate a generalized linear model with a linear bias function is useful for several reasons. It makes the underlying statistical assumptions explicit so the user can judge their appropriateness for a given application. It provides an alternative method to solve for the model parameters, which is computationally more efficient than using the minimum bias iterative method. In fact not all linear bias functions allow an iterative solution; in these cases, solving a generalized linear model using maximum likelihood provides an ef-

fective way to determine model parameters. Finally, it opens up the possibility of using statistical techniques for parameter estimates, analysis of residuals and model fit, significance of effects, and comparison of different models.

ACKNOWLEDGEMENT

I would like to thank Bill Emmons, John Huddleston, and particularly David Bassi and Thom McDaniel for their help and support while I was writing this paper. I am very grateful to David Fennell and Sue Groshong for patiently explaining numerous statistical concepts to me. Debra McClenahan gave me the final impetus to finish the paper. Finally, the comments from the three reviewers, Stuart Klugman, Steve Groeschen and Peter Wu were invaluable, and they led to significant changes from the first draft.

1. INTRODUCTION

History and Background

Bailey and Simon [2, 3], first considered bias in classification ratemaking and introduced minimum bias models. Since classification plans use fewer variables than underwriting cells and impose an arithmetic structure on the data, fitted rates in some cells may be biased, that is, not equal to the expected rate. Bias is a feature of the structure of the classification plan and not a result of a small overall sample size; bias could still exist even if there were sufficient data for all the cells to be individually credible. Of course, in such a situation an actuary would not use a classification plan.

Bailey and Simon [3] proposed their famous list of four criteria for an acceptable set of relativities:

BaS1: It should reproduce experience for each class and overall (balanced for each class and overall).

BaS2: It should reflect the relative credibility of the various groups.

BaS3: It should provide the minimum amount of departure from the raw data for the maximum number of people.

BaS4: It should produce a rate for each sub-group of risks which is close enough to the experience so that the differences could reasonably be caused by *chance*.

Condition BaS1 means that classification rates for each class should be balanced, that is, have zero bias. Obviously, zero bias by class implies zero bias overall.

Bailey points out that since more than one set of rates can be unbiased in the aggregate, it is necessary to have a method for comparing them. The average bias has already been set to zero, by criteria BaS1, and so it cannot be used. Bailey suggests the average absolute deviation and the chi-square statistic, particularly if cells are large enough to assume normality. He mentions that neither of these statistics has a known theoretical distribution and stresses that they should be used for comparison between models and not for tests of significance. This paper shows there is a natural correspondence between linear bias functions and generalized linear models. The theory of generalized linear models can then be used to define and analyze various measures of fit statistically, improving upon Bailey's more ad hoc methods.

In 1988, Brown [5] revisited minimum bias. His approach was to replace the bias function with an expression from the likelihood function and then solve for parameters to maximize its value. By assuming a distribution for the underlying quantity being modeled, he converts the problem to "an exercise in statistical modeling." This paper takes the opposite approach and goes *from* a particular class of bias functions *to* a statistical distribution. Brown also comments that "[t]o this point we have not been able to use GLIM [generalized linear models] to reproduce results obtained by Bailey's additive model"; see Section 4 below for such a reconciliation.

Venter's review [26] of Brown considers four alternatives to Bailey's methods:

V1: Alternatives to the balance principle.

V2: More general arithmetic functions to determine classification rates.

V3: Allow individual cells to vary from an arithmetically defined base.

V4: Do not use an arithmetic function to determine classification rates.

Venter comments that Brown's paper is mainly concerned with V1. This paper is largely concerned with V1 and V2, but also has comments on V3 and V4. Link functions, introduced below, allow more general arithmetic functions. The Box–Cox transformation, which Venter mentions, is an example of a link function. Section 10 mentions a method related to mixed models which is exactly what Venter proposed in V3 to determine unbiased rates.

Venter also comments that "the connection with general linear models does not seem to be the primary emphasis of [Brown's] paper." This paper builds on Brown's initial work by focusing on the connection between the minimum bias methods and generalized linear models, and by providing a more in-depth explanation of generalized linear models based on ideas already familiar to actuaries. Showing how they provide a unified treatment of minimum bias models will give actuaries another reason to learn more about generalized linear models. Other actuarial applications of generalized linear models have been proposed in McCullagh and Nelder [17], Renshaw [23], Haberman and Renshaw [9], and Wright [27].

Contents

Section 2 recalls some familiar material about linear models and sets up the progression from general linear models to generalized linear models by analyzing the three components of a general linear model.

Section 3 explains the non-uniqueness of solutions to a classification plan and how to get around the problem.

Section 4 explains the elementary, but unfamiliar, relationship between the cross classification ratemaking notation used in minimum bias models and the standard statistical, matrix notation used in linear models. It derives a matrix version of Bailey's minimum bias equations, and shows how Bailey's additive model is a simple linear model. The section ends with a general matrix formulation of balance and introduces a numerical example.

Section 5 introduces a family of linear bias functions and an associated measure of model fit called deviance, both related to a variance function. By construction, minimum linear bias corresponds to the minimum deviance best-fit model. It also shows how, in some cases, the minimum bias solution can be obtained using iterative equations.

Section 6 defines the exponential family of distributions and gives several examples. It explains the relationship between variance functions and distributions, which is then used to convert the minimum bias models of Section 5 into fully defined statistical models.

Section 7 introduces generalized linear models and their connection with minimum linear bias. This correspondence holds regardless of whether an iterative method can be used to solve the minimum bias problem, so generalized linear models extend the existing family of models. A detailed set of examples, comparing different linear bias assumptions, is also given.

Section 8 discusses measures of model fit associated with generalized linear models. Fit is discussed at several different levels, ranging from selection of covariates to selection of link functions and variance functions.

Section 9 is concerned with numerical computations. It explains how and when the iterative equations obtained using Bailey's minimum bias equations converge. It also discusses how to solve generalized linear models using iteratively re-weighted least squares. Appendix B gives SAS computer code illustrating a hands-on example of this approach.

Section 10 gives some suggestions for future work. It touches on some recent work of Lee and Nelder [19] on mixed models and hierarchical generalized linear models, which can be regarded as an extension of the work in this paper and which provides unbiased predictors for all cells.

The theory is illustrated throughout with simple examples the reader can reproduce.

In the first seven sections of the paper, most concepts are developed from first principles and very little background in statistics is assumed. Sections 8 and 9 make greater demands on the reader, assuming more statistical and mathematical background, respectively.

Notation

Random variables will be denoted by capitals and realized values in lower case. Vectors will be denoted by bold lower case letters. Matrices will be denoted by bold upper case letters, typically **A**, **B**, **X** and **W**. The (ij)th element of a matrix **X** will be denoted x_{ij}, $x_{i,j}$ or $\mathbf{X}(i,j)$. Some matrices will be given in block form. If **W** is a block matrix, then $\mathbf{W}_{ij}$ will denote the block in the (i,j)th place. Superscript t denotes transpose. Random observations are denoted r, r_i, r_{ij}; Greek letters typically refer to model parameters or fitted values. Matrix dimensions are denoted $m \times n$.

2. LINEAR MODELS

A statistical model is defined by specifying a probability distribution for the quantity being modeled. Fitted values, predicted by the model, can then be determined from the relevant probability distribution, usually as the mean. The goal of using a model is to replace the data, which may have many thousands of observations, with a far smaller set of parameters without losing too much information. A good model helps the actuary better understand the data and make reasonable predictions from it. Models can be designed to facilitate the construction of classification ratemaking tables.

In a basic linear model the fitted values are linear combinations of the model parameters. Examples of linear models include analyses of variance (ANOVA), linear regression, and general linear regression.

In order to find model parameter values, it is necessary to select an objective function. The objective function can measure the deviance between the underlying data and the fitted values for different parameter choices, or it can be based on other criteria such as minimum variance amongst unbiased estimators. Least squares and maximum likelihood are two common examples of the former type of objective function. A single statistical model can give rise to different parameter solutions depending upon the objective function used. Therefore it is necessary to include the objective function in an effective description of the model.

The input data for all models considered here can be given as a two-dimensional array. The rows correspond to the different observations or units. The first column corresponds to the response variate which can be continuous (such as pure premium, frequency or severity) or discrete (such as claim count).

The remaining columns correspond to the explanatory variates, or covariates, whose values are supposed to explain the values of the response. Covariates can be qualitative or quantitative. A qualitative covariate, called a factor, takes on non-numerical values called levels, such as vehicle-use, vehicle type or sex. Quantitative covariates have numeric values. Examples include age, time, weight of vehicle, or price of vehicle. Age group is a qualitative covariate. If the covariates are all factors, then the rows of the input can be labeled by the levels of the factors (as in Example 2.1 below). Classification ratemaking naturally uses these coordinates. However, they are generally not used if some of the some covariates are continuous, as in Example 2.2.

EXAMPLE 2.1 A two-way analysis of variance with no interactions assumes each observation r_{ij} is a realization of an independent, normally distributed random variable with mean $a_i + b_j$ and variance σ^2. Parameters are selected using either maximum likelihood, minimum square error, or minimum variance amongst unbiased estimators; the three are equivalent for this model. The a_i and b_j are the effects corresponding to the different levels of the two factors (classification variables). In texts on linear models this example is often presented in the equivalent form $r_{ij} = a_i + b_j + e_{ij}$, where the errors e_{ij} are independent, normally distributed random variables with mean 0 and variance σ^2. For example, r_{ij} could be the observed pure premium in cell i,j of an auto classification plan, with a_i the factor for age of operator group i and b_j the factor for vehicle use group j. If r_{ij} is the average of w_{ij} exposures, then it is a realization of a variable with variance σ^2/w_{ij} and w_{ij} is called the weight of the i,jth cell.

EXAMPLE 2.2 A linear regression model assumes each observation r_i is a realization of an independent, normally distributed random variable with mean $a + bx_i$ and variance σ^2. There is a single continuous covariate whose values are given by x_i. The same three objectives can be used to solve for a and b. The model can also be written $r_i = a + bx_i + e_i$, where e_i are independent, normally distributed random variables with mean 0 and variance σ^2. Actuaries use linear regression to compute trends, in which case r_i is the observed pure premium, or log of pure premium, at time i and $x_i = i$.

The input data for a linear model can be compactly described using vectors and matrices. Suppose there are n observations. The responses can be put into an $n \times 1$ column vector $\mathbf{r} = (r_1, \ldots, r_n)^t$. The covariates can be arranged into a design matrix $\mathbf{X}$ which has one row for each observation and one column for each parameter of the model. Let p be the number of parameters and let $\mathbf{x}_i$ be the ith row of $\mathbf{X}$, so $\mathbf{x}_i$ is a $1 \times p$ row vector. If all the covariates are factors, then the design matrix has one column for each level of each factor and consists of 0's and 1's. In Example 2.1, if there are three age groups and three vehicle use classes, then the design matrix would have six columns. In Example 2.2, $\mathbf{X}$ has two columns, corresponding to a and b. The first column is all 1's, corresponding to the constant term; the second is given by $(x_1, \ldots, x_n)^t$.

The parameters of a linear model can be arranged into a $p \times 1$ column vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^t$. Finally, let $\mu_i = \mathrm{E}(R_i)$ be the fitted value of the ith response and let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^t$. A general linear model, which includes both analysis of variance and linear regression as special cases, assumes

$$\mathbf{r} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \qquad \boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}, \tag{2.1}$$

where the error term $\mathbf{e} = (e_1, \ldots, e_n)^t$ has e_i independent, normally distributed with mean 0 and variance σ^2. Thus R_i is assumed to be independent, normally distributed with mean $\mu_i = \mathbf{x}_i\boldsymbol{\beta}$ and variance σ^2.

Three important assumptions underlie a general linear model:

1. Constant variance: the σ^2 term does not vary between different responses. When the ith response is an average

of w_i individual responses, each with variance σ^2, then the variance is σ^2/w_i, and again σ^2 does not vary between observations. The w_i are prior weights.

2. Normality of errors: the errors e_i are independent, identically distributed normal random variables.

3. Linear: the fitted value $\mu_i = \mathbf{x}_i\boldsymbol{\beta} = \sum_j \mathbf{x}_{ij}\beta_j$ is a linear combination of the parameters, so the systematic effects are additive.

In actuarial work it is common that the responses are averages from populations with different sizes. In Example 2.1, there are typically more exposures in the mature operator classes than in youthful and senior operator classes. General linear models allow for such differences in variance by using prior weights which vary by observation—as in assumption (1) above.

The second assumption, normal errors, is frequently a problem in actuarial applications. Losses, severities, pure premiums and frequencies are all positive and generally positively skewed; they are therefore not normally distributed. The log transformation is often applied to the data prior to using a linear model in order to improve normality. The log transformation is also applied in order to convert multiplicative effects into additive ones.

EXAMPLE 2.3 Example 2.1 modeled R_{ij} as normally distributed with mean $a_i + b_j$ and variance σ^2/w_{ij}, where w_{ij} is the number of exposures in the i,jth cell. Applying the log transformation to the response, we can consider the same model for $\log(R_{ij})$. On the untransformed scale, the model for R_{ij} is lognormal with parameters $a_i + b_j$ and σ^2/w_{ij} (see the Appendix to Hogg and Klugman [10] or Appendix A of Klugman, Panjer and Willmot [16]). The systematic effects are now multiplicative. Also $\mathrm{E}(R_{ij}) = \exp(a_i + b_j + \sigma^2/2w_{ij})$, and the variance depends on the fitted mean because $\mathrm{Var}(R_{ij}) = (\mathrm{E}(R_{ij}))^2(\exp(\sigma^2/2w_{ij}) - 1)$.

Generalized Linear Models

In order to set up *generalized* linear models, consider a general linear model as split into three components:

GLM1: A random component: observations r_i are assumed to come from an independent normal distribution R_i with $\mathrm{E}(R_i) = \mu_i$.

GLM2: A systematic component: the covariates $\mathbf{x}_i = (x_{i1},\ldots, x_{ip})^t$ give a linear predictor

$$\eta_i = \sum_j x_{ij}\beta_j.$$

GLM3: A link between the random and systematic components:

$$\eta = \mu.$$

The parameters are selected using the maximum likelihood objective.

A generalized linear model allows extensions to GLM1 and GLM3. GLM2 is retained since the model is still assumed to be *linear*.

Assumption GLM1 is generalized to allow the R_i to have a distribution from the exponential family, defined in Section 6. The exponential family includes the normal, Poisson, binomial, gamma and inverse Gaussian distributions. The lognormal distribution is not a member of the exponential family. The recent book by Jørgensen [13] is a good reference on exponential distributions.

In GLM3, the identity link $\eta_i = \mu_i$ between the random and systematic components is generalized to allow $\eta_i = g(\mu_i)$ for any strictly monotonic, differentiable function g. Three common choices are $g(x) = x$, $g(x) = \log(x)$ and $g(x) = 1/x$. The log-link has been discussed above. The reciprocal-link can be understood as representing rates: premium is the dollar rate per year; the reciprocal of premium is therefore years of coverage per dollar of premium. While not something that has been tried to date in actuarial applications, there is no reason why the systematic effects should not be additive on the reciprocal scale. McCullagh and Nelder [17, Section 8.4] gives an insurance example.

In a general linear model, scale transformations may be applied to the responses prior to fitting in order to increase the validity of GLM1-3. However, the three assumptions may be mutually incompatible and so the question of an appropriate scale can be very problematic (see [17, Section 2.1] for an example). For a generalized linear model, normality and constant variance are no longer required. The choice of link-function (scale) is therefore driven solely by the need to ensure additivity of effects. Since transformations in generalized linear models are used to achieve one end, rather than three in a general linear model, there is more flexibility in the modeling process.

EXAMPLE 2.4 The next three items illustrate how generalized linear models include, extend, and differ from general linear models:

(a) A generalized linear model with identity link function and normal errors is a general linear model.

(b) A generalized linear model version of Example 2.1 with gamma error distribution and a reciprocal link, would model R_{ij} as an independent gamma random variable with $\mathrm{E}(R_{ij}) = \mu_{ij} = 1/(a_i + b_j)$ and $\mathrm{Var}(R_{ij}) = \mu_{ij}^2\phi/w_{ij}$. The constant ϕ acts like σ in Example 2.1.

(c) A generalized linear model with log link and normal errors is *not* the same as applying a general linear model to the log responses. The generalized linear model assumes R_{ij} is *normally* distributed with mean $\exp(a_i + b_j)$ and variance σ^2/w_{ij}. The general linear model applied to the log transformed data (Example 2.3) assumes that R_{ij} is *lognormally* distributed and that $\log(R_{ij})$ has mean $a_i + b_j$ and variance σ^2/w_{ij}. In the generalized linear model the log-link is only trying to achieve additivity of effects; the error distribution is specified separately. Exhibit 8, described fully in Section 7, shows the differences between these models applied to an example dataset.

3. UNIQUENESS OF PARAMETERS

Going back to Example 2.1, it is clear that the parameters of a linear model need not be unique. If $\mu_{ij} = a_i + b_j$, then

$$\mu_{ij} = (a_i + \alpha) + (b_j - \alpha) \tag{3.1}$$

for all constants α. Similarly, if $\mu_{ij} = a_i b_j$ then $\mu_{ij} = (\alpha a_i)(b_j/\alpha)$ for all constants $\alpha \neq 0$. If the model is $\mu_{ijk} = a_i + b_j + c_k$, then the situation is even worse: there are two degrees of freedom because $\mu_{ijk} = (a_i + \alpha_1 + \alpha_2) + (b_j - \alpha_1) + (c_k - \alpha_2)$ for all α_1 and α_2. In general, it is easy to see there are $q - 1$ degrees of freedom when there are q classification variables. Therefore it is necessary to select $q - 1$ base classes in order to have unique parameters. This is familiar from setting up rate classification plans. For example, the personal auto plan has one base rate for the married, aged 25–50, pleasure-use, single standard vehicle, zero-points class, and deviations for all other classes.

There is no canonical method for selecting the base classes needed to ensure unique parameters. Here is one possible approach. First select one classification *cell* as a base. Then, select one classification *variable* which will not have a base. Finally, set the parameters corresponding to the base class in all the other classification variables to zero (additive models) or one (multiplicative models). This specifies the values of $q - 1$ parameters and so removes all degrees of freedom. Now the parameters for all the non-base classification variables are deviations from the base class for that variable. Picking different base classes leads to different parameters, but the fitted values remain the same.[1]

In Example 2.1, we could select mature drivers and pleasure-use as the base cell, and age as the base classification. This forces pleasure-use to be the base class in the vehicle-use classification, and so the parameter for pleasure-use is set to zero. Since b_1 corresponds to pleasure use, this choice is the same as selecting $\alpha = b_1$ in Equation 3.1.

In conclusion, a linear model or minimum bias method which uses all the available parameters will generally not have a unique solution. However, the non-uniqueness is of a trivial nature and the fitted values will be unique. After making an arbitrary selection of base classes, the remaining parameters will be unique. This is what Bailey and Simon [3] mean when they say "[the parameters] can only be regarded in relationship to the coordinate system in which they find themselves."

4. MATRIX FORMULATION

As Venter [26] noted, it is not clear to those unfamiliar with linear models how they are related to minimum bias methods. Moreover, the translation from statistical linear models to minimum bias methods is hampered by different uses of the same notation. We will follow Brown's notation as much as possible, since actuaries are probably most familiar with his approach. This section explains the relationship between linear models and minimum bias methods and provides a dictionary to translate between the two. In order to keep difficulties of notation in the

[1] Selecting base classes corresponds to deleting columns from the design matrix. Selecting $q-1$ base classes ensures that the resulting design matrix $\hat{X}$ has maximal rank. This in turn implies $\hat{X}'\hat{X}$ is invertible and so the normal equations can be solved uniquely for the remaining parameters. In general linear models, non-uniqueness is handled by computing the generalized inverse of $X'X$. The generalized inverses can be regarded as a method for picking base classes. See Rao [22, Chapter 1b.5], for more details.

background, we consider only a simple additive model with two variables. Extensions to more general models are easy to work out—indeed the point of this section is to convince the reader they will work out just as expected. The auto classification plan will be used to provide examples.

Minimum Bias Method Language

The generic minimum bias method attempts to explain a collection of observed values r_{ij} with two sets of parameters x_i and y_j, $i = 1,\ldots,n_1$, $j = 1,\ldots,n_2$. For example, r_{ij} could be the pure premium in the (i,j)th cell, x_i may correspond to the ith age classification, and y_j to the jth vehicle use classification such as pleasure, drive to work, or business. Let w_{ij} denote the number of exposures in the (i,j)th cell. Minimum bias methods then give iterative equations to solve for the x_i's and y_j's.

For example, Bailey's additive method models r_{ij} as $x_i + y_j$ (hence the appellation "additive") in such a way that, for all i,

$$\sum_j w_{ij}(r_{ij} - (x_i + y_j)) = 0, \tag{4.1}$$

and similarly for j. Equation 4.1 means that the model is balanced (i.e., has zero weighted bias) for each class i and in total (summing over i), and so minimizes bias. Rearranging Equation 4.1 gives the familiar form of Bailey's additive method:

$$x_i = \sum_j w_{ij}(r_{ij} - y_j) \Big/ \sum_j w_{ij}, \tag{4.2}$$

and similarly

$$y_j = \sum_i w_{ij}(r_{ij} - x_i) \Big/ \sum_i w_{ij}. \tag{4.3}$$

This notation is shorthand for an iterative procedure, where the transition from the lth to $l+1$st iteration is

$$x_i^{(l+1)} = \sum_j w_{ij}(r_{ij} - y_j^{(l)}) \bigg/ \sum_j w_{ij},$$

and similarly for $y_j^{(l+1)}$ in terms of $x_i^{(l+1)}$. The final result of the iterative procedure is given by $x_i = \lim_{l\to\infty} x_i^{(l)}$, and similarly for y.

Translation

The key to translating from minimum bias notation to linear model notation is how the observations are indexed. In linear models they are indexed by one parameter, whereas in the minimum bias method they are indexed by two parameters (or more generally, by the number of classification variables). The translation is described by the following correspondences. In all cases the left hand side gives the minimum bias notation and the right hand side the linear model notation. Also, in this section commas are inserted between subscript indices for clarity. The difference in how observations are indexed is illustrated by the following two correspondences between $n_1 n_2 \times 1$ column vectors:

$$\begin{pmatrix} r_{1,1} \\ r_{1,2} \\ \vdots \\ r_{1,n_2} \\ r_{2,1} \\ \vdots \\ r_{n_1,1} \\ \vdots \\ r_{n_1,n_2} \end{pmatrix} \leftrightarrow \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{n_2} \\ r_{n_2+1} \\ \vdots \\ r_{(n_1-1)n_2+1} \\ \vdots \\ r_{n_1 n_2} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} w_{1,1} \\ w_{1,2} \\ \vdots \\ w_{1,n_2} \\ w_{2,1} \\ \vdots \\ w_{n_1,1} \\ \vdots \\ w_{n_1,n_2} \end{pmatrix} \leftrightarrow \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{n_2} \\ w_{n_2+1} \\ \vdots \\ w_{(n_1-1)n_2+1} \\ \vdots \\ w_{n_1 n_2} \end{pmatrix}.$$

The different levels of the two classifications (or effects) correspond as

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{n_1} \\ y_1 \\ \vdots \\ y_{n_2} \end{pmatrix} \leftrightarrow \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_{n_1} \\ \beta_{n_1+1} \\ \vdots \\ \beta_{n_1+n_2} \end{pmatrix}. \tag{4.4}$$

Let $n = n_1 n_2$ be the number of observations, and $p = n_1 + n_2$ be the number of model parameters. Our translation assumes there are observations for each of the $n = n_1 n_2$ possible combinations of x_i and y_j. If necessary, the model can be brought into this form by using zero weights in any empty cells.

Linear Model Language

A statistical linear model attempts to explain a collection of observed values r_i using linear combinations of a smaller number of parameters. In our setting, the model explains pure premiums r_i, $i = 1,\ldots,n$, using linear combinations of parameters $\beta_1,\ldots,\beta_p$ given by

$$r_i = \sum_j x_{ij}\beta_j + e_i,$$

where e_i is a random error term. In matrix language this can be written

$$\mathbf{r} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e},$$

where $\mathbf{X} = (x_{ij})$ is the $n \times p$ design matrix of covariates, and $\mathbf{r} = (r_1,\ldots,r_n)^t$, $\mathbf{e} = (e_1,\ldots,e_n)^t$, and $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)^t$ are column vectors.

The design matrix corresponding to the two-variable additive linear model is the $n \times p$ matrix

$$\mathbf{X} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{I} \\ \vdots & \vdots \\ \mathbf{A}_{n_1} & \mathbf{I} \end{pmatrix}, \tag{4.5}$$

where

$$\mathbf{A}_i = \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}, \tag{4.6}$$

with dimension $n_2 \times n_1$, has zero entries except for 1's in the ith column, and $\mathbf{I}$ is the $n_2 \times n_2$ identity matrix. Using the block matrix form of $\mathbf{X}$, and the translation Equation 4.4, it is easy to see that

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{X} \begin{pmatrix} x_1 \\ \vdots \\ x_{n_1} \\ y_1 \\ \vdots \\ y_{n_2} \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_1 + y_{n_2} \\ x_2 + y_1 \\ \vdots \\ x_2 + y_{n_2} \\ \vdots \\ x_{n_1} + y_1 \\ \vdots \\ x_{n_1} + y_{n_2} \end{pmatrix},$$

dimensions $(n \times p)(p \times 1) = n \times 1$, demonstrating the translation between minimum bias notation and linear model notation.

Solution of Linear Models

It is well known that the maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ satisfies the following normal equations under the assumption of independent and identically distributed normal errors (see Rao [22, Section 4a.2])

$$\mathbf{X}^t\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^t\mathbf{r}. \tag{4.7}$$

If observation i has weight w_i, the solution satisfies

$$\mathbf{X}^t\mathbf{W}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^t\mathbf{W}\mathbf{r}, \tag{4.8}$$

where $\mathbf{W} = \operatorname{diag}(w_1,\ldots,w_n)$ is the diagonal matrix of weights.

Next we compute Equation 4.8, for the two-variable additive model using the definitions and translations introduced above. In minimum bias notation, the matrix of weights can be written as a block matrix

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{W}_{n_1} \end{pmatrix} \quad \text{dimension } n \times n, \tag{4.9}$$

where

$$\mathbf{W}_i = \begin{pmatrix} w_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{in_2} \end{pmatrix} \quad \text{dimension } n_2 \times n_2.$$

Using the block matrix form of $\mathbf{X}$ and $\mathbf{W}$, it is a simple computation to show

$$\mathbf{X}^t\mathbf{W}\mathbf{X} = \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C}^t & \mathbf{D} \end{pmatrix},$$

dimension $p \times p$, where **B**, **C** and **D** are given by

$$\mathbf{B} = \begin{pmatrix} \sum_j w_{1j} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sum_j w_{n_1 j} \end{pmatrix} \quad \text{dimension } n_1 \times n_1,$$

$$\mathbf{C} = \begin{pmatrix} w_{11} & \cdots & w_{1n_2} \\ \vdots & & \vdots \\ w_{n_1 1} & \cdots & w_{n_1 n_2} \end{pmatrix} \quad \text{dimension } n_1 \times n_2,$$

and

$$\mathbf{D} = \begin{pmatrix} \sum_i w_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sum_i w_{in_2} \end{pmatrix} \quad \text{dimension } n_2 \times n_2.$$

Therefore

$$\begin{aligned} \mathbf{X}^t\mathbf{W}\mathbf{X}\boldsymbol{\beta} &= \mathbf{X}^t\mathbf{W}\mathbf{X}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C}^t & \mathbf{D} \end{pmatrix}\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{Bx} + \mathbf{Cy} \\ \mathbf{C}^t\mathbf{x} + \mathbf{Dy} \end{pmatrix}, \end{aligned}$$

giving the $p \times 1$ vector equality

$$\mathbf{X}^t\mathbf{W}\mathbf{X}\boldsymbol{\beta} = \begin{pmatrix} x_1 \sum_j w_{1j} + \sum_j w_{1j} y_j \\ \vdots \\ x_{n_1} \sum_j w_{n_1 j} + \sum_j w_{n_1 j} y_j \\ \sum_i w_{i1} x_i + y_1 \sum_i w_{i1} \\ \vdots \\ \sum_i w_{in_2} x_i + y_{n_2} \sum_i w_{in_2} \end{pmatrix}. \qquad (4.10)$$

On the other hand,

$$\mathbf{X}^t\mathbf{W}\mathbf{r} = \begin{pmatrix} \sum_j w_{1j} r_{1j} \\ \vdots \\ \sum_j w_{n_1 j} r_{n_1 j} \\ \sum_i w_{i1} r_{i1} \\ \vdots \\ \sum_i w_{in_2} r_{in_2} \end{pmatrix} \quad \text{dimension } p \times 1. \qquad (4.11)$$

Equating corresponding rows of Equation 4.10 and Equation 4.11—the normal equations—gives exactly Equation 4.2 and Equation 4.3, respectively, demonstrating that the results of a two-effect additive general linear model are the same as the Bailey additive method.

This is a significant result for several reasons. First, it shows the minimum bias parameters are the same as the maximum likelihood parameters assuming normal errors, which the user may or may not regard as a reasonable assumption for his or her application. Second, it is much more efficient to solve the normal equations than perform the minimum bias iteration, which typically converges quite slowly (see Section 9). Third, knowing that the result is the same as a linear model allows the statistics developed to analyze linear models to be applied. For example, information about residuals and influence of outliers can be used.

General Theory and a Matrix Formulation of Balance

It is easy to generalize the preceding discussion to the case of a general linear model with q classification variables. Let the ith classification variable have n_i levels, $i = 1,\ldots,q$. Thus there are $p = n_1 + \cdots + n_q$ different parameters and, assuming no empty cells, $n = n_1 \cdots n_q$ observations.

The minimum bias notation associates an $n \times n_i$ design matrix $\mathbf{A}_i$ and an $n_i \times 1$ parameter vector $\mathbf{a}_i$ with the ith classification variable. The $n \times 1$ vector of modeled rates $\boldsymbol{\mu} = (\mu_{1,\ldots,1},\ldots,\mu_{n_1,\ldots,n_q})^t$ is

$$\boldsymbol{\mu} = (\mathbf{A}_1 \cdots \mathbf{A}_q) \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_q \end{pmatrix} = \mathbf{A}_1\mathbf{a}_1 + \cdots + \mathbf{A}_q\mathbf{a}_q. \qquad (4.12)$$

In linear model language, the design matrix $\mathbf{X}$ has dimension $n \times p$ and equals the horizontal concatenation $(\mathbf{A}_1 \cdots \mathbf{A}_q)$. The parameter vector $\boldsymbol{\beta}$ has dimension $p \times 1$ and equals $(\mathbf{a}_1, \ldots, \mathbf{a}_q)^t$, and Equation 4.12 becomes $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$.

Note that the linear model notation makes it possible to use two-dimensional matrix notation to describe models with any number of classification variables.

Using this notation and the same approach used to derive Equation 4.10 and Equation 4.11 shows that the normal equation condition

$$\mathbf{X}^t\mathbf{W}(\mathbf{r} - \boldsymbol{\mu}) = \mathbf{0} \tag{4.13}$$

is exactly a matrix formulation of condition BaS1—that relativities be balanced by class. This interpretation of Equation 4.13 is important and will be used repeatedly below.

To see why Equation 4.13 is the balance condition, first use the translation of Equation 4.2 to write it as:

$$\mathbf{X}^t\mathbf{W}(\mathbf{r} - \boldsymbol{\mu}) = \begin{pmatrix} \mathbf{A}_1^t \\ \vdots \\ \mathbf{A}_q^t \end{pmatrix} \mathbf{W}(\mathbf{r} - \boldsymbol{\mu})$$

$$= \begin{pmatrix} \mathbf{A}_1^t\mathbf{W}(\mathbf{r} - \boldsymbol{\mu}) \\ \vdots \\ \mathbf{A}_q^t\mathbf{W}(\mathbf{r} - \boldsymbol{\mu}) \end{pmatrix} = \mathbf{0}. \tag{4.14}$$

Consider balance over the first level of the first classification variable. By permuting columns of $\mathbf{X}$, this can be done without loss of generality. Similarly, by permuting the observations, assume that $\mathbf{A}_1$ has the form:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

the vertical concatenation of n_1 different matrices each with $n_2 \cdots n_q$ rows and n_1 columns and one column of ones. Then the first row of Equation 4.14 is given by the sum product of the first column $\mathbf{A}_1$ (i.e., the first *row* of $\mathbf{A}_1^t$) with $\mathbf{W}(\mathbf{r} - \boldsymbol{\mu})$, which gives

$$\sum_{j_2,\ldots j_q} w_{1,j_2,\ldots j_q}(r_{1,j_2,\ldots j_q} - \mu_{1,j_2,\ldots j_q}) = 0,$$

exactly the sum over all other classes required by the balance condition.

Numerical Example

We now introduce a numerical example which will be used throughout the paper to illustrate the theory. The data, shown in Exhibit 1, gives average claim severity for private passenger auto collision.[2] The severities have been adjusted for severity trend. There are $n = 32$ observations and two classification variables: age group and vehicle-use. Age group has eight levels and

[2]The data is derived from McCullagh and Nelder's example [17, Section 8.4].

vehicle-use four. The response variable r is average claim severity. The weights w are given by the number of claims underlying the average severity. Exhibit 2 gives the one-way weighted average severities.

Exhibit 3 gives the design matrix **A** corresponding to the age group classification. **A** has the block form shown in Equation 4.6. Exhibit 4 gives the design matrix **B** corresponding to the vehicle-use classification. Pleasure-use has been selected as the base (as in Section 3) and the corresponding column of the design matrix has been deleted; this accounts for the rows of zeros. The design matrix for the whole model is $\mathbf{X} = (\mathbf{A}\ \mathbf{B})$. Except for the deleted column in **B**, **X** has the form given in Equation 4.5.

Exhibit 5 uses the iterative method, Equation 4.2, to fit an additive minimum bias model to the data. There are 50 iterations shown (column 1). Column 2 shows the length of the change in the parameter vector from one iteration to the next. Columns 3–13 show how the parameters change with each iteration. Columns 14–17 will be explained in Section 9. Exhibit 6 shows the solution to the normal equations Equation 4.8. The resulting parameters are all within 2 cents of the values in the last row of Exhibit 5 as expected. Had more iterations been performed the results would have been closer.

This example will be continued in Sections 7, 8 and 9.

5. BIAS FUNCTIONS AND DEVIANCE FUNCTIONS

Bailey's first criterion for a set of classification relativities, that rates be balanced (unbiased) for each class and in total, makes it necessary for the actuary to be able to measure the bias in a set of rates. Bailey's third and fourth conditions, which require a minimum departure from the raw data and a departure that could be caused by chance, make it necessary to measure the deviance between the fitted rates and the data and to quantify its likelihood.

In the papers on minimum bias discussed in the Introduction, none of the authors differentiated between a measure of bias and a measure of deviance. A measure of bias should be proportional to the predicted value minus the observed value, and can be positive or negative. A measure of deviance, or model goodness of fit, should be like a distance: always positive with a minimum of zero for an exact fit (zero bias). Deviance need not be symmetric; we may care more about negatively biased estimates than positively biased ones or, vice versa.

This section will introduce three concepts: variance functions, linear bias functions and deviance functions, and then show how they are related. All three concepts have to do with specifying distributions—a key part of a statistical model. However, they are independent of the choice of covariates.

In this section r denotes the response, with individual units being r_i, or r_{ij} in the example. The fitted means are μ or μ_i.

Ordinary bias is the difference $r - \mu$ between an observation r and a fitted value μ. When adding the biases of many observations and fitted values, there are two reasons why it may be desirable to give more or less weight to different observations. First, if the observations come from cells with different numbers of exposures then their variances will be different. As explained in Section 2, this possibility is handled by using prior weights for each observation.

The second reason to weight the biases of individual observations differently is if the variance of the underlying distribution is a function of its mean (the fitted value). This is a very important departure from normal distribution models where the prior weights do not depend on the fitted values. In Example 2.1, r_{ij} is a sample from R_{ij} which is normally distributed with mean $\mu_{ij} = a_i + b_j$ and variance σ^2. The variance is independent of the mean. In Example 2.4(b), r_{ij} is a sample from R_{ij} which has a gamma distribution with mean μ_{ij} and variance $\phi\mu_{ij}^2$ (assuming all weights are 1). Now the variance of an individual observation is a function of the fitted cell mean μ_{ij}. Clearly, large biases from a cell with a large mean are more likely, and should be weighted less, than those from a cell with a small mean. In this situation we will use variance functions to give appropriate weights to each cell when adding biases. Once again, it is important to realize that variance functions are not a feature of normal distribution models and that they represent a substantial generalization.

A variance function, typically denoted V, is any strictly positive function of a single variable. Three examples of variance functions are $V(\mu) \equiv 1$ for $\mu \in (-\infty, \infty)$, $V(\mu) = \mu$ for $\mu \in (0, \infty)$, and $V(\mu) = \mu^2$ also for $\mu \in (0, \infty)$. It should not be a surprise that the first can arise from the normal distribution, and the last can arise from the gamma distribution.

Combining variance functions and prior weights—the two reasons to weight biases from individual cells differently—we define a linear bias function to be a function of the form

$$b(r;\mu) = \frac{w(r-\mu)}{V(\mu)},$$

where V is a variance function and w is a prior weight. The weight may vary between observations, but is not a function of the observation or of the fitted value.

In applications there would be many observations r_i, each with a fitted value μ_i and possibly different weights w_i. The total bias would then be

$$\sum_i b(r_i;\mu_i) = \sum_i \frac{w_i(r_i-\mu_i)}{V(\mu_i)}.$$

The functions $r-\mu$, $(r-\mu)/\mu$, and $(r-\mu)/\mu^2$ are three examples of linear bias functions, each with $w = 1$, corresponding to the variance functions given above.

A deviance function is some measure of the distance between an observation r and a fitted value μ. The deviance $d(r;\mu)$ should satisfy the following two conditions common to a distance:

Dev1: $d(r;r) = 0$ for all r, and

Dev2: $d(r;\mu) > 0$ for all $r \neq \mu$.

The weighted squared difference $d(r;\mu) = w(r-\mu)^2$, $w > 0$, is an example of a deviance function.

An important difference between bias and deviance is that deviance, which corresponds to distance, is always positive while bias can be positive or negative. Deviance can be regarded as a value judgment: "how concerned am I that r is this far from μ?" Deviance functions need not be symmetric about $r = \mu$.

It is possible to associate a deviance function with a linear bias function by defining

$$d(r;\mu) = 2w \int_{\mu}^{r} \frac{(r-t)}{V(t)}\, dt. \tag{5.1}$$

Clearly this definition satisfies Dev1 and Dev2. Note that by the Fundamental Theorem of Calculus,

$$\frac{\partial d}{\partial \mu} = -2w\frac{(r-\mu)}{V(\mu)}.$$

Examples of Deviance Functions

(a) If $b(r;\mu) = r - \mu$ is ordinary bias, then

$$d(r;\mu) = 2\int_{\mu}^{r}(r-t)\,dt = (r-\mu)^2$$

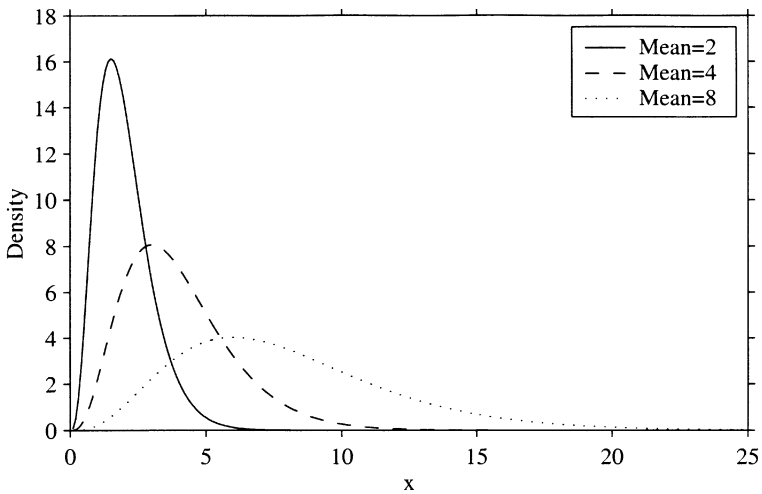
is the squared distance deviance, with weight $w = 1$.

(b) If $b(r;\mu) = (r-\mu)/\mu^2$ corresponds to $V(\mu) = \mu^2$ for $\mu \in (0,\infty)$, then

$$\begin{aligned} d(r;\mu) &= 2\int_{\mu}^{r}\frac{(r-t)}{t^2}\,dt \\ &= 2\left(\frac{r-\mu}{\mu} - \log\left(\frac{r}{\mu}\right)\right), \end{aligned}$$

FIGURE 1

GAMMA DISTRIBUTION DENSITY



again with weight $w = 1$. In this case the deviance is not symmetric about $r = \mu$. Figures 1 and 2 show plots of the gamma density and corresponding deviance function for three different means μ.

(c) The deviance $d(r;\mu) = w|r-\mu|$, $w > 0$, is an example which does not correspond to a linear bias function.

Returning to the case of multiple observations r_i with fitted values μ_i, the total deviance is

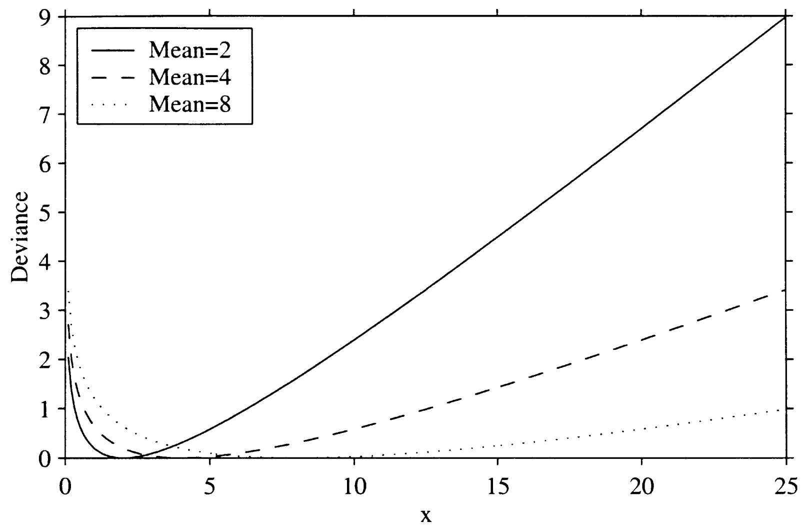
$$D = \sum_i d_i = \sum_i d(r_i;\mu_i).$$

Suppose $\mu_i = h(\mathbf{x}_i\boldsymbol{\beta})$ is a function of a linear combination of covariates $\mathbf{x}_i = (x_{i1},\ldots,x_{ip})$ and parameters $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)$, as it would be in the generalized linear model setting.[3] We find the

[3]The function h is the inverse of the link function which will be introduced in Section 6. The link function g relates the linear predictor to the mean: $\mathbf{x}_i\beta = g(\mu_i)$.

FIGURE 2

GAMMA DISTRIBUTION DEVIANCE



minimum deviance over the parameter vector $\boldsymbol{\beta}$ by solving the system of p equations

$$\frac{\partial D}{\partial \beta_j} = 0, \tag{5.2}$$

for $j = 1, \ldots, p$. Using the chain rule and assuming the deviance function is related to a linear bias function as in Equation 5.1 gives:

$$\begin{aligned}
\frac{\partial D}{\partial \beta_j} &= \sum_i \frac{\partial d_i}{\partial \beta_j} \\
&= \sum_i \frac{\partial d_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j} \\
&= -2\sum_i \frac{w_i(r_i - \mu_i)}{V(\mu_i)} h'(\mathbf{x}_i\boldsymbol{\beta})x_{ij}.
\end{aligned} \tag{5.3}$$

Let $\mathbf{X}$ be the design matrix with rows $\mathbf{x}_i$, $\mathbf{W}$ be the diagonal matrix of weights with iith element $w_i h'(\mathbf{x}_i\boldsymbol{\beta})/V(\mu_i)$, and $\boldsymbol{\mu}$ equal $(h(\mathbf{x}_1\boldsymbol{\beta}),\ldots,h(\mathbf{x}_n\boldsymbol{\beta}))^t$. Then Equation 5.3 can be written as

$$\mathbf{X}^t\mathbf{W}(\mathbf{r}-\boldsymbol{\mu}) = \mathbf{0} \tag{5.4}$$

which by Equation 4.13 is the zero bias equation. This shows that Bailey and Simon's balance criteria, BaS1, is equivalent to a minimum deviance criteria when bias is measured using a linear bias function, and weights are adjusted for the link function and form of the model using $h'(\mathbf{x}_i\boldsymbol{\beta})$.

The adjustment in Equation 5.3, given by $h'(\mathbf{x}_i\boldsymbol{\beta})x_{ij}$, depends upon the form of the underlying statistical model. This shows clearly how the bias function (which is related to the underlying distribution) and the form of the linear model (link and covariates) both impact the minimum bias parameters. The separation mirrors that between the error distribution and the link function exhibited in GLM1 and GLM3.

Examples of Minimum Bias Models

(a) $V \equiv 1$ and $h(x) = x$ reproduces the familiar additive minimum bias model which has already been considered in Section 4.

(b) Let $V(\mu) = \mu$ and $h(x) = e^x$. Using the minimum bias notation from Section 4, the minimum deviance condition Equation 5.3, which sets the bias for the ith level of the first classification variable to zero, is

$$\sum_{j=1}^{n_2} \frac{w_{ij}(r_{ij} - e^{a_i+b_j})}{e^{a_i+b_j}} e^{a_i+b_j} = \sum_{j=1}^{n_2} w_{ij}(r_{ij} - e^{a_i+b_j}) = 0,$$

including the link-related adjustment. Therefore

$$e^{a_i} = \sum_j w_{ij} r_{ij} \bigg/ \sum_j w_{ij} e^{b_j}$$

and similarly

$$e^{b_j} = \sum_i w_{ij} r_{ij} \Bigg/ \sum_i w_{ij} e^{a_i}$$

giving Bailey's multiplicative model.

See Section 7 for many more examples.

Summary

The definitions of linear bias function and deviance function have set up a natural correspondence:

$$\text{Deviance} \xrightarrow{\frac{\partial}{\partial \mu}} \text{Linear Bias Function},$$

$$d(y;\mu) \longrightarrow \frac{\partial d}{\partial \mu},$$

$$\int_\mu^r b(r;t)\,dt \longrightarrow -b(r;\mu), \qquad \text{and}$$

$$\text{Minimum Deviance} \longrightarrow \text{Zero bias by class.}$$

It follows from these definitions that the balance criterion sets the average bias to zero. However, except in trivial cases, the total minimum deviance is non-zero and is available as a model-fit statistic which can be used to select between models. This is an important step, especially since deviance has a reasonably well understood distribution. It is developed in Section 8.

Many minimum bias equations can now be derived using different link functions and linear bias functions, several of which lead to iterative equations. Everything in this section has been developed with no explicitly defined statistical model—since no probability distributions have been mentioned. Leaving out the statistical model makes the presentation more elementary and focuses on the intuitively reasonable roles of bias and deviance. In order to put minimum bias methods onto a firm statistical footing, a goal of the paper, we turn next to the theory of generalized linear models and exponential distributions and its relation to linear bias functions and deviance.

6. EXPONENTIAL DISTRIBUTIONS

The following diagram gives a schematic of Section 5 for the normal distribution.

$$\begin{array}{ccc}
\text{Balanced by class} & \longleftarrow & \text{Linear Bias Function} = r-\mu \\
{\scriptstyle\text{Differentiation}}\uparrow & & \uparrow{\scriptstyle\text{Differentiation}} \\
\text{Least Squares} & \longleftarrow & \text{Deviance Function} = (r-\mu)^2 \\
\| & & \downarrow \\
\text{Maximum Likelihood Parameters} & \longleftarrow & \text{Normal Distribution.}
\end{array}$$

To generalize to arbitrary linear bias functions, we need a family of distributions extending the normal which fills out the lower right hand corner of the diagram. It should have a likelihood function related to the given deviance function in the same way as the normal likelihood is related to the square distance deviance. Solving maximum likelihood for μ should correspond to minimum deviance, and will give balanced (according to the appropriate notion of bias) classification factors. The required family of distributions is called the exponential family. This section will define it and derive some of its important properties.

The exponential family of distributions[4] is the two-parameter family whose density functions can be written in the form

$$f(r;\mu,\phi) = c(r,\phi)\exp\left(-\frac{1}{2\phi}d(r;\mu)\right), \tag{6.1}$$

[4] This definition is slightly different from that in McCullagh and Nelder [17] and other sources on generalized linear models. See Appendix A for a reconciliation with the usual definition. The approach here is derived from Jørgensen [13] and McCullagh and Nelder [17, Chapter 9].

where d is a deviance function derived from a linear bias function using Equation 5.1. Using the squared distance deviance, unit weights $w = 1$, and $\phi \equiv \sigma^2$ shows that the normal distribution is in the exponential family, and that it corresponds to $V(\mu) = 1$. The gamma, binomial, Poisson and inverse Gaussian[5] distributions are also members of the exponential family. The exponential distribution, being a special case of the gamma, is also in the exponential family. It is important in Equation 6.1 that the function c depends only on r and ϕ; the same constant has to hold for all values of μ. This is a hard condition to satisfy. For example, it can be shown there is no such c when the deviance is derived from the variance function $V(\mu) = \mu^\zeta$ with $0 < \zeta < 1$.

Equation 6.1 and the definition of linear bias functions in terms of variance functions imply that an exponential family distribution is determined by the variance function.

If a random variable R has an exponential family distribution given by Equation 6.1 then

$$\mathrm{E}(R) = \mu \tag{6.2}$$

and

$$\mathrm{Var}(R) = \frac{\phi}{w} V(\mu), \tag{6.3}$$

which helps to explain the choice of μ as the first parameter and also why V is called the variance function. Because of its role in Equation 6.3, ϕ is called the dispersion parameter. Equations 6.2 and 6.3 follow immediately from two well-known results about the loglikelihood function $l = l(\mu,\phi;r) = \log f(r;\mu,\phi)$. The first is

$$\mathrm{E}\left(\frac{\partial l}{\partial \mu}\right) = 0, \tag{6.4}$$

[5]For more information on the inverse Gaussian, see Johnson, Kotz and Balakrishnan [12] and Panjer and Willmot [21]. It is similar to the lognormal distribution and can be used to model severity distributions.

$(\mathrm{E}(\partial l/\partial\mu) = \mathrm{E}(f'/f) = \int f' = \partial/\partial\mu \int f = \partial/\partial\mu(1) = 0)$. Equation 6.4 implies Equation 6.2. The second is

$$\mathrm{E}\left(\frac{\partial^2 l}{\partial\mu^2}\right) + \mathrm{E}\left[\left(\frac{\partial l}{\partial\mu}\right)^2\right] = 0, \tag{6.5}$$

which is derived similarly and which implies Equation 6.3.

The next two subsections derive the deviance functions associated with the gamma distribution and the inverse Gaussian distribution. The gamma example starts with the density and derives the variance function. The inverse Gaussian example goes in the opposite direction and starts with a variance function. In both cases the reader may (correctly) suspect the calculations are easier if one knows what the answer is going to be! Similar calculations can be performed for the Poisson and binomial distributions.

Gamma Distribution in the Exponential Family

The usual parameterization of the gamma density is

$$f(r;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} r^{\alpha-1} e^{-\beta r},$$

which has mean α/β and variance α/β^2. Since the parameter of interest is the mean, it makes sense to reparameterize to $\mu = \alpha/\beta$ and $\nu = \alpha$. The variance becomes μ^2/ν and the density becomes

$$f(r;\mu,\nu) = \left(\frac{\nu}{\mu}\right)^\nu \frac{1}{\Gamma(\nu)} r^{\nu-1} e^{-\nu r/\mu}.$$

Assuming weight w, Equation 6.3 gives $\phi\mu^2/w = \mu^2/\nu$, so $\nu = w/\phi$. Rearranging the density gives:

$$\begin{aligned} f(r;\mu,\nu) &= \frac{\nu^\nu r^{-1}}{\Gamma(\nu)} \exp(\nu \log(r/\mu) - \nu r/\mu) \\ &= \frac{\nu^\nu r^{-1} e^{-\nu}}{\Gamma(\nu)} \exp\left(-\frac{\nu}{2} 2\left(\left(\frac{r-\mu}{\mu}\right) - \log\left(\frac{r}{\mu}\right)\right)\right). \end{aligned} \tag{6.6}$$

Since the deviance $d = 2((r-\mu)/\mu - \log(r/\mu))$ corresponds to the variance function $V(\mu^2)$—see Example 5.1(b)—the gamma distribution is in the exponential family.

Exponential Density Corresponding to the Variance Function $V(\mu) = \mu^3$

The deviance function corresponding to $V(\mu) = \mu^3$ is given by

$$\begin{aligned} d(r;\mu) &= 2\int_{\mu}^{r} \frac{r-t}{t^3}\,dt \\ &= \frac{1}{r} + \frac{r}{\mu^2} - \frac{2}{\mu} \\ &= \frac{(r-\mu)^2}{\mu^2 r}. \end{aligned}$$

The corresponding exponential family distribution when $w = 1$ is

$$f(r;\mu,\phi) = c(r,\phi)\exp\left(-\frac{1}{2\phi}\frac{(r-\mu)^2}{\mu^2 r}\right),$$

which is exactly the inverse Gaussian distribution. The term $c(r,\phi)$ is given by

$$\sqrt{\frac{1}{2\pi\phi r^3}}.$$

The usual parameters for the inverse Gaussian are $1/\phi$ and $1/\mu$.

The variance function corresponding to the Poisson distribution is $V(\mu) = \mu$; for the binomial distribution it is $V(\mu) = \mu(1-\mu)$.

The modeling interpretation of V is clear from its role in linear bias functions. Now that we know how some variance functions and distributions match up we can make some further observations. The normal distribution model assumes constant variance, which is why the second important adjustment in Section 5 is not present in normal theory models. The Poisson model assumes the variance is proportional to the mean. The gamma model assumes the variance is proportional to the square of the mean, that is, that the coefficient of variation is constant. The inverse Gaussian assumes that the variance is proportional to the cube of the mean. The form of the variance function is very important in modeling, since the modeler will generally attempt to give smaller weights to observations with larger variances. Allowing the variance to be a function of the fitted mean gives generalized linear models a significant advantage over normal, constant variance, models.

Section 8 and Jørgensen [13] discuss other members of the exponential family. In particular see Jørgensen's Chapter 4 and Table 4.1.

7. GENERALIZED LINEAR MODELS AND THEIR CONNECTION WITH MINIMUM LINEAR BIAS

This section will explain how to solve generalized linear models using a maximum likelihood objective function, and show the connection between such solutions and solutions of minimum deviance models using linear bias functions. A thorough understanding of generalized linear models requires a more detailed treatment than can be given in this paper. The book by McCullagh and Nelder [17] is an excellent source for those desiring more information.

Section 2 divided general linear models into three components. The components were a random part, a systematic part and a link between the two—see GLM1-3. The random component can be any member of the exponential family, rather than just the normal distribution. The link function can be any monotonic function. Common choices include $\eta = \mu$, $\eta = \log(\mu)$, $\eta = 1/\mu$, $\eta = 1/\mu^2$ and the logit function $\eta = \log(\mu/(1-\mu))$. The link in a generalized linear model is a function of the predicted mean, $\eta = g(\mu)$, as opposed to the inverse link functions h used in Section 5 which are functions of the linear predictor $\mu = h(\eta)$.

Specification of a Generalized Linear Model

The full specification of a generalized linear model consists of:

- input data,
- model and distribution assumptions, and
- an objective function.

The input data comprises n observations $\mathbf{r} = (r_1, \ldots, r_n)^t$, n prior weights $\mathbf{w} = (w_1, \ldots, w_n)^t$, and p covariates $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})$ for each observation $i = 1, \ldots, n$. The covariates are the rows of the design matrix $\mathbf{X}$.

The model and distribution assumptions mirror the description GLM1-3. Observations r_i are assumed to be sampled from an exponential family distribution with mean μ_i and second parameter ϕ/w_i. The mean is related to the linear predictor using a link function

$$\mu_i = h(\eta_i), \qquad \eta_i = g(\mu_i),$$

and the linear predictor is related to the covariates by

$$\eta_i = \sum_j x_{ij}\beta_j = \mathbf{x}_i\boldsymbol{\beta}$$

for parameters $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^t$. Finally, the parameters are selected using the maximum likelihood objective.

The differences between a generalized and general linear model are the link function and the exponential family error distribution.

Maximum Likelihood Equations for a Generalized Linear Model

Let d be the deviance function associated with the exponential distribution used to define the model. From the definition of

TABLE 1
PARAMETERS FOR EXPONENTIAL FAMILY DISTRIBUTIONS

Quantity	Normal	Gamma	Inverse Gaussian
$V(\mu)$	1	μ^2	μ^3
Deviance, $d(r;\mu)$	$(r-\mu)^2$	$2\left(\frac{r-\mu}{\mu} - \log\left(\frac{r}{\mu}\right)\right)$	$\frac{(r-\mu)^2}{\mu^2 r}$
Dispersion, ϕ	ϕ	$\phi = 1/\nu$	ϕ
c	$(2\pi\phi)^{-1/2}$	$\nu^\nu r^{-1} e^{-\nu}/\Gamma(\nu)$	$(2\pi\phi r^3)^{-1/2}$

the exponential family, Equation 6.1, the loglikelihood is given by

$$l = l(\boldsymbol{\beta};\mathbf{r}) = \sum_{i=1}^{n} -\frac{1}{2\phi} d(r_i;\mu_i) + \log(c(r_i,\phi)). \tag{7.1}$$

To help the reader work through some explicit examples, Table 1 gives a summary of the functions introduced so far for the normal, gamma and inverse Gaussian distributions. If the weights $w \neq 1$, then replace ϕ with ϕ/w.

We find the maximum likelihood parameters $\hat{\boldsymbol{\beta}}$ by solving the system of p equations

$$\frac{\partial l}{\partial \beta_j} = 0$$

for $j = 1,\ldots,p$. Calculating from Equation 7.1 gives:

$$\begin{aligned}
\frac{\partial l}{\partial \beta_j} &= -\sum_i \frac{1}{2\phi} \frac{\partial d(r_i;\mu_i)}{\partial \beta_j} \\
&= -\sum_i \frac{w_i}{2\phi} \frac{\partial}{\partial \mu_i} \left(2 \int_{\mu_i}^{r_i} \frac{r_i - t}{V(t)} dt \right) \frac{\partial \mu_i}{\partial \beta_j} \\
&= \sum_i \frac{w_i}{\phi} \frac{r_i - \mu_i}{V(\mu_i)} \frac{\partial h(\mathbf{x}_i \boldsymbol{\beta})}{\partial \beta_j} \\
&= \sum_i \frac{w_i}{\phi} \frac{r_i - \mu_i}{V(\mu_i)} h'(\mathbf{x}_i \boldsymbol{\beta}) x_{ij}
\end{aligned}$$

since $\mathbf{x}_i\boldsymbol{\beta} = \sum_j x_{ij}\beta_j$. Equating to zero, the ϕ cancels out (just as σ cancels out of normal error linear models) giving the maximum likelihood equations for β_j:

$$\sum_{i=1}^{n} \hat{w}_i(r_i - \mu_i)x_{ij} = 0, \tag{7.2}$$

where the adjusted weight is defined as

$$\hat{w}_i = \frac{w_i h'(\mathbf{x}_i\boldsymbol{\beta})}{V(\mu_i)}. \tag{7.3}$$

Let $\mathbf{W}$ be the $n \times n$ diagonal matrix of adjusted weights $\hat{w}_i$. Then writing Equation 7.2 in matrix notation gives

$$\mathbf{X}^t\mathbf{W}(\mathbf{r} - \boldsymbol{\mu}) = \mathbf{0}. \tag{7.4}$$

As expected from the definition of exponential densities, Equation 7.4 is the same as the minimum deviance equations Equation 5.4. We have shown that the solution to the generalized linear model specified above is the same as the solution to the minimum bias model with the same covariates, link function, and associated variance function.

Special cases of the correspondence between generalized linear models and minimum linear bias models include:

$$\begin{aligned} \text{Normal} &\leftrightarrow V(\mu) = 1, \\ \text{Binomial} &\leftrightarrow V(\mu) = \mu(1-\mu), \\ \text{Poisson} &\leftrightarrow V(\mu) = \mu, \\ \text{Gamma} &\leftrightarrow V(\mu) = \mu^2, \qquad \text{and} \\ \text{Inverse Gaussian} &\leftrightarrow V(\mu) = \mu^3. \end{aligned}$$

The correspondence holds for all link functions. It also holds regardless of whether the minimum linear bias problem can be converted into a set of iterative equations. If the iterative equations exist, they can be used to solve for the parameters. In all cases, the theory of generalized linear models can be used to find the model parameters.

Canonical Link

If $\hat{w}_i = w_i$ in Equation 7.3, then h is called the canonical link corresponding to the variance function V. Clearly the canonical link satisfies the differential equation $V(h(\eta)) = h'(\eta)$. For example, if $V(\mu) = \mu$, then $h(\eta) = e^\eta$ is the canonical link. It is easier to find the maximum likelihood parameters using the canonical link because the weight matrix **W** is independent of the fitted values. If the canonical link is used, then adjusted balance is the same as balance in Bailey's definition. Despite its name, there is no need to use the canonical link associated with a particular variance function.

Explicit Examples

This subsection presents some explicit forms of the correspondence laid out above, including six of the eight different minimum bias models given by Brown [5].

Assume there are two classification variables and use the minimum bias notation from Section 4. Thus i and j are used to label both the observations and the parameters. Equation 7.4 translates into

$$\mathbf{0} = \mathbf{X}^t\mathbf{W}(\mathbf{r} - \boldsymbol{\mu}) = \begin{pmatrix} \sum_j \hat{w}_{1j}(r_{1j} - \mu_{1j}) \\ \vdots \\ \sum_j \hat{w}_{n_1 j}(r_{n_1 j} - \mu_{n_1 j}) \\ \sum_i \hat{w}_{i1}(r_{i1} - \mu_{i1}) \\ \vdots \\ \sum_i \hat{w}_{in_2}(r_{in_2} - \mu_{in_2}) \end{pmatrix}, \qquad \text{dimension } p \times 1,$$

(compare with Equation 4.9 and Equation 4.10). An equation from the first block gives

$$\sum_{j=1}^{n_2} \hat{w}_{ij}(r_{ij} - \mu_{ij}) = 0, \qquad i = 1,\ldots,n_1, \tag{7.5}$$

while one from the second block gives

$$\sum_{i=1}^{n_1} \hat{w}_{ij}(r_{ij} - \mu_{ij}) = 0, \qquad j = 1,\ldots,n_2. \tag{7.6}$$

The basic symmetry of the minimum bias method is already clear in the above equations.

a) Identity Link Function

For the identity link function, $\eta_{ij} = \mu_{ij}$ and $d\eta/d\mu = 1$, so

$$\hat{w}_{ij} = \frac{w_{ij}}{V(\mu_{ij})}.$$

Moreover, using an additive model, $\eta_{ij} = x_i + y_j$, and so $\mu_{ij} = x_i + y_j$. Substituting into the maximum likelihood equation Equation 7.5 gives

$$\begin{aligned} 0 &= \sum_{j=1}^{n_2} \hat{w}_{ij}(r_{ij} - \mu_{ij}) \\ &= \sum_{j=1}^{n_2} \frac{w_{ij}}{V(\mu_{ij})}(r_{ij} - (x_i + y_j)) \\ &= \sum_{j=1}^{n_2} \frac{w_{ij}}{V(\mu_{ij})}(r_{ij} - y_j) - x_i \sum_{j=1}^{n_2} \frac{w_{ij}}{V(\mu_{ij})}, \end{aligned}$$

for $i = 1,\ldots,n_1$. Hence

$$x_i = \sum_{j=1}^{n_2} w_{ij}(r_{ij} - y_j)/V(\mu_{ij}) \Bigg/ \sum_{j=1}^{n_2} w_{ij}/V(\mu_{ij}), \tag{7.7}$$

and similarly

$$y_j = \sum_{i=1}^{n_1} w_{ij}(r_{ij} - x_i)/V(\mu_{ij}) \Bigg/ \sum_{i=1}^{n_1} w_{ij}/V(\mu_{ij}), \qquad (7.8)$$

for $j = 1,\ldots,n_2$.

For the normal distribution, $V(\mu) = 1$. Substituting into Equation 7.7 gives

$$x_i = \sum_j w_{ij}(r_{ij} - y_j) \Bigg/ \sum_j w_{ij}, \qquad (7.9)$$

which is Bailey's additive model discussed in Section 4.

For the Poisson distribution, $V(\mu) = \mu$, and so Equation 7.7 gives

$$x_i = \sum_j w_{ij}(r_{ij} - y_j)/\mu_{ij} \Bigg/ \sum_j w_{ij}/\mu_{ij}, \qquad (7.10)$$

which is a new minimum bias method. For the gamma distribution, $V(\mu) = \mu^2$, and so Equation 7.7 gives

$$x_i = \sum_j w_{ij}(r_{ij} - y_j)/\mu_{ij}^2 \Bigg/ \sum_j w_{ij}/\mu_{ij}^2, \qquad (7.11)$$

which is another new method. Finally, for the inverse Gaussian distribution, $V(\mu) = \mu^3$, and so Equation 7.7 gives

$$x_i = \sum_j w_{ij}(r_{ij} - y_j)/\mu_{ij}^3 \Bigg/ \sum_j w_{ij}/\mu_{ij}^3, \qquad (7.12)$$

which is a third new method. The binomial distribution, with $V(\mu) = \mu(1 - \mu)$, also gives a new method.

The models in Equations 7.10 to 7.12 give progressively less and less weight to observations with higher predicted means μ_{ij}.

b) Log Link Function

For the log link function, $\eta = \log(\mu)$, so $d\eta/d\mu = 1/\mu$, which gives

$$\hat{w}_{ij} = \frac{w_{ij}\mu_{ij}}{V(\mu_{ij})}.$$

In this case, $\mu_{ij} = \exp(\eta_{ij}) = \exp(x_i + y_j) =: a_i b_j$. As expected the log link converts an additive model into a multiplicative one. Substituting into Equation 7.5 gives:

$$\begin{aligned} 0 &= \sum_{j=1}^{n_2} \hat{w}_{ij}(r_{ij} - \mu_{ij}) \\ &= \sum_{j=1}^{n_2} \frac{w_{ij}a_i b_j}{V(\mu_{ij})}(r_{ij} - a_i b_j) \\ &= \sum_{j=1}^{n_2} \frac{w_{ij}r_{ij}b_j}{V(\mu_{ij})} - a_i \sum_{j=1}^{n_2} \frac{w_{ij}b_j^2}{V(\mu_{ij})}, \end{aligned}$$

for $i = 1,\ldots,n_1$. Hence

$$a_i = \sum_{j=1}^{n_2} w_{ij}r_{ij}b_j/V(\mu_{ij}) \Bigg/ \sum_{j=1}^{n_2} w_{ij}b_j^2/V(\mu_{ij}), \qquad (7.13)$$

and similarly for b_j.

Now substituting V's for the normal, Poisson, gamma and inverse Gaussian distributions gives the following four minimum bias methods:

$$a_i = \sum_j w_{ij}r_{ij}b_j \Bigg/ \sum_j w_{ij}b_j^2, \qquad (7.14)$$

$$a_i = \sum_j w_{ij}r_{ij} \Bigg/ \sum_j w_{ij}b_j, \qquad (7.15)$$

$$a_i = \sum_j w_{ij} r_{ij}/b_j \Big/ \sum_j w_{ij}, \qquad \text{and} \tag{7.16}$$

$$a_i = \sum_j w_{ij} r_{ij}/b_j^2 \Big/ \sum_j w_{ij}/b_j. \tag{7.17}$$

Equation 7.15 is Bailey's multiplicative model, and Equation 7.16 is Brown's exponential model—which Venter comments also works for the gamma. Equations 7.14 and 7.17 appear to be new. Again, going from Equation 7.14 to Equation 7.17, the models give less and less weight to observations with high predicted means.

c) Ad Hoc Methods

Other functions besides the identity, log, and logit can be used as links. Two common choices are $\eta = 1/\mu$ and $\eta = 1/\mu^2$. For the inverse function, $d\eta/d\mu = -1/\mu^2$, so $\hat{w} = w\mu^2/V(\mu)$, and $\mu_{ij} = 1/\eta_{ij} = 1/(x_i + y_j)$. Substituting into Equation 7.5, it is easy to see that only the inverse Gaussian produces a tractable minimum bias method. For the inverse Gaussian, $V(\mu) = \mu^3$, so Equation 7.5 gives

$$\begin{aligned}
0 &= \sum_{j=1}^{n_2} \hat{w}_{ij}(r_{ij} - \mu_{ij}) \\
&= \sum_j \frac{w_{ij}}{\mu_{ij}}(r_{ij} - \mu_{ij}) \\
&= \sum_j \frac{w_{ij} r_{ij}}{\mu_{ij}} - w_{ij} \\
&= \sum_j w_{ij} r_{ij}(x_i + y_j) - w_{ij} \\
&= x_i \sum_j w_{ij} r_{ij} + \sum_j w_{ij} r_{ij} y_j - \sum_j w_{ij},
\end{aligned}$$

and hence we get another new iterative method:

$$x_i = \sum_j w_{ij}(1 - r_{ij}y_j) \bigg/ \sum_j w_{ij}r_{ij}. \qquad (7.18)$$

In this method one set of parameters will be negative and the other positive.

Other Variance Assumptions

Brown proposes two models where the variance of r_{ij} is proportional to $1/w_{ij}^2$ rather than $1/w_{ij}$. Although, as Venter points out, the latter is a more natural choice, the former assumption can be handled within our framework by simply using weights w_{ij}^2 rather than w_{ij}. For example, Equation 7.9 becomes

$$x_i = \sum_j w_{ij}^2(r_{ij} - y_j) \bigg/ \sum_j w_{ij}^2, \qquad (7.19)$$

which is Brown's model 7.

Correspondence with Brown's Models

For the reader's convenience, this subsection identifies our models with the nine models in Brown's paper:

B1: Poisson, multiplicative, Equation 7.15.

B2: Normal, additive, Equation 7.9.

B3: Bailey–Simon, multiplicative—see [3, Equation 7] for derivation. This method comes from minimizing a χ^2-statistic, rather than maximizing a likelihood function. Since generalized linear models rely on maximum likelihood, we would not expect to be able to reproduce it. Unlike B4, it does not use the Newton method.

B4: Bailey–Simon, additive—see [3, p. 12] for derivation. This method (which certainly puzzled the author as a

Part 9 exam candidate!) also minimizes a χ^2-statistic. Its derivation uses the Newton method.

B5: Gamma, multiplicative, Equation 7.16; note the exponential is a special case of the gamma.

B6: Normal, multiplicative with variance proportional to $1/w^2$, Equation 7.14, upon replacing w with w^2.

B7: Normal, additive, with variance proportional to $1/w^2$, Equation 7.19.

B8: The same as B1.

B9: Normal, multiplicative, Equation 7.14. Brown derives B9 using least squares and Venter uses maximum likelihood. The two approaches agree because the likelihood of a normally distributed observation is proportional to its squared distance from the mean.

Numerical Example, Continued

We now present the results of fitting ten generalized linear models to the data presented in Section 4. The models are described in Table 2 below.

So far we have not been concerned with the value of the parameter ϕ. It is well known that in general linear models, parameter estimates and predicted values are independent of the variance of the error term (usually labeled σ^2 rather than ϕ). Since ϕ does not appear in Equation 7.4, the same is true of generalized linear models. However, just as for general linear models, it is necessary to estimate ϕ in order to determine statistics such as standard errors of predicted values. In general linear models,

$$\hat{\sigma}^2 = \sum_i w_i (r_i - \mu_i)^2/(n-p)$$

is used as an estimator of σ^2, where n is the number of observations and p is the number of parameters. In generalized linear

TABLE 2

DESCRIPTION OF MODELS

Model Number	Error Distribution	Link Function	Variance Function
1	Normal	Identity	$V(\mu) = 1$
2	Normal	Log	$V(\mu) = 1$
3	Normal	Inverse	$V(\mu) = 1$
4	Gamma	Identity	$V(\mu) = \mu^2$
5	Gamma	Log	$V(\mu) = \mu^2$
6	Gamma	Inverse	$V(\mu) = \mu^2$
7	Inverse Gaussian	Identity	$V(\mu) = \mu^3$
8	Inverse Gaussian	Log	$V(\mu) = \mu^3$
9	Inverse Gaussian	Inverse	$V(\mu) = \mu^3$
10	Inverse Gaussian	Inverse Square	$V(\mu) = \mu^3$

models, ϕ is estimated using the moment estimator

$$\hat{\phi} = \frac{1}{n-p}\sum_i w_i \frac{(r_i - \mu_i)^2}{V(\mu_i)}. \qquad (7.20)$$

It can also be estimated using

$$\hat{\phi} = \frac{D}{n-p} = \frac{1}{n-p}\sum_i d(r_i, \mu_i), \qquad (7.21)$$

where D is the total deviance (see McCullagh and Nelder [17]). Note that the weights are included in the deviance d in Equation 7.21. Another way to estimate ϕ is to use the maximum likelihood estimate. Equation 7.1 ensures that the maximum likelihood parameters are unchanged whether or not ϕ is estimated. SAS's "proc genmod" uses maximum likelihood by default (see [24]), and the statistics reported below are based on it unless otherwise noted.

Exhibit 7 gives the parameters corresponding to the ten models in Table 2. Each panel of Exhibit 7 shows the parameter estimates, the standard error of the estimate, the χ^2-statistic to test if the parameter is significantly different from zero, and the corresponding p-value from the χ^2-distribution. (See Section 8 for more discussion of the χ^2-statistic.) When the link function is not the identity, Exhibit 7 also shows the parameter estimates transformed by the inverse link. For example, in the first row of Exhibit 7-2, $265.22 = e^{5.5806}$. The final row gives the scale function, which is equal to $\sqrt{\phi}$ for the normal and inverse Gaussian distributions, and $1/\phi$ for the gamma distribution. Again, maximum likelihood is used to estimate ϕ.

Examining Exhibit 7 shows that all parameters except "drive to work (DTW) less than 10 miles" are significantly different from zero for all models. All models indicate there is not a statistically significant difference between "drive to work less than 10 miles" and pleasure-use. The other two use classifications are significantly different from one another. The estimates and standard errors within the age classifications show there is not a statistically significant difference among all levels. For example, the 35–39 and 40–49 classes are not significantly different for most models, although exact results depend on the choice of ϕ. In the gamma model with identity link using maximum likelihood gives the estimate $\hat{\phi} = 0.9741$, and the contrast between these two classes has a χ^2-statistic of 4.07, which is significant at the 5% level ($p = 4.4\%$). However, using Equation 7.20 results in an estimate $\hat{\phi} = 1.4879$ with a χ^2-statistic of 2.839 which is not significant at the 5% level ($p = 9.2\%$). In the first case the standard error of the 35–39 class is 8.13 (Exhibit 7-4); in the second it is 10.04.

Exhibit 8 compares the fitted values from three models: the standard linear model (column 5), a general linear model applied to log(severity) (column 6), and a generalized linear model with normal errors and log link (column 7). As pointed out in Example 2.4(c), the three are distinct and give different answers.

Exhibit 9 summarizes the predicted severities by class, by model. The choice of link function and error distribution has a considerable impact on the predicted means in some cells. Using a gamma or inverse Gaussian error term generally results in a greater range of estimates, as does the log or reciprocal link function. Since this is only illustrative data we will not comment on the specific results. See Renshaw [23] for a more detailed analysis of similar data, together with other suggestions for modeling and assessing model fit.

Exhibit 10 gives the average bias

$$\sum_i w_{ij}(r_{ij} - \mu_{ij}) \Big/ \sum_i w_{ij} \tag{7.22}$$

for each j and

$$\sum_j w_{ij}(r_{ij} - \mu_{ij}) \Big/ \sum_j w_{ij}$$

for each i, for each model. For the normal/identity model, the average bias is zero, since this model is Bailey's additive model. The gamma/inverse model and inverse Gaussian/inverse square models are also balanced because the respective link functions are the canonical links (as discussed earlier in this section), and so the adjustment to the weights in Equation 7.3 equals 1, reducing Equation 7.4 to Equation 7.22. In the other cases, the parameters are zero bias according to the relevant adjusted bias function, but not according to that given by Equation 7.22. This provides an interesting example of Venter's V1—alternatives to bias functions.

Exhibit 11 gives the average absolute bias suggested by Bailey [2]:

$$\sum_i w_{ij}|r_{ij} - \mu_{ij}| \Big/ \sum_i w_{ij} \tag{7.23}$$

for each i, and similarly for j. The gamma/identity model has the lowest average absolute bias. Finally, the value of the likelihood is available as a fit statistic, since these models were fit using maximum likelihood over all parameters (including ϕ). The re-

TABLE 3
MODEL LOGLIKELIHOODS

Model	Distribution	Link	Loglikelihood
1	Normal	Identity	−144.303
2	Normal	Log	−144.435
3	Normal	Inverse	−145.792
4	Gamma	Identity	−140.753
5	Gamma	Log	−141.055
6	Gamma	Inverse	−143.267
7	Inverse Gaussian	Identity	−141.078
8	Inverse Gaussian	Log	−141.347
9	Inverse Gaussian	Inverse	−143.343
10	Inverse Gaussian	Sqr Inverse	−147.224

sults are shown in Table 3. Other statistics that can be used to select between models are discussed in Section 8.

These examples hint at the power of the statistical viewpoint. Using a minimum bias approach not within the statistical framework, it would be impossible to discuss the standard error of predicted values and parameters, or to ask whether two parameters are statistically significantly different. Having the tools to answer such questions can provide useful information to help in designing and parameterizing classification plans. The statistical model also gives information on model fit, discussed in the next section, which helps select covariates, as well as link and variance functions within parameterized families. Again, these tools are not available with the minimum bias approach. Fundamentally it is the connection between variance functions and exponential family distributions that makes the statistical viewpoint possible.

8. MODEL FIT STATISTICS

Generalized linear model and minimum bias methods allow the actuary to consider a large number of models: different choices of covariates, different link functions and different variance functions. It is obviously important to be able to determine if one model fits the data better than the others. The specification of a generalized linear model in Section 7 shows there are at least four distinct model fit questions:

1. comparing different sets of covariates for a given link function and variance function (error distribution),
2. comparing different link functions and covariates for a given variance function,
3. comparing different variance functions for a given set of covariates and link function, and
4. simultaneously comparing different link and variance functions and covariates.

In this section we will discuss some of the available statistical tests of model fit. These methods extend the earlier work of Bailey and Simon.

Comparing Sets of Covariates

The simplest test of model fit looks for information about the best set of covariates assuming given link and variance functions. In the numerical example, is anything really gained from adding a vehicle-use classification? Analysis of variance is used in normal-error model theory to assess the significance of effects and answer such questions. For generalized linear models, we look at an analysis of deviance table, obtained from a nested sequence of models. Unfortunately, unlike the normal-error theory where the χ^2- and F-distributions give exact results, only approximations and asymptotic results are available for generalized linear models. McCullagh and Nelder [17] recommend analysis of deviance as a screening device for models and regard this as an area where more work is required.

Consider the gamma distribution model with identity link. With two explanatory variables available, we can consider a nested sequence of four models: intercept only, age only, age and vehicle type with no interaction, and age and vehicle type with interaction. The last model is complete—it has as many

TABLE 4
ANALYSIS OF DEVIANCE

Model	Deviance	Δ Deviance	Degrees of Freedom	Mean Deviance
Intercept	347.0331			
Age	264.8553	82.1778	7	11.74
Age + Vehicle	31.2453	233.6100	3	77.87
Complete	0	31.2453	21	1.49

parameters as there are observations and so fits perfectly. Table 4 shows the resulting analysis of deviance. For each model, it shows the deviance, the reduction in deviance from adding covariates, the number of incremental degrees of freedom, and the mean incremental deviance per degree of freedom. The degrees of freedom are computed as the incremental number of parameters from one model to the next. The model with an intercept has only one parameter. Including age variables adds seven more parameters, and so on. The complete model has one parameter for each of the 32 observations.

The mean deviance has an approximate χ^2-distribution. Adding the age variable and then the vehicle type variable both significantly improve the model fit. When more explanatory variables are available, an analysis of deviance is helpful in deciding which to use in a model, and in particular, in assessing which interaction effects are significant and should be included.

Comparing Link Functions

The models discussed in Section 7 used the identity, inverse and log links, all of which belong to the power-link family[6]

$$\eta = \begin{cases} \mu^\lambda & \text{for} \quad \lambda \neq 0, \\ \log(\mu) & \text{for} \quad \lambda = 0. \end{cases}$$

[6]Considering $(\mu^\lambda - 1)/\lambda$ instead of μ^λ makes the family appear more natural, because $(\mu^\lambda - 1)/\lambda \to \log(\mu)$ as $\lambda \to 0$. This form of the power-link function is called the Box–Cox transformation. It is mentioned in Venter's review [26].

TABLE 5
DEVIANCE VS. LINK POWER λ

λ	Deviance
−1.800	43.828
−1.300	38.966
−0.800	35.190
−0.300	32.724
0.200	31.464
0.700	31.129
1.200	31.418
1.450	31.717

According to Nelder and Lee [18, Section 2.3], we can use the deviance to compare different link functions as well as different covariates. Table 5 shows the deviance for various values of λ, again using the gamma distribution.

The deviance is relatively flat across the range $0.325 \leq \lambda \leq 1.075$, which includes the identity link. The deviance for the inverse link $\lambda = -1$ is substantially greater than for λ in this range.

More on Variance Functions

Before discussing tests over sets of variance functions, we must mention a few facts about them. Jørgensen, [13] and [14], discusses the exponential families corresponding to variance functions beyond the simple examples we have considered so far. His results include the following which are of interest to actuaries:

1. $V(\mu) = \mu^{\zeta}$ for $1 < \zeta < 2$ corresponds to the Tweedie distribution, which is a compound distribution with Poisson frequency component and gamma severity component. It is a mixed distribution with a non-zero probability of taking the value zero, which makes it useful in modeling aggregate distributions. Jørgensen and deSouza [15] fit the Tweedie model to Brazilian auto data.

2. $V(\mu) = \mu^\zeta$ for $\zeta < 0$ corresponds to an extreme stable distribution. Non-normal stable distributions are thick tailed distributions which may be useful in fitting loss data.

3. $V(\mu) = \mu^\zeta$ for $2 < \zeta < \infty$, $\zeta \neq 3$ corresponds to a positive stable distribution.

4. $V(\mu) = \mu^\zeta$ for $0 < \zeta < 1$ does not give an exponential family distribution.

5. $V(\mu) = \mu(1 + \mu/\nu)$ corresponds to the negative binomial distribution.

6. $V(\mu) = \mu(1 + \tau\mu^2)$ corresponds to the Poisson-inverse Gamma distribution. Renshaw [23] gives the deviance functions for both of the last two distributions.

The power variance function family leads naturally to the question of determining the best estimate for ζ, to which we now turn.

Comparing Variance Functions

The deviance cannot be used to select an optimal ζ because the deviance of an individual observation $(r - \mu)/\mu^\zeta \to 0$ as $\zeta \to \infty$ for $\mu > 1$. This means a deviance-based objective would generally claim ζ should be very large and that the model fit was excellent. Clearly it is necessary to include some measure of the likelihood of ζ in the objective function to counter-balance the effect of the variance function on the deviance. In general, according to Nelder and Lee [18], deviance cannot be used to compare different variance functions on the same data.

One way to include the likelihood of ζ would be to use the full likelihood function for the corresponding density. This method was used in the examples shown in Section 7 for the normal, gamma and inverse Gaussian distributions—where the densities are known. Unfortunately, for most exponential family distributions, including the Tweedie and stable distributions, there is no simple closed form expression for the density or distribution function. It is therefore not possible to write down the likelihood function.

The way out of this impasse is to use a tractable approximation to the density function, such as the saddlepoint approximation. Details of the derivation are beyond the scope of this paper, but the result is to replace the deviance function

$$d(r_i;\mu) = 2w_i \int_{\mu}^{r_i} \frac{r_i - t}{V(t)}\, dt \tag{8.1}$$

with an extended deviance function (extended quasi-likelihood in the literature)

$$d(r_i;\mu) = 2\frac{w_i}{\phi} \int_{\mu}^{r_i} \frac{r_i - t}{V(t)}\, dt + \log(\phi V(r_i)). \tag{8.2}$$

The added term grows with V, thus providing the desired counter-balance to the first term, which shrinks. Note that V is evaluated at the responses r_i rather than the fitted means μ_i. Including the scale parameter ϕ allows Equation 8.2 to be used both for inference over parameterized families of variance functions and for different values of ϕ. Jørgensen [13, Example 3.1 p. 104] explains the saddlepoint approximation for a gamma distribution, which is just Stirling's formula for the gamma function. See McCullagh and Nelder [17, Chapter 9], Nelder and Lee [18], and Renshaw [23] for more about extended deviance functions. [18] also defines and compares other extensions of deviance.

Table 6 shows the extended deviances for various values of ζ modeled with the identity link function. The table shows a reasonable range $1.95 \le \zeta \le 3.45$, which includes both the gamma distribution $\zeta = 2$ and inverse Gaussian distribution $\zeta = 3$. Combining the results of Tables 5 and 6 shows the gamma or inverse Gaussian distribution with identity link is still a reasonable choice even if we are free to select from the power link family and power variance function family. These conclusions are in line with the full likelihood results in Table 3 and the average

TABLE 6
EXTENDED DEVIANCE VS. VARIANCE FUNCTIONS $V(\mu) = \mu^{\zeta}$

ζ	Deviance
1.20	372.740
1.45	372.020
1.70	371.422
1.95	370.946
2.20	370.597
2.45	370.374
2.70	370.282
2.95	370.321
3.20	370.494
3.45	370.800

absolute deviations in Exhibit 11, where the gamma/identity and inverse Gaussian/identity models show the best results.

Deviance Profiles and Comparing Link and Variance Functions

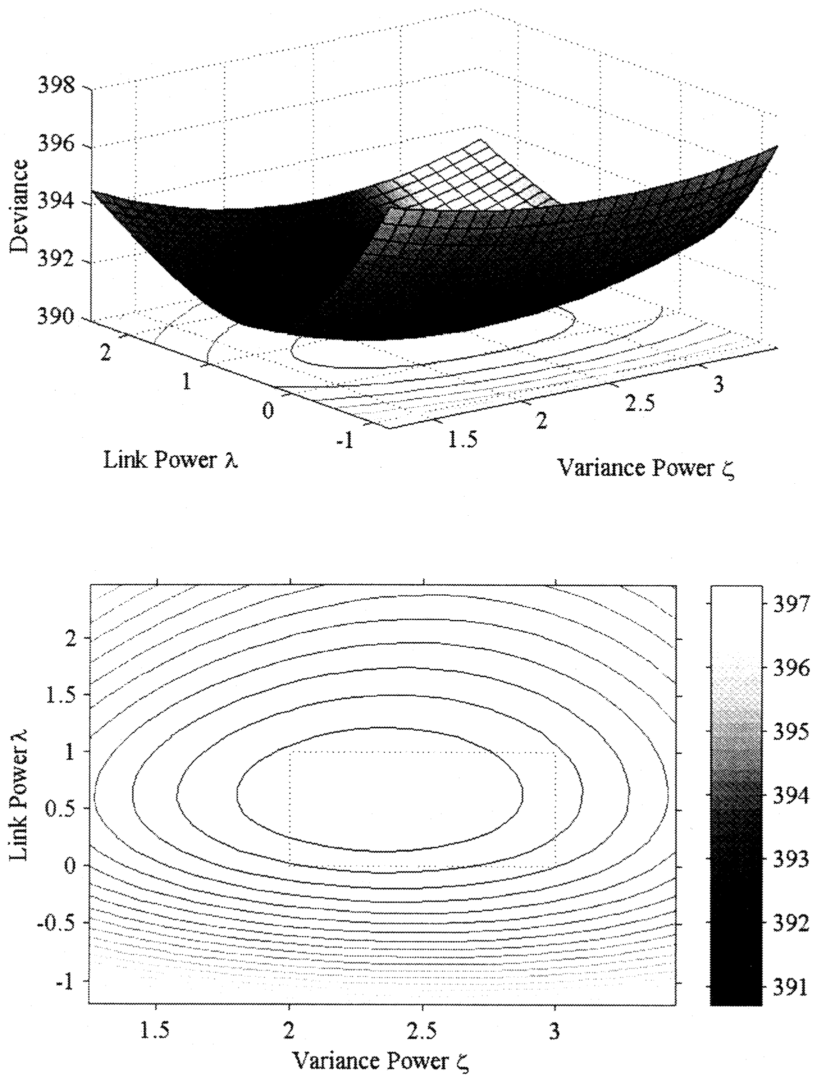
The last step we will consider combines the power link and variance functions and looks for the overall minimum extended deviance estimators. Figure 3 shows a contour plot of extended deviance over ζ and λ. The results are as expected from the one dimensional calculations. The dotted rectangle shows a range of λ from log link to the identity and ζ from gamma distribution to inverse Gaussian.

9. COMPUTATIONS

Section 9 is in two parts. The first discusses the iterative method for solving minimum bias models. For the additive model with identity link, it gives a sufficient condition for the iterative method to converge (no matter the initial conditions), explains precisely how it converges in terms of the eigenvectors of a particular matrix, and gives a telescoping argument that jumps to the solution of the iterative process once the first iteration has been computed.

FIGURE 3

DEVIANCE PROFILE POWER LINK AND VARIANCE FUNCTIONS



The second section discusses how to find the maximum likelihood parameters in a generalized linear model. Even though commercial software exists to solve generalized linear models, it is instructive to perform the calculations by hand, and we explain how to do this. Examples of SAS code to solve generalized linear models using both the SAS/Stat procedure "genmod" and a "bare hands" approach using matrix algebra are given in Appendix B.

At several points this section discusses a notion of computational efficiency. Two algorithms are of similar computational efficiency if they will run in about the same time for *all sizes of input*. (Technically, if n is the problem size, and $f(n)$ and $g(n)$ are the number of elementary operations required to solve the problem using two methods, then they are of the same computational efficiency if $f = O(g)$ and $g = O(f)$, Borwein and Borwein [4, Chapter 6]. Recall $f = O(g)$ means there is a constant K so that $f(n) \leq Kg(n)$ for all n.)

Iterative Methods

Bailey's original paper [2] introduces the additive and multiplicative models and suggests the iterative method for finding parameters:

> Using a predetermined set of estimators for each territory, construction, and protection, we can solve the [minimum bias] formula for the estimator for each occupancy. We can then use these calculated estimators for each occupancy to calculate a revised set of estimators for each territory using a similar formula, and continue this process until the estimators stabilize.

Since Bailey's paper, it has become common for actuaries to use this iterative method. For example, ISO [11] explicitly describes the three-way minimum bias model for the personal auto classification plan as iterative.

Just because the minimum bias model *suggests* using an iterative method to solve for the parameters, it does not follow that such a method is the best method to use. Section 4 showed that the usual additive model is simply a general linear model; and so it is far more computationally efficient to solve the normal equations (no iterations, few matrix multiplications and one inverse) than it is to use the iterative method. Any actuaries still using iterative methods should investigate whether the generalized linear model approach outlined in this paper would speed up their calculations—as well as providing them with more useful diagnostic information.

This section considers the iterative method for the additive model with identity link which is used by ISO for the personal auto class plan. The iterative method is considered in detail despite its shortcomings, because many actuaries may have tried the method (perhaps as Part 9 students) and may have wondered what initial conditions are required for convergence and may also have noted the strange way the models converge. We explain the convergence in detail and also show it is not necessary to perform many iterations, even if the iterative paradigm is followed. However, the final message of this section is *do not use the iterative method for Bailey's additive model*—solve the normal equations instead!

We will use the notation of Section 4 and consider two classification variables—extensions are immediate. Assume that base classes have been selected so that the sum-of-squares and products matrix $\mathbf{X}^t\mathbf{W}\mathbf{X}$ is invertible, $\mathbf{a}$ has dimension $n_1 \times 1$ and $\mathbf{b}$ has dimension $n_2 \times 1$. Finally assume $n_2 \leq n_1$; if this is not the case then swap $\mathbf{a}$ and $\mathbf{b}$. For this example the adjusted weights $\hat{w} = w$ (see Equation 7.3).

From Equation 4.14 the minimum bias equations can be written as

$$(\mathbf{A} \quad \mathbf{B})^t\mathbf{W}\left(\mathbf{r} - (\mathbf{A} \quad \mathbf{B})\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix}\right)$$

$$= \begin{pmatrix}\mathbf{A}^t\mathbf{W}(\mathbf{r} - \mathbf{A}\mathbf{a} - \mathbf{B}\mathbf{b})\\ \mathbf{A}^t\mathbf{W}(\mathbf{r} - \mathbf{A}\mathbf{a} - \mathbf{B}\mathbf{b})\end{pmatrix} = \mathbf{0}_{(n_1+n_2)\times 1}. \qquad (9.1)$$

Re-arranging Equation 9.1 gives

$$\mathbf{a} = (\mathbf{A}^t\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^t\mathbf{W}(\mathbf{r} - \mathbf{B}\mathbf{b}), \qquad \text{and} \tag{9.2}$$

$$\mathbf{b} = (\mathbf{B}^t\mathbf{W}\mathbf{B})^{-1}\mathbf{B}^t\mathbf{W}(\mathbf{r} - \mathbf{A}\mathbf{a}). \tag{9.3}$$

The iterative solution starts with some initial choice $\mathbf{b}^{(0)}$ and uses Equation 9.2 to solve for $\mathbf{a}^{(1)}$. Substituting $\mathbf{a}^{(1)}$ into the Equation 9.3 gives an expression for $\mathbf{b}^{(1)}$. Iterating gives $\mathbf{a}^{(2)}$, $\mathbf{b}^{(2)}$, and so forth. The procedure stops when the difference between successive iterations is sufficiently small. Set $\mathbf{v}^{(m)} = \mathbf{b}^{(m)} - \mathbf{b}^{(m-1)}$ equal to the difference in the m and $(m-1)$th iterations for $\mathbf{b}$. Note there is an asymmetry between the $\mathbf{a}$-iterations and the $\mathbf{b}$-iterations based on where we choose to start.

Set

$$\mathbf{M} = (\mathbf{B}^t\mathbf{W}\mathbf{B})^{-1}\mathbf{B}^t\mathbf{W}\mathbf{A}(\mathbf{A}^t\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^t\mathbf{W}\mathbf{B}, \tag{9.4}$$

an $n_2 \times n_2$ matrix. A straightforward telescoping argument shows that

$$\mathbf{b} = (\mathbf{I} - \mathbf{M})^{-1}\mathbf{v}^{(1)} + \mathbf{b}^{(0)}, \tag{9.5}$$

provided $\mathbf{M}^m \to 0$. We can guarantee that $\mathbf{M}^m \to 0$ as $m \to \infty$ if all the eigenvalues of $\mathbf{M}$ have absolute value less than 1. This gives a necessary condition for the iterative method to converge, and, moreover, Equation 9.5 shows how to "jump" straight to the final solution after computing only one iteration, $\mathbf{a}^{(1)}$ and $\mathbf{b}^{(1)}$. This method of solving the minimum bias problem will run much faster than the iterative method, but will still be slower than solving the normal equations (computing $\mathbf{M}$ alone involves eleven matrix multiplications and two inverses).

It is also possible to show that $\mathbf{v}^{(m)}$ tends to a scalar multiple of the eigenvector associated with the largest eigenvalue of $\mathbf{M}$, and the iterative method converges along the direction of that eigenvector. Moreover, the distance between subsequent iterations of

$\mathbf{b}^{(m)}$ decreases by approximately the absolute value of the largest eigenvector for m large enough.

Numerical Example, Continued

Exhibit 5 illustrates the above theory. Column 2 gives the length of $\mathbf{v}$, Column 14 gives the ratio of successive iterations of $\mathbf{v}$, and Columns 15–17 give the three components of $\mathbf{v}^{(m)}$. The ratio of lengths of $\mathbf{v}$ should converge to the largest eigenvalue of the matrix $\mathbf{M}$ defined by Equation 9.4. For the data underlying Exhibit 5:

$$\mathbf{M} = \begin{pmatrix} 0.457572 & 0.300972 & 0.118435 \\ 0.431800 & 0.306347 & 0.122798 \\ 0.428349 & 0.309565 & 0.126608 \end{pmatrix}, \qquad (9.6)$$

which has eigenvalues 0.000541, 0.010541 and 0.859445. This explains the 0.85944's that appear in Exhibit 5; their appearance is quick since the largest eigenvalue is so much greater than the other two. The overall convergence of the model is quite slow, since 0.859 is close to 1.0.

Exhibits 12 and 13 show how the iterative method converges for two other models: gamma/identity and gamma/inverse, respectively. Convergence is particularly slow for the latter; after 25 iterations the parameters are nowhere near their final values. The methods of this section do not apply to non-canonical link functions because the weight matrix $\mathbf{W}$ must be re-evaluated between each iteration, and so the telescoping argument will not hold.

Solving Generalized Linear Models

One conclusion of this paper is that many useful minimum linear bias models correspond in a natural way with generalized linear models. However, not all minimum linear bias models have a tractable iterative solution. It is therefore useful to know how to solve generalized linear models. Since there are pre-programmed routines for generalized linear models[8] we give only a brief overview here. This section follows McCullagh and Nelder [17]. The notation is the same as the first part of Section 7.

From Equation 7.4 the maximum likelihood equations for the generalized linear model are given by

$$\mathbf{X}^t\mathbf{W}(\mathbf{r}-\boldsymbol{\mu}) = \mathbf{0},$$

where $\mathbf{W}$ is the diagonal matrix with entries $\hat{w}_i = w_i h'(\mathbf{x}_i\boldsymbol{\beta})/V(\mu_i)$.

To find the maximum likelihood it is necessary to solve $\partial l/\partial\beta_j = 0$, for $j = 1,\ldots,p$. This can be done using a method related to the Newton–Raphson method. In one dimension the Newton–Raphson method solves an equation $f(x) = 0$ by iterating $x_{n+1} = x_n - f(x_n)/f'(x_n)$. We are trying to solve the vector equation $\mathbf{u}(\boldsymbol{\beta}) = \mathbf{0}$, where

$$\mathbf{u}(\boldsymbol{\beta}) = \mathbf{u} = \partial l/\partial\boldsymbol{\beta} = (\partial l/\partial\beta_1,\ldots,\partial l/\partial\beta_p)^t.$$

Looking at Newton–Raphson suggests trying $\boldsymbol{\beta}_{n+1} = \boldsymbol{\beta}_n - (\partial\mathbf{u}/\partial\boldsymbol{\beta})^{-1}\mathbf{u}$. The term $\partial\mathbf{u}/\partial\boldsymbol{\beta}$ is called the Hessian. The negative Hessian is called the observed information matrix (see Hogg and Klugman [10, p. 121]); it is generally a random quantity. Fisher's scoring method simplifies the Newton–Raphson method by using the expected value of the Hessian rather than the Hessian itself; it often results in more staightforward calculations.

To apply Fisher's scoring method, let

$$\mathbf{H} = -\mathrm{E}\left(\frac{\partial^2 l}{\partial\beta_j\partial\beta_k}\right) = -\mathrm{E}\left(\frac{\partial\mathbf{u}}{\partial\beta}\right)$$

[8] As well as GLIM, mentioned by Brown, SAS now includes a procedure, "proc genmod" to solve generalized linear models in its SAS/Stat package. "Proc genmod" has the same syntax as "proc glm".

be the negative expected value of the Hessian matrix. Given an estimate $\boldsymbol{\beta}_n$ of $\boldsymbol{\beta}$, we find the next adjustment $\mathbf{a}$ by solving $\mathbf{Ha} = \mathbf{u}$. (The adjustment term in the Newton–Raphson method, $a = f(x_n)/f'(x_n)$, satisfies $f'(x_n)a = f(x_n)$. Here, $f \leftrightarrow \mathbf{u}$ and $f' \leftrightarrow H$.) From Equation 7.4, $\mathbf{u} = \mathbf{X}^t\mathbf{W}(\mathbf{r} - \boldsymbol{\mu})$, and so

$$\begin{aligned}
\mathbf{H} &= -\mathrm{E}\left(\frac{\partial \mathbf{u}}{\partial \boldsymbol{\beta}}\right) \\
&= -\mathrm{E}\left(\frac{\partial}{\partial \boldsymbol{\beta}}\mathbf{X}^t\mathbf{W}(\mathbf{r} - \boldsymbol{\mu})\right) \\
&= -\mathrm{E}\left(\frac{\partial(\mathbf{X}^t\mathbf{W})}{\partial \boldsymbol{\beta}}(\mathbf{r} - \boldsymbol{\mu}) + (\mathbf{X}^t\mathbf{W})\frac{\partial}{\partial \boldsymbol{\beta}}(\mathbf{r} - \boldsymbol{\mu})\right) && (9.7) \\
&= \mathrm{E}\left(\mathbf{X}^t\mathbf{W}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\beta}}\right) && (9.8) \\
&= \mathrm{E}\left(\mathbf{X}^t\mathbf{W}\frac{d\boldsymbol{\mu}}{d\boldsymbol{\eta}}\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\beta}}\right) && (9.9) \\
&= \mathbf{X}^t\tilde{\mathbf{W}}\mathbf{X}, && (9.10)
\end{aligned}$$

which is the weighted sums of squares and products matrix for the model with weights

$$\tilde{\mathbf{W}} = \mathrm{diag}\left(\frac{w_i(h'(\mathbf{x}_i\boldsymbol{\beta})^2}{V(\mu_i)}\right).$$

Equation 9.7 uses the chain rule; Equation 9.8 uses the fact that $\mathrm{E}(\mathbf{r}) = \boldsymbol{\mu}$ and $\partial\mathbf{r}/\partial\boldsymbol{\beta} = \mathbf{0}$ ($\mathbf{r}$ is a vector of numbers); Equation 9.9 uses the chain rule and the fact that $\boldsymbol{\eta} = X\boldsymbol{\beta}$; and, finally, Equation 9.10 uses the fact that $\mathbf{X}$ is constant. Since $\boldsymbol{\beta}_{n+1} = \boldsymbol{\beta}_n + \mathbf{a}$, $\mathbf{H}\boldsymbol{\beta}_{n+1} = \mathbf{H}\boldsymbol{\beta}_n + \mathbf{Ha} = \mathbf{H}\boldsymbol{\beta}_n + \mathbf{u}$, and hence

$$\begin{aligned}
\mathbf{X}^t\tilde{\mathbf{W}}\mathbf{X}\boldsymbol{\beta}_{n+1} &= \mathbf{X}^t\tilde{\mathbf{W}}\mathbf{X}\boldsymbol{\beta}_n + \mathbf{u} \\
&= \mathbf{X}^t\tilde{\mathbf{W}}\boldsymbol{\eta}_n + \mathbf{X}^t\tilde{\mathbf{W}}\frac{d\boldsymbol{\eta}_n}{d\boldsymbol{\mu}}(\mathbf{r} - \boldsymbol{\mu}) \\
&= \mathbf{X}^t\tilde{\mathbf{W}}\left(\boldsymbol{\eta}_n + \frac{d\boldsymbol{\eta}_n}{d\boldsymbol{\mu}}(\mathbf{r} - \boldsymbol{\mu})\right). && (9.11)
\end{aligned}$$

Equation 9.11 is the normal equation for a linear weighted least-squares model of the data $\boldsymbol{\eta}_n + (d\boldsymbol{\eta}_n/d\boldsymbol{\mu})(\mathbf{r} - \boldsymbol{\mu})$ using design matrix $\mathbf{X}$ and weights $\tilde{\mathbf{W}}$.

Note that

$$g(\mu) + (r - \mu)g'(\mu) = \eta + (r - \mu)\frac{d\eta}{d\mu}$$

is a linear approximation to $g(r) = h^{-1}(r)$, and that

$$\mathrm{Var}(g(\mu) + (r - \mu)g'(\mu)) = V(\mu)\left(\frac{d\boldsymbol{\eta}}{d\boldsymbol{\mu}}\right)^2 = \tilde{\mathbf{W}}^{-1}$$

up to a factor involving ϕ.

In order to implement this iteratively re-weighted least squares method we can start by taking $\boldsymbol{\mu} = \mathbf{r}$. Certain observations may need to be adjusted, for example zero values when the log or inverse power links are used. The method is easy to implement in a matrix programming language such as MATLAB, APL or SAS IML. Annotated SAS IML code is given in Appendix B.

10. FUTURE RESEARCH

Bailey [2] points out that in statistics the best estimator is a minimum variance unbiased estimator, but that in classification ratemaking there are typically no unbiased estimators.

Venter's third suggestion, of allowing individual cells to vary from an arithmetically defined base, gives a way to produce unbiased estimators. Credibility weighting the model pure premium with the experience would give asymptotically unbiased rates, because in a large enough sample each cell would be fully credible. Venter notes such an approach was used in the 1981 Massachusetts auto rate hearings. The credibility factor used was Bühlmann credibility

$$Z = \frac{n}{n + K}, \qquad K = \frac{\text{Expected process variance}}{\text{Variance of hypothetical means}}, \tag{10.1}$$

where n is the number of exposures in the cell.

A credibility approach was also hinted at by Bailey, who discusses the problem of combining information about youthful drivers and business classes into youthful business drivers: "[The data] may be insufficient to be fully reliable but it will always provide *some information*."

The statistical theory of mixed models provides a method of credibility weighting fitted values and raw data. The details of mixed models are beyond the scope of this paper; the interested reader should consult Searle, et al. [25]. In fact, Equation 42 on page 57 of Searle uses mixed models to give an unbiased predictor for a cell pure premium as

$$(1-Z)\times \text{model fit} + Z \times \text{cell average},$$

where credibility Z is given by Equation 10.1. A very nice recent paper by Nelder and Verrall [20] extends the same result to a certain family of generalized linear mixed models and discusses some possible actuarial applications. Lee and Nelder [19] gives a more detailed description of the theory, together with some (non-actuarial) examples. Aside from their application to credibility theory, mixed models could also be used in territorial ratemaking, just as they are currently used in geophysical statistics (see Cressie [6]).

11. CONCLUSION

We have introduced generalized linear models by making a connection between them and minimum bias models, with which actuaries are already familiar. The connection is made possible by using variance functions to define linear bias functions and then relating them to the exponential family of distributions. The definitions imply that minimum bias corresponds to the maximum likelihood solution of the associated generalized linear model. By starting with the known and familiar we have provided an introduction to generalized linear models, which is easier to understand than descriptions which start from abstract definitions. We have also explained how generalized linear models extend the well known ANOVA and regression analyses. Two by-products of the exposition were to clarify uniqueness of parameters for class plans, and to explain the different notations used in linear models and minimum bias methods. Finally, the iterative paradigm for solving minimum bias models is shown not to be useful given the more efficient algorithms available for solving generalized linear models. Actuaries should not implement the iterative method. Whenever possible, they should use explicit statistical models.

Linear bias functions are an alternative to the usual measure of bias and so extend Venter's first alternative to Bailey's methods. Link functions, introduced as part of the definition of generalized linear models, allow for more general arithmetic functions to determine classification rates. However, since the models are still linear they do not allow functions such as $r_{ijk} = x_i y_j + z_k$ suggested by Venter.

In jumping from actuaries of the second kind, who use risk theory and probabilistic models, to actuaries of the third kind who use stochastic models and financial tools [see 7, p. 45], I believe the profession may have overlooked an important intermediate step: the statistical actuary—perhaps actuary of the 5/2nds kind? A statistical approach is perfect for data-intensive lines, such as personal auto and homeowners. I hope this and other statistical papers which have appeared recently will encourage actuaries working in data-intensive lines to take statistics beyond that which is required for an Associateship in either North American actuarial society, and to start taking advantage of its power in their work.

REFERENCES

[1] Abraham, Bovas, and Johannes Ledolter, *Statistical Methods for Forecasting*, John Wiley and Sons, New York, 1983.

[2] Bailey, Robert A., "Insurance Rates with Minimum Bias," *PCAS* L, pp. 4–13.

[3] Bailey, Robert A., and LeRoy J. Simon, "Two Studies in Automobile Insurance Ratemaking," *PCAS* XLVII, pp. 1–19.

[4] Borwein, Jonathan M., and Peter B. Borwein, *Pi and the AGM*, John Wiley and Sons, New York, 1987.

[5] Brown, Robert L., "Minimum Bias with Generalized Linear Models," *PCAS* LXXV, pp. 187–217.

[6] Cressie, Noel A. C., *Statistics for Spatial Data*, Revised Edition, John Wiley and Sons, New York, 1993.

[7] D'Arcy, Stephen, "On Becoming an Actuary of the Third Kind," *PCAS* LXXVI, pp. 45–76.

[8] Graves, Nancy C., and Richard Castillo, *Commercial General Liability Ratemaking for Premises and Operations*, 1990 CAS Discussion Paper Program II, pp. 631–696.

[9] Haberman, Steven, and A. R. Renshaw, "Generalized Linear Models and Actuarial Science," *The Statistician* 45, 4, 1996, pp. 407–436.

[10] Hogg, Robert V., and Stuart A. Klugman, *Loss Distributions*, John Wiley and Sons, New York, 1983.

[11] Minutes of Personal Lines Advisory Panel, *Personal Auto Classification Plan Review*, Insurance Services Office PLAP-96-18, 1996.

[12] Johnson, Norman L., Samuel Kotz, and N. Balakrishnan, *Continuous Univariate Distributions, Volume 1*, Second Edition, John Wiley and Sons, New York, 1994.

[13] Jørgensen, Bent, *The Theory of Dispersion Models*, Chapman and Hall, London, 1997.

[14] Jørgensen, Bent, "Exponential Dispersion Models," *J. R. Statist. Soc. B* 49, 2, 1987, pp. 127–162.

[15] Jørgensen, Bent, and Marta C. Paes de Souza, "Fitting Tweedie's Compound Poisson Model to Insurance Claims Data," *Scand. Actuarial J.* 1, 1994, pp. 69–93.

[16] Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot, *Loss Models, From Data to Decisions*, John Wiley and Sons, New York, 1998.

[17] McCullagh, P., and J. A. Nelder, *Generalized Linear Models*, Second Edition, Chapman and Hall, London, 1989.

[18] Lee, Y., and J. A. Nelder, "Likelihood, Quasi-Likelihood and Pseudolikelihood: Some Comparisons," *J. R. Statist. Soc. B* 54, 1992, pp. 273–284.

[19] Lee, Y., and J. A. Nelder, "Hierarchical Generalized Linear Models," *J. R. Statist. Soc. B* 58, 1996, pp. 619–678.

[20] Nelder, J. A., and R. J. Verrall, "Credibility Theory and Generalized Linear Models," *ASTIN Bulletin* 27, (1), pp. 71–82.

[21] Panjer, H. H., and G. E. Willmot, *Insurance Risk Models*, Society of Actuaries, Schaumburg, IL, 1992.

[22] Rao, C. Radhakrishna, *Linear Statistical Inference and Its Applications*, Second Edition, John Wiley and Sons, New York, 1973.

[23] Renshaw, A. E., "Modeling the Claims Process in the Presence of Covariates," *ASTIN Bulletin* 24, 2, pp. 265–286.

[24] SAS/STAT Software, Changes and Enhancements through Release 6.11, SAS Institute Inc., Cary, NC, 1996.

[25] Searle S. R., George Casella, and C. E. McCulloch, *Variance Components*, John Wiley and Sons, New York, 1992.

[26] Venter, Gary G., "Discussion of Minimum Bias with Generalized Linear Models," *PCAS* LXXVII, pp. 337–349.

[27] Wright, Thomas S., "Stochastic Claims Reserving When Past Claim Numbers are Known," *PCAS* LXXIX, pp. 255–361.

EXHIBIT 1

UNDERLYING DATA FOR NUMERICAL EXAMPLES

Observation	Age Group	Vehicle-Use	Severity	Claim Count
1	17–20	Pleasure	250.48	21
2	17–20	Drive to Work < 10 miles	274.78	40
3	17–20	Drive to Work > 10 miles	244.52	23
4	17–20	Business	797.80	5
5	21–24	Pleasure	213.71	63
6	21–24	Drive to Work < 10 miles	298.60	171
7	21–24	Drive to Work > 10 miles	298.13	92
8	21–24	Business	362.23	44
9	25–29	Pleasure	250.57	140
10	25–29	Drive to Work < 10 miles	248.56	343
11	25–29	Drive to Work > 10 miles	297.90	318
12	25–29	Business	342.31	129
13	30–34	Pleasure	229.09	123
14	30–34	Drive to Work < 10 miles	228.48	448
15	30–34	Drive to Work > 10 miles	293.87	361
16	30–34	Business	367.46	169
17	35–39	Pleasure	153.62	151
18	35–39	Drive to Work < 10 miles	201.67	479
19	35–39	Drive to Work > 10 miles	238.21	381
20	35–39	Business	256.21	166
21	40–49	Pleasure	208.59	245
22	40–49	Drive to Work < 10 miles	202.80	970
23	40–49	Drive to Work > 10 miles	236.06	719
24	40–49	Business	352.49	304
25	50–59	Pleasure	207.57	266
26	50–59	Drive to Work < 10 miles	202.67	859
27	50–59	Drive to Work > 10 miles	253.63	504
28	50–59	Business	340.56	162
29	60+	Pleasure	192.00	260
30	60+	Drive to Work < 10 miles	196.33	578
31	60+	Drive to Work > 10 miles	259.79	312
32	60+	Business	342.58	96

EXHIBIT 2

ONE-WAY SUMMARY OF UNDERLYING DATA

Age Group	Vehicle-Use	Severity	Claim Count
All	All	241.46	8,942
All	Pleasure	206.00	1,269
All	Drive to Work $<$ 10 miles	213.62	3,888
All	Drive to Work $>$ 10 miles	259.50	2,710
All	Business	338.54	1,075
17–20	All	290.61	89
21–24	All	291.60	370
25–29	All	278.74	930
30–34	All	271.32	1,101
35–39	All	215.03	1,177
40–49	All	234.45	2,238
50–59	All	230.21	1,791
60+	All	222.59	1,246

EXHIBIT 3

DESIGN MATRIX FOR AGE GROUP CLASSIFICATION

	Age Group							
Observation	17–20	21–24	25–29	30–34	35–39	40–49	50–59	60+
1	1	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0
5	0	1	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0
7	0	1	0	0	0	0	0	0
8	0	1	0	0	0	0	0	0
9	0	0	1	0	0	0	0	0
10	0	0	1	0	0	0	0	0
11	0	0	1	0	0	0	0	0
12	0	0	1	0	0	0	0	0
13	0	0	0	1	0	0	0	0
14	0	0	0	1	0	0	0	0
15	0	0	0	1	0	0	0	0
16	0	0	0	1	0	0	0	0
17	0	0	0	0	1	0	0	0
18	0	0	0	0	1	0	0	0
19	0	0	0	0	1	0	0	0
20	0	0	0	0	1	0	0	0
21	0	0	0	0	0	1	0	0
22	0	0	0	0	0	1	0	0
23	0	0	0	0	0	1	0	0
24	0	0	0	0	0	1	0	0
25	0	0	0	0	0	0	1	0
26	0	0	0	0	0	0	1	0
27	0	0	0	0	0	0	1	0
28	0	0	0	0	0	0	1	0
29	0	0	0	0	0	0	0	1
30	0	0	0	0	0	0	0	1
31	0	0	0	0	0	0	0	1
32	0	0	0	0	0	0	0	1

EXHIBIT 4

Design Matrix for Vehicle-Use Classification

	Vehicle-Use Classification		
Observation	Drive to Work < 10 miles	Drive to Work > 10 miles	Business
1	0	0	0
2	1	0	0
3	0	1	0
4	0	0	1
5	0	0	0
6	1	0	0
7	0	1	0
8	0	0	1
9	0	0	0
10	1	0	0
11	0	1	0
12	0	0	1
13	0	0	0
14	1	0	0
15	0	1	0
16	0	0	1
17	0	0	0
18	1	0	0
19	0	1	0
20	0	0	1
21	0	0	0
22	1	0	0
23	0	1	0
24	0	0	1
25	0	0	0
26	1	0	0
27	0	1	0
28	0	0	1
29	0	0	0
30	1	0	0
31	0	1	0
32	0	0	1

EXHIBIT 5

Minimum Bias Iterations for Bailey's Additive Model

(1) Itn	(2) Len	(3) 17–20	(4) 21–24	(5) 25–29	(6) 30–34	(7) 35–39	(8) 40–49	(9) 50–64	(10) 64+	(11) DTW < 10	(12) DTW > 10	(13) Business	(14) Chg Len	(15) v(1)	(16) v(2)	(17) v(3)
1	100.35	290.61	291.60	278.74	271.32	215.02	234.45	230.21	222.59	−26.98	17.41	95.08	100.35	−0.2688	0.17346	0.94743
2	9.22740	292.89	288.43	269.55	262.00	206.96	227.64	229.65	223.42	−22.29	22.76	100.95	0.09195	0.50853	0.58051	0.63593
3	7.59761	289.07	284.23	265.17	257.43	202.49	223.09	225.36	219.45	−17.93	27.15	105.36	0.82337	0.57392	0.57752	0.58060
4	6.52655	285.73	280.60	261.45	253.54	198.67	219.19	221.64	215.99	−14.18	30.92	109.14	0.85903	0.57473	0.57744	0.57987
5	5.60918	282.86	277.48	258.25	250.20	195.39	215.84	218.44	213.01	−10.96	34.16	112.39	0.85944	0.57474	0.57744	0.57986
6	4.82078	280.39	274.80	255.51	247.33	192.57	212.96	215.69	210.45	−8.18	36.94	115.19	0.85944	0.57474	0.57744	0.57986
7	4.14319	278.27	272.50	253.15	244.86	190.15	210.49	213.32	208.26	−5.80	39.34	117.59	0.85944	0.57474	0.57744	0.57986
8	3.56084	276.44	270.51	251.12	242.73	188.07	208.36	211.29	206.37	−3.76	41.39	119.66	0.85944	0.57474	0.57744	0.57986
9	3.06035	274.88	268.81	249.37	240.91	186.28	206.53	209.54	204.74	−2.00	43.16	121.43	0.85944	0.57474	0.57744	0.57986
10	2.63020	273.53	267.35	247.87	239.34	184.74	204.96	208.04	203.35	−0.49	44.68	122.96	0.85944	0.57474	0.57744	0.57986
11	2.26051	272.37	266.09	246.58	238.00	183.42	203.61	206.75	202.15	0.81	45.98	124.27	0.85944	0.57474	0.57744	0.57986
12	1.94279	271.38	265.01	245.48	236.84	182.28	202.45	205.64	201.12	1.93	47.10	125.39	0.85944	0.57474	0.57744	0.57986
13	1.66972	270.52	264.08	244.53	235.84	181.31	201.45	204.69	200.23	2.89	48.07	126.36	0.85944	0.57474	0.57744	0.57986
14	1.43503	269.79	263.28	243.71	234.99	180.47	200.59	203.87	199.47	3.71	48.90	127.19	0.85944	0.57474	0.57744	0.57986
15	1.23333	269.16	262.60	243.00	234.25	179.75	199.86	203.16	198.82	4.42	49.61	127.91	0.85944	0.57474	0.57744	0.57986
16	1.05998	268.61	262.01	242.40	233.62	179.13	199.22	202.56	198.26	5.03	50.22	128.52	0.85944	0.57474	0.57744	0.57986
17	0.91099	268.15	261.50	241.88	233.08	178.59	198.68	202.04	197.77	5.56	50.75	129.05	0.85944	0.57474	0.57744	0.57986
18	0.78295	267.75	261.06	241.43	232.61	178.14	198.21	201.59	197.36	6.01	51.20	129.51	0.85944	0.57474	0.57744	0.57986
19	0.67290	267.40	260.69	241.05	232.21	177.74	197.81	201.21	197.00	6.39	51.59	129.90	0.85944	0.57474	0.57744	0.57986
20	0.57832	267.10	260.37	240.72	231.87	177.40	197.46	200.88	196.69	6.72	51.92	130.23	0.85944	0.57474	0.57744	0.57986
21	0.49703	266.85	260.09	240.44	231.57	177.11	197.17	200.59	196.43	7.01	52.21	130.52	0.85944	0.57474	0.57744	0.57986
22	0.42717	266.63	259.85	240.19	231.32	176.86	196.91	200.35	196.20	7.26	52.46	130.77	0.85944	0.57474	0.57744	0.57986
23	0.36713	266.44	259.65	239.99	231.10	176.65	196.69	200.14	196.01	7.47	52.67	130.98	0.85944	0.57474	0.57744	0.57986
24	0.31553	266.28	259.47	239.81	230.91	176.46	196.50	199.96	195.84	7.65	52.85	131.16	0.85944	0.57474	0.57744	0.57986
25	0.27118	266.14	259.32	239.65	230.75	176.31	196.34	199.81	195.70	7.80	53.01	131.32	0.85944	0.57474	0.57744	0.57986
26	0.23306	266.02	259.19	239.52	230.61	176.17	196.20	199.67	195.57	7.94	53.14	131.46	0.85944	0.57474	0.57744	0.57986
27	0.20031	265.92	259.08	239.40	230.49	176.05	196.08	199.56	195.47	8.05	53.26	131.57	0.85944	0.57474	0.57744	0.57986

28	0.17215	265.83	258.99	239.31	230.39	175.95	195.98	199.46	195.38	8.15	53.36	131.67	0.85944	0.57474	0.57744	0.57986
29	0.14796	265.76	258.90	239.22	230.30	175.86	195.89	199.38	195.30	8.24	53.44	131.76	0.85944	0.57474	0.57744	0.57986
30	0.12716	265.69	258.83	239.15	230.22	175.79	195.82	199.30	195.23	8.31	53.51	131.83	0.85944	0.57474	0.57744	0.57986
31	0.10929	265.64	258.77	239.09	230.16	175.73	195.75	199.24	195.17	8.37	53.58	131.89	0.85944	0.57474	0.57744	0.57986
32	0.09393	265.59	258.72	239.03	230.10	175.67	195.70	199.19	195.12	8.43	53.63	131.95	0.85944	0.57474	0.57744	0.57986
33	0.08072	265.55	258.68	238.99	230.05	175.62	195.65	199.14	195.08	8.47	53.68	132.00	0.85944	0.57474	0.57744	0.57986
34	0.06938	265.51	258.64	238.95	230.01	175.58	195.61	199.10	195.04	8.51	53.72	132.04	0.85944	0.57474	0.57744	0.57986
35	0.05963	265.48	258.60	238.91	229.98	175.55	195.57	199.07	195.01	8.55	53.75	132.07	0.85944	0.57474	0.57744	0.57986
36	0.05125	265.45	258.58	238.88	229.95	175.52	195.54	199.04	194.98	8.58	53.78	132.10	0.85944	0.57474	0.57744	0.57986
37	0.04404	265.43	258.55	238.86	229.92	175.49	195.51	199.01	194.96	8.60	53.81	132.13	0.85944	0.57474	0.57744	0.57986
38	0.03785	265.41	258.53	238.84	229.90	175.47	195.49	198.99	194.94	8.62	53.83	132.15	0.85944	0.57474	0.57744	0.57986
39	0.03253	265.40	258.51	238.82	229.88	175.45	195.47	198.97	194.92	8.64	53.85	132.17	0.85944	0.57474	0.57744	0.57986
40	0.02796	265.38	258.50	238.80	229.86	175.44	195.45	198.96	194.91	8.66	53.87	132.18	0.85944	0.57474	0.57744	0.57986
41	0.02403	265.37	258.48	238.79	229.85	175.42	195.44	198.94	194.90	8.67	53.88	132.20	0.85944	0.57474	0.57744	0.57986
42	0.02065	265.36	258.47	238.78	229.83	175.41	195.43	198.93	194.89	8.68	53.89	132.21	0.85944	0.57474	0.57744	0.57986
43	0.01775	265.35	258.46	238.77	229.82	175.40	195.42	198.92	194.88	8.69	53.90	132.22	0.85944	0.57474	0.57744	0.57986
44	0.01525	265.34	258.45	238.76	229.81	175.39	195.41	198.91	194.87	8.70	53.91	132.23	0.85944	0.57474	0.57744	0.57986
45	0.01311	265.33	258.45	238.75	229.81	175.38	195.40	198.91	194.86	8.71	53.92	132.24	0.85944	0.57474	0.57744	0.57986
46	0.01127	265.33	258.44	238.75	229.80	175.38	195.39	198.90	194.85	8.72	53.92	132.24	0.85944	0.57474	0.57744	0.57986
47	0.00968	265.32	258.43	238.74	229.79	175.37	195.39	198.89	194.85	8.72	53.93	132.25	0.85944	0.57474	0.57744	0.57986
48	0.00832	265.32	258.43	238.73	229.79	175.37	195.38	198.89	194.85	8.73	53.93	132.25	0.85944	0.57474	0.57744	0.57986
49	0.00715	265.32	258.43	238.73	229.78	175.36	195.38	198.88	194.84	8.73	53.94	132.26	0.85944	0.57474	0.57744	0.57986
50	0.00615	265.31	258.42	238.73	229.78	175.36	195.37	198.88	194.84	8.74	53.94	132.26	0.85944	0.57474	0.57744	0.57986

EXHIBIT 6

General Linear Model Parameters

Parameter	Value
Age 17–20	265.29
Age 21–24	258.40
Age 25–29	238.71
Age 30–34	229.76
Age 35–39	175.34
Age 40–49	195.35
Age 50–59	198.86
Age 60+	194.82
Drive to Work $<$ 10 miles	8.76
Drive to Work $>$ 10 miles	53.96
Business	132.28

EXHIBIT 7-1

PARAMETER VALUES AND STATISTICS FOR NORMAL MODEL, IDENTITY LINK

Parameter	Level	Estimate	Standard Error	Chi Squared	*p* value
Age	17–20	265.29	31.536	70.769	0.000
Age	21–24	258.40	16.797	236.658	0.000
Age	25–29	238.71	12.144	386.375	0.000
Age	30–34	229.76	11.773	380.880	0.000
Age	35–39	175.34	11.459	234.139	0.000
Age	40–49	195.35	9.992	382.258	0.000
Age	50–59	198.86	10.197	380.313	0.000
Age	60+	194.82	10.814	324.582	0.000
Vehicle-Use	DTW < 10	8.76	9.418	0.865	0.353
Vehicle-Use	DTW > 10	53.96	9.936	29.498	0.000
Vehicle-Use	Business	132.28	12.124	119.041	0.000
Scale		291.56	36.32		

EXHIBIT 7-2

PARAMETER VALUES AND STATISTICS FOR NORMAL MODEL, LOG LINK

Parameter	Level	Estimated	Transformed Estimate	Standard Error	Chi Squared	*p* value
Age	17–20	5.581	265.22	0.108	2,650.726	0.000
Age	21–24	5.514	248.21	0.063	7,611.630	0.000
Age	25–29	5.444	231.37	0.051	11,339.710	0.000
Age	30–34	5.421	226.18	0.050	11,635.699	0.000
Age	35–39	5.186	178.76	0.055	9,038.202	0.000
Age	40–49	5.289	198.19	0.047	12,821.908	0.000
Age	50–59	5.301	200.49	0.048	12,326.811	0.000
Age	60+	5.286	197.55	0.051	10,887.200	0.000
Vehicle-Use	DTW < 10	0.041	1.04	0.045	0.822	0.365
Vehicle-Use	DTW > 10	0.231	1.26	0.045	26.090	0.000
Vehicle-Use	Business	0.495	1.64	0.048	107.072	0.000
Scale		291.75		36.47		

EXHIBIT 7-3

PARAMETER VALUES AND STATISTICS FOR NORMAL MODEL, INVERSE LINK

Parameter	Level	Estimated	Transformed Estimate	Standard Error	Chi Squared	p value
Age	17–20	3.7615e-03	265.85	0.0004	110.100	0.000
Age	21–24	4.2575e-03	234.88	0.0003	273.197	0.000
Age	25–29	4.4685e-03	223.79	0.0002	376.913	0.000
Age	30–34	4.5015e-03	222.15	0.0002	394.962	0.000
Age	35–39	5.4337e-03	184.04	0.0003	433.329	0.000
Age	40–49	4.9521e-03	201.93	0.0002	498.633	0.000
Age	50–59	4.9256e-03	203.02	0.0002	470.954	0.000
Age	60+	4.9756e-03	200.98	0.0002	432.339	0.000
Vehicle-Use	DTW < 10	−1.8560e-04	N/A	0.0002	0.690	0.406
Vehicle-Use	DTW > 10	−9.7374e-04	N/A	0.0002	20.430	0.000
Vehicle-Use	Business	−1.8592e-03	N/A	0.0002	75.130	0.000
Scale		304.39		38.05		

EXHIBIT 7-4

PARAMETER VALUES AND STATISTICS FOR GAMMA MODEL, IDENTITY LINK

Parameter	Level	Estimate	Standard Error	Chi Squared	p value
Age	17–20	257.79	29.673	75.475	0.000
Age	21–24	261.08	15.670	277.578	0.000
Age	25–29	241.05	10.114	568.048	0.000
Age	30–34	228.18	9.442	584.056	0.000
Age	35–39	179.60	8.126	488.552	0.000
Age	40–49	194.89	7.016	771.627	0.000
Age	50–59	198.46	7.195	760.810	0.000
Age	60+	193.04	7.600	645.146	0.000
Vehicle-Use	DTW < 10	8.63	6.571	1.727	0.189
Vehicle-Use	DTW > 10	53.74	7.482	51.590	0.000
Vehicle-Use	Business	131.44	11.629	127.745	0.000
Scale		1.03	0.256		

EXHIBIT 7-5

Parameter Values and Statistics for Gamma Model, Log Link

Parameter	Level	Estimate	Transformed Estimate	Standard Error	Chi Squared	p value
Age	17–20	5.541	254.89	0.108	2,624.365	0.000
Age	21–24	5.536	253.70	0.058	9,118.370	0.000
Age	25–29	5.460	235.18	0.041	17,358.976	0.000
Age	30–34	5.418	225.37	0.040	18,045.458	0.000
Age	35–39	5.201	181.47	0.040	17,179.268	0.000
Age	40–49	5.280	196.33	0.034	23,972.547	0.000
Age	50–59	5.295	199.34	0.035	23,125.662	0.000
Age	60+	5.273	195.00	0.037	20,170.519	0.000
Vehicle-Use	DTW < 10	0.041	1.04	0.032	1.610	0.205
Vehicle-Use	DTW > 10	0.234	1.26	0.034	47.306	0.000
Vehicle-Use	Business	0.497	1.64	0.042	143.183	0.000
Scale		1.007		0.25		

EXHIBIT 7-6

Parameter Values and Statistics for Gamma Model, Inverse Link

Parameter	Level	Estimate	Transformed Estimate	Standard Error	Chi Squared	p value
Age	17–20	3.9881e-03	250.75	0.000	97.894	0.000
Age	21–24	4.1205e-03	242.69	0.000	319.554	0.000
Age	25–29	4.3830e-03	228.15	0.000	557.756	0.000
Age	30–34	4.5016e-03	222.14	0.000	598.102	0.000
Age	35–39	5.4096e-03	184.86	0.000	733.982	0.000
Age	40–49	5.0241e-03	199.04	0.000	867.415	0.000
Age	50–59	4.9727e-03	201.10	0.000	813.669	0.000
Age	60+	5.0559e-03	197.79	0.000	736.713	0.000
Vehicle-Use	DTW < 10	−1.8995e-04	N/A	0.000	1.312	0.252
Vehicle-Use	DTW > 10	−1.0005e-03	N/A	0.000	36.478	0.000
Vehicle-Use	Business	−1.8767e-03	N/A	0.000	115.193	0.000
Scale		8.7803e-01		0.2189		

EXHIBIT 7-7

Parameter Values and Statistics Inverse Gaussian Model, Identity Link

Parameter	Level	Estimate	Standard Error	Chi Squared	p value
Age	17–20	255.91	30.274	71.460	0.000
Age	21–24	261.83	16.277	258.748	0.000
Age	25–29	241.72	9.987	585.880	0.000
Age	30–34	227.34	9.032	633.606	0.000
Age	35–39	180.52	7.341	604.698	0.000
Age	40–49	194.90	6.246	973.654	0.000
Age	50–59	198.27	6.428	951.478	0.000
Age	60+	192.28	6.747	812.169	0.000
Vehicle-Use	DTW $<$ 10	8.72	5.862	2.211	0.137
Vehicle-Use	DTW $>$ 10	53.77	7.000	59.003	0.000
Vehicle-Use	Business	131.24	12.620	108.154	0.000
Scale		0.0616	0.0077		

EXHIBIT 7-8

Parameter Values and Statistics for Inverse Gaussian Model, Log Link

Parameter	Level	Estimate	Transformed Estimate	Standard Error	Chi Squared	p value
Age	17–20	5.532	252.65	0.112	2,440.618	0.000
Age	21–24	5.544	255.68	0.060	8,532.551	0.000
Age	25–29	5.466	236.62	0.040	18,319.266	0.000
Age	30–34	5.416	224.87	0.039	19,731.991	0.000
Age	35–39	5.205	182.20	0.036	20,832.767	0.000
Age	40–49	5.277	195.85	0.031	29,507.021	0.000
Age	50–59	5.293	198.91	0.031	28,397.992	0.000
Age	60+	5.268	193.96	0.033	24,828.173	0.000
Vehicle-Use	DTW $<$ 10	0.041	1.04	0.029	2.025	0.155
Vehicle-Use	DTW $>$ 10	0.236	1.27	0.032	55.575	0.000
Vehicle-Use	Business	0.499	1.65	0.043	134.646	0.000
Scale		0.062		0.008		

EXHIBIT 7-9

Parameter Values and Statistics for Inverse Gaussian Model, Inverse Link

Parameter	Level	Estimate	Transformed Estimate	Standard Error	Chi Squared	p value
Age	17–20	4.0365e-03	247.74	4.2843e-04	88.766	0.000
Age	21–24	4.0590e-03	246.36	2.3219e-04	305.595	0.000
Age	25–29	4.3454e-03	230.13	1.7757e-04	598.851	0.000
Age	30–34	4.5071e-03	221.87	1.7487e-04	664.311	0.000
Age	35–39	5.4073e-03	184.94	1.8031e-04	899.289	0.000
Age	40–49	5.0537e-03	197.87	1.5568e-04	1,053.840	0.000
Age	50–59	4.9939e-03	200.24	1.5895e-04	987.095	0.000
Age	60+	5.0932e-03	196.34	1.6946e-04	903.299	0.000
Vehicle-Use	DTW < 10	−1.9182e-04	N/A	1.4885e-04	1.661	0.198
Vehicle-Use	DTW > 10	−1.0146e-03	N/A	1.5237e-04	44.340	0.000
Vehicle-Use	Business	−1.9018e-03	N/A	1.7072e-04	124.096	0.000
Scale		6.6167e-02		8.2708e-03		

EXHIBIT 7-10

Parameter Values and Statistics for Inverse Gaussian Model, Inverse Square Link

Parameter	Level	Estimate	Transformed Estimate	Standard Error	Chi Squared	p value
Age	17–20	1.7319e-05	240.29	3.1178e-06	30.858	0.000
Age	21–24	1.8382e-05	233.24	1.9533e-06	88.561	0.000
Age	25–29	2.0061e-05	223.27	1.6907e-06	140.795	0.000
Age	30–34	2.0853e-05	218.98	1.6894e-06	152.372	0.000
Age	35–39	2.8057e-05	188.79	1.9110e-06	215.555	0.000
Age	40–49	2.4743e-05	201.04	1.6212e-06	232.938	0.000
Age	50–59	2.4391e-05	202.48	1.6577e-06	216.488	0.000
Age	60+	2.5133e-05	199.47	1.7745e-06	200.608	0.000
Vehicle-Use	DTW < 10	−1.7550e-06	N/A	1.6060e-06	1.194	0.275
Vehicle-Use	DTW > 10	−8.6033e-06	N/A	1.5708e-06	30.000	0.000
Vehicle-Use	Business	−1.4323e-05	N/A	1.5899e-06	81.153	0.000
Scale		0.0747		9.3373e-03		

EXHIBIT 8

Parameter Values and Statistics for Generalized Linear Model With Log Link and Normal Errors, and General Linear Model Applied to Log Responses

(1) Age	(2) Vehicle-Use	(3) Severity	(4) Claim Count	(5) Normal/ Identity Link	(6) General Linear Model on Log(severity)	(7) Normal/ Log Link
17–20	Pleasure	250.48	21	265.29	248.57	265.22
17–20	DTW < 15	274.78	40	274.05	259.50	276.34
17–20	DTW > 15	244.52	23	319.26	314.74	334.23
17–20	Business	797.80	5	397.58	407.54	435.21
21–24	Pleasure	213.71	63	258.40	251.48	248.21
21–24	DTW < 15	298.60	171	267.16	262.54	258.61
21–24	DTW > 15	298.13	92	312.37	318.43	312.79
21–24	Business	362.23	44	390.68	412.31	407.29
25–29	Pleasure	250.57	140	238.71	234.64	231.37
25–29	DTW < 15	248.56	343	247.46	244.96	241.06
25–29	DTW > 15	297.90	318	292.67	297.10	291.57
25–29	Business	342.31	129	370.99	384.70	379.65
30–34	Pleasure	229.09	123	229.76	225.07	226.18
30–34	DTW < 15	228.48	448	238.52	234.97	235.66
30–34	DTW > 15	293.87	361	283.72	284.98	285.04
30–34	Business	367.46	169	362.04	369.01	371.15
35–39	Pleasure	153.62	151	175.34	180.50	178.76
35–39	DTW < 15	201.67	479	184.09	188.44	186.25
35–39	DTW > 15	238.21	381	229.30	228.55	225.27
35–39	Business	256.21	166	307.62	295.94	293.33
40–49	Pleasure	208.59	245	195.35	195.89	198.19
40–49	DTW < 15	202.80	970	204.11	204.50	206.50
40–49	DTW > 15	236.06	719	249.32	248.04	249.76
40–49	Business	352.49	304	327.63	321.17	325.22
50–59	Pleasure	207.57	266	198.86	199.02	200.49
50–59	DTW < 15	202.67	859	207.62	207.77	208.90
50–59	DTW > 15	253.63	504	252.82	252.00	252.66
50–59	Business	340.56	162	331.14	326.30	328.99
60+	Pleasure	192.00	260	194.82	194.61	197.55
60+	DTW < 15	196.33	578	203.58	203.17	205.83
60+	DTW > 15	259.79	312	248.78	246.42	248.95
60+	Business	342.58	96	327.10	319.08	324.16

EXHIBIT 9

MODEL SEVERITIES

				Model Number, Distribution, and Link									
Age	Vehicle Use	Claim Count	Severity	1 Normal Identity	2 Normal Log	3 Normal Inverse	4 Gamma Identity	5 Gamma Log	6 Gamma Inverse	7 Inv Gs Identity	8 Inv Gs Log	9 Inv Gs Inverse	10 Inv Gs Inv Sqr
17–20	Pleasure	21	250.48	265.29	265.22	265.85	257.79	254.89	250.75	255.91	252.65	247.74	240.29
17–20	DTW < 10M	40	274.78	274.05	276.34	279.65	266.42	265.56	263.29	264.63	263.29	260.1	253.47
17–20	DTW > 10M	23	244.52	319.26	334.23	358.71	311.53	322.17	334.72	309.68	319.82	330.92	338.72
17–20	Business	5	797.80	397.58	435.21	525.67	389.23	419.06	473.63	387.15	416.17	468.44	577.68
21–24	Pleasure	63	213.71	258.40	248.21	234.88	261.08	253.7	242.69	261.83	255.68	246.36	233.24
21–24	DTW < 10M	171	298.60	267.16	258.61	245.58	269.71	264.32	254.42	270.54	266.44	258.58	245.24
21–24	DTW > 10M	92	298.13	312.37	312.79	304.53	314.82	320.66	320.51	315.59	323.65	328.47	319.79
21–24	Business	44	362.23	390.68	407.29	416.96	392.51	417.10	445.67	393.06	421.16	463.55	496.35
25–29	Pleasure	140	250.57	238.71	231.37	223.79	241.05	235.18	228.15	241.72	236.62	230.13	223.27
25–29	DTW < 10M	343	248.56	247.46	241.06	233.49	249.69	245.02	238.49	250.44	246.58	240.75	233.73
25–29	DTW > 10M	318	297.90	292.67	291.57	286.14	294.80	297.26	295.64	295.49	299.52	300.23	295.43
25–29	Business	129	342.31	370.99	379.65	383.25	372.49	386.66	398.99	372.96	389.77	409.22	417.46
30–34	Pleasure	123	229.09	229.76	226.18	222.15	228.18	225.37	222.14	227.34	224.87	221.87	218.98
30–34	DTW < 10M	448	228.48	238.52	235.66	231.70	236.82	234.80	231.93	236.06	234.33	231.74	228.82
30–34	DTW > 10M	361	293.87	283.72	285.03	283.47	281.93	284.85	285.63	281.11	284.64	286.33	285.72
30–34	Business	169	367.46	362.04	371.15	378.46	359.62	370.52	380.97	358.58	370.41	383.83	391.31
35–39	Pleasure	151	153.62	175.34	178.76	184.04	179.6	181.47	184.86	180.52	182.2	184.94	188.79
35–39	DTW < 10M	479	201.67	184.09	186.25	190.54	188.23	189.06	191.59	189.24	189.87	191.74	194.99

EXHIBIT 9

MODEL SEVERITIES

(Continued)

				Model Number, Distribution, and Link									
Age	Vehicle Use	Claim Count	Severity	1 Normal Identity	2 Normal Log	3 Normal Inverse	4 Gamma Identity	5 Gamma Log	6 Gamma Inverse	7 Inv Gs Identity	8 Inv Gs Log	9 Inv Gs Inverse	10 Inv Gs Inv Sqr
35–39	DTW > 10M	381	238.21	229.30	225.27	224.22	233.34	229.37	226.81	234.29	230.64	227.65	226.73
35–39	Business	166	256.20	307.62	293.33	279.76	311.04	298.35	283.06	311.76	300.13	285.26	269.84
40–49	Pleasure	245	208.59	195.35	198.19	201.93	194.89	196.33	199.04	194.90	195.85	197.87	201.04
40–49	DTW < 10M	970	202.80	204.11	206.50	209.80	203.52	204.54	206.86	203.62	204.09	205.68	208.57
40–49	DTW > 10M	719	236.06	249.32	249.76	251.36	248.63	248.15	248.54	248.67	247.91	247.58	248.91
40–49	Business	304	352.49	327.63	325.22	323.32	326.33	322.78	317.73	326.14	322.61	317.26	309.78
50–59	Pleasure	266	207.57	198.86	200.49	203.02	198.46	199.34	201.1	198.27	198.91	200.24	202.48
50–59	DTW < 10M	859	202.67	207.62	208.90	210.97	207.10	207.68	209.09	206.98	207.29	208.24	210.18
50–59	DTW > 10M	504	253.63	252.82	252.66	253.05	252.20	251.95	251.75	252.03	251.79	251.30	251.68
50–59	Business	162	340.56	331.14	328.99	326.12	329.90	327.72	323.00	329.51	327.66	323.40	315.15
60+	Pleasure	260	192.00	194.82	197.55	200.98	193.04	195.00	197.79	192.28	193.96	196.34	199.47
60+	DTW < 10M	578	196.33	203.58	205.83	208.77	201.67	203.16	205.51	200.99	202.12	204.02	206.82
60+	DTW > 10M	312	259.79	248.78	248.95	249.88	246.78	246.47	246.59	246.04	245.52	245.18	245.96
60+	Business	96	342.58	327.10	324.16	320.88	324.48	320.6	314.55	323.52	319.5	313.34	304.14

EXHIBIT 10

AVERAGE BIAS BY MODEL

Age Class	Vehicle Use	Normal Identity	Normal Log	Normal Inverse	Gamma Identity	Gamma Log	Gamma Inverse	Inv Gs Identity	Inv Gs Log	Inv Gs Inverse	Inv Gs Inv Sqr
All	All	0	−0.03	−0.13	0.02	−0.04	0	0.04	−0.14	−0.25	0
All	Business	0	0.15	1.11	0.43	−0.50	0	0.61	−1.12	−2.55	0
All	DTW < 10M	0	−0.23	−0.82	0	0.18	0	−0.02	0.28	0.35	0
All	DTW > 10M	0	0.27	0.74	−0.07	−0.26	0	−0.08	−0.56	−0.49	0
All	Pleasure	0	−0.19	−0.90	−0.08	0.16	0	0	0.30	0.37	0
17–20	All.	0	−6.99	−20.04	7.67	4.31	0	9.51	6.63	3.42	0
21–24	All	0	3.61	12.80	−2.46	−3.09	0	−3.23	−5.63	−6.66	0
25–29	All	0	2.64	7.93	−2.11	−2.31	0	−2.79	−4.31	−4.12	0
30–34	All	0	−0.27	1.18	1.83	0.33	0	2.66	0.66	−0.56	0
35–39	All	0	2.00	1.83	−4.02	−1.52	0	−4.96	−2.61	−0.66	0
40–49	All	0	−1.16	−3.25	0.70	0.74	0	0.67	1.09	1.01	0
50–59	All	0	−0.62	−1.84	0.59	0.46	0	0.76	0.76	0.62	0
60+	All	0	−1.43	−3.49	1.96	1.23	0	2.69	2.25	1.44	0
17–20	Business	400.22	362.59	272.13	408.57	378.74	324.17	410.65	381.63	329.36	220.12
17–20	DTW < 10M	0.72	−1.56	−4.87	8.35	9.22	11.49	10.14	11.49	14.68	21.30
17–20	DTW > 10M	−74.74	−89.71	−114.20	−67.01	−77.65	−90.20	−65.16	−75.29	−86.40	−94.20
17–20	Pleasure	−14.82	−14.75	−15.37	−7.31	−4.42	−0.27	−5.44	−2.18	2.74	10.19
21–24	Business	−28.46	−45.06	−54.73	−30.29	−54.87	−83.45	−30.84	−58.93	−101.30	−134.10
21–24	DTW < 10M	31.44	39.99	53.02	28.89	34.29	44.19	28.06	32.16	40.02	53.36
21–24	DTW > 10M	−14.23	−14.66	−6.40	−16.69	−22.53	−22.38	−17.46	−25.52	−30.34	−21.66
21–24	Pleasure	−44.69	−34.49	−21.16	−47.36	−39.99	−28.97	−48.11	−41.96	−32.65	−19.53
25–29	Business	−28.68	−37.34	−40.94	−30.18	−44.35	−56.68	−30.65	−47.46	−66.91	−75.15
25–29	DTW < 10M	1.09	7.49	15.07	−1.13	3.53	10.07	−1.88	1.98	7.80	14.83

EXHIBIT 10

AVERAGE BIAS BY MODEL

(Continued)

Age Class	Vehicle Use	Normal Identity	Normal Log	Normal Inverse	Gamma Identity	Gamma Log	Gamma Inverse	Inv Gs Identity	Inv Gs Log	Inv Gs Inverse	Inv Gs Inv Sqr
25–29	DTW > 10M	5.23	6.33	11.75	3.10	0.64	2.26	2.41	−1.62	−2.33	2.47
25–29	Pleasure	11.87	19.21	26.78	9.52	15.39	22.42	8.85	13.95	20.45	27.30
30–34	Business	5.42	−3.69	−11.00	7.84	−3.06	−13.51	8.88	−2.94	−16.37	−23.85
30–34	DTW < 10M	−10.04	−7.18	−3.22	−8.34	−6.32	−3.45	−7.58	−5.85	−3.26	−0.34
30–34	DTW > 10M	10.14	8.83	10.40	11.94	9.02	8.24	12.76	9.22	7.54	8.15
30–34	Pleasure	−0.67	2.91	6.94	0.91	3.72	6.95	1.75	4.22	7.22	10.10
35–39	Business	−51.41	−37.12	−23.55	−54.83	−42.15	−26.85	−55.56	−43.93	−29.06	−13.63
35–39	DTW < 10M	17.58	15.42	11.12	13.43	12.60	10.08	12.43	11.80	9.93	6.68
35–39	DTW > 10M	8.91	12.94	13.99	4.87	8.84	11.40	3.92	7.57	10.56	11.48
35–39	Pleasure	−21.71	−25.14	−30.41	−25.98	−27.85	−31.24	−26.90	−28.58	−31.31	−35.17
40–49	Business	24.86	27.28	29.18	26.17	29.71	34.77	26.35	29.89	35.23	42.71
40–49	DTW < 10M	−1.31	−3.70	−6.99	−0.72	−1.74	−4.06	−0.82	−1.29	−2.88	−5.76
40–49	DTW > 10M	−13.26	−13.71	−15.30	−12.58	−12.09	−12.48	−12.62	−11.86	−11.52	−12.86
40–49	Pleasure	13.24	10.40	6.66	13.70	12.26	9.55	13.69	12.74	10.72	7.56
50–59	Business	9.41	11.56	14.44	10.65	12.83	17.55	11.05	12.90	17.15	25.40
50–59	DTW < 10M	−4.95	−6.23	−8.30	−4.43	−5.01	−6.42	−4.31	−4.62	−5.57	−7.51
50–59	DTW > 10M	0.81	0.97	0.58	1.43	1.69	1.88	1.60	1.84	2.33	1.96
50–59	Pleasure	8.71	7.07	4.55	9.11	8.23	6.47	9.30	8.65	7.32	5.09
60+	Business	15.48	18.42	21.71	18.11	21.98	28.04	19.07	23.09	29.24	38.44
60+	DTW < 10M	−7.24	−9.50	−12.44	−5.34	−6.83	−9.18	−4.66	−5.79	−7.69	−10.49
60+	DTW > 10M	11.01	10.84	9.91	13.01	13.32	13.21	13.75	14.27	14.61	13.83
60+	Pleasure	−2.82	−5.55	−8.98	−1.04	−3.01	−5.79	−0.28	−1.97	−4.34	−7.47

EXHIBIT 11

Average Absolute Deviation by Model

Age Class	Vehicle Use	Normal Identity	Normal Log	Normal Inverse	Gamma Identity	Gamma Log	Gamma Inverse	Inv Gs Identity	Inv Gs Log	Inv Gs Inverse	Inv Gs Inv Sqr
All	All	10.62	11.66	13.07	10.19	10.83	12.34	10.16	10.67	12.25	13.88
All	Business	25.09	25.42	26.15	27.08	28.62	32.98	27.64	29.58	35.93	40.73
All	DTW < 10M	7.30	8.87	10.88	6.02	6.75	8.38	5.76	6.04	7.29	9.39
All	DTW > 10M	9.27	10.06	11.23	8.87	9.00	9.67	8.91	9.12	9.70	9.89
All	Pleasure	11.38	11.98	12.63	11.47	12.14	12.69	11.47	12.15	12.84	13.37
17–20	All	45.62	47.74	50.61	45.75	46.53	46.75	45.75	46.57	48.07	48.69
21–24	All	29.06	33.36	36.21	29.17	34.78	40.84	29.17	35.36	43.65	49.32
25–29	All	7.96	13.00	19.29	7.10	9.99	15.72	7.10	9.97	16.03	20.85
30–34	All	8.32	6.71	7.18	8.61	6.41	6.96	8.82	6.33	7.11	7.60
35–39	All	20.07	18.92	16.28	18.11	17.51	15.59	17.61	17.11	15.57	12.87
40–49	All	9.65	10.85	12.64	9.41	10.02	11.54	9.48	9.82	10.91	13.26
50–59	All	4.75	5.36	6.13	4.84	5.26	6.15	4.90	5.18	5.97	7.21
60+	All	7.90	9.70	11.80	7.35	8.83	10.93	7.13	8.45	10.39	12.85

EXHIBIT 12

Iterative Method for Gamma Distribution with Identity Link

Iteration	Len	17–20	21–24	25–29	30–34	35–39	40–49	50–64	64+	DTW < 10	DTW > 10	Business
1	103.99	261.58	277.28	270.56	255.06	203.67	220.51	218.97	209.50	−14.59	29.37	99.82
2	12.09	276.43	282.51	261.04	249.68	201.25	215.90	218.40	211.88	−11.85	33.21	110.95
3	5.4360	273.53	279.24	257.82	246.10	198.02	212.51	215.40	209.07	−8.71	36.38	114.05
4	4.7355	271.04	276.40	255.18	243.29	195.13	209.74	212.72	206.52	−5.98	39.11	116.79
5	4.0687	268.90	273.96	252.91	240.88	192.65	207.36	210.42	204.33	−3.63	41.46	119.14
6	3.4776	267.08	271.87	250.97	238.81	190.52	205.32	208.46	202.47	−1.63	43.47	121.15
7	2.9545	265.53	270.09	249.33	237.06	188.71	203.59	206.80	200.89	0.08	45.17	122.86
8	2.4967	264.23	268.58	247.93	235.57	187.18	202.13	205.39	199.56	1.52	46.62	124.30
9	2.1001	263.13	267.31	246.76	234.32	185.90	200.90	204.21	198.44	2.73	47.83	125.52
10	1.7594	262.21	266.24	245.78	233.27	184.82	199.87	203.22	197.51	3.74	48.84	126.53
11	1.4689	261.44	265.35	244.96	232.39	183.91	199.00	202.40	196.73	4.59	49.69	127.38
12	1.2227	260.81	264.61	244.28	231.66	183.16	198.29	201.71	196.09	5.29	50.40	128.09
13	1.0152	260.28	263.99	243.72	231.05	182.54	197.69	201.14	195.55	5.88	50.98	128.68
14	0.8412	259.84	263.48	243.25	230.54	182.02	197.20	200.67	195.11	6.36	51.47	129.16
15	0.6957	259.48	263.05	242.86	230.13	181.59	196.79	200.28	194.74	6.77	51.87	129.57
16	0.5746	259.18	262.70	242.54	229.78	181.24	196.45	199.96	194.44	7.10	52.20	129.90
17	0.4739	258.93	262.41	242.28	229.50	180.95	196.17	199.69	194.19	7.37	52.48	130.17
18	0.3905	258.73	262.18	242.06	229.27	180.71	195.95	199.47	193.98	7.60	52.70	130.40
19	0.3215	258.56	261.98	241.88	229.07	180.51	195.76	199.29	193.81	7.78	52.89	130.58
20	0.2645	258.42	261.82	241.73	228.91	180.35	195.60	199.14	193.68	7.93	53.04	130.74
21	0.2175	258.31	261.69	241.61	228.78	180.21	195.47	199.02	193.56	8.06	53.17	130.86
22	0.1788	258.21	261.58	241.51	228.68	180.10	195.37	198.92	193.47	8.16	53.27	130.97
23	0.1468	258.14	261.49	241.43	228.59	180.01	195.28	198.84	193.39	8.25	53.35	131.05
24	0.1206	258.08	261.41	241.36	228.52	179.94	195.21	198.77	193.33	8.32	53.42	131.12
25	0.0990	258.02	261.35	241.31	228.46	179.88	195.15	198.72	193.27	8.37	53.48	131.18
200	0.0000	257.79	261.08	241.05	228.18	179.60	194.89	198.46	193.04	8.63	53.74	131.44

EXHIBIT 13

ITERATIVE METHOD FOR GAMMA DISTRIBUTION WITH INVERSE LINK

Iteration	Len	17–20	21–24	25–29	30–34	35–39	40–49	50–64	64+	DTW < 10	DTW > 10	Business
1	0.207137	−0.793187	−0.871778	−0.861086	−0.901986	−0.903692	−0.898338	−0.861744	−0.815516	0.878257	0.880340	0.882672
2	0.183232	−0.697337	−0.766409	−0.757331	−0.793385	−0.794597	−0.790027	−0.757540	−0.716818	0.772807	0.774567	0.776530
3	0.161249	−0.613156	−0.673934	−0.665915	−0.697623	−0.698580	−0.694602	−0.666018	−0.630171	0.680031	0.681481	0.683102
4	0.141896	−0.539080	−0.592561	−0.585470	−0.613353	−0.614086	−0.610629	−0.585483	−0.553926	0.598390	0.599567	0.600887
5	0.124865	−0.473894	−0.520954	−0.514680	−0.539196	−0.539734	−0.536734	−0.514613	−0.486832	0.526547	0.527485	0.528540
6	0.109879	−0.416533	−0.457941	−0.452386	−0.473941	−0.474306	−0.471708	−0.452249	−0.427791	0.463327	0.464054	0.464876
7	0.096691	−0.366055	−0.402492	−0.397568	−0.416517	−0.416730	−0.414486	−0.397371	−0.375836	0.407695	0.408236	0.408853
8	0.085086	−0.321637	−0.353697	−0.349330	−0.365985	−0.366065	−0.364133	−0.349078	−0.330116	0.358740	0.359117	0.359553
9	0.074874	−0.282549	−0.310759	−0.306882	−0.321518	−0.321480	−0.319823	−0.306582	−0.289884	0.315661	0.315894	0.316171
10	0.065888	−0.248153	−0.272974	−0.269528	−0.282388	−0.282247	−0.280831	−0.269187	−0.254481	0.277752	0.277858	0.277996
11	0.057980	−0.217885	−0.239724	−0.236657	−0.247955	−0.247722	−0.246519	−0.236279	−0.223326	0.244393	0.244387	0.244402
12	0.051021	−0.191250	−0.210465	−0.207732	−0.217654	−0.217341	−0.216325	−0.207321	−0.195911	0.215038	0.214934	0.214841
13	0.044897	−0.167811	−0.184718	−0.182278	−0.190990	−0.190607	−0.189755	−0.181839	−0.171786	0.189206	0.189016	0.188827
14	0.039509	−0.147186	−0.162060	−0.159880	−0.167526	−0.167081	−0.166373	−0.159415	−0.150557	0.166474	0.166208	0.165936
15	0.034767	−0.129036	−0.142123	−0.140169	−0.146879	−0.146379	−0.145799	−0.139683	−0.131876	0.146471	0.146138	0.145792
16	0.030594	−0.113064	−0.124578	−0.122824	−0.128709	−0.128161	−0.127693	−0.122318	−0.115437	0.128868	0.128476	0.128065
17	0.026922	−0.099010	−0.109139	−0.107561	−0.112720	−0.112130	−0.111761	−0.107038	−0.100971	0.113378	0.112935	0.112466
18	0.023691	−0.086642	−0.095553	−0.094130	−0.098651	−0.098023	−0.097741	−0.093592	−0.088241	0.099747	0.099259	0.098740
19	0.020848	−0.075759	−0.083597	−0.082311	−0.086270	−0.085609	−0.085403	−0.081760	−0.077039	0.087753	0.087224	0.086661
20	0.018345	−0.066182	−0.073077	−0.071910	−0.075375	−0.074686	−0.074546	−0.071347	−0.067181	0.077197	0.076633	0.076031
21	0.016144	−0.057754	−0.063819	−0.062758	−0.065787	−0.065073	−0.064993	−0.062185	−0.058507	0.067909	0.067314	0.066678
22	0.014206	−0.050338	−0.055672	−0.054704	−0.057350	−0.056614	−0.056586	−0.054122	−0.050873	0.059736	0.059113	0.058447
23	0.012501	−0.043812	−0.048503	−0.047617	−0.049926	−0.049170	−0.049188	−0.047027	−0.044156	0.052543	0.051896	0.051204
24	0.011001	−0.038069	−0.042194	−0.041380	−0.043393	−0.042619	−0.042678	−0.040783	−0.038245	0.046214	0.045546	0.044830
25	0.009680	−0.033015	−0.036643	−0.035892	−0.037644	−0.036855	−0.036949	−0.035289	−0.033044	0.040644	0.039958	0.039221
200	0.000000	0.004037	0.004059	0.004345	0.004507	0.005407	0.005054	0.004994	0.005093	−0.000192	−0.001015	−0.001902
Inverse:		247.74	246.37	230.13	221.87	184.94	197.87	200.24	196.34	−5,208.33	−985.22	−525.76

APPENDIX A

RECONCILIATION OF NOTATION WITH THE LITERATURE

McCullagh and Nelder [17] define the exponential as a two-parameter family of distributions whose density functions can be written in the form:

$$f(r;\theta,\phi) = \exp((r\theta - b(\theta))/a(\phi) + c(r,\phi)). \qquad \text{(A.1)}$$

Generally $a(\phi) = \phi/w$ where w is a known prior weight. We will assume a has this form. Thus to reconcile Equation A.1 with 6.1 it is enough to explain what is meant by the identity

$$r\theta - b(\theta) = -\frac{1}{2}d(r;\mu) = \int_r^{\mu} \frac{(r-t)}{V(t)}dt. \qquad \text{(A.2)}$$

We must define the function b. Differentiating Equation A.2 with respect to θ gives

$$r - b'(\theta) = \frac{r-\mu}{V(\mu)}\frac{d\mu}{d\theta},$$

because r is a constant. Taking expected values over r shows $\mu = b'(\theta)$ since $\mathrm{E}(r) = \mu$ by Equation 6.2, and so the right hand side vanishes. Substituting for μ and canceling $r - \mu$ shows $V(\mu) = b''(\theta)$. Thus the function b satisfies the differential equation

$$V(b'(\theta)) = b''(\theta), \qquad \text{(A.3)}$$

which is enough to determine b; θ is simply an argument.

Example 6.1 Revisited

Example 6.1 showed that the gamma distribution belongs to the exponential family by deriving the deviance function from the density function. We now assume the form of the variance function and derive the density using the function b. $V(\mu) = \mu^2$ corresponds to the gamma distribution, so Equation A.3 gives

$$(b'(\theta))^2 = b''(\theta),$$

whence

$$\mu = b'(\theta) = -\frac{1}{\theta},$$

and

$$b(\theta) = -\log(-\theta).$$

Plugging into Equation A.1 gives exactly Equation 6.6 with $\phi = 1/\nu$.

Connection with Generalized Linear Models

To solve for the parameters of a generalized linear model using maximum likelihood directly from Equation A.1, it is necessary to differentiate the log likelihood of an observation r_i:

$$l(\theta,\phi;r_i) = l = w_i(r_i\theta - b(\theta))/\phi + c(r_i,\phi)$$

with respect to β_j. Using the chain rule and substituting $\mu = b'(\theta)$, $d\mu/d\theta = b''(\theta) = V(\mu)$ gives

$$\begin{aligned}\frac{\partial l}{\partial \beta_j} &= \frac{\partial l}{\partial \theta}\frac{d\theta}{d\mu}\frac{d\mu}{d\eta}\frac{\partial \eta}{\partial \beta_j}\\ &= \frac{w_i(r_i - b'(\theta))}{\phi}\frac{1}{b''(\theta)}\frac{d\mu}{d\eta}x_{ij}\\ &= \frac{w_i(r_i-\mu)}{\phi}\frac{1}{V(\mu)}\frac{d\mu}{d\eta}x_{ij},\end{aligned}$$

which is Equation 5.3 for one observation $r_i = r$, up to a factor of ϕ which cancels out.

APPENDIX B

COMPUTER SOLUTION OF GENERALIZED LINEAR MODELS

This section contains annotated SAS IML code to compute the parameters for a generalized linear model with log link and gamma errors.

The dataset CARDATA contains the following variables:

1. AGE, the age group classification
2. VUSE, the vehicle-use classification
3. LOSS, the average severity
4. NUMBER, the number of claim counts,

as shown in Exhibit 1.

Comments in SAS are enclosed between * and ; . In IML the statement * denotes matrix multiplication, # denotes componentwise multiplication, and ## denotes componentwise exponentiation.

The SAS IML code is as follows:

```
DATA CARDATA;
INPUT AGE VUSE LOSS NUMBER;
CARDS;
data lines
;

PROC IML;

* READ ALL VARIABLES INTO IML VARIABLES AGE, VUSE, R AND W ;
USE CARDATA;
READ ALL VAR AGE INTO AGE;
READ ALL VAR VUSE INTO VUSE;
READ ALL VAR LOSS INTO R;
READ ALL VAR NUMBER INTO W;
```

```
* COMPUTE DESIGN MATRICES ;
A = DESIGN(AGE);
B = DESIGN(VUSE);
* SELECT A BASE CLASS BY DELETING A COLUMN OF B ;
* [,1:3] MEANS SELECT COLUMNS 1 THRU 3 ;
B = B[,1:3];
* MODEL DESIGN MATRIX = HORIZONTAL CONCATENATION OF A AND B ;
X = A||B;

* DEFINE A FUNCTION TO COMPUTE THE VARIANCE FUNCTION FOR A ;
* GAMMA DISTRIBUTION ;
START VARFUN(MUIN);
RETURN(MUIN# MUIN); * COMPONENTWISE MULTIPILCATION ;
FINISH;

* WEIGHTS FOR THE LOG LINK, PER Equation 7.3 ;
START W(MUIN);
ANS = MUIN# # 2 / VARFUN(MUIN);
RETURN(ANS);
FINISH;

* INITIALIZE WITH DATA ;
MU = R;
ETA = LOG(MU);

* SET UP HOLDERS FOR CURRENT AND PREVIOUS PARAMETERS ;
* J(NCOL(X),1,10) RETURNS A NCOL(X) x 1 MATRIX WITH VALUE 10, ETC ;
LASTBETA = J(NCOL(X),1,10);
BETA = J(NCOL(X),1,0);

* WHILE SQUARED DISTANCE BETWEEN BETA AND LAST BETA IS LARGE DO ;
DO WHILE((BETA-LASTBETA)‘ * (BETA-LASTBETA) > 1E-9);
    * COMPUTE AUXILLARY VARIABLE ;
    Z = ETA + (R - MU) # DETADMU(MU);

    * SAVE LAST BETA VECTOR ;
    LASTBETA = BETA;

    * DO WEIGHTED LEAST SQUARES;
    * NOTE: GINV = INVERSE ;
```

```
    WEIGHT = W(MU) # W;
    BETA = GINV(X‘ * ( WEIGHT # X)) * X‘ * (WEIGHT # Z);

    * COMPUTE PREDICTED VALUES ;
    ETA = X * BETA;
    MU = EXP(ETA);

END;

* PRINT OUT PARAMETERS ;
PRINT I BETA[F = 8.4];

* NOW COMPUTE THE VARIOUS STATS, DEVIANCE AND SO FORTH ;
* MU AND ETA ALREADY HOLD THE LAST ESTIMATES OF PRED VALUES ETC;

* COMPUTE VAR;
VAR = VARFUN(MU);

* COMPUTE GAMMA DEVIANCE ;
DEV = 2 # W # (-LOG(R / MU) + ((R-MU) / MU));

* PEARSON RESIDUAL AND DEVIANCE RESIDUALS ;
PEARES = (R - MU ) / SQRT(VAR);
DEVRES = SIGN(R - MU) # SQRT(DEV);

NOBS = NROW(X); * NUMBER OF OBSERVATIONS ;
NPARAM = NCOL(X); * NUMBER OF PARAMETERS ;
DF = NOBS - NPARAM; * NUMBER OF DEGREES OF FREEDOM ;

PEARSON = (PEARES# PEARES)[+]; * [+] = SUM OVER COMPONENTS ;
DEVIANCE = DEV[+];
PRINT PEARSON, (PEARSON / DF)[LABEL = ”DISPERSION = PEARSON/DF”],
DEVIANCE, (DEVIANCE / DF)[LABEL =”DEVIANCE/DF”];

* LOGLIKELIHOOD FOR GAMMA DISTRIBUTION ;
PHI = DEVIANCE / DF; * ESTIMATE FOR PHI ;
LLH = (W/PHI) # LOG(W # R / (PHI # MU))
- W # R / (PHI # MU) - LOG(R) - LGAMMA(W/PHI);
* LGAMMA = LOG(GAMMA FUNCTION) ;

PRINT (LLH[+]);
```

```
** ABOVE CODE WILL GIVE THE SAME RESULT AS THE FOLLOWING CODE ;
** USING THE BUILT-IN SAS GENERALIZED LINEAR MODEL ROUTINE, PROC ;
** GENMOD ;

PROC GENMOD DATA = CARDATA;
CLASS AGE VUSE;
SCWGT NUMBER;
MODEL R = AGE VUSE / NOINT DIST = GAMMA LINK = LOG DSCALE;
RUN;
```

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXX

SURPLUS—CONCEPTS, MEASURES OF RETURN, AND DETERMINATION

RUSSELL E. BINGHAM

DISCUSSION BY ROBERT K. BENDER
VOLUME LXXXIV

DISCUSSION BY DAVID RUHM AND CARLETON GROSE

1. INTRODUCTION

Dr. Bender has made the results obtained in Mr. Bingham's paper more accessible by focusing on the essential elements that influence measurement of return, and by providing a variety of detailed examples. In addition, Dr. Bender has extended the work in several directions. Several of the results obtained in Dr. Bender's discussion paper are fundamental to the study of surplus and return on equity (ROE). In particular, Dr. Bender describes two basic tests of reasonableness that can be applied to any rate-of-return model in order to check the model's soundness. Because of their universal applicability, these tests are a major contribution.

Two major results presented in the discussion are: 1) the three measures of return discussed in the paper are equal to each other under a specific earnings release pattern, and 2) one of the measures (the NPV ratio) is constant with respect to the earnings release pattern. As will be discussed below, there are some accounting issues that must be dealt with in order to make use of these results, and they do not generally hold true for a model that does not include reserve margin (or some equivalent mechanism). Despite this caveat, Dr. Bender's paper contains other important findings, and represents a substantial contribution to the actuarial literature on surplus and profitability measurement.

2. CALENDAR YEAR MEASURES

In his introduction, Dr. Bender states: “When evaluating the return earned by a particular product line, it is this long-term investment of surplus that must be considered. This is in sharp contrast to calendar year measures in which it is assumed that all of the company surplus supports the currently written exposure.”

Dr. Bender correctly points out that surplus supports exposures from all accident years that have not yet been closed, as well as current writings. In particular, it should not be assumed that surplus supports only the current year’s written premium. Although this is a common interpretation of calendar year profitability measures, no such assumption necessarily exists, even when a premium-to-surplus ratio is used in profitability measurement. Such calculations use premium as a measure of the volume of business (including prior years’ exposures), while the surplus serves as a measure of internal capitalization.

The premium-to-surplus ratio measure must be used with caution because the current year’s written premium is a very imperfect measure of the volume of the business. A hidden assumption is that the ratio of outstanding liabilities plus the expected future liabilities arising from the current writings to the current year’s written premium is a constant. This is generally not true because current writings will fluctuate according to many factors such as entry into and exit from lines of business, pricing adequacy, market conditions, etc. Additionally, it is affected through changes on the liability side such as loss payout characteristics, inflation, etc. It can be shown that in a steady-state situation (with no underlying price, exposure, or loss characteristic changes) that the current year’s written premium is an accurate measure of the volume of liabilities of the business. Although the premium-to-surplus ratio measure has problems, it is a convenient way to allocate capital in a model. If a premium-to-surplus ratio measure is used, it must reflect current premiums and past liabilities and the volatility inherent in both.

For example, suppose a company plans to write $1,000,000 of premium during a given calendar year and has $2,500,000 of loss reserves at the beginning of the year. (For simplicity, unearned premium reserves will be omitted from this example.) Also suppose that the company performs a comprehensive analysis of risk for its portfolio, determining that $250,000 of surplus should be allocated to support the expected future liabilities from the writing of the premium and $550,000 of surplus should be allocated to supporting the outstanding loss reserves. The total surplus commitment is $800,000, which can be construed to produce a written premium to surplus ratio of 1.25. This does not imply that $100 of surplus supports each $125 of premium written.

3. PRODUCT ACCOUNT AND SURPLUS ACCOUNT

Dr. Bender discusses a useful perspective that was developed in the Bingham paper. In his overview of Bingham's methodology, Dr. Bender writes, "The world can be divided into three parts … the insurance product, shareholder funds [surplus], and everything that is external to the other two parts." The conceptual distinction between product account and surplus account can either be directly incorporated into a model or at least kept in mind by an actuary while developing and testing a model.

An application of this paradigm occurs later in the paper, when Dr. Bender observes that the ROE must equal the investment rate of return if the insurance product account generates an operating gain of zero. If one imagines the product account generating no outflow or inflow of funds, and the surplus account generating the investment rate, then the result becomes readily apparent without the need for calculations. This test can be employed by an actuary to check the soundness of a return model being considered for use. Dr. Bender's conclusion that the calendar year steady-state model fails the test, and is therefore inherently inaccurate, appears correct. Both the reasonableness test and this conclusion are noteworthy contributions.

4. SURPLUS ACCOUNTING AND RETURN MEASUREMENT

An accounting problem arises when Dr. Bender discusses income that is generated from funds in the product account (generally known as "income from insurance operations"). This includes earned premium and investment income on underwriting funds (but not investment income on surplus) minus incurred losses and expenses. Dr. Bender writes, "While reserves and supporting surplus are clearly identified as 'belonging' to the insurance product, the time at which other funds that arise from the insurance product are released to the surplus account is somewhat arbitrary."

The problem is that there is no such action as "releasing funds to the surplus account." Surplus by definition is the amount of assets in excess of liabilities and is thus the balancing item on the balance sheet. Assuming that liabilities are consistently stated without bias (which is generally assumed in models of this kind), the only way surplus can be deliberately increased or decreased is through transactions with external shareholders. Operating gain cannot remain in the insurance product account, even if generated by funds in the product account: as soon as any such gain is recognized, it immediately and automatically becomes surplus, by the definitions of income and surplus.

The model shown in Dr. Bender's exhibits allows income to accumulate as "retained earnings" in the product account, rather than as an increase in surplus. But these "retained earnings" are actually additional surplus and must either be distributed to shareholders or counted as surplus in the denominator of ROE. Either way, the actual surplus levels and flows differ from those shown in Dr. Bender's exhibits. Although his demonstration and proof of the equality of the three return measures is mathematically sound, this equality is not a true representation of ROE because the surplus is inaccurately stated.

That said, Dr. Bender's analysis and results are valid when reserve margin is included in the model. Reserve margin is the amount by which a reserve (the stated value of a liability) exceeds the unbiased estimate of the liability's value. Reserve margins have an important, legitimate use that has been documented in the literature [1].

Reserve margin neatly fills the role of "retained earnings" in the paper's exhibits. Since reserve margin is part of total reserves, it is in the product account. A reserve margin can be viewed as an asset or "operating gain" that has not yet been recognized as an increase to surplus, which is exactly what "retained earnings" are. Dr. Bender notes that retained earnings act as "… an additional buffer against insolvency risk." A positive reserve margin does act as an additional buffer, absorbing the impact of adverse results before surplus is affected. Finally, the level of reserve margin can be selected to increase or decrease the surplus level, providing a mechanism for releasing funds to the surplus account.

If we substitute the label "Reserve Margin" for "Retained Earnings" in the paper's exhibits, all of the paper's results hold. The only question is whether it is reasonable to include reserve margin in a return model. This is a question to be decided by the individual model designer, based in part on the particular application for which the model is being developed.

A minor remaining problem is that the paper's exhibits often show a negative value for retained earnings. Negative reserve margin implies inadequate nominal reserves, which would inflate the calculated return. A negative reserve margin condition may not be acceptable in some return modeling applications.

5. NOMINAL VS. DISCOUNTED RESERVES

Dr. Bender makes an important point: if a company calculates required supporting surplus based on nominal unpaid losses so that a performance criterion (e.g., probability of ruin less than 2%) is met, then the result is a surplus requirement for the future (when the loss payments are to be made). A lesser amount of surplus is sufficient at the time of the evaluation, since the surplus can accumulate investment income during the interim. The question that Dr. Bender then addresses is how much surplus is required at the time of evaluation to meet the performance criterion.

Dr. Bender advocates calculating the surplus requirement based on discounted loss reserves. His method is to apply a leverage ratio to the discounted reserves. The leverage ratio is calculated from the probability distribution of discounted future payments, so that timing risk and investment return risk are accounted for in the distribution. The resulting surplus meets the performance criterion with respect to the discounted reserves at the time of evaluation.

For example, suppose nominal loss reserves are \$10,000 and discounted reserves are \$8,000. Suppose also that ultimate paid losses will be less than \$15,000 with 98% probability, and that the distribution of discounted unpaid losses has its 98th percentile at \$9,600 (considering all possible interest rate and payout scenarios). To meet the performance criterion of $P(\text{ruin}) < 2\%$ using nominal loss reserves, the supporting surplus would be $\$15{,}000 - \$10{,}000 = \$5{,}000$, which corresponds to a 2.00 reserves-to-surplus leverage ratio. Using discounted reserves, the surplus required would be $\$9{,}600 - \$8{,}000 = \$1{,}600$ for a 5.00 leverage ratio. Although the 5.00 leverage ratio seems high, there is a 98% probability that the \$9,600 fund will accumulate sufficient investment income to pay all claims as they come due.

This method meets the performance criterion on discounted reserves at the date of evaluation and simultaneously provides proper funding to meet the performance criterion at the future payment dates. It is a mathematically correct answer to the question that was posed.

There are two notable objections to using Dr. Bender's discounted reserves approach: 1) it is presently impossible to accurately quantify the probability distributions of future interest rate levels and claims payment patterns, both of which are fundamental elements for determining the distribution of discounted unpaid losses; 2) if claims develop adversely as of a later evaluation, more surplus may have to be obtained to continue to meet the performance criterion. If additional surplus is available at each evaluation point (as could be the case for an insurance company within a holding company group), this is not a problem. If not (as could be the case for a small stand-alone company), there is no margin for such a contingency.

Both of these objections are addressed by using nominal loss reserves. The only distribution to be considered is the aggregate loss distribution, which can usually be estimated reasonably. If additional surplus should be required at a later evaluation, a portion of the investment income earned on surplus can be retained, rather than released as earnings.

Future developments in financial analysis may eventually provide solutions to the first objection. The second objection could be addressed by setting the surplus level a little higher, so as to provide a prescribed cushion on top of the surplus level that is dictated by the performance criterion. The amount of cushion would thus be selected more precisely than the somewhat arbitrary investment income cushion provided by using nominal loss reserves.

Dr. Bender did raise the possibility of adverse loss development and the consequent need for additional surplus. He treated this issue in Section 6 of his paper, using the following example: expected nominal losses of \$44 are initially allocated \$22 of surplus (using a 2 : 1 rule), for a total funding requirement of $66. Two years later, the losses are re-evaluated, and the best estimate is $60. Dr. Bender offered three possible solutions:

1. Allow the surplus level to drop as a result of the adverse loss development. In the example, the additional $16 of adverse development would be absorbed by the original surplus allocation, and the new surplus level would be $6. The total funding requirement is still $66.

2. Restore surplus to its original level. For the example, this would mean increasing the surplus level to $22, for a total funding requirement of $82.

3. Increase the surplus level, following the original surplus rule. In this example, the rule was a 2 : 1 ratio, so the new surplus level would be $30, and total funds would be $90.

Which of these alternatives is used may depend on the application. For example, the first approach is often implicit in a pricing model, where surplus is set with the knowledge that worse or better results will be achieved over the sample space of lines and years. In fact, a total exhaustion of the surplus ("ruin") is actually expected to occur a certain percentage of the time, if a probability of ruin method is used to set surplus.

None of these three alternatives corresponds to the surplus calculation method that Dr. Bender proposes. The new information that produced the higher reserve valuation should be incorporated into the leverage ratio. We propose a fourth alternative: calculate a new leverage ratio in the same way that the original 2 : 1 ratio was calculated, perhaps based on variability of outstanding losses (nominal or discounted). Apply the leverage ratio to the current valuation of outstanding losses to determine current surplus requirements. This alternative resembles the third approach, but is more consistent with the surplus calculation ideas that Dr. Bender puts forth.

Dr. Bender indicates that using discounted reserves to calculate required surplus allows one to account for timing risk and investment return risk. A caution is in order: simply applying a leverage ratio to discounted reserves to calculate required surplus does not account for either timing risk or investment return risk. Both of these risks are higher for long payment patterns, but discounted reserves are lower for longer patterns. Applying a fixed leverage ratio to discounted reserves would result in less surplus being assigned to a longer pattern, but the increased timing and investment risks would warrant more surplus (all else being equal). If a leverage ratio is used with discounted reserves, then the ratio must be explicitly calculated based on the variability of the discounted future payments, as Dr. Bender advises.

6. INACCURACY OF THE CALENDAR YEAR RETURN MEASURE

Dr. Bender provides excellent explanations and exhibits to show that calendar year accounting distorts the measurement of return. For Dr. Bender's first "reasonableness test," the insurance product is priced at break-even so that the total return should equal the investment rate obtained on surplus. In the paper's example, the calendar year return (under statutory accounting) is 8.1%, much higher than the 5.0% investment rate. We constructed our own model and independently verified the accuracy of this result, assuming the surplus levels presented in the paper's exhibit.

Dr. Bender continues with a discussion of the calendar year distortion, explaining the result from several perspectives. His lucid explanations make it possible for readers to understand how the calendar year measure fails to produce the proper result. Dr. Bender then notes that the exposure growth rate assumption influences the calendar year return, so that if the growth rate is assumed to be equal to the investment rate, the calendar year return will then produce the correct result. Finally, another example is given in which the insurance product clearly loses money, but the calendar year return is erroneously higher than the investment rate.

The case that is made against calendar year return is so compelling that the unavoidable conclusion seems to be that calendar year return is (in general) an inaccurate measure of actual return. But what if calendar year return is used to measure a company's performance, either by internal management or external parties? An actuary who is building a return model for, say, pricing purposes will probably still have to include calendar year return in the model (perhaps alongside another return measure). The actuary also will have to consider the calendar year return in the decision-making process, while at the same time recognizing that the calendar year result does not accurately depict profitability.

The fact is that calendar year ROE is currently a prevalent method of calculating return. Dr. Bender's findings should motivate us to conduct research into alternative return measures.

7. SELF-SUPPORTING PREMIUM AND INFINITE RETURN

Dr. Bender's second "reasonableness test" considers the situation where premium is large enough to produce its own supporting surplus as it earns. Surplus allocation formulas often allocate surplus to a policy or line before any premium is earned, on the theory that risk is related to the unearned premium and is present from the time a policy is written. Another perspective is that losses are incurred as premium earns, so the surplus associated with a portion of premium is not needed until the moment that premium is earned, because that's the time when the insurer is actually exposed to loss (not before). After the premium earns, some of the surplus then remains associated with the corresponding loss reserves and runs off accordingly.

Both perspectives are useful. A surplus allocation formula can be used to budget needed surplus for a line of business at annual intervals, based on upward variability of losses from the expected level. The earning perspective can then be used to reduce the amount of budgeted surplus by the profit that the line is expected to generate as the premium earns. This expected profit will accrue to surplus if actually realized, so it is "future surplus:" not available at the time of budgeting but also not needed until realized and available. If losses are greater than expected, the impact will first be a reduction in this "future surplus," before budgeted surplus is impacted.

As Dr. Bender states, if premium is high enough, the budgeted surplus requirement becomes zero, because the entire surplus need is met by the earning of the premium. Therefore, no investment is required up-front, and the return (under the expected losses scenario) is infinite. Dr. Bender then compares the three return measures as the premium rises to the infinite-return value and observes that only the internal rate of return (IRR) measure yields the correct result. The other two measures produce finite values for return, even when the premium is high enough to generate its own supporting surplus "on-the-fly."

The problem again is that surplus is not being calculated according to the correct formula. The liabilities are discounted at the investment rate, and there is no recognition of the unearned premium reserve liability at the beginning. (In earlier exhibits, it appears that the concepts of "invested capital" and "surplus" are being confused with each other.) In spite of this, the IRR results that are presented in the paper can be reproduced under correct accounting by setting assets equal to Dr. Bender's funding requirement at each point in time.

In any case, the other two return measures (Calendar Year ROE and net present value (NPV) Ratio) will not produce values that approach infinity, no matter how high the premium is. This is because both of these measures are ratios, with total surplus in the denominator. Calendar Year ROE equals Total Income/Total Surplus, and NPV Ratio equals NPV(Total Income)/NPV(Total Surplus). The only way either ratio could be infinite is if the surplus level is kept at zero for the entire period, which would not make any sense since some supporting surplus must be held until losses are completely paid. Dr. Bender states that the NPV Ratio measure would approach infinity if surplus requirements were reduced "in recognition of the retained operating gain," but again this "retained operating gain" is actually surplus. The NPV Ratio simply cannot produce the infinite return result.

Exhibit 1 shows a simple example that compares the three return measures. The premium has been set to a high level, so that the policy generates its own surplus (and then some) as premium earns. As the exhibit shows, IRR is infinite because there is zero initial investment and all the cash flows to the investors are positive. The other two return measures produce values that are finite, though large.

The IRR measure produces an infinite return in this example because it is focused on the flows between the company and the shareholders (or the "surplus surplus" account, to use Dr. Bender's terminology), rather than on the company's internal surplus. The other measures implicitly identify the company's internal surplus as invested funds, and measure the return against those funds. Ironically, the Internal Rate of Return (IRR) is distinguished here by its reliance on the company's external transactions with shareholders, versus the alternative return measures, which are based on internal company surplus.

8. CONCLUSION

In summary, Dr. Bender has written a discussion paper that stands on its own. All of Dr. Bender's findings discussed above are essential to a complete understanding of return measurement, and many of them can be directly incorporated into return modeling applications.

REFERENCES

[1] Balcarek, Rafal J., "Effect of Loss Reserve Margins in Calendar Year Results," *PCAS* LIII, 1966, pp. 1–17.

EXHIBIT 1

A SELF-SUPPORTING LINE

Premium = $2,000
Loss = $1,000 paid 2 years after inception
Surplus = 50% of Nominal Loss Reserves
Investment Income = 5% per year
Taxes are omitted

Underwriting Quantities

Time, yrs	Written Premium	Earned Premium	Incurred Loss	Paid Loss	Unearned Premium Reserve	Loss Reserve
Inception	2,000	0	0	0	2,000	0
1	0	2,000	1,000	0	0	1,000
2	0	0	0	1,000	0	0
Total	2,000	2,000	1,000	1,000		

Assets, Liabilities, and Surplus

Time, yrs	UEP Reserve	Loss Reserve	Total Liabilities	Surplus	Total Assets
Inception	2,000	0	2,000	0	2,000
1	0	1,000	1,000	500	1,500
2	0	0	0	0	0

Investment Income Calculation

Time, yrs	Total Assets	Assets Not Investable	Investable Assets	5.00% Investment Income
Inception	2,000	0	2,000	0
1	1,500	0	1,500	100
2	0	0	0	75
Total				175

EXHIBIT 1

PAGE 2

A SELF-SUPPORTING LINE

Calculation of Total Income

Time, yrs	Earned Premium	Incurred Loss	Net U/W Income	Investment Income	Total Income
Inception	0	0	0	0	0
1	2,000	1,000	1,000	100	1,100
2	0	0	0	75	75
Total	2,000	1,000	1,000	175	1,175

Calculation of Flows to Shareholder

Time, yrs	Surplus	Change in Surplus	Total Income	Flows To/(From) Shareholder
Inception	0	0	0	0
1	500	500	1,100	600
2	0	−500	75	575
Total	500		1,175	
NPV	476		1,171	

NPV(Income)/NPV(Surplus) = 1,171/476 = 246%
Calendar Year Average Return = 1,175/500 = 235%
IRR = Infinity

ADDRESS TO NEW MEMBERS—MAY 17, 1999

M. STANLEY HUGHEY

As a representative of the rather distant past, it is my privilege to welcome all the new Fellows and new Associates into membership in the Casualty Actuarial Society. At the same time I want to both compliment and congratulate each of you for reaching this very significant milepost in your career.

Some of us oldsters can still remember the hours upon hours of concentrated study, and the sacrifice of burning the midnight oil to build actuarial knowledge, rather than reading light novels or becoming a couch potato—and perhaps even more important, the sacrifice of quality time with your family, while you hit the actuarial books with the aim of long term benefit to that family.

Yes, this is a great and important milepost, and you are all to be congratulated on reaching it.

In fact, as a sort of turning point in your lives, this occasion takes on many of the characteristics of a graduation, and whether or not you appreciate it, I am in the position of being asked to make a sort of "Graduation Speech." This is both good and bad. The bad part is that you have undoubtedly had your fill of graduation speeches, and can pretty well predict what I am going to say. But the good part is that if I don't finish within 8 or so minutes from now, Steve Lehmann will open the trap door I'm standing on, banishing forever any remaining words of wisdom.

Many of you are parents, and a very wise source, "anonymous," once said that parents should supply their children with two things—"roots" and "wings." I'm going to adapt this to the goals of the CAS, as expressed in the CAS Mission Statement. As "roots," the CAS is an organization designed "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."

As new Fellows and Associates, you have embraced these very meaningful words and goals as part of your lives, but by your acceptance into the CAS this morning, you have in turn been embraced by these same meaningful words and goals, as they make up the roots and goals of the Casualty Actuarial Society.

We will not spend a lot of time on history in these comments, but our actuarial roots include ratemaking, credibility theory, loss reserves, financial measurement, reinsurance, self-insurance and classification systems. Many actuarial principles have been established, tested and written into our standards. Others have not stood the test of time, and we have had to grow a new root structure.

To save you the trouble of looking it up, I became a Fellow in 1947, and served as President in 1974. That means that I have been around for 50+ years, and have been a witness to the forming of these roots, as well as the mushrooming membership—from about 200 in my early years to over 3,000 currently.

In 1989, I was privileged to summarize the CAS history up to that date. In that effort, I used a quote from Carl Hubbell, the great baseball pitcher from more years back than most of you remember. I am taking the liberty of repeating it here, because it so appropriately introduces the second part of my remarks—the "Wings" part of the "Roots and Wings" theme.

Quoting Carl Hubbell:

> A fellow doesn't last long on what he has done. He's got to keep delivering as he goes along.

This is the challenge part of this graduation ceremony. Stated simply, you have your roots in the CAS, but you can't stop where you are, and you must forge ahead into new horizons.

Referring back to the CAS Mission Statement, the CAS is shouting at you to unlimber your wings and soar into the unknown—in a disciplined way, of course. Referring now to "wings," let me quote: "The purposes of the Casualty Actuarial Society are to advance the body of knowledge of actuarial science applied to property/casualty and similar business and financial risks."

The challenge of the CAS to new Fellows and new Associates to "—keep delivering as you move along" is crystal clear.

Incidentally, I am delighted to be lifting these quotes about the CAS from the March 15, 1999 letter and supporting material Steve Lehmann sent to the CAS membership as a report on the organization's Strategic Plan. So, these quotes are both authoritative and recent.

Further on the subject of new frontiers, and spreading your wings, is the program material for this meeting. Most of it looks forward and not back. Let me emphasize by listing some of the discussion subjects. Your Program Committee is obviously looking forward:

- European Union's Impact
- Y2K Update
- Financial Markets
- Securitization of Risk
- Loss Portfolio Transfers
- Auto Insurance in the New Millennium
- Actuaries in Non-traditional Roles
- DFA in the Real World
- Emerging Financial Markets

Speaking from a 50-year vantage point, I'm impressed with the new subjects. Thirty, twenty, and even ten years ago, these subjects simply were not there.

Now in closing, I want to get around to the subject that I expect most of you have been wondering about since I first stood up here. "Why in the world would Stan wear a jacket like that for a serious business presentation?" Well, it was not an accident, and I want to use it to make a point. In business days in Chicago, I wore dark suits like everyone else, and today I would not wear this jacket to a business meeting in Chicago. (Or New York or Boston or Atlanta.)

But in case you hadn't noticed, Florida is different—far more casual and far more colorful. And so, in your business careers, yes even as actuaries, you must learn to use your wings to adapt to new and different situations. Changes come, and we must lead or at least keep pace with any new solutions which are helpful in solving both old and new problems.

In summary, I would like to emphasize two kernels of wisdom from whatever store of knowledge I have accumulated over 50 years of experience:

Keep your roots deep in the CAS fundamentals.

Soar with the wings of new developments which provide better solutions.

MINUTES OF THE 1999 SPRING MEETING

May 16–19, 1999

DISNEY'S CONTEMPORARY RESORT AT THE
WALT DISNEY WORLD RESORT

LAKE BUENA VISTA, FLORIDA

Sunday, May 16, 1999

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 1999 Executive Council discussed their roles in the Society with the new members. In addition, Robert A. Anker, who is a past president of the CAS, gave a short talk on the American Academy of Actuaries' (AAA) Casualty Practice Council.

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 17, 1999

Registration continued from 7:00 a.m. to 8:00 a.m.

The 1999 Business Session, which was held from 8:00 a.m. to 9:00 a.m., started off the first full day of activities for the 1999 Spring Meeting. Mr. Lehmann 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), Charles A. Bryan (1990), David P. Flynn (1992), Charles C. Hewitt Jr. (1972), M. Stanley Hughey (1974), Allan M. Kaufman (1994), Michael L. Toothman

(1991), Mavis A. Walters (1997), and Michael A. Walters (1986).

Mr. Lehmann also recognized special guests in the audience: Howard Bolnick, president of the Society of Actuaries; Stephen P. D'Arcy, president-elect of the American Risk and Insurance Association; Linda Lamel, executive director of the Risk and Insurance Management Society; and Michael L. Toothman, president-elect of the Conference of Consulting Actuaries.

Curtis Gary Dean, Robert S. Miccolis, and Kevin B. Thompson announced the 160 new Associates and Alice H. Gannon announced the 13 new Fellows. The names of these individuals follow.

NEW FELLOWS

Mustafa Bin Ahmad
Betsy A. Branagan
Elliot Ross Burn
Brian Harris Deephouse
Alana C. Farrell
Bruce Daniel Fell
Claudine Helene Kazanecki
Deborah M. King
Eleni Kourou
Dawn M. Lawson
Richard Borge Lord
Michael Shane
Christopher C. Swetonic

NEW ASSOCIATES

Jason R. Abrams
Michael Bryan Adams
Anthony L. Alfieri
Silvia J. Alvarez
Gwendolyn Anderson
Paul D. Anderson
Amy Petea Angell
Anju Arora
Nathalie J. Auger
Amy Lynn Baranek
Patrick Beaudoin
David James Belany
Kristen Maria Bessette
John T. Binder
Mario Binetti
Christopher David Bohn
Mark E. Bohrer
David R. Border
Thomas S. Botsko
Stephane Brisson
Karen Ann Brostrom
Conni Jean Brown
Paul Edward Budde
Julie Burdick
Derek D. Burkhalter
Anthony Robert Bustillo
Allison F. Carp
Daniel George Charbonneau
Nathalie Charbonneau
Todd Douglas Cheema
Yvonne W. Y. Cheng
Julia Feng-Ming Chu
Jeffrey Alan Clements
Jeffrey J. Clinch
Eric John Clymer
Carolyn J. Coe
Steven A. Cohen
Larry Kevin Conlee
Peter J. Cooper

Sean Oswald Curtis Cooper
Sharon R. Corrigan
David Ernest Corsi
Jose R. Couret
John Edward Daniel
Mujtaba H. Datoo
Catherine L. DePolo
Jean A. DeSantis
Timothy Michael DiLellio
Sophie Duval
James Robert Elicker
Gregory James Engl
Brian Michael Fernandes
Kenneth D. Fikes
Janine Anne Finan
Sean Paul Forbes
Ronnie Samuel Fowler
Mark R. Frank
Serge Gagne
James M. Gallagher
Anne M. Garside
Justin Gordon Gensler
Emily C. Gilde
Theresa Giunta
Todd Bennett Glassman
Paul E. Green Jr.
Joseph Paul Greenwood
Michael S. Harrington
Bryan Hartigan
Jeffery Tim Hay
Qing He
Amy Louise Hicks
Jay T. Hieb
Glenn R. Hiltpold
Glenn Steven Hochler
Brook A. Hoffman
Todd Harrison Hoivik
Terrie Lynn Howard
Paul Jerome Johnson
Bryon Robert Jones
Burt D. Jones
Derek A. Jones
Ung Min Kim
Thomas F. Krause
Isabelle La Palme
Travis J. Lappe
Borwen Lee
Christian Lemay
Brendan Michael Leonard
Karen N. Levine
Sally Margaret Levy
Sharon Xiaoyin Li
Dengxing Lin
Kelly A. Lysaght
Kevin M. Madigan
Vahan A. Mahdasian
Atul Malhotra
Albert Maroun
Jason Aaron Martin
Laura Smith McAnena
Timothy L. McCarthy
Rasa Varanka McKean
Sarah Kathryn McNair-Grove
Kirk Francis Menanson
Ain Milner
Michael W. Morro
John-Giang L. Nguyen
Michael Douglas Nielsen
Randall William Oja
Sheri L. Oleshko
Leo Martin Orth Jr.
Gerard J. Palisi
Prabha Pattabiraman
Michael A. Pauletti
Fanny C. Paz-Prizant
Rosemary Catherine Peck
John Michael Pergrossi
Sylvain Perrier
Christopher Kent Perry
Anthony J. Pipia
Jordan J. Pitz
Thomas LeRoy Poklen Jr.
William Dwayne Rader Jr.
Sara Reinmann
Sylvain Renaud
Mario Richard
David C. Riek
Kathleen Frances Robinson
Joseph Francis Rosta Jr.
Janelle Pamela Rotondi
Robert Allan Rowe
Joseph John Sacala
James C. Santo

Frances Ginette Sarrel
Jason Thomas Sash
Jeremy Nelson Scharnick
Jeffery Wayne Scholl
Annmarie Schuster
Peter Abraham Scourtis
David Garrett Shafer
Vladimir Shander
Seth Shenghit
Mark Richard Strona
Jayme P. Stubitz
Stephen James Talley
Jo Dee Thiel-Westbrook
Robert M. Thomas II
Jennifer L. Throm
Gary Steven Traicoff
Andrea Elisabeth Trimble
Brian K. Turner
Jon S. Walters
Douglas M. Warner
David W. Warren
Kevin Earl Weathers
Trevar K. Withers
Meredith Martin Woodcock
Jonathan Stanger Woodruff
Perry Keith Wooley
Yin Zhang
Steven Bradley Zielke

Mr. Lehmann then introduced M. Stanley Hughey, a past president of the Society, who presented the Address to New Members.

David R. Chernick, CAS vice president-programs and communications, spoke to the meeting participants about the highlights of this meeting and what was planned in the program.

James Surrago, vice chairperson of the Continuing Education Committee, announced that three *Proceedings* papers and two discussions of *Proceedings* papers would be presented at this meeting. In all, five papers were accepted for publication in the 1999 *Proceedings of the Casualty Actuarial Society*.

Mr. Surrago also gave a brief description of this year's Call Paper Program on Securitization of Risk. He announced that all of the call papers would be presented at this meeting. In addition, the papers were published in the 1999 CAS *Discussion Paper Program* and could be found on the CAS Web Site. Mr. Surrago presented the Michelbacher Prize to Richard W. Gorvett for his paper, "Insurance Securitization: The Development of a New Asset Class," and to Donald F. Mango for his paper, "Risk Load and the Default Rate of Surplus." The Michelbacher Prize commemorates the work of Gustav F. Michelbacher and honors the authors of the best paper(s) 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.

Mr. Lehmann then began the presentation of other awards. He 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. Mr. Lehmann announced that Jingsu Pu is the recipient of the 1999 CAS Harold W. Schloss Memorial Scholarship Fund. Pu will be presented with a $500 scholarship.

Mr. Lehmann then concluded the business session of the Spring Meeting by calling for a review of *Proceedings* papers.

Mr. Lehmann next introduced the featured speaker, Lawrence Kudlow, who is chief economist, director of research, and senior vice president of American Skandia Life Assurance, as well as a business commentator and noted economist.

The first General Session was held from 10:30 a.m. to noon.

"European Union's Impact on the Insurance Industry"

Moderator:	Terry G. Clarke Managing Principal Tillinghast-Towers Perrin
Panelists:	Catherine Cresswell Chief Actuary Government Actuary's Department
	Robert P. Hartwig Vice President & Chief Economist Insurance Information Institute
	Jay B. Morrow Vice President & Actuary American International Underwriters

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. "Actuarial and Economic Aspects of Securitization of Risk"

   Authors: Samuel H. Cox
   Georgia State University

   Joseph R. Fairchild
   Georgia State University

   Hal W. Pedersen
   Georgia State University

2. "Property/Liability Insurance Risk Management and Securitization"

   Author: Trent R. Vaughn
   GRE Insurance

3. "Eliminating Mortgage Insurance through Risk-Adjusted Interest Rates (The Securitization of Mortgage Default Risk)"

   Authors: Bruce D. Fell
   Arthur Andersen LLP

   William S. Ober
   Arthur Andersen LLP

4. "Risk Load and the Default Rate of Surplus"

   Author: Donald F. Mango
   Zurich Centre ReSource, Ltd.

The concurrent sessions presented from 1:15 p.m. to 2:45 p.m. were:

1. The New Actuarial Exam Structure: Responses by Universities and Implications for Recruiting

   Moderator: Richard W. Gorvett
   Assistant Professor of Finance and Insurance
   The College of Insurance

Panelists: David M. Elkins
Senior Actuary
Allstate Insurance Company

Curtis E. Huntington
Professor of Mathematics
Director of Actuarial Program
University of Michigan

Dale S. Porfilio
Associate Actuary
Allstate Insurance Company

2. The Outlook for Automobile Insurance in the New Millennium

Moderator: Michael A. Walters
Principal
Tillinghast-Towers Perrin

Panelists: J. Parker Boone
Senior Vice President
InsWeb

Charles A. Bryan
Senior Vice President & Chief Actuary
Nationwide Insurance Company

Anne E. Kelly
Chief Casualty Actuary
New York State Insurance Department

Shirley Grogan
Assistant Vice President–Auto Pricing
The Hartford

3. What is the Right Mix: Blending of Governmental and Private Funding for Catastrophic Exposure

Moderator: David R. Chernick
Assistant Vice President & Actuary
Allstate Insurance Company

Panelists: Rade T. Musulin
Vice President–Actuary
Florida Farm Bureau Insurance Companies

Tim R. Richison
Chief Financial Officer
California Earthquake Authority

4. Professionalism Continuing Education

Moderator/Facilitator: Gregory L. Hayward
Actuary
State Farm Mutual Automobile Insurance Company

Facilitators: C. Gary Dean
Assistant Vice President & Actuary
SAFECO/American States Business Insurance

Thomas C. Griffin
Staff Attorney
American Academy of Actuaries

William J. VonSeggern
Assistant Vice President
Fireman's Fund Insurance Companies

Proceedings papers presented during this time were:

1. "Workers Compensation Reserve Uncertainty"

Authors: Douglas M. Hodes
Liberty Mutual Group

Sholom Feldblum
Liberty Mutual Group

Gary Blumsohn
F & G Re, Inc.

2. "Levels of Determinism in Workers Compensation Reinsurance Commutations"

   Author: Gary Blumsohn
   F & G Re, Inc.

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. "Pricing Catastrophe Risk: Could CAT Futures Have Coped with Andrew?"

   Authors: Stephen P. D'Arcy
   University of Illinois

   Virginia G. France
   University of Illinois

   Richard W. Gorvett
   The College of Insurance

2. "Insurance Securitization: The Development of a New Asset Class"

   Author: Richard W. Gorvett
   The College of Insurance

The additional concurrent sessions presented from 3:15 p.m. to 4:45 p.m. were:

1. Actuaries in Nontraditional Roles

   Moderator/Panelist: Paul J. Brehm
   Vice President
   St. Paul Fire & Marine Insurance Company

   Panelists: Richard R. Anderson
   Actuary
   Risk Management Solutions

Gregory V. Ostergren
President & CEO
American National Property & Casualty Company

David G. Walker
Director/Associate Actuary
Allianz Insurance Company

2. The Path to Fellowship

Moderator: Patrick K. Devlin
Senior Consultant
PricewaterhouseCoopers LLP

Panelists: Daniel L. Hogan Jr.
Assistant Vice President
Hartford Financial Services

Steven A. Kelner
Vice President
American Re-Insurance Company

Dee Dee Mays
Senior Regional Actuary
National Council on Compensation Insurance

Christopher Tait
Consulting Actuary
Milliman & Robertson, Inc.

3. Questions and Answers with the CAS Board of Directors

Moderator: Alice H. Gannon
Vice President
United Services Automobile Association

Panelists: Jerome A. Degerness
President
Degerness Consulting Services, Inc.

Gail M. Ross
Vice President
Am-Re Consultants, Inc.

Michael L. Toothman
Partner
Arthur Andersen LLP

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 18, 1999

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:

"Year 2000 Update"

Moderator/ Panelist:	Philip D. Miller Consulting Actuary Tillinghast-Towers Perrin
Panelists:	Dan Romito Vice President, Year 2000 CNA Insurance Companies
	Martin P. Sheffield Vice President A.M. Best Company

"Financial Markets and Their Impact on the Property/Casualty Industry"

Moderator:	David P. Flynn Consultant
Panelists:	Peter Bouyoucos Principal Morgan Stanley Dean Witter

Kevin T. Cronin
President
International Insurance Council

Robert Klein
Director of the Center for Risk
Management and
Insurance Research
Georgia State University

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. Additional call papers presented during this time were:

1. "Index Heritage Performance: A Bootstrap Study of Hurricane Fran"

   Authors: Xin Cao
   IndexCo, LLP

   Bruce Thomas
   IndexCo, LLP

2. "Uncertainty in Hurricane Risk Modeling and Implications for Securitization"

   Author: David Miller
   Guy Carpenter

The additional concurrent sessions presented from 10:00 a.m. to 11:30 a.m. were:

1. The Future of Workers Compensation Ratemaking

   Moderator/Panelist: Timothy L. Wisecarver
   President
   Pennsylvania and Delaware
   Compensation Rating Bureaus

   Panelists: Michael Lamb
   Casualty Actuary
   Oregon Insurance Division

	Pamela Sealand Reale Assistant Vice President & Actuary Orion Capital/EBI Companies

2. DFA in the Real World

Moderator:	Robert A. Daino Vice President Am-Re Consultants, Inc.
Panelists:	Manuel Almagro Vice President Swiss Re Investors
	Charles C. Emma Consulting Actuary Miller, Rapp, Herbers & Terry, Inc.
	John W. Gradwell Associate Actuary Sedgwick Re Insurance Strategy, Inc.

3. Emerging Financial Products

Moderator:	William F. Dove Vice President Centre Solutions
Panelists:	Michael K. Curry Senior Vice President Capital Reinsurance Company
	Eugene O'Keane Vice President American Re Financial Products
	Scott M. Sanderson Senior Vice President J&H Marsh & McLennan

4. Loss Portfolio Transfers

Moderator/ Panelist:	Chris E. Nelson Vice President CNA Re

Panelists: Jean A. Connolly
Director
PricewaterhouseCoopers LLP

Elizabeth E. L. Hansen
Senior Vice President
E. W. Blanch Company, Inc.

Various CAS committees met from 12:00 p.m. to 5:00 p.m. Presentation of call papers, concurrent sessions, and *Proceedings* papers continued from 12:30 p.m. to 2:00 p.m. Certain call papers and concurrent sessions presented earlier were repeated. The additional call paper presented during this time was:

1. "Catastrophe Risk Securitization: Insurer and Investor"

   Authors: Glenn G. Meyers
   Insurance Services Office, Inc.

   John J. Kollar
   Insurance Services Office, Inc.

Proceedings papers presented during this time were:

1. Discussion of the discussion of "Surplus—Concepts, Measures of Return, and Determination"

   (by Russell E. Bingham, *PCAS*, LXXX, 1993, p. 55)

   (Discussion by Robert K. Bender, *PCAS*, LXXXIV, 1997, p. 44)

   Discussion by: David L. Ruhm
   AIG Risk Finance

   Carlton R. Grose
   Universal Underwriters Group

2. "A Systematic Relationship Between Minimum Bias and Generalized Linear Model"

   Author: Stephen J. Mildenhall
   CNA Re

3. Discussion of "The Complement of Credibility"
   (by Joseph A. Boor, *PCAS*, LXXXIII, 1996, p. 1)
   Discussion by: Sholom Fledblum
   Liberty Mutual Group

All members and guests enjoyed dinner at a Coney Island Beach Party from 6:30 p.m. to 10:00 p.m.

Wednesday, May 19, 1999

Certain call papers and concurrent sessions that had been presented earlier during the meeting were repeated this morning from 8:00 a.m. to 9:30 a.m. Additional concurrent sessions presented were:

1. Joint Code of Professional Conduct
   Moderator: Jack M. Turnquist
   Totidem Verbis
   Panelists: Lauren M. Bloom
   General Counsel
   American Academy of Actuaries

   C. Gary Dean
   Assistant Vice President & Actuary
   SAFECO/American States Business Insurance

   Roger A. Schultz
   Assistant Vice President
   Allstate Insurance Company
2. Social Security
   Moderator: Ron Gebhardtsbauer
   Senior Pension Fellow
   American Academy of Actuaries
   Panelists: Gareth Davis
   Policy Analyst
   The Heritage Foundation

Cori E. Uccello
Research Associate
The Urban Institute

3. Pricing Unique Exposures

Moderator/ Panelist:	Beth E. Fitzgerald Assistant Vice President Insurance Services Office, Inc.
Panelists:	Paul C. Martin Consulting Actuary
	Bernard H. Gilden Property Actuary The Hartford

4. Managing the “Managed”—Changing Liabilities Within the Health Care System

Moderator:	Bernard Horovitz Actuary Executive Risk Indemnity
Panelists:	Susan Huntington Director of Healthcare Risk Management Executive Risk Indemnity
	Richard B. Lord Assistant Actuary Milliman & Robertson, Inc.

After a refreshment break, the final General Session was held from 10:00 a.m. to 11:30 a.m.:

“A Chief Actuary Discussion on Market Behavior”

Moderator:	Phillip N. Ben-Zvi Principal-In-Charge PricewaterhouseCoopers LLP
Panelists:	John C. Burville Chief Actuary ACE Limited

Charles H. Dangelo
President
AIG Risk Management, Inc.

Frederick O. Kist
Senior Vice President and Corporate Actuary
CNA Insurance Companies

Steven G. Lehmann officially adjourned the 1999 CAS Spring Meeting at 11:45 a.m. after closing remarks and an announcement of future CAS meetings.

Attendees of the 1999 CAS Spring Meeting

The 1999 CAS Spring Meeting was attended by 319 Fellows, 264 Associates, and 62 Guests. The names of the Fellows and Associates in attendance follow:

FELLOWS

Mark A. Addiego
Stephanie J. Albrinck
Terry J. Alfuth
Manuel Almagro Jr.
Larry D. Anderson
Richard R. Anderson
Robert A. Anker
Lawrence J. Artes
Timothy J. Banick
W. Brian Barnes
Gregory S. Beaulieu
Douglas L. Beck
Allan R. Becker
Stephen A. Belden
David M. Bellusci
Phillip N. Ben-Zvi
Douglas S. Benedict
Regina M. Berens
Michele P. Bernal
William P. Biegaj
Richard A. Bill
Gavin C. Blair
Jean-François Blais
Ralph S. Blanchard III
Gary Blumsohn
J. Parker Boone
Joseph A. Boor
Ronald L. Bornhuetter
Wallis A. Boyd Jr.
George P. Bradley
Betsy A. Branagan
Paul J. Brehm
Charles A. Bryan
James E. Buck
George Burger
Mark E. Burgess
Elliot Ross Burn
John F. Butcher II
J'ne Elizabeth Byckovski
Christopher S. Carlson
Kenneth E. Carlton
Lynn R. Carroll
Michael J. Caulfield
Dennis K. Chan
Scott K. Charbonneau
David R. Chernick
Gary C.K. Cheung
Rita E. Ciccariello
Gregory J. Ciezadlo
Kay A. Cleary
Eugene C. Connell
Martin L. Couture
Catherine Cresswell
Frederick F. Cripe

Alan M. Crowe
Michael K. Curry
Robert J. Curry
Michael T. Curtis
Stephen P. D'Arcy
Robert A. Daino
Joyce A. Dallessio
Charles H. Dangelo
Thomas J. DeFalco
Curtis Gary Dean
Brian H. Deephouse
Jerome A. Degerness
Howard V. Dempster
Patrick K. Devlin
Edward D. Dew
Stephen R. DiCenso
Jeffrey F. Deigl
James L. Dornfeld
Victor G. dos Santos
William F. Dove
Michael C. Dubin
Brian Duffy
Thomas J. Duffy
M. L. Butch Dye
Jeffrey Eddinger
Dale R. Edlefson
Douglas D. Eland
David M. Elkins
Thomas J. Ellefson
Charles C. Emma
Glenn A. Evans
Philip A. Evensen
John S. Ewert
Alana C. Farrell
Dennis D. Fasking
Sholom Feldblum
Bruce D. Fell
Carole M. Ferrero
Ginda Kaplan Fisher
Russell S. Fisher
Beth E. Fitzgerald
David P. Flynn
Edward W. Ford
Christian Fournier
Bruce F. Friedberg
Patricia A. Furst
Scott F. Galiardo
Alice H. Gannon
Louis Gariepy
Eric J. Gesick
Robert A. Giambo
John F. Gibson
Gregory S. Girard
Bradley J. Gleason
Daniel C. Goddard
Leonard R. Goldberg
Irwin H. Goldfarb
Charles T. Goldie
Richard W. Gorvett
Linda M. Goss
Gregory S. Grace
Steven A. Green
Russell H. Greig Jr.
Carleton R. Grose
Terry D. Gusler
David N. Hafling
Greg M. Haft
Robert C. Hallstrom
Elizabeth E. L. Hansen
Christopher L. Harris
Roger M. Hayne
David H. Hays
Gregory L. Hayward
Barton W. Hedges
Dennis R. Henry
Teresa J. Herderick
Steven C. Herman
Charles C. Hewitt Jr.
Daniel L. Hogan Jr.
Beth M. Hostager
Brian A. Hughes
M. Stanley Hughey
Robert P. Irvan
Christopher D. Jacks
Ronald W. Jean
Andrew P. Johnson
Daniel K. Johnson
Eric J. Johnson
Kurt J. Johnson
Mark R. Johnson
Thomas S. Johnston
Ira Mitchell Kaplan
Frank J. Karlinski III
Janet S. Katz
Allan M. Kaufman
Claudine H. Kazanecki
Hsien-Ming Keh
Brandon D. Keller
Tony J. Kellner
Anne E. Kelly
Steven A. Kelner
Kevin A. Kesby
Joe C. Kim
Deborah M. King
Frederick O. Kist
Charles D. Kline Jr.
Fredrick L. Klinker
Terry A. Knull

John J. Kollar
Mikhael I. Koski
Eleni Kourou
Thomas J. Kozik
Gary R. Kratzer
John R. Kryczka
Ronald T. Kuehn
David R. Kunze
Paul E. Lacko
Blair W. Laddusaw
David A. Lalonde
Dean K. Lamb
John A. Lamb
R. Michael Lamb
Michael A. LaMonica
Nicholas J. Lannutti
Michael D. Larson
Paul W. Lavrey
Dawn M. Lawson
Robert H. Lee
Marc-Andre Lefebvre
Steven G. Lehmann
John J. Lewandowski
John J. Limpert
Richard A. Lino
Richard W. Lo
Deborah E. Logan
Richard Borge Lord
Stephen P. Lowe
Robert G. Lowery
Aileen C. Lyle
Mark J. Mahon
Gary P. Maile
Donald F. Mango
Anthony L. Manzitto
Paul C. Martin
Isaac Mashitz
Steven E. Math
Dee Dee Mays
Heidi J. McBride
Michael G. McCarter
Charles W. McConnell
Richard T. McDonald
Liam Michael McFarlane
Stephen J. McGee
Dennis T. McNeese
Robert E. Meyer
Glenn G. Meyers
Robert S. Miccolis
Stephen J. Mildenhall
David L. Miller
Philip D. Miller
Susan M. Miller
Jay B. Morrow
Raymond D. Muller
Timothy O. Muzzey
Chris E. Nelson
Richard T. Newell Jr.
Peter M. Nonken
G. Chris Nyce
Marc F. Oberholtzer
Kevin Jon Olsen
Layne M. Onufer
Marlene D. Orr
Joanne M. Ottone
Rudy A. Palenik
Joseph M. Palmer
Chandrakant C. Patel
Bruce Paterson
Sarah L. Petersen
Steven Petlick
Dale S. Porfilio
Jeffrey H. Post
Virginia R. Prevosto
Deborah W. Price
David S. Pugel
Richard A. Quintano
Christine E. Radau
Rajagopalan K. Raman
Kiran Rasaretnam
Ralph L. Rathjen
Pamela Sealand Reale
John J. Reynolds III
Andrew S. Ribaudo
Brad M. Ritter
Kevin B. Robbins
A. Scott Romito
Deborah M. Rosenberg
Gail M. Ross
Richard J. Roth Jr.
Bradley H. Rowe
James B. Rowland
Jean Roy
Stuart G. Sadwin
Stephen Paul Sauthoff
Peter J. Schultheiss
Roger A. Schultz
Peter R. Schwanke
Robert F. Scott Jr.
Terry M. Seckel
Alan R. Seeley
Margaret E. Seiter
Peter Senak
Michael Shane
Derrick D. Shannon
Jerome J. Siewert
David Skurnick

John Slusarski
Lee M. Smith
Bruce R. Spidell
David Spiegler
Elisabeth Stadler
Douglas W. Stang
Brian M. Stoll
Edward C. Stone
Kevin D. Strous
James Surrago
Russel L. Sutter
Collin J. Suttie
Jeanne E. Swanson
Ronald J. Swanstrom
Christopher C. Swetonic
Christopher Tait
Kathleen W. Terrill
Richard D. Thomas
Kevin B. Thompson
Michael L. Toothman
Janet A. Trafecanty
Patrick N. Tures
Gail E. Tverberg
James F. Tygh
Timothy J. Ungashick
Jeffrey A. Van Kley
Trent R. Vaughn
Joseph L. Volponi
William J. VonSeggern
Gregory M. Wacker
Robert H. Wainscott
Mavis A. Walters
Michael A. Walters
Bryan C. Ware
Dominic A. Weber
John P. Welch
Geoffrey T. Werner
David C. Westerholm
Charles S. White
David L. White
Mark Whitman
Kevin L. Wick
Chad C. Wischmeyer
Timothy L. Wisecarver
Paul E. Wulterkens
Floyd M. Yager
Richard P. Yocius
Heather E. Yow
James W. Yow
Doug A. Zearfoss

ASSOCIATES

Jason R. Abrams
Michael B. Adams
Stephen A. Alexander
Anthony L. Alfieri
Silvia J. Alvarez
Athula Alwis
Gwendolyn Anderson
Paul D. Anderson
Nancy L. Arico
Nathalie J. Auger
Glenn R. Balling
Joanne Balling
Phillip W. Banet
Amy L. Baranek
Patrick Beaudoin
David J. Belany
Kristen M. Bessette
John T. Binder
Mario Binetti
Christopher D. Bohn
Raju Bohra
Mark E. Bohrer
John P. Booher
David R. Border
Sherri L. Border
Thomas S. Botsko
Erik R. Bouvin
Steven A. Briggs
Stephane Brisson
Karen A. Brostrom
Conni J. Brown
Paul E. Budde
Julie Burdick
Hugh E. Burgess
Derek D. Burkhalter
William E. Burns
Anthony R. Bustillo
Sandra L. Cagley
Allison F. Carp
Paul A. Chabarek
Daniel G. Charbonneau
Nathalie Charbonneau
Debra S. Charlop
Todd D. Cheema
Yvonne W. Y. Cheng
Michael J. Christian
Theresa A. Christian
Julia Feng-Ming Chu
Christopher J. Claus
Jeffrey A. Clements
Jeffrey J. Clinch

Eric J. Clymer
Carolyn J. Coe
Steven A. Cohen
Larry K. Conlee
Peter J. Cooper
Sean O. Cooper
Sharon R. Corrigan
David E. Corsi
William F. Costa
Jose R. Couret
Kathleen T. Cunningham
John E. Daniel
Todd H. Dashoff
Mujtaba H. Datoo
Catherine L. DePolo
Timothy M. DiLellio
Kevin F. Downs
Sara P. Drexler
Sophie Duval
Brian M. Fernandes
Kenneth D. Fikes
Robert F. Flannery
Sean Paul Forbes
Sarah Jane Fore
Ronnie S. Fowler
Mark R. Frank
Serge Gagne
James M. Gallagher
Donald M. Gambardella
Anne M. Garside
Lynn A. Gehant
Christine A. Gennett
Justin G. Gensler
Emily C. Gilde
Bernard H. Gilden
Steven B. Goldberg
Jay C. Gotelaere
John W. Gradwell
Gary Granoff
Paul E. Green Jr.
Joseph P. Greenwood
David J. Gronski
Jacqueline Lewis Gronski
William A. Guffey
Nasser Hadidi
John A. Hagglund
Aaron Halpert
Michael S. Harrington
Bryan Hartigan
Jeffery T. Hay
Qing He
Joseph A. Herbers
Amy L. Hicks
Jay T. Hieb
Glenn R. Hiltpold
Gary P. Hobart
Glenn S. Hochler
Brook A. Hoffman
Todd H. Hoivik
Eric J. Hornick
Bernard R. Horovitz
Terrie L. Howard
Gloria A. Huberman
John J. Javaruski
Brian E. Johnson
Paul J. Johnson
Bryon R. Jones
Burt D. Jones
Derek A. Jones
Daniel R. Kamen
Ung Min Kim
Martin T. King
Kirk L. Kutch
Isabelle La Palme
Travis J. Lappe
Betty F. Lee
Borwen Lee
Ramona C. Lee
Todd W. Lehmann
Christian Lemay
Bradley H. Lemons
Charles Letourneau
Karen N. Levine
Craig A. Levitz
John N. Levy
Sally M. Levy
Philip Lew
Sharon Xiaoyin Li
Dengxing Lin
Elizabeth Long
Ronald P. Lowe Jr.
Kelly A. Lysaght
Daniel P. Maguire
Cornwell H. Mah
Vahan A. Mahdasian
Atul Malhotra
Sudershan Malik
Albert Maroun
Joseph Marracello
Jason Aaron Martin
Tracey L. Matthew
Laura Smith McAnena
Timothy L. McCarthy
Phillip E. McKneely
Kirk F. Menanson

William A. Mendralla
Ain Milner
Michael W. Morro
Michael J. Moss
Robert J. Moss
Rade T. Musulin
John-Giang L. Nguyen
Michael D. Nielsen
James D. O'Malley
Randall W. Oja
Sheri L. Oleshko
Douglas W. Oliver
Richard A. Olsen
Leo M. Orth Jr.
Gregory V. Ostergren
John A. Pagliaccio
Gerard J. Palisi
Prabha Pattabiraman
Michael A. Pauletti
Fanny C. Paz-Prizant
Rosemary C. Peck
Jeremy P. Pecora
Claude Penland
John M. Pergrossi
Sylvain Perrier
Christopher K. Perry
Anthony J. Pipia
Jordan J. Pitz
Thomas L. Poklen Jr.
Kathy Popejoy
Matthew H. Price
Patricia A. Pyle
Sasikala Raman
James E. Rech
Sara Reinmann
Sylvain Renaud
Karin M. Rhoads
W. Vernon Rice
Mario Richard
Christopher R. Ritter
Kathleen F. Robinson
Rebecca L. Roever
Christine R. Ross
Sandra L. Ross
Joseph F. Rosta Jr.
Scott J. Roth
Janelle P. Rotondi
Robert A. Rowe
David L. Ruhm
Joanne E. Russell
Stephen P. Russell
Maureen S. Ruth
John P. Ryan
Joseph J. Sacala
Asif M. Sardar
Frances G. Sarrel
Jason T. Sash
Susan C. Schoenberger
Jeffery W. Scholl
Annmarie Schuster
Peter A. Scourtis
Michael L. Scruggs
David G. Shafer
Vladimir Shander
Seth Shenghit
James J. Smaga
Katherine R. S. Smith
David C. Snow
Klayton N. Southwood
Mark R. Strona
Jayme P. Stubitz
Lisa M. Sukow
Brian K. Sullivan
Stephen J. Talley
Robert M. Thomas II
Jennifer L. Throm
Nanette Tingley
Gary S. Traicoff
Andrea E. Trimble
Brian K. Turner
Karen P. Valenti
Phillip C. Vigliaturo
Jerome F. Vogel
David G. Walker
Jon S. Walters
Gregory S. Wanner
Stephen D. Warfel
Douglas M. Warner
David W. Warren
Kevin E. Weathers
Lynne K. Wehmueller
Robert G. Weinberg
Russell B. Wenitsky
Jo Dee Westbrook
Michael W. Whatley
Lawrence White
Thomas J. White
Mary E. Wills
William F. Wilson
Bonnie S. Wittman
Robert F. Wolf
Meredith M. Woodcock
Jonathan S. Woodruff
Perry K. Wooley
Robert S. Yenke
Yin Zhang
Steven B. Zielke
Edward J. Zonenberg

Volume LXXXVI, Part 2 **No. 165**

PROCEEDINGS
November 14, 15, 16, 17, 1999

RESIDUAL MARKET PRICING

RICHARD B. AMUNDSON

Abstract

Residual market plans often review their rates based on the experience of the plans themselves. The typical result is an indication for a large increase, which the regulator then judgmentally reduces. To the extent that equilibrium exists between voluntary and residual markets, it results from ignoring the indications. Plans' experience can call for rate decreases as well as increases, especially with no allowance for profit. Indications for decreases are politically harder to ignore and could destroy the voluntary market if followed.

Break-even residual market pricing, if truly followed, has unpredictable consequences on prices and market shares for the residual and the voluntary markets. This paper proposes an alternative to break-even pricing. With input from all concerned, a state should first

establish specific goals for the residual market plan in terms of market share, burden on insureds in the voluntary market, and maximum surcharge for insureds in the plan. Regulators can then set plan prices at a consistent level above voluntary prices to meet the established goals.

1. INTRODUCTION

1.1. A Paradox

In 1983, the State of Minnesota merged its departments of insurance, banking, and securities into a single Department of Commerce. The first commissioner of the newly created Department was determined to keep consumer prices down wherever possible. Among the duties of the Department was to review the rates of the assigned risk plan (ARP). During the seven years ending with 1989, despite many requests for rate increases, the Department allowed only a single increase in the state's private passenger automobile assigned risk plan. At the beginning of that period, ARP judged its rates to be adequate; at the end, ARP calculated a needed increase of 10.3%, with the one increase in the interim being 14.8%. That implies an average annual needed increase of 3.4% during those seven years ($1.034 = (1.103 \times 1.148)^{1/7}$). Annual increases of 3.4% were modest at the time, so the commissioner's strategy of holding down ARP rates appeared to be successful.

A change of commissioners in 1989 brought a new philosophy, one that permitted ARP rates to rise. Between 1989 and 1994, ARP took increases of 12.0%, 20.4%, 19.5%, and 13.8%; and there was still an indication of 33.7% at the end of that period. That implies an average annual needed increase of 17.3% during those five years. ARP was smaller at the end of the period, but the goal that it be self-supporting was as far away as ever. Loss ratios stayed high as rates went up, and the drivers that remained insured with ARP had little to celebrate. External economic indices did nothing to explain the sudden shift from annual cost increases of 3.4% to increases of 17.3%. The only obvious change was the Department's change in attitude toward change itself: the culprit appeared to be the strategy of letting ARP follow its own indicated rate increases.

1.2. An Actuarial Explanation

None of this was hard to explain. In the beginning, insurers rejected only the very most unwanted drivers—the worst of the worst. They were happy to write a borderline driver for $1,000. But when inflation pushed the voluntary market price for that driver's policy up to $1,100, ARP, whose rates had not budged, might write the driver at $1,050. These borderline drivers moving into ARP were the best of the worst, and they improved the quality of ARP's book of business as it grew. Exactly the opposite occurred when ARP shrank. When ARP's prices began increasing faster than those of the voluntary market, ARP's insureds began moving to the voluntary market to get better prices. The voluntary market was interested only in the best of ARP's business, of course; and, when ARP lost its best customers, its loss ratio began to climb.

After years of increases, when things were back to the original balance between voluntary and assigned risk, the indications for ARP were as high as ever. The actuary at the Department wrote a memo explaining why this was and what one might have to do in the future to keep everything in balance. To continue following indications blindly seemed sure to lead to the disappearance of ARP—not a bad idea in the eyes of some, but not politically viable in this case. The presence of a contingency factor in the analysis posed a problem; it added to the price of each policy, not unlike a profit margin, even though this was non-profit business. ARP rates tended to rise mercilessly; and the contingency factor only exacerbated the tendency, pushing rates for the dwindling number of policyholders to truly unaffordable levels. It seemed a good idea to get rid of the contingency factor.

1.3. A Second Paradox

The Department also regulates the workers compensation assigned risk plan. In 1995, something surprising began to occur: this ARP, whose rates were already low, needed rate *decreases*. Whether this was just random noise or a true reflection of the risks in ARP, it seemed unwise for the rates to get too close to the voluntary market rates. The voluntary market charges for the same coverages as ARP but, in addition, charges for profit because of the risk of writing business. The ARP analysis had no charge for risk even though, of course, the ARP business is just as risky as the voluntary business. This gave ARP a rate advantage—it could pick up market share and constantly improve its book, and the voluntary market could eventually disappear. The Department actuary reasoned that one might prevent that disaster by including a contingency factor in the analysis to keep rates from falling too low.

All this was strangely familiar. The same actuary (who happens to be the author of this paper) had argued, not so long before, against a contingency factor in the case of auto assigned risk. What was wrong? What was the truth?

1.4. The Scales Fall From Our Eyes

The truth is all of the above. Both of these scenarios can happen, even though they are complete opposites. A residual market that bases its prices on its own experience has no certainty of reaching an acceptable equilibrium, as this paper will demonstrate. To achieve the goals normally desired for an assigned risk plan, the state should base the plan's rates on voluntary market rates and not on the plan's own experience.

2. A MODEL OF RESIDUAL MARKET PRICING

2.1. Some Assumptions

We will look at residual market plans that set prices to break even based on their own experience. Of course, with break-even pricing, a plan may still realize profits or losses. The plan design may or may not give the profit to insurers, but it will virtually always give insurers the loss. The examples in this paper assume that insurers get the profit as well as the loss. The conclusions of the paper are still valid if insurers do not get the profit, but the examples are a bit more complex.

We will ignore self-insurance. Assume that all employers, drivers, etc., must buy insurance and that they have two options: an insurance company in the voluntary market or ARP, our surrogate for all residual markets. Assume further that within each classification there is a continuum of expected losses per exposure: there are insureds with very few losses expected for each exposure unit, there are others with very high expected losses, and there is everything in between.

Let us look at a simplified financial model that illustrates some important relationships between the residual and voluntary markets. First suppose there is no ARP. Now imagine an insurer that needs a \$100 investment in surplus to take on \$200 of expected loss at the end of the coming year and that there are no expenses. Further suppose that one can earn 5% risk-free on invested assets and that, given the uncertainty in the expected losses, the insurer needs a 15% return on the venture. Thus, if it collects \$200 in premium up front and invests it along with the \$100 of surplus, it will earn \$15 during the year. Then if losses materialize as expected, the insurer will pay out \$200 at year-end and will keep the original \$100 plus \$15 of investment income—the expected return is exactly what the insurer needs.

From the extreme where a for-profit, voluntary market collects all the premium, let us go to the opposite extreme where the non-profit ARP collects all the premium and pays the entire \$200 of loss. The voluntary insurer now has no premium, but it has continued responsibility for potential bottom line losses of ARP. Even with no premium, the insurer still needs the entire \$100 in surplus that it needed when it was the one collecting premium and paying claims. That \$100 was to protect against insolvency, and all the risks that it protected against still exist. Not only do they still exist, but they are all on the back of the insurer. ARP carries no surplus and assesses the insurer for any losses at the end of the day, whether they arise from excessive claims or from investments or from anything else.

Remember, moreover, that one can get a return of \$5 with no risk. An insurer might want to add some risk in exchange for an increased return. In the extreme case where the insurer has no premium, though, if the insurer did not share in ARP's profit, it would be taking on risk in exchange for a *decreased* return. The insurer will be interested in assuming ARP's risk only if it gets the full profit that it would have gotten in the absence of ARP. In order to realize the full profit, ARP must charge the full \$200 of premium. Thus, *no matter what market share ARP has*, the system still needs the full \$100 of surplus and the full \$200 of premium.

The preceding argument assumed that private insurers are at risk for residual market losses, so one might be tempted to assume that the result does not hold in the absence of private insurance. By eliminating private insurance, might premiums be reduced? No. The argument did not rely on the private status of the insurers; the risk remains whether or not private investors are bearing it. The risk takers, whether taxpayers or policyholders, will put up the surplus and reap the rewards explicitly or implicitly.

Let us turn our attention away from the extremes and consider the more usual case. Typically, ARP will have part of the market and insurers will have the rest. Consider a single premium group: all insureds of like size in a single class. Suppose ARP charges a premium of R for a member of this group. ARP may vary its rate somewhat due to merit rating; but, unlike the voluntary market, it does not do any underwriting, so it will not charge the variety of rates typical of the voluntary market. Assume that ARP charges the same rate to all insureds in the group. The voluntary market by contrast, through the forces of underwriting and competition, charges a rate proportional to expected losses. This will result from a combination of schedule rating, experience rating, retrospective rating, and underwriting by companies with differing rates and differing niches. Remember that there is a whole spectrum of expected losses. For the moment, assume that the underwriting cost is negligible; it will not change the result to assume it is significant, but it clutters the argument. Let the market price be ax, where x is the expected loss. In order to attract any business the market must charge less than ARP.

2.2. A Natural Limit: Assigned Risk Must Charge Strictly More Than Market Average

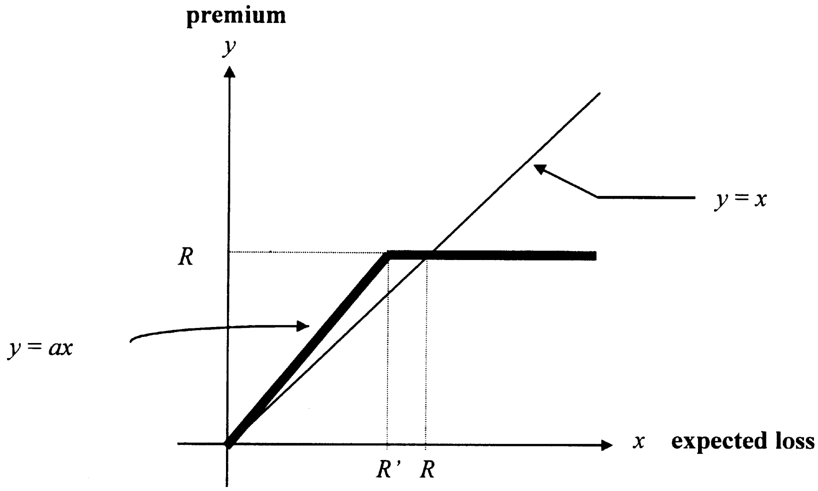
The graph in Figure 1 illustrates the market in equilibrium. The x-axis represents expected losses; the y-axis, premium. We continue to ignore expenses and to assume that investment income alone will generate appropriate profit for insurers. In an unfettered market, the line $y = x$ represents the appropriate relationship between premium and loss. With ARP charging a premium of R and the voluntary market charging ax, the bold curve represents actual prices charged. If the expected loss is greater than R', where $R' = R/a$, ARP will write the risk. If the expected loss is less than R', the voluntary market will write the risk.

Insureds whose expected losses are less than R pay more than they would in a completely free market, while insureds whose expected losses are greater than R pay less.

We will show that if $R \leq L$, where L is the average expected loss, there is no solution to the pricing problem of insurers. That is, there is no premium they can charge that would attract customers and would give them enough to pay claims and adequately reward them for the risks they would be taking.

FIGURE 1

THE MARKET IN EQUILIBRIUM



If $R > L$, there is a solution but it is not necessarily stable. If R increases or decreases depending on ARP's own experience, ARP will most likely not be in equilibrium: it will grow or shrink depending on the distribution of expected losses.

For the case $R \leq L$, it is almost self-evident that insurers can not compete. If there are n insureds, the total premium needed is nL. If ARP has m insureds, its premium will be mR. The voluntary market must then collect a total of $nL - mR$ from the remaining $n - m$ policyholders. If $R \leq L$, then $(nL - mR)/(n - m) \geq (nR - mR)/(n - m) = R$; that is, the voluntary market would have to charge on average *at least* as much as ARP.

Given that there is a continuum of expected losses, one can prove the stronger result that the insurers' pricing problem is solvable if and only if R is strictly greater than L. Furthermore the solution, when it exists, is unique.

This is easy to visualize with the help of Figure 1. The bold curve on the graph represents the prices of the combined voluntary and residual markets. The line $y = x$ represents prices in the absence of a residual market. The voluntary market seeks a value of a for which the overall average price of the bold curve is exactly the same as for the line $y = x$. For small values of a, the entire bold curve will be below the line $y = x$. As a increases, the bold curve approaches the horizontal line $y = R$. The average price will increase from 0 to R as a increases from 0 to ∞, but the average will never quite reach R for any finite value of a. Thus if $R \leq L$, the average price generated by the bold curve can never be as great as L, the average generated by the line $y = x$. If $R > L$, there must be some point at which, as a approaches ∞, the average price represented by the bold curve equals L. (Appendix A provides a more complete proof of this result.)

What this has demonstrated so far is that, however ARP sets its rates, it should not simply gear them to the average risk. They must be higher; otherwise the voluntary market will deconstruct. The danger that ARP will gear its rates to the average risk increases as ARP's market share increases. Because the argument above applies to a single class, the danger is not limited to the case where average ARP rates are higher than the overall market rates—ARP can take over the market segment by segment. If ARP sets its rates for the average risk and, in addition, includes no allowance for profit, the voluntary market has no choice but to abandon the segment in question.

3. THE ELUSIVE SEARCH FOR EQUILIBRIUM

3.1. The Rate Review

Let us suppose that $R > L$ and that the market has spent some time in equilibrium in the sense that the relative prices and market shares of ARP and the voluntary market have remained stable. Now the time has arrived for ARP to review its rates. What happens? Look back to the graph in Figure 1. ARP has been overcharging insureds with expected losses between R' and R and undercharging those with expected losses greater than R. The net effect is an undercharge, which the voluntary market makes up by overcharging all its insureds.

Because ARP has been undercharging, shouldn't its experience indicate that it needs an increase? Not necessarily. ARP has been undercharging when one considers the need for profit, but ARP does not include a profit margin in its rate analysis. It is possible that ARP has charged enough to pay claims and that its analysis on a non-profit basis will show a need for a rate decrease. This is not the normal course of events with residual market plans, but it is possible, especially for individual segments of the market. Whether ARP's analysis will show the need for an increase or for a decrease is a function of the distribution of expected losses. One can construct distributions that go both ways, as the examples in Tables 1 through 4 (discussed later in this paper) will illustrate.

If ARP uses a market-level profit margin in its analysis, it will generally see the need for an increase. Residual market plans often do include a "contingency" allowance, which serves somewhat the same purpose and does increase the probability that the analysis will indicate the need for a rate increase. For just the right distribution, just the right value of R, and just the right contingency factor, equilibrium may occur; but it will be precarious.

The tendency is rather for continual indications for rate increases, or continual indications for decreases. In the first case, if ARP follows the indications, it will eventually price itself out of existence; in the second case, it is the voluntary market that will disappear if ARP follows the indications. The more common scenario is the first; and equilibrium usually occurs only because ARP ignores the indications: ARP takes lesser increases at the insistence of the regulator. Because this is an inherently unpredictable road to equilibrium, it opens the door to many problems.

The more serious scenario, and fortunately the more rare so far, is the one in which ARP sees a need for a decrease. It is more serious because if ARP follows its indications under this scenario, the voluntary market may well disappear. As in the case where increases are indicated, the only sure way to remain in equilibrium is to ignore the indications; but that is not easy in the face of political pressures to lower rates. Let us look at some simple, finite examples that show the two possibilities.

3.2. Assigned Risk Plans That Follow Their Own Experience May Grow

First, continuing with our earlier assumptions, imagine a distribution of expected losses with ten equally likely possible outcomes: the integers ranging from 20 to 29. The voluntary market with its diversity of players and underwriting capabilities distinguishes among policies with different expectations and charges accordingly, while ARP takes all comers at the same price. The voluntary market sets its prices for a break-even underwriting return, getting its profit from investment income. ARP prices at a 5% discount in order to break even *after* investment income (i.e., ARP is non-profit). Table 1 summarizes this situation.

X is the random variable representing a policy's expected losses, with its ten possible outcomes (in column 1) each having a probability of 0.10 (column 2). The data in columns 3, 4, 5 and 6 assume that ARP writes all risks with expected losses greater than the value of x in column 1. If ARP writes all the risks with expected losses greater than 20, for example, it will have to charge 23.81 per risk in order to break even (column 4, first row). With investment income, it will have $25.00 = 23.81 \times 1.05$ to pay claims (25.00 is the average value of expected losses for policies whose expected losses are greater than 20).

The first entry in the third column, 30.71, is what the voluntary market would have to charge for a risk with expected losses of 20, given that ARP writes everything with greater expected losses. The voluntary market must collect not only the

TABLE 1
AN EXPANDING ARP WITH LIMITED EQUILIBRIUM

(1)	(2)	(3)	(4)	(5)	(6)
		If ARP writes all risks with expected losses greater than x			
x	$P[X = x]$	voluntary market rate for x	ARP rate (for all)	voluntary market rate for $x+1$	1: ARP gains −1: ARP loses 0: equilibrium
20	0.10	30.71	23.81	32.25	1
21	0.10	25.98	24.29	27.21	1
22	0.10	25.03	24.76	26.16	1
23	0.10	25.02	25.24	26.11	0
24	0.10	25.40	25.71	26.46	0
25	0.10	25.97	26.19	27.01	0
26	0.10	26.65	26.67	27.67	0
27	0.10	27.39	27.14	28.40	1
28	0.10	28.18	27.62	29.19	1
29	0.10	29.00			

24.5 = average expected loss = $E[X]$

Column (4): ARP rate = $A(x) = E[X \mid X > x]/1.05$
Column (3): vol mkt rate for $x = V(x) = ax$,
where $a = (E[X] - A(x) \cdot P[X > x])/(E[X \mid X \leq x] \cdot P[X \leq x])$
Column (5): vol mkt rate for $x + 1 = V(x) \cdot (x+1)/x = a \cdot (x+1)$

20 needed to pay the claims and provide for the profit for the risks that it writes, but it must also collect enough to provide for the profit on all the risks that ARP writes, since it (and not ARP) is taking on the risk. The combined premium that the voluntary market and ARP collect would then be, on average, 24.5 ($0.1 \times 30.71 + 0.9 \times 23.81$). The overall expected loss is 24.5 and exactly what is needed to keep the voluntary market in the game. That forces the voluntary market to charge more than ARP ($30.71 > 23.81$), so the voluntary market would lose the risks with expected losses of 20 to ARP in this situation. The 1 in the sixth column of the first row is a flag to indicate that ARP would capture this risk, too, once it had all the larger risks.

We assume that the voluntary market uniformly loads its expected ARP assessment by applying the multiplier, *a* to the rate that it would otherwise charge. The voluntary market would then charge 32.25 (column 5) for a risk with expected losses of 21, again given that ARP writes everything with expected losses greater than 20. If the voluntary market rate in column 5 were less than the ARP rate in column 4, then ARP would lose the risks with expected losses of 21 to the voluntary market; in that case, the flag in column 6 would be set to -1. A zero in column 6 indicates equilibrium, and occurs when the voluntary market rate for x is less than the ARP rate, which is in turn less than the voluntary market rate for $x+1$ (i.e., column 3 < column 4 < column 5).

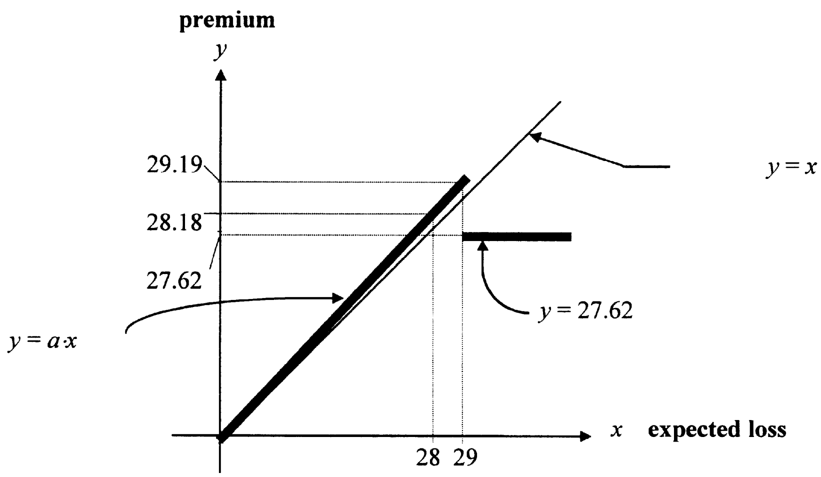
Each row represents a distinct rating scenario: the columns of voluntary market rates for x and $x+1$ are not lists of rates all of which would be available at the same time. For example, the table contains two voluntary market rates for risks with expected losses of 21: 32.25 in row 1, column 5, and 25.98 in row 2, column 3. 32.25 is the voluntary market rate if ARP writes everything greater than 20, while 25.98 is the voluntary market rate if ARP writes everything greater than 21. The full schedule of voluntary market rates is not displayed for every ARP rate; the table displays only the two rates (in columns 3 and 5), which lie at the boundary of ARP's book of business for the row in question. To know if ARP will grow or shrink or remain in equilibrium, we need only look at the boundary.

For each row of Table 1, one could construct a graph similar to that in Figure 1. Figure 2, for example, corresponds to the row $x = 28$ of Table 1. As in Figure 1, the bold line segment through the origin represents the premium that the voluntary market charges, while the bold horizontal segment represents ARP's premium. The premiums represented by the bold line segments generate an average premium of $L = E[X]$, just as in the case of Figure 1. The obvious difference is that the graph in Figure 2 is discontinuous.

For Figure 1, we required the two segments to join at (R', R); and we varied R' (by varying a) to obtain an adequate total premium, without regard for the adequacy of ARP by itself. We

FIGURE 2

THE MARKET IN DISEQUILIBRIUM



showed that, for $R > L$, there is always an R' that solves this problem.

For Figure 2, we fix the left end point of the horizontal ARP segment at 29 on the x-axis and allow the segment to move up or down until ARP's premium balances its own discounted expected losses. The voluntary market segment then pivots at the origin to attain the desired total premium. The discontinuity in the graph represents a state of disequilibrium between ARP and the voluntary market. ARP is momentarily in balance but the system is not: ARP sets its rates for one group of insureds, but the rates themselves will cause that group to change.

If ARP starts out writing only risks with expected losses greater than 28, it will charge 27.62 (29.00/1.05). Because the voluntary market must then charge 28.18 for a risk with expected losses = 28, ARP, with its lower price, will take over this level as well. ARP's price (based on its own new experience for the

TABLE 2
A VANISHING VOLUNTARY MARKET

(1)	(2)	(3)	(4)	(5)	(6)
		If ARP writes all risks with expected losses greater than x			
x	$P[X = x]$	voluntary market rate for x	ARP rate (for all)	voluntary market rate for $x + 1$	1: ARP gains −1: ARP loses 0: equilibrium
20	0.0028	429.57	23.35	451.05	1
21	0.0095	115.79	23.38	121.30	1
22	0.0316	47.96	23.46	50.14	1
23	0.1053	29.86	23.65	31.16	1
24	0.3508	25.23	24.21	26.28	1
25	0.3508	25.23	25.14	26.24	1
26	0.1053	26.06	26.04	27.07	1
27	0.0316	27.02	26.89	28.02	1
28	0.0095	28.01	27.62	29.00	1
29	0.0028	29.00			

24.5 = average expected loss

risks with expected losses of 28 and 29) will drop to 27.14 (row $x = 27$ of Table 1). The voluntary market then needs to charge 27.39 for a risk with expected losses of 27, but that still exceeds ARP's rate, so ARP will capture the risks with expected losses of 27 too. Now, based on the experience of risks with expected losses of 27, 28 and 29, ARP will again lower its rate, this time to 26.67 (row $x = 26$ of Table 1). This time however, because the voluntary market will need only 26.65 for risks with expected losses of 26, it will keep risks with that level of expectation or better; and the market will be in equilibrium.

There is nothing robust or inevitable about this equilibrium. Table 2 presents the same scenario as Table 1, except that the probabilities have changed. The overall expected loss is still 24.5, but the distribution is more concentrated. In this case, if ARP starts with risks whose expected losses are greater than 28 and bases its future rates on its own experience, it will capture the entire market before reaching equilibrium. ARP will undercut the voluntary market at the high-priced end of the voluntary market's book, causing the high-priced business to move to ARP. This will improve ARP's experience, and ARP will lower its price. The voluntary market will have a higher risk load, which will increase the voluntary market's price. After the price adjustments, ARP will undercut the voluntary market at the next level. With the distribution shown in Table 2, the cycle will continue until ARP has all the business.

One needs to take care with the conclusions that one draws from these examples. It is true that as a distribution becomes more dispersed ARP is less likely to take over, but not all uniform distributions result in a balanced equilibrium between ARP and the voluntary market. Since one can construct examples where nearly anything happens, the only firm conclusion that one can draw is that the evolution of ARP is sensitive to the distribution of expected losses among insureds. There is no mathematical certainty of equilibrium or even of the direction that the evolution will take.

3.3. Assigned Risk Plans That Follow Their Own Experience May Shrink

Let us look at some examples where ARP's experience will lead to a rate increase. The distribution of the random variable X in Table 3 is essentially a shifted, truncated Poisson. (Think of X as defined by $X = \min(1 + Y, 10)$, where Y has a Poisson distribution with $\lambda = 2.74$. We concentrate the probabilities of the tail at 10 simply to make a readable table.) Now we see negative flags in column 6, meaning that ARP will be increasing rates and losing business to the voluntary market if it follows its own indications—even with non-profit pricing. If it starts out writing everything with expected losses greater than 2, it will have a beginning rate of 4.17. The voluntary market will undercut it with a rate of 4.13 for risks with expected losses of 3. ARP's market share will drop, ARP's rate will increase, and the voluntary market will then beat ARP's price for risks with expected losses of

TABLE 3
A SHRINKING ARP WITH EQUILIBRIUM ONLY AT TWO EXTREMES

(1)	(2)	(3)	(4)	(5)	(6)
		If ARP writes all risks with expected losses greater than x			
x	$P[X = x]$	voluntary market rate for x	ARP rate (for all)	voluntary market rate for $x+1$	1: ARP gains −1: ARP loses 0: equilibrium
1	0.0646	3.71	3.74	7.42	0
2	0.1769	2.76	4.17	4.13	−1
3	0.2424	3.32	4.79	4.43	−1
4	0.2214	4.16	5.52	5.20	−1
5	0.1516	5.08	6.32	6.10	−1
6	0.0831	6.04	7.16	7.05	−1
7	0.0379	7.02	8.02	8.02	−1
8	0.0149	8.01	8.85	9.01	0
9	0.0051	9.00	9.52	10.00	0
10	0.0021	10.00			

3.74 = average expected loss

4. The cycle will continue until the market reaches equilibrium, with ARP writing only risks with expected losses of 9 and 10 at a rate of 8.85.

This is an interesting example not just because it illustrates that ARP's experience can cause it to lose, as well as gain, market share; it also illustrates that equilibrium, even within a single distribution, can occur at extremely different points. ARP and the voluntary market can be in equilibrium if ARP writes all risks with expected losses larger than 1 at a rate of 3.74, or if ARP writes all risks with expected losses larger than 8 with a rate of 8.85. In the first case ARP will have a market share of 93.6%; in the second, 1.7% (see Table 3A of Appendix B for calculation of market shares). ARP and the voluntary market will not be in equilibrium anywhere in-between these two extremes.

A market share of 1.7% for ARP is certainly not extreme, but there is no guarantee that ARP will stop at 1.7%. Look at one

TABLE 4
A VANISHING ARP

(1)	(2)	(3)	(4)	(5)	(6)
		If ARP writes all risks with expected losses greater than x			
x	$P[X=x]$	voluntary market rate for x	ARP rate (for all)	voluntary market rate for $x+1$	1: ARP gains −1: ARP loses 0: equilibrium
1	0.5400	1.12	2.71	2.23	−1
2	0.2484	2.07	3.66	3.11	−1
3	0.1143	3.05	4.61	4.07	−1
4	0.0526	4.03	5.56	5.04	−1
5	0.0242	5.02	6.49	6.02	−1
6	0.0111	6.01	7.40	7.01	−1
7	0.0051	7.01	8.26	8.01	−1
8	0.0024	8.00	9.01	9.00	−1
9	0.0011	9.00	9.52	10.00	0
10	0.0009	10.00			

1.85 = average expected loss

last example: Table 4 shows a truncated geometric distribution. For x less than 10, $P[X=x] = 0.54 \times 0.46^{x-1}$; the balance of the distribution is concentrated at $x = 10$. In this case, there is no equilibrium for the voluntary market at the small end of the market; ARP has either all of the market or nearly none of it. Equilibrium can occur with ARP writing risks with expected losses of 10, at a rate of 9.52, and a market share of 0.5% (Table 4A, Appendix B). Even this equilibrium occurs only because the distribution is truncated; if it were not truncated, equilibrium would not occur until ARP's market share was less than 0.01% and its rate nearly 17, more than 9 times the average market rate (Table 4B, Appendix B). By tweaking the parameters a little, one can push this equilibrium market share to any extreme.

The above examples assume that the voluntary market operates freely. If regulatory constraint becomes too severe, none of these examples will bear much resemblance to the real behavior of the market. They are still relevant though—just as the force of gravity is relevant to an engineer—because they show the natural forces at work against the barriers of regulation.

4. HOW TO SET THE RATES

4.1. An Alternative to Break-Even Pricing

One might be tempted to argue that because the above examples are filled with instances of equilibrium, it is reasonable for assigned risk plans to base their prices on their own experience. Unfortunately, the equilibrium is capricious—one never knows where or whether it will occur. Equilibrium, moreover, desirable as it is, is not an end in itself. Society will probably not accept an equilibrium that leaves insurers a tiny fraction of the market, or that charges assigned risk plan members ten times the voluntary market rates. In any case, ARP's pricing strategy should be consistent with public goals. The public may accept letting some residual markets price themselves out of existence and may be well served by so doing. In those cases break-even pricing with a contingency factor may work well, provided ARP really follows the indications. Where the consensus is in favor of keeping and controlling the residual market, however, the break-even approach is not a good one.

So how should ARP set its rates? If one starts with the assumptions that there should (or in any case *will*) be an assigned risk plan, that it should not be overly burdensome on the insureds in the voluntary market and that it should not have wild swings in market share, there is a reasonable solution to the rate problem. The solution is to base ARP's rates on total industry experience, but set at a level consistently higher than that which a typical insurer would need to charge in the voluntary market. One can start with industrywide pure premiums, for example, and load them with an expense and profit factor which is 25% above that of the industry average (or whatever percentage seems reasonable in line with studies of the market and the philosophy of a given state). The market will seek its own equilibrium; in the typical case ARP will lose money, but the burden on voluntary insureds will not be excessive. At the same time ARP's rates will be high, but not intolerably high. Thus a start-up employer who truly has a contribution to make to society, for example, will have a chance.

4.2. Setting Specific Goals

Words such as *reasonable* and *excessive* are rather vague; one must define them in order to use them in actually setting rates. Their definitions may vary from state to state and from line to line, and probably with the passage of time as well. They will come through compromise and consensus—there is no optimal solution that everyone will accept. The key is to have specific goals and to structure the pricing to accomplish those goals.

The voluntary market attempts to identify true costs underlying whatever it is insuring; and, by varying its prices according to those costs, it steers production of goods and services toward those that are most efficient. This feature of insurance is very beneficial to society. A state should choose a goal for residual market share that guarantees the continuation of a large voluntary market so as to give society the benefit of an efficient economy, with the ideal being a totally voluntary market.

On the other hand, rightly or wrongly, the government has constrained the operation of the insurance market for many decades. Workers compensation statutes are a prime example: despite the benefits of the statutes, they raise a high hurdle for many small employers. Residual market plans often enable such employers to enter and compete in the marketplace, something that could occur naturally in the absence of the workers compensation statutes. One could view residual markets as intervention needed because states interfered with the natural flow of the marketplace when they first created laws such as the workers compensation statutes. Residual markets will almost surely continue to have their adherents and, if their prices are unaffordable for virtually everyone, consumers will revolt and probably revolt successfully.

So in determining the parameters of the pricing problem, one has two somewhat conflicting goals: the bigger the voluntary market the better, and residual market rates should not be unaffordable for all. A third guiding factor is consideration for the voluntary market insureds—the expected assessment of residual market losses on these innocent bystanders should not be punitive. A fourth guiding factor is the *status quo*. Too abrupt a change can be harmful—partly because it might unleash unexpected and uncontrollable consequences, and partly because it would be in some sense a change of the rules under which many people have been operating in good faith.

Reasonable goals for a residual market plan might be a market share of under 1%, a rate of under 150% of the voluntary market, an expected assessment on the voluntary market of under 0.5%, and (during the catch-up period if one is needed) annual price adjustments of under 10% relative to the voluntary market. This paper is not trying to suggest the exact parameters to use; it is merely suggesting a way to approach them.

Of course, the voluntary market does not charge a single rate that one can use as a basis for the ARP rates. In the above example where ARP rates are under 150% of the voluntary market, what is "voluntary market?" A reasonable starting place is to use statewide pure premiums loaded with average industry expenses and profits. In place of statewide pure premiums one might also use the pure premiums or rates generated by a large ratemaking bureau operating in the state, provided that the bureau's members represent a significant enough market share.

It will be helpful to look not only at the *average* voluntary market rates, but also at the *spread* of rates. In particular, some companies specialize in non-standard business and provide a valuable service to the marketplace. Before arbitrarily selecting an upper bound of say 150% of average, it will be helpful to know where the rates of the non-standard writers fall relative to the overall average. A state could do its citizens a disservice if it sets a limit that cuts out the non-standard carriers.

Finally, although this paper suggests abandoning break-even pricing, ARP's own experience still has an important role in ARP pricing. In order to measure the expected assessment on the voluntary market, ARP still must analyze its own experience. If ARP's experience indicates excessive future assessments, ARP will need to adjust its rates within the constraints of the other goals. The state may even need to change the goals if all the goals are already at the limit of their constraints. In addition, an analysis of ARP's experience can be helpful to the voluntary market in identifying opportunities to depopulate ARP.

4.3. Using the Goals to Set Prices

With a set of specific residual market goals in hand, a state does not need to fight the unpredictability of break-even pricing. It can take the more stable path of setting residual market prices as a direct multiple of voluntary market prices, and it can measure its success directly from its goals.

Suppose that a state sets ARP rates by looking at ARP's own experience, judgmentally modifying the indication (essentially ignoring it), and finally ending up with rates that currently average 105% of voluntary rates. Now consider the following alternative. Having first set specific goals for ARP, the state gathers all the data it needs to monitor the goals. What are the market shares of the residual and voluntary markets? What are the average rates of voluntary writers (paying separate attention to companies specializing in non-standard business)? What are the average expense ratios? What are the underlying loss costs? Then the state measures its goals against the data. Are all the goals met? If so, the state leaves the prices at 105% of voluntary (as measured by loss costs and average expense ratios) and the job is done.

Probably, though, 105% of voluntary will not achieve the goals. So the state increases the rates to 110% or 115% of voluntary, depending on the "catch-up" parameter. Next year it looks again at the experience and market data. Gradually the state adjusts the ARP-to-voluntary ratio until it meets its goals—not break-even goals with all their unpredictability, but goals based directly on society's specific expectations of ARP.

Once the state finds the multiplier that meets its goals, it sets future rates using the same multiplier. As long as the goals are met, ARP's own experience will have no effect on ARP's rates. For example, if the goals call for a market share of under 1% and a burden on the voluntary market of under 0.5%, ARP could consistently lose 50 cents or more on each dollar of premium provided its market share remains sufficiently small. Its market share will remain sufficiently small as long as the multiplier is sufficiently large. By the same token, a fortuitous ARP profit will have no effect on the rates either; ARP's insureds will be rewarded for good experience not by ARP rate decreases but rather by movement into the voluntary market.

The advantage of this market-based pricing approach is not necessarily to reduce the overall losses of the residual market, but rather to enable more conscious control over the residual market. Rather than having an official ratemaking procedure (break-even pricing) that is not actually followed and that could lead to totally unacceptable results if it were followed, states would articulate their true goals and consciously manage them. Some residual markets might very well shrink as a result and would probably produce fewer losses, but that is not a necessary consequence of moving to market-based pricing. What will happen will depend on the goals of the individual states. In any case, one can not measure the true cost of a residual market by its bottom-line losses alone. Voluntary market insureds bear the risk charge for the residual market even when the residual market is profitable, and all of society pays for the loss of diversity when a residual market gets too big.

5. FINAL THOUGHTS

The original impetus for this paper sprang from real-life observation of the outcomes that this simple model predicts; the predictions are not merely theoretical. Of course, the worst examples of residual market problems arise not from using break-even pricing, but rather from suppressing rates and ignoring the effects. What appears to be an easy solution to that problem—namely basing residual market rates directly on residual market experience—is in general not a solution at all.

This paper demonstrates that under break-even residual market pricing, regardless of the goal that one sets for residual market share, one can find a loss distribution that leads to a market share very different from the goal. The paper does not look at empirical loss distributions to predict how specific residual markets would behave under them. That is an interesting area for additional research, but the paper's thesis is that such research is not essential if there is an approach to residual market pricing whose success is independent of loss distribution. It turns out that there is such an approach; namely, to base residual market prices on total market experience, at a level consistently above that of voluntary market prices. That approach not only solves the market-share problem, but it also enables focusing on and achieving all of the other goals of the residual market to the extent that the goals are achievable.

APPENDIX A

PROOF OF EXISTENCE AND UNIQUENESS OF SOLUTION TO PRICING PROBLEM

The insurers' pricing problem—to solve for a in Equation (A.1) below—has a solution if and only if $L < R$, where L is the average expected loss and R is the average ARP premium. The solution, when it exists, is unique.

Proof Let F be the distribution function of the expected losses. As a distribution function, F is right-continuous. Assume furthermore that $F(0) = 0$. To allow $F(0) > 0$ would be to assume that for some insureds not even the *possibility* of a loss exists; F, remember, is the distribution of *expected* losses, not of actual losses. We have:

$$L = \int_0^\infty x\,dF = \int_0^{R/a} ax\,dF + \int_{R/a}^\infty R\,dF. \qquad \text{(A.1)}$$

Equation (A.1) merely says that the expected losses are equal to the premium of the voluntary market plus the premium of ARP. The insurers' pricing problem is to solve for a. Set

$$g(a) = L - \int_0^{R/a} ax\,dF - \int_{R/a}^\infty R\,dF. \qquad \text{(A.2)}$$

Solving equation (A.1) for a is equivalent to finding a zero of the function g defined by equation (A.2). g is a continuous, monotonically decreasing function on the interval $(0,\infty)$, so it has at most one zero. If it ever changes sign, it has exactly one zero

$$g(1) = \int_0^\infty x\,dF - \int_0^R x\,dF - \int_R^\infty R\,dF = \int_R^\infty (x-R)\,dF > 0.$$

Thus $g(a)$ is positive for $a \le 1$. Now look at $g(a)$ as a increases. For $0 \le x \le R/a$, $ax \le R$, so

$$\int_0^{R/a} ax\,dF \le \int_0^{R/a} R\,dF = R(F(R/a) - F(0)).$$

Since F is right-continuous, $\lim_{a\to\infty} R(F(R/a) - F(0)) = 0$, so also

$$\lim_{a\to\infty}\int_0^{R/a} ax\,dF = 0. \tag{A.3}$$

Because $F(0) = 0$ and again because F is right-continuous,

$$\lim_{a\to\infty}\int_{R/a}^{\infty} R\,dF = R. \tag{A.4}$$

Finally, combining equations (A.2), (A.3), and (A.4) we have

$$\lim_{a\to\infty} g(a) = L - R,$$

which is negative if and only if $L < R$. Thus if $L \geq R$, there is no a for which $g(a) = 0$, and equation (A.1) has no solution. If $L < R$, there is a unique solution.

If we removed the requirement that there exist insureds with arbitrarily large expected losses, our conclusion would not change. For values of R greater than the largest expected loss, the solution would be $a = 1$ and all the business would be in the voluntary market. If we removed the requirement that there exist insureds with arbitrarily small expected losses, there might be some degenerate solutions. In that case, g would no longer be monotonically decreasing on the entire interval $(0,\infty)$, but only on $(0,R/b)$, where b is the smallest possible expected loss—more precisely, $b = \inf\{x : F(x) > 0\}$. For all $a > R/b$, we'd have $g(a) = L - R$, so that for $R = L$ there would be infinitely many solutions of the equation $g(a) = 0$. These solutions are rather trivial; they are simply all multipliers, a, large enough to charge the tiniest risk more than R, so that ARP writes all of the business.

APPENDIX B

ARP MARKET SHARE CALCULATIONS

This appendix contains Tables 3A, 4A and 4B; these tables extend Tables 3 and 4 to show calculations of ARP market shares. In addition, Table 4B extends the truncation point of the geometric distribution from 10 to 20 to show a more extreme example of diminishing ARP market share. The data in the first six columns of Tables 3A and 4A come directly from the corresponding Tables 3 and 4 of the paper. The reader will find explanations of the additional columns (columns 7 through 11) in the tables themselves.

TABLE 3A

MARKET SHARE CALCULATIONS FOR SHRINKING ARP

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		If ARP writes all risks with expected losses greater than x								
x	$P[X=x]$	voluntary market rate for x	ARP rate (for all)	voluntary market rate for $x+1$	1: ARP gains −1: ARP losses 0: equilibrium	cost for expected loss = x (1) × (2)	cost for expected loss > x	ARP premium (8)/1.05	voluntary market premium $(7)_{tot} - (9)$	ARP market share $(9)/(7)_{tot}$
1	0.0646	3.71	3.74	7.42	0	0.0646	3.6747	3.4997	0.2396	93.6%
2	0.1769	2.76	4.17	4.13	−1	0.3538	3.3208	3.1627	0.5766	84.6%
3	0.2424	3.32	4.79	4.43	−1	0.7272	2.5937	2.4702	1.2691	66.1%
4	0.2214	4.16	5.52	5.20	−1	0.8855	1.7082	1.6268	2.1124	43.5%
5	0.1516	5.08	6.32	6.10	−1	0.7582	0.9500	0.9047	2.8345	24.2%
6	0.0831	6.04	7.16	7.05	−1	0.4986	0.4514	0.4299	3.3094	11.5%
7	0.0379	7.02	8.02	8.02	−1	0.2656	0.1857	0.1769	3.5624	4.7%
8	0.0149	8.01	8.85	9.01	0	0.1188	0.0669	0.0637	3.6756	1.7%
9	0.0051	9.00	9.52	10.00	0	0.0458	0.0211	0.0201	3.7192	0.5%
10	0.0021	10.00				0.0211			3.7393	0.0%
Total	1.0000					3.7393				

Column (8): row x = sum of Column (7) from row $x+1$ through row 10

TABLE 4A

Market Share Calculations for Vanishing ARP

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		If ARP writes all risks with expected losses greater than x								
x	$P[X = x]$	voluntary market rate for x	ARP rate (for all)	voluntary market rate for $x+1$	1: ARP gains −1: ARP losses 0: equilibrium	cost for expected loss $= x$ $(1) \times (2)$	cost for expected loss $> x$	ARP premium (8)/1.05	voluntary market premium $(7)_{tot} - (9)$	ARP market share $(9)/(7)_{tot}$
1	0.5400	1.12	2.71	2.23	−1	0.5400	1.3111	1.2486	0.6024	67.4%
2	0.2484	2.07	3.66	3.11	−1	0.4968	0.8143	0.7755	1.0756	41.9%
3	0.1143	3.05	4.61	4.07	−1	0.3428	0.4715	0.4490	1.4020	24.3%
4	0.0526	4.03	5.56	5.04	−1	0.2102	0.2612	0.2488	1.6023	13.4%
5	0.0242	5.02	6.49	6.02	−1	0.1209	0.1403	0.1337	1.7174	7.2%
6	0.0111	6.01	7.40	7.01	−1	0.0667	0.0736	0.0701	1.7810	3.8%
7	0.0051	7.01	8.26	8.01	−1	0.0358	0.0378	0.0360	1.8151	1.9%
8	0.0024	8.00	9.01	9.00	−1	0.0188	0.0190	0.0181	1.8330	1.0%
9	0.0011	9.00	9.52	10.00	0	0.0097	0.0092	0.0088	1.8423	0.5%
10	0.0009	10.00				0.0092			1.8511	0.0%
Total	1.0000					1.8511				

Column (8): row x = sum of Column (7) from row $x+1$ through row 10

TABLE 4B

EXTENDED MARKET SHARE CALCULATIONS FOR VANISHING ARP

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		If ARP writes all risks with expected losses greater than x								
x	$P[X = x]$	voluntary market rate for x	ARP rate (for all)	voluntary market rate for $x+1$	1: ARP gains −1: ARP losses 0: equilibrium	cost for expected loss = x (1) × (2)	cost for expected loss > x	ARP premium (8)/1.05	voluntary market premium $(7)_{tot} - (9)$	ARP market share $(9)/(7)_{tot}$
1	0.5400	1.12	2.72	2.23	−1	0.5400	1.3119	1.2486	0.6025	67.47%
2	0.2484	2.07	3.67	3.11	−1	0.4968	0.8151	0.7755	1.0756	41.92%
3	0.1143	3.05	4.62	4.07	−1	0.3428	0.4723	0.4490	1.4021	24.29%
4	0.0526	4.03	5.57	5.04	−1	0.2102	0.2620	0.2488	1.6023	13.47%
5	0.0242	5.02	6.53	6.02	−1	0.1209	0.1411	0.1337	1.7174	7.26%
6	0.0111	6.01	7.48	7.01	−1	0.0667	0.0744	0.0701	1.7810	3.83%
7	0.0051	7.01	8.43	8.01	−1	0.0358	0.0386	0.0360	1.8151	1.98%
8	0.0024	8.00	9.38	9.00	−1	0.0188	0.0198	0.0181	1.8330	1.02%
9	0.0011	9.00	10.33	10.00	−1	0.0097	0.0100	0.0088	1.8423	0.51%
10	0.0005	10.00	11.29	11.00	−1	0.0050	0.0050	0.0048	1.8471	0.26%
11	0.0002	11.00	12.24	12.00	−1	0.0025	0.0025	0.0024	1.8495	0.13%
12	0.0001	12.00	13.19	13.00	−1	0.0013	0.0012	0.0012	1.8507	0.06%
13	0.0000	13.00	14.14	14.00	−1	0.0006	0.0006	0.0006	1.8513	0.03%
14	0.0000	14.00	15.08	15.00	−1	0.0003	0.0003	0.0003	1.8516	0.02%
15	0.0000	15.00	16.01	16.00	−1	0.0002	0.0001	0.0001	1.8517	0.01%
16	0.0000	16.00	16.92	17.00	0	0.0001	0.0001	0.0001	1.8518	0.00%
17	0.0000	17.00	17.78	18.00	0	0.0000	0.0000	0.0000	1.8518	0.00%
18	0.0000	18.00	18.53	19.00	0	0.0000	0.0000	0.0000	1.8518	0.00%
19	0.0000	19.00	19.05	20.00	0	0.0000	0.0000	0.0000	1.8518	0.00%
20	0.0000	20.00				0.0000		0.0000	1.8519	0.00%
Total	1.0000					1.8519				

Column (8): row x = sum of Column (7) from row $x + 1$ through row 20

DIRTY WORDS: INTERPRETING AND USING EPA DATA IN AN ACTUARIAL ANALYSIS OF AN INSURER'S SUPERFUND-RELATED CLAIM COSTS

STEVEN J. FINKELSTEIN

Abstract

A significant amount of liability exposure for many insurers stems from pollution-related claims. Many of these pollution-related claims, in turn, stem from the implementation of the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) of 1980, also known as Superfund. This paper discusses adjustments necessary to properly use the EPA's records of decision (RoDs) and Comprehensive Environmental Response, Compensation and Liability Information System (CERCLIS) data in actuarial analyses of Superfund costs. Background on the Superfund process and an approach to using the data in an exposure-type analysis suitable to insurers with significant potential exposure to environmental losses are also presented. The paper also discusses the difficulties typically facing an actuary in non-Superfund site cleanup cost evaluations, and concludes with some comments on environmental liability discounting considerations.

ACKNOWLEDGEMENT

The author thankfully acknowledges the assistance of several people, without whom this paper might have been far less interesting: Mike Goldstein (Environmental Protection Agency), Kate Siggerud (General Accounting Office), and Sandy Susten (Agency for Toxic Substances and Disease Registry). The author would also like to thank Orin Linden and Christopher Diamantoukos for their thoughts on this write-up. Finally, the author greatly appreciates the input received from the Casualty Actuarial Society's Committee on Review of Papers (CORP) for its significant commentary and assistance in preparing this paper for presentation in the Proceedings.

This paper should be read with the understanding that the opinions expressed herein represent the views of the author, and do not necessarily represent the views of the Casualty Actuarial Society, Ernst & Young LLP, or anyone else.

1. FROM THE GROUND UP: AN INTRODUCTION

A significant amount of liability exposure for many insurers stems from pollution-related claims. Many of these pollution-related claims, in turn, stem from the implementation of the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) of 1980, also known as Superfund.

Currently, there are two primary sources of Superfund cost-related information available for use in an environmental analysis: Records of Decision (RoDs) published by the Environmental Protection Agency (EPA), and the Comprehensive Environmental Response, Compensation and Liability Information System (CERCLIS). While data from these sources is readily available from the EPA,[1] information on the appropriate use of that data is not as easily found. Given the importance of reasonably estimating these liabilities in connection with acquisitions, commutations and financial reporting, a thorough understanding of the data underlying many of these analyses is vital. This paper is an attempt to fill the gap in CAS literature relating to environmental cost data and its use in environmental analyses.

2. DIGGING IN: AN OVERVIEW OF THE SUPERFUND PROCESS

The Superfund process begins with the discovery of a location which represents *either a current or potential future* health

[1]Most readily through *WWW.EPA.GOV/Superfund/*, which is the EPA's Superfund web site. In addition to the EPA, the Agency of Toxic Substances and Disease Registry (ATSDR) also maintains a database accessible through the Internet at *http://atsdr1.atsdr.cdc.gov:8080/hazdat.html* with information on public health hazard levels (discussed in Appendix C).

hazard. The potential for future hazard is generally based on (1) the potential for current contamination levels to spread at a particular site, (2) plausible future uses of that site, and (3) plausible estimates of the future size of the population at and adjacent to that site. If this discovery is reported to the EPA, information on that "site" is put into the CERCLIS database.

An off-site preliminary assessment is then performed to characterize the site as a potentially imminent, serious, or non-serious threat. Imminent threats are addressed through emergency removal actions, designed to reduce the threat to a serious or non-serious level. Serious (but not imminent) threats are addressed through site inspections, which include on-site evaluations to better characterize whether or not the site requires further EPA attention (including an emergency removal action not already initiated, due to insufficient information at the preliminary assessment phase). A site determined to pose no serious threat receives no further attention by the EPA.

The EPA then uses a hazard ranking system (HRS) to prioritize those sites that still pose a potentially serious threat. The HRS is a quantitative assessment, on a scale of 1 to 100, of the level of hazard to human health via several "exposure pathways." These pathways represent different ways that a hazard can expose human beings to a health risk—for example, through ground and surface water, the soil and the air. If the HRS is high enough (currently, 28.5[2] or greater), the EPA "proposes" that the site be included in the National Priorities List (NPL), representing those sites which, in the EPA's estimation, represent the greatest potential hazard to human health, past, present

[2]The 28.5 threshold score was derived "because it would yield an initial NPL of at least 400 sites as suggested by CERCLA, not because of any determination that it represented a threshold in the significance of risks presented by sites." [1] This apparent need to initially list at least 400 sites on the NPL may somewhat mitigate the argument that the hazard level of the average site listed early in the program exceeds the hazard level of the average site listed more recently. This is discussed further later in this paper, as well as in Appendix B.

or future.[3] Community discussions are then held, and after some additional work, these sites may be listed on the "final" NPL.[4] It is worth noting some of the events that have impacted past, and may impact future, site listings:

- As noted earlier, CERCLA appeared to suggest that at least four hundred sites should be listed on the initial NPL in 1983.
- Federal facilities started showing up with some regularity in 1987, after the Superfund Amendments and Reauthorization Act of 1986 (SARA) gave the EPA a level of control over remedy selection at Federal facilities.
- Between the mid-1980s and early to mid-1990s, the capabilities of the states' individual Superfund programs grew, perhaps leading to a shift in emphasis from Federal to State enforcement.
- In December of 1990, the HRS was revised, leading to fewer annual NPL site listings per year.
- "Governor's Concurrence" legislation enacted in July of 1995 required the EPA to seek approval from a state before listing a site located there on the NPL. Since then, more than 30 sites were not listed, at the request of the relevant states' governors.

It is also worth noting two additional means by which a site may be listed on the NPL. First, each state is entitled to select a single site and include it on the NPL, regardless of that site's HRS score, if the state feels that the site represents a significant

[3]The preliminary nature of the data used to derive the HRS is believed to be useful for determining whether or not a site represents a potentially significant hazard, but it is not necessarily useful for ranking the relative hazard levels of those sites which exceed the HRS threshold. In addition, if the HRS reaches this threshold before all pathways are scored, the remaining pathways might not be scored. For these reasons, the author recommends not using the HRS to estimate the relative hazard levels of Superfund sites.

[4]There are actually two NPLs—one for Federal sites (i.e., federally owned), and one for non-Federal sites. Only the non-Federal sites are usually considered relevant to estimating an insurer's potential environmental liabilities. Information on whether a particular site is a Federal facility is available in CERCLIS.

danger to public health. Second, a site may be listed if all of the following conditions (taken from [2]) are met:

- The Agency for Toxic Substances and Disease Registry (ATSDR) of the U.S. Public Health Service has issued a health advisory that recommends dissociation of individuals from the site.
- EPA determines that the site poses a significant threat to public health.
- EPA anticipates that it will be more cost-effective to use its remedial authority (available only at NPL sites) than to use its removal authority to respond to the site.[5]

Sites that were reviewed and subsequently not listed on the NPL remained in the CERCLIS database for many years, leaving them with a stigma stemming from the belief that there was a strong possibility they might still become NPL sites at some later date. To alleviate this concern, the EPA created a new database in March of 1995 which would store these "archived" sites. The database was called NFRAP, which stands for "No Further Remedial Action Planned," and, by September 30, 1996, it contained 25,000–30,000 sites no longer being considered for NPL status. However, these sites remain within the purview of the state and local governments, who may require further action.

How to Remedy a Bad Situation: An Introduction to Records of Decision

For sites listed on the NPL, the next step is to determine what actions would constitute an appropriate remedy. The EPA publishes the details relating to these "remedial actions" (RAs), addressing the potential contamination at a particular location in a "record of decision" (RoD). These RoDs typically include a

---

[5]A removal action is a mechanism whereby the EPA can take immediate action to "remove" hazardous substances posing an immediate threat to public health and the environment, rather than allowing the threat to linger until that site is listed on the NPL, making it eligible for a more extensive (but likely less timely) cleanup effort.

description of the problem that is being addressed, the remedy selected to address the problem, and the expected cost associated with the selected remedy.

There are two types of costs usually addressed in the RoDs—those related to the construction of the selected remedy (capital costs) and those related to the implementation, operation and on-going maintenance of the selected remedy over time (operation and maintenance, or O&M costs). Once issued, RoD cost estimates are not typically updated to reflect new information, except in the event of a fundamental change in the approach required or technology to be used.

There are three types of RoDs issued: interim RoDs, which address either a partial remedy or a "quick fix" to prevent the further spread of contamination that will be addressed in a later RoD; final RoDs, which represent either the complete remedy at a particular location or the completion of a remedy begun earlier in an interim RoD; and amendment RoDs, which supplant previous RoDs due to a change in scope, cost or both. These amendment RoDs can be either interim amendment RoDs or final amendment RoDs, though interim amendment RoDs are rare.

A single RoD need not address the remedy required for an entire site. Sometimes, multiple RoDs are issued. This is done because an NPL site may have several problems needing to be addressed, such as groundwater and soil contamination. These problems may be addressed as two separate "operable units" (OUs) of that site, in different RoDs. It is worth noting that these RoDs are not necessarily issued at the same time—the EPA (or any other party responsible for site cleanup) may address the groundwater issue at a site (which might soon contaminate an adjacent town's drinking water if unchecked), but forego cleanup efforts relating to the soil contamination. This might happen if the contaminated soil is felt to be a less immediate risk to human health than exists currently at another site. In this case, the EPA might divert its resources toward that other site, and return to the first site later. It is also worth noting that multiple OUs at a site typically relate to different contaminated media at that site (e.g., groundwater and soil), which may or may not be present at different locations of the site. In other words, two OUs at a site should not automatically imply two geographic areas requiring attention at that site. Similarly—and adding to the confusion—a RoD may also address a single OU comprised of multiple contaminated media (e.g., groundwater and soil together). Also, remember that a single OU may be addressed through multiple RoDs (i.e., an interim, a final and/or an amendment RoD).

Digging Deeper: Remedial Design Costs

As technical as they might appear to be, RoDs only address the *general* approach to be used in implementing the selected remedy. After the RoD is issued, the "remedial design" (RD) phase provides the *specific* approach to be used in implementing the general remedy outlined in the RoD. The RD cost estimate and the costs included in the RoD are intended to represent the same items (i.e., capital and O&M costs); since the approach is more detailed in the RD phase, however, the RD cost estimates are expected to be more refined. EPA guidance indicates that the actual costs incurred for cleanup activities should be between 70% and 150% of the RoD cost estimate, but only between 95% and 115% of the RD cost estimate.[6]

It is possible that the cost or approach of the RA selected in the RD phase may be significantly different from the cost or approach of the RA as outlined in the RoD, perhaps as a result of unforeseen conditions encountered at a given site. If these significant differences do not result in a fundamental change to the general remedy selected in the RoD, the EPA would typically issue an "Explanation of Significant Differences" memorandum (ESD), outlining the nature and cause of the differences. This

[6]The RD documentation relating to each RA is generally made available for public viewing near the area to be remediated. To the best of this author's knowledge, the RD documents are not consolidated in a single, publicly-available database.

differs from an amendment RoD, which results from significant differences in the approach of the RA that *do* result in fundamental changes to the general remedy selected in the RoD.

Once remedial construction activities have been completed, a site or OU can be labeled “construction complete.” This does not mean that the selected remedy has been put into operation yet; only that the necessary construction required to do so has been completed. Additionally, significant O&M activities may be required after the remedy is enacted.

After all necessary construction is completed, the selected remedy is instituted and O&M activities (if any) are concluded, that site, OU or particular formerly contaminated media may be “deleted” from the NPL, indicating that no further action is deemed necessary. *Not all deleted sites represent completed cleanups, however*. Resource Conservation and Recovery Act (RCRA) sites may be deleted from the NPL before cleanup activities have been completed “if the site is being, or will be, adequately addressed under the RCRA corrective action program under an existing permit or order.” [3] A short introduction to RCRA, for those not familiar with it, is included in Appendix D.

It’s a Dirty Job, but Someone’s Gotta Do It: Cleanup Cost Liability Allocation

At any point along the way in the Superfund process, the EPA may uncover leads on people and companies they believe to be potentially responsible for a given site’s polluted status. A list of these potentially responsible parties (PRPs), which has previously been available through the EPA’s SETS database (Site Enforcement Tracking System), is now included in CERCLIS. Allocation of liability among PRPs at any given site is considered by many as the single most difficult aspect of estimating Superfund liability. The count of PRPs at a given site changes over time. In addition, a PRP’s share of liability might not correlate well with the number of PRPs potentially sharing the cleanup cost at that site (in part because the group of PRPs connected by the EPA to many sites can be characterized as a small number of large polluters and a large number of smaller ones, skewing the proportions).

To help distinguish the possibly responsible from the probably responsible, the actuary should consider looking at other types of communications between the EPA and parties that may be liable at NPL sites. The following is a list that, in the author's opinion, might be used to form a "Superfund Liability Pyramid," in the sense that the items in the list are ordered from least to most likely responsible for activities at a Superfund site:

- *general notice letter recipients*—the EPA sends this letter to parties to inform them of their potential responsibility for site cleanup-related activities.
- *special notice letter recipients*—the EPA sends this letter to parties to inform them of their right to offer to conduct the cleanup efforts at a site.
- *unilateral administrative order (UAO) recipients*—the EPA uses UAOs to "unilaterally order" parties to undertake activities at a site.
- *parties to an administrative order on consent (AoCs) or consent decree*—these documents formalize agreements reached between the EPA and other parties relating to Superfund-related actions those parties have agreed to undertake.

Once in communication with the EPA, an entity involved in a cleanup effort may seek out additional parties to share the responsibility for cleanup-related costs, in addition to those other parties already in communication with the EPA. These additional parties—sued for cooperation not by the EPA, but by those already responsible for cleanup-related costs—are called "collateral suit defendants." Since the EPA is unconnected to the search for these additional PRPs, they would not be included in the EPA's data when and if they are found. To this author's knowledge, there are no good publicly-available data sources for information on collateral suit defendants.

Superfund Action Figures: EPA Expenditure Data

While RoDs contain estimated prospective remedial action costs, EPA's actual costs incurred to date relating to remedial and pre-remedial activities can be found in the CERCLIS and NFRAP databases. The information contained in them is identical, except that NFRAP contains information on sites where no further EPA activity is planned, and CERCLIS contains information on all other sites reported to the EPA. Throughout this paper, reference to CERCLIS should be understood to include NFRAP.

Users of CERCLIS information must be cautious since *only* those costs incurred to date directly by the EPA (referred to as "fund-financed" costs) are included in CERCLIS.[7] As a result, the cost information in CERCLIS is only potentially complete and up to date for activities with a fund-financed cleanup effort.[8] In other situations (i.e., a PRP-financed activity), CERCLIS only includes costs relating to the EPA's oversight of that activity—the cost of *performing* that activity must still be quantified, perhaps based on the average cost of similar, fund-financed activities.

In evaluating how the costs of PRP-financed activities may relate to corresponding, historical fund-financed activities, the reader should note that the General Accounting Office (GAO) had the following to say about the EPA's cost controls [5]:

> "…our recent review found that in spite of the [EPA's] actions, several problems persist: (1) EPA's regions are

---

[7]Note that no O&M costs are to be incurred by the EPA under Superfund. These costs are intended to be the responsibility of either the states or PRPs. However, since the definition of O&M activities differs between CERCLIS and the RoDs (as will be discussed later), some O&M costs arguably are fund-financed.

[8]Even these fund-financed efforts require that some of the capital costs be borne by the states, implying that CERCLIS might not have complete cost information on even these sites. For example, "The President shall not provide any remedial actions pursuant to this section unless…the state will pay or assure payment of (i) 10 per centum of the costs of the remedial action, including all future maintenance…" [3]

> still too dependent upon the contractors' own cost proposals to establish the price of cost-reimbursable work, (2) EPA continues to pay its contractors a high percentage of total contract costs to cover administrative expenses rather than ensuring the maximum amount of available moneys is going toward the actual cleanup work, and (3) little progress has been made in improving the timeliness of audits to verify the accuracy of billions of dollars in Superfund contract charges."

Working with the cost information in CERCLIS is not straightforward. Even for fund-financed activities, the costs cannot always simply be added up to derive a given activity's total incurred cost. For example, some activities are funded by the Superfund program but overseen by a state instead of the EPA. For some of these "state-led" activities, the state is responsible for its own share of the cost from the outset, which would not be included in CERCLIS. A detailed schematic of the cost data included in CERCLIS is shown in Exhibit 1. Exhibit 2 compares and contrasts the data contained in CERCLIS and RoDs.

3. GETTING DOWN AND DIRTY: WHAT ARE SUPERFUND'S COSTS?

Litigation and other transaction costs aside, what are the costs incurred under the Superfund program? Exhibit 3 displays a list of the activities that have typically been included in the EPA's review of an NPL site, with estimates of the average duration and cost for each type of action.

Intent on improving the process, the EPA introduced the Superfund Accelerated Cleanup Model (SACM), designed to streamline the process by (1) combining the preliminary assessment and site inspection steps into a single step (Site Screening and Assessment), eliminating much duplication of assessment-related effort, (2) instituting consistent remedy selections for similar sites rather than assuming site-specific remedies were always required, yielding more efficient and cost-effective cleanups, and (3) creating regional decision teams to more effectively prioritize the cleanup efforts of Superfund sites in each region. The EPA's consistent remedy selection strategy, as well as another recent initiative—increased remedy selection updating through RoD amendments—will be revisited later.

4. MUDDYING THE WATERS: "BROAD" VS. "NARROW" REMEDIAL ACTIONS

Before beginning a discussion on cleanup costs, a note about terminology is in order. The term "remedial action" as used so far has referred to the costs associated with all aspects of the cleanup process (capital costs plus all O&M costs), as is typically done when discussing cleanup (remedial) vs. other-than-cleanup (non- or pre-remedial) actions. Within the context of discussing cleanup costs only, however, the phrase "remedial action" has two different meanings. When used in a RoD or other engineering costing study, it typically relates to those costs incurred only to *construct* the remedy (i.e., the *capital costs*)—the actual *implementation* of the remedy and any other O&M-related activities would be considered when estimating O&M costs. Alternatively, to determine which costs are eligible for Superfund funding, the EPA considers RA costs as those which must be incurred to safeguard the environment from the contamination at an environmentally-impaired site—clearly, a broader definition, incorporating both the *construction* and (at least partial) *implementation* of the remedy. Therefore, the capital costs displayed in the RoDs (usually representing construction costs only) typically should not be compared to the RA costs in the EPA's CERCLIS database without first adjusting for the percentage of total RA costs included in CERCLIS (see Exhibit 1) and the addition of a portion of the RoD's O&M costs. The appropriate portion of the RoD's O&M costs to include in this comparison is up to ten years when groundwater or surface water restoration is included, and up to one year in other cases.

5. SPARE THE ROD: WHAT IS (AND IS NOT) INCLUDED IN A RECORD OF DECISION

RoD costs typically represent the sum of undiscounted capital costs (relating to remedial actions) and discounted O&M costs, yielding a total which is neither fully discounted nor undiscounted. Unwinding the discount in the O&M estimate requires three items: an estimate of O&M expenditures by year, the discount rate used and the expected duration of O&M activities in years. There are three issues relating to these items:

- Annual O&M costs do not represent estimates of O&M expenditures by year since they do not include a provision for inflation. As an example of the magnitude of this issue, an annual inflation rate of only 3% over an eighteen year period (an estimate of the average duration of O&M activities where no groundwater issues are present [6]) increases the total O&M cost estimate by approximately one-third. Over a thirty-year period (the maximum duration included in RoD O&M cost estimates), the estimated total O&M cost would increase by approximately 60%.[9]

- The discount rate used to calculate the present value of total O&M costs is not always included in the RoDs. Exhibit 4 provides a list of the discount rates likely applicable to this calculation, according to RoD-related guidance and other documentation in effect during each period. Note that the inflationary impact excluded from the annual O&M costs above is included here as a reduction to the nominal discount rate selected—hence the term "pre-tax, after inflation" discount rate, as shown in Exhibit 4. The reader should be aware, however, that this discount rate is reduced by the overall inflation level of the economy. It may be possible that these O&M-

[9]The increase of 32% can be calculated as the summation of $j = 1$ to 18 over the expression $(1.03)^{(j-0.5)}/18$. The increase of 61% can be calculated similarly, using a summation of $j = 1$ to 30, and dividing by 30.

related costs, which are largely construction and labor-related, are subject to a different degree of inflation than the average inflationary level of the economy as a whole.

An example should help to clarify the issues above and simultaneously explain how the O&M cost information in RoDs has frequently been misinterpreted. Assume, for example, an inflation rate of 3%, a nominal discount rate of 10%, and an expected first O&M payment (as indicated in the RoD) of $1,000, with O&M activities expected to continue for 30 years. The present value of the first O&M payment—assuming it is expected to occur during the second year of cleanup activities—might be calculated either as $\$1{,}000 * (1.03)/(1.10)$, or as $1,000/1.07 (where 1.07 is the rounded result of $1.10/1.03 = 1.067$). Similarly, the present value of the second payment would be either $\$1{,}000 * (1.03)^2/(1.10)^2$, or simply $\$1{,}000/(1.07)^2$. It should be clear from these examples that it is easier and faster to simply work with the 7% “after inflation” discount rate and the constant $1,000 starting value than to use both the inflation rate of 3% and the nominal, pre-inflation discount rate of 10%. Unfortunately, the fact that the first year’s payment is frequently referred to as the “annual” O&M cost, has led to the traditional approach of estimating undiscounted O&M costs as this allegedly “annual” O&M cost, multiplied by the number of years of O&M activities—in this case, yielding $30,000 ($= 30$ years $* \$1{,}000$ per year). However, applying the 3% inflation and 30 year duration assumptions to the $1,000 first year O&M cost yields an undiscounted cost estimate of $49,003—more than 60% greater than the $30,000 estimate. In addition, if it is believed that O&M cost inflation is 5% per year, rather than the 3% general inflation rate, the undiscounted O&M cost estimate becomes $69,761—more than double the $30,000 estimate typically derived. This is especially important in evaluating the extent to which RoD cost estimates have historically over- or understated actual costs incurred. If the actual O&M costs incurred for this RoD's O&M activities were between $50,000 and $70,000, the traditional approach would indicate that the actual O&M costs are in the neighborhood of 67%–133% greater than the expected costs. In reality, however, we can see that correct estimation of the undiscounted O&M cost would imply that our estimate was right on target, assuming a 3%–5% inflation rate over the thirty-year period applied.

- 1EPA guidance documents [7] note that for the purpose of estimating the total O&M discounted cost, the maximum duration of O&M activities permitted is thirty years. This is because the EPA is only concerned with providing a discounted estimate of O&M costs, and the EPA believes that there is little gained on that basis by continuing beyond thirty years.[10] As a practical matter, many of the cleanup efforts requiring thirty-year O&M costs are actually expected to continue forever.[11]

In addition to the above, two additional considerations regarding RoD cost adjustments are noteworthy:

- Although the focus of the above was primarily on O&M costs, for construction efforts expected to require more than a year to complete, there may be some level of capital cost inflation as well.

[10]Readers of [8] may recall the comment that "there was a clear pattern of 30 years as the standard duration (of O&M costs)," (p. A-10) consistent with the EPA's maximum allowable O&M duration for RoD costing purposes.

[11]From [9], the following is offered with regard to O&M activity durations: "The federal government, states, and responsible parties must perform some long-term operations and maintenance at almost two-thirds, or 173, of the 275 sites we reviewed that were formerly or are currently on the National Priorities List and where the cleanup remedy has been constructed. These activities—which include controlling the erosion of landfill covers, treating contaminated groundwater, or implementing and enforcing restrictions on the use of land or water on or adjacent to the sites—will continue for decades, and, in some cases, indefinitely." Also, from the EPA's own documentation [7], "Remedial action alternatives requiring perpetual care should not be costed beyond thirty years, for the purpose of feasibility analysis. The present worth of costs beyond this period become negligible and have little impact on the total present worth alternative."

- When included in the RoDs, both capital and annual O&M costs are typically stated in "current dollars," where "current" refers to the year in which the RoD was written—not necessarily the year either construction or O&M activities are expected to begin.

Appendix A includes a sample RoD Summary taken from the EPA's web site, and an approach which can be used to calculate the undiscounted cleanup cost estimate implied by information included in that sample RoD, adjusting for the above issues. (Note the assumption that the duration of O&M activities will not extend beyond thirty years, which may not be reasonable.) Row 16 of Appendix A, Exhibit 1 displays the undiscounted total cost estimate for this RoD ($81,178,343). This amount is between two and three times greater than the estimate of present worth total costs actually displayed in the RoD ($30,720,300, from Row 1). The magnitude of this difference emphasizes the importance of properly interpreting the RoD data prior to its use in actuarial analyses.

6. SUM IN-SITE: ESTIMATING INDIVIDUAL SITE COSTS BY ADDING ROD COST ESTIMATES

There are several issues which hamper the use of RoD data for estimating individual undiscounted Superfund site cost estimates, including the following:

- There are many sites for which no RoDs have been issued.
- The most recently issued RoDs may not yet be readily available.
- A site may have two or more OUs, but currently only one RoD addressing only one of them.
- A RoD need not address the final remediation for an OU (or combination of OUs). As noted above, interim RoDs may be stop-gap measures designed merely to contain the spread of contamination, rather than reduce or eliminate it. A subsequent

RoD would address the completion of the clean-up effort at that OU.

- RoDs represent up-front estimates of long-term costs. As a result, it may be necessary to include an average Superfund RoD cost redundancy/deficiency factor in the actuary's analysis.[12]
- Some RoDs relate to remedies which may continue indefinitely, yielding an infinite ultimate cost on an undiscounted basis. However, information provided in the RoD usually shows activities limited to a specified duration (typically, up to thirty years for O&M). In the remainder of this paper, the phrase "*adjusted* RoD cost" will be used to represent the undiscounted RoD cost derived using the information provided in the RoD. We avoid using the phrase "*undiscounted* RoD cost," since it may be infinite, as noted above.

The model described in the following sections is an attempt to address at least some of the above issues by modeling RoD costs directly, rather than site costs. It is not proposed as "the" environmental model, but one of several different frameworks which are available to the actuary for modeling Superfund liabilities. Additionally, the reader should note that, as much as possible, the author has assumed that little if any data from the insurer is available to assist in performing this analysis. Clearly, the actuary should consider all data that may be available from an insurer in performing this type of study. However, to the extent that different insurers may have different levels of Superfund data available for this type of study, the author felt that this assumption would hopefully provide a model useful to the widest possible audience.

[12]From a practical perspective, this may be impossible. First, capital costs in the RoDs and CERCLIS may have differing definitions, as noted earlier. Second, the EPA cannot collect O&M expenditure information from the PRPs, so actual O&M costs incurred are not available publicly. Therefore, no true "actual to expected" total RoD cost comparisons may be made for RoDs calling for O&M activities, short of independently gathering large quantities of proprietary data from numerous sources.

7. SACM (A SUPERFUND *ACTUARIAL* CLEANUP MODEL): INCORPORATING RODS IN AN ANALYSIS OF THE TOTAL, SUPERFUND-RELATED COSTS OF AN INSURER

First, we define a claim in this model as an insured's cost relating to a single site[13], subject to the applicable coverage terms, policy periods, and insurer defenses against incurring environmental liability. The model described here estimates an insurer's total Superfund liability as the sum of the liability stemming from claims at current NPL sites and the liability stemming from claims at future NPL sites. Each of these aspects is addressed separately below, followed by an introduction to the concept of policy buybacks and known site settlements.

The general approach used in this model to estimate the liability at current NPL sites is as follows:

1. Estimate the cleanup cost on each current NPL site. For each site, this includes three components: actual, historical costs from (or perhaps based on data in) CERCLIS; previously-estimated future costs, from current RoDs; and not-yet-estimated future costs, if any, from future RoDs. The first two items have already been discussed; we address the third item in the next section of this paper.
2. Estimate each insured's share of liability at each relevant NPL site. An introduction to this topic was discussed earlier in Section 2 (*It's a Dirty Job, but Someone's Gotta Do It: Cleanup Cost Liability Allocation*).
3. Multiply items (1) and (2) together for each insured with a current NPL-based claim to estimate that insured's share of the relevant NPL site cost.
4. Apply any relevant cost add-on factors, such as for allocated loss adjustment expenses (ALAE), to the insured's share of the relevant NPL site cost.

---

[13] Adjustments to this assumption may be made by the actuary as appropriate. For example, some insureds may attempt to aggregate all Superfund sites into a single claim to mitigate the impact of multiple, large retentions.

5. Apply the relevant coverage factors (e.g., attachment point, limit, share of layer), coverage triggers, cost allocation scheme (e.g., pro-rated over several years using total limits by year), and other claim-specific factor adjustments (such as the probability of successfully denying coverage for the claim) to derive the estimated cost to the insurer of that particular claim.[14]
6. Sum the estimated costs to the insurer of the current claims on current NPL sites (based on the application of steps 1–5 above).
7. Adjust this total to include a provision for future claims on current NPL sites.

The primary focus of this paper is on those items which relate to the use of EPA data in an exposure analysis. Therefore, items (4) and (5) above—though unquestionably important concepts—will not be addressed in this paper.

The Hole is Greater than the Sum of its Parts: Estimating Record of Decision and Relevant Operable Unit Counts by Site

So how can RoDs be used to estimate the total cost of a Superfund site? This model divides that task into three components:

1. estimating the number of RoDs per OU at the site,
2. estimating the number of OUs per site, and
3. estimating the cost indicated in each current and future RoD at that site.

An analysis of the estimated number of RoDs per OU at a site is included in Exhibit 5. Many OUs do not and will not have RoDs associated with them, and therefore will not be considered in this remedial action cost analysis. These OUs represent among

[14]Note that, depending on the terms of the insurance agreement, Steps 4 and 5 may need to be reversed. For example, if ALAE is covered in proportion to the amount of loss covered, Step 5 would need to be performed prior to Step 4.

other things, site-wide preliminary assessments (typically, OU 00) and emergency removal actions. There are costs associated with these removal action OUs, which are discussed later. However, at this point, we only want to consider those OUs which do (or will) have RoDs. To accomplish this, we can develop the ratio of the number of RoDs issued to date to the number of OUs with at least one RoD issued to date, by NPL site listing year, as displayed in Exhibit 5.[15]

An analysis of the estimated number of OUs per site is included in Exhibit 6. Once again, we circumvent the issue of OUs which will not have RoDs by developing the ratio of operable units with at least one RoD to NPL sites with at least one RoD. While there is variation in the results, note that the ultimate expected number of OUs per site for the 1987–1994 years is 1.47, almost identical to the estimate of 1.48 OUs per site from [8, p. 48]. Although potentially reasonable based on this comparison, however, research into approaches to estimate the tail factor for this type of analysis is left open as a topic for future study.

The specific approach used by the actuary to incorporate future RoDs at current NPL sites is at his or her discretion; the important point is that some form of development is necessary. Even on known sites, there may be future OUs planned. And, even on known OUs, there may be future RoDs planned (or not planned, but which will later be required). At the very least, an OU with an interim remedy RoD issued will likely require a follow-up RoD, describing any subsequently required cleanup efforts.

Note that this approach estimates RoDs per OU and OUs per site separately, rather than estimating RoDs per site directly. This is because a RoD cost typically relates to a given OU, rather than to the total site. Once we estimate the number of additional

[15]Note that an amendment RoD should not automatically be counted as an additional RoD for a given OU, since it can supplant, rather than just supplement the original. “No action remedy” RoDs with no (or minimal) associated costs should also be removed, unless the analysis’ average RoD cost(s) reflect them.

RoDs required at a current OU (using the ultimate RoD/OU ratio determined above), we can estimate the cost of these future RoDs by looking at the costs of RoDs relating to OUs with similar characteristics (i.e., similar types of contaminated media) at other sites. Similarly, when we estimate the number of future OUs at a given site (using the ratio of OUs with RoDs to sites with RoDs, also discussed above), we can estimate the characteristics of these additional OUs by looking at the characteristics of other OUs at similar sites (e.g., chemical plants, manufacturing plants, etc.). Then, once the characteristics of these future OUs have been determined, estimating the future RoD counts and costs on those future OUs is similar to estimating the future RoD counts and costs on current OUs.

The estimations referred to above are achieved in this model through simulation, based on the expected values derived previously. Simulation is also used to estimate the cost of future RoDs, which is addressed in the next section of this paper. The idea of simulating costs is especially important when estimating the cost for excess policy limits. As an example, suppose a particular site cleanup will cost either $500,000 or $1.5 million, depending on which of two equally-likely cleanup alternatives outlined in the relevant RoD is selected. The expected cost of this cleanup would be $1 million ($= 50\% * \$500{,}000 + 50\% * \$1.5$ million). If you are a reinsurer covering losses in excess of $1 million, you might not establish a reserve for this claim, since its expected cost only reaches, but does not pierce, the attachment point. However, there is a 50% chance that the reinsurer may be asked for $500,000 (since there is a 50% chance that the cost will be $1.5 million), and a 50% chance that the reinsurer may not be asked for any reinsurance recovery (if the cost is only $500,000). Under this scenario, then, a reasonable reserve for the reinsurer might be $250,000 ($= 50\% * \$500{,}000 + 50\% * \$0$), rather than the $0 reserve that might be established using the expected value method. From the primary insurance company viewpoint, an insurer protected by this reinsurance coverage would have booked

$1 million using the expected cost approach, but only $750,000 (= $1 million total expected cost, less the $250,000 ceded to the reinsurer) by incorporating variability into the site cost estimates.

No Clean Break from the Past: Estimating Future RoD Costs Using Environmental Characteristics

At this point, we have simulated the number and characteristics of future OUs at current sites, and simulated the number of future RoDs on those OUs. We now turn our attention to estimating the costs to be included in these future RoDs. First, we must differentiate between interim and final RoDs. This is infrequently discussed, but can be vitally important. *An "average RoD cost" multiplied by the current average number of RoDs per site yields a biased-low estimate of the average cleanup cost per site, if any of the sites contain interim RoDs for which the final RoDs have not yet been issued.* As a simple example, suppose only one Superfund site exists, with one operable unit and one (interim) RoD issued to address it. The average cost to clean that site using this approach would be the cost of that interim RoD, despite the fact that a final RoD will follow at some point in the future.

But even this level of detail—where interim and final RoDs are separately reviewed—can be further refined by selecting a set of *environmental characteristics* that best subdivides both the interim and final remedial action costs into even more homogeneous categories. The author believes that the more important, readily quantifiable characteristics are the remedy selected (for example, treatment vs. containment of the contamination), presence or absence of groundwater issues, and the process lead (i.e., whether the EPA or PRP was responsible to create the RoD). Additional characteristics based on the EPA's decision to promote consistency in remedy selections (discussed shortly) may also be considered. Other characteristics, such as the size and accessibility of the contaminated area, as well as current "policy" regarding preferred remedies are also highly relevant—but can be difficult to ascertain consistently and objectively via the RoDs.

Once the groundwater status and selected remedy values for a RoD are determined, they are fixed from that point forward for the remedial action relating to that RoD. The process lead, however, may change over time, as the EPA may turn over the responsibility for a site's cleanup to other parties during the remediation efforts. To the extent that the actuary believes that an EPA-led effort and non-EPA-led effort may differ in cost, some analyses of the past and future likelihood and timing of these (potential) changeovers is appropriate. Alternatively, one might try modeling based on an assumed frequency of changeovers for EPA-led activities at Superfund sites.

Another possibly relevant and measurable characteristic is the year the RoD was issued. These might be segregated into four groups:

1. *1986 and prior.* These RoDs were written in the program's infancy and addressed some of the most hazardous sites addressed through the Superfund program. The worst of these sites represents the most volatile and variable costs in recorded, historical RoDs.
2. *1987–1989.* The Superfund Amendments and Reauthorization Act of 1986 (SARA) directed the EPA to ensure that cleanups would be adequately protective of human health and the environment through the selection of more permanent remedies (i.e., emphasizing treatment, rather than containment).
3. *1990–1994.* An "enforcement first" policy, issued in 1989, led to a strong shift from EPA-led to PRP-led cleanup efforts.
4. *1995–Present.* The EPA begins phasing in new administrative reforms, intended to speed up cleanup efforts, improve cost-effectiveness and cut down on litigation. Costs included in RoDs issued since 1995 will likely be based on these initiatives, and should therefore be grouped accordingly.

The above should be considered in addition to the previously mentioned characteristics (plus any others the actuary feels are appropriate) with respect to the ever-present credibility trade-off: increasing the homogeneity of the data by breaking it up into additional pieces may simultaneously decrease the credibility of the data, since each piece would have less data included in it.[16]

We have now established the level of detail to be incorporated in this model to estimate the cost of a claim at a current Superfund site. The current RoD costs can be taken directly from the data in the Adjusted RoD Cost Database established earlier. The number of future RoDs required has also been determined. The characteristics of the additional RoDs required for a given OU can be simulated, based on the characteristics of RoDs relating to other OUs with similar OU characteristics. Once each future RoD's characteristics are simulated, the future RoD costs can be simulated based on the average and variance of costs in similar, current RoDs.

Several considerations relating to the simulation of these future RoD costs are noteworthy. First, which RoDs should be used, and why? The actuary may be able to allow for future legal, social and technological changes in future RoD cost estimates by only using the mean and variance of costs from similar RoDs issued during the most recent years. Two specific EPA initiatives prompt this suggestion. First, the EPA expects to reduce future costs by approximately $500 million based on its review and updates to more than 90 previously issued RoDs from the early years of the program.[17] In other words, the past will be

---

[16]In addition to helping quantify the cost of Superfund sites, environmental characteristics are also useful in helping an insurer's claim department evaluate the reasonableness of the insured's requested amount. For example, suppose a claim submitted by a policyholder relates to a site with contaminated soil being addressed by a containment remedy. The cleanup cost underlying the insured's claim can be benchmarked using the cost from RoDs that address contaminated soil through containment remedies at other sites.

[17]The reform guidance relating to these cost reductions was issued September 27, 1996. A significant portion of this savings is a result of three RoD cost adjustments: the Western Processing Site in Washington, the Norwood PCB Site in Massachusetts, and Metamora Site in Michigan have seen RoD cost reductions of $82 million, $47 million, and $28 million, respectively.

adjusted to look more like the present. Second, the EPA has set in place "presumptive remedies" for certain types of sites. According to Carol M. Browner, Administrator of the EPA:

> "Presumptive remedies are based on scientific and engineering analyses performed at similar Superfund sites and are used to eliminate duplication of effort, facilitate site characterization, and simplify analysis of cleanup options. EPA issued presumptive remedy guidances for the following: municipal landfill sites; sites with volatile organic compounds in the soil; wood treater sites; and a groundwater presumptive response strategy." [10]

In other words, the future will also be adjusted to look more like the present and the (adjusted) past. Therefore, limiting the data used to only the most recent data (which is not currently being adjusted) may reasonably address this issue. Then, after the average and variance of each combination of RoD characteristics is calculated using the most recent data, future RoD costs may be simulated.

Why use only recent RoDs to predict future RoDs on current sites? Exhibit 7 displays a graph of the history of RoD remedy selections from 1982 to the present. Note that from 1982 to 1986, containment-only remedies were the most prevalent. From 1987 to 1991, consistent with SARA's expressed preference for permanent remedies, treatment-oriented remedies predominated. From 1992 to the present, however, there is a slow but steady increase in "other" remedies. This grouping includes no-action remedies, site monitoring, site access restriction, and other such non-containment or treatment-based approaches. On average, these remedies cost less than containment or treatment remedies, and have yielded a decreasing average RoD cost in recent years. However, the majority of RoDs issued in recent years actually relate to sites listed on the NPL in the earlier years of the program, which have already had their more serious threats addressed in previous RoDs. It may be reasonable, therefore, to estimate the cost of future RoDs relating to these "mature" current sites using recent RoDs (which also likely relate to other "mature" sites).

However, many recently-listed (and some not-so-recently-listed) Superfund sites have not yet had their most serious threats addressed by any RoD. For these sites, using this overall current average RoD cost (relating primarily to mature sites) may not be appropriate. The author recommends instead simulating initial RoDs at these sites using the average cost of similar, initial RoDs recently issued at other sites. If it is necessary to simulate additional RoDs on these sites, the approach described in the previous paragraph may be appropriate.

A second consideration relating to the simulation of future RoD costs is that not all current sites should have the need for future RoDs randomly determined. It may be reasonable to expect that no additional RoDs will be required on sites which have either been deleted from the NPL or labeled construction-complete.

Third, an additional adjustment might be made to the data reflecting those few sites whose total costs are a multiple of the overall average. These sites are frequently referred to by actuaries as "megasites."[18] Insurers should be aware of their insureds with claims relating to these sites (which include, for example, Love Canal and Stringfellow), and should separate their potential liability at these sites from any analysis of their potential liability at the more "standard" Superfund sites, the same way that an actuary would typically segregate large losses from development triangles.[19] The actuary should remain alert to the possibility of new megasites, however, like the General Electric

---

[18]Interestingly enough, according to the RCRA/Superfund Hotline (1-800-424-9346), the EPA's original use of the term "megasite" did not refer to sites with high cleanup costs, but to sites with high remedial investigation and feasibility study (RI/FS) costs (in excess of $3 million).

[19]The presence of these megasites may invalidate the use of unadjusted average Superfund site cost estimates in an actuarial analysis. Since megasites would be included in an estimate of the average Superfund site cleanup cost, an insurer (or insured) not potentially

Pittsfield, Massachusetts Plant/Housatonic River site, currently estimated to cost more than $200 million and require more than ten years to clean—and only proposed for inclusion on the NPL in September of 1997!

Finally, we must account for the variability between a given effort's expected and actual cost, in addition to the variability of a given effort's expected cost alone. As noted earlier, according to the EPA, the actual cost of remediation should be between 70% and 150% of the RoD's expected cost. If the actuary considers the RoD cost as a "best estimate" with, say, a 95% probability that the actual cost will be between 70% and 150% of that best estimate, then the actual cost associated with each RoD could be simulated based on the expected cost and other relevant parameters.[20]

Now that we can estimate the cost of current claims on current NPL sites, we turn our attention to estimating the number of future claims on current NPL sites. The number of current claims on current NPL sites is readily available to the insurer; the estimate of future claims on current NPL sites requires some additional work, as described in the following section.

The Fly in the Ointment: Estimating Future Claims on Current Sites

One way to estimate the number of future claims on current NPL sites is to estimate the ultimate number of claims relating

---

liable at these megasites should likely use a lower estimate. Conversely, for an insurer (or insured) with liability at one or more megasites, the overall average is likely too low to apply. In those cases where the insurer doesn't know if an insured is or will become linked to a megasite, the actuary might decide in those cases that the overall average may be appropriate. Conversely, given the time that has elapsed since these megasites have been listed, the actuary may decide that, if the insured hasn't notified the insurer by now, there is likely no link present, and the average excluding the megasites may be used. This is, of course, at the discretion of each individual actuary's judgment.

[20]There is a question as to whether it is the nominal or discounted actual cost that should be between 70% and 150% of the expected RoD cost. In the case of a site requiring perpetual care, however, a range of 70%–150% of the expected undiscounted cost is almost meaningless. As a result, the actuary may want to adjust the model to reflect the likelihood that the costs fall within 70%–150% of the discounted RoD cost.

to current NPL sites, and subtract out the number of claims reported to date on those sites. Estimating the ultimate claim count for current sites can be done using a variation on the standard, actuarial triangle format and (ideally) internal company data. In the approach outlined in this paper, each row represents a different NPL listing year (i.e., sites listed on the NPL in 1983, sites listed on the NPL in 1984, etc.) and each column represents the amount of time (in years) between when a site was listed on the NPL and when a claim relating to that site was reported to the insurer (or reinsurer). This approach allows us to develop to ultimate the number of claims which will be presented to an insurer/reinsurer relating to sites listed on the NPL in each site listing year. Unlike typical development approaches, however, many PRPs will have reported claims to their insurers prior to the year a given site achieved NPL status. This is not a problem, since the triangle need not and should not have a "0" or "1" as its first column heading. Under this approach, the left-most column should be a negative number representing the greatest time lag between when an insured first notified its insurer of its PRP status at a site and when that site was subsequently listed on the NPL. The goal here is to develop to ultimate the number of claims relating to current Superfund sites.

If company data at this level of detail is not available (and usually it is not), an alternative is to use the EPA's data on PRP counts and notification dates (formerly in SETS, currently in CERCLIS) and NPL site listing dates (in CERCLIS) to estimate the ultimate number of PRPs linked to current NPL sites. As an example of how this approach would work, the reader is referred to Appendix B.

The resulting PRP notification pattern can then be lagged to reflect the expected average additional time between the EPA notifying a PRP of its potential liability at a site, and the PRP notifying its insurer.[21] To estimate this additional time lag, the

[21] This lag should also consider an adjustment for notification to reinsurers (and excess carriers) if appropriate, as well as collateral suit defendants, who by definition cannot

actuary should consider differences in the manner in which data has historically been reported to the insurance company. In the early days of pollution coverage disputes, many insureds reported multiple claims all at once, as part of declaratory judgment ("DJ") actions. These simultaneous, multiple reportings stemmed from the sudden recognition of possible insurance coverage availability. If the policyholder subsequently received notice of its potential liability at other sites, however, these additional claims would usually be reported to the insurer even in the midst of DJ proceedings to avoid possible late notice issues on those new claims. As a result, an insurer reviewing its data may notice an initial "flood" of claims from its insureds (during which there was likely no relationship between PRP and insurer notification dates), followed by a more stable relationship between PRP and insurer notifications. Since a new "flood" of initial claim reportings from an insurer's policyholders is unlikely to occur in the future, the author suggests that the time lag between PRP and insurer notifications relevant to future claim reportings may be estimated using PRP notification and corresponding claim report dates, excluding the policyholders' initial, multiple-claim reportings from the late 1980s to the early 1990s. Multiple claim reportings by insureds after this time period may either be included or excluded, depending on the actuary's judgment as to whether they should be considered part of future expectations or aberrational.

The actuary may also want to separately review policyholders according to their relative likelihood of liability for Superfund-related costs. (See the "Superfund Liability Pyramid" discussion in *It's a Dirty Job, but Someone's Gotta Do It: Cleanup Cost Liability Allocation* in Section 2.) These splits were not included in this paper, as it would complicate the description of the approach. Also, it is possible that a single policyholder linked to a

---

notify their insurers until after another PRP seeks them out. Estimating these time lags—which will no doubt differ for insurers and reinsurers—may be a very worthwhile area for future research.

single site may yield claims in multiple policy years. Adjustments to reflect this issue, if any are desired, may be made based on a review of insurance company claims data and discussions with legal counsel.

Other factors possibly impacting the time lag between NPL site listing date and insurer notification include CERCLA-related legislative or administrative changes, major coverage-related court decisions and insurer settlement procedures. While these are significant issues, the author believes that they may only have a modest impact with regard to this particular time lag issue. First, the author is not aware of any recent CERCLA legislation that might have significantly impacted this time lag. In addition, litigation over the question of whether or not insurance coverage is applicable to Superfund-related cleanup costs has slowed, with recent decisions in the environmental area focusing more on the allocation of costs among the insured and insurers (where applicable) than the determination of coverage. As a result, focusing on the more recent development factors in the parallelogram (and possibly any trends in those factors) may diminish any potential concern regarding these issues. Finally, though insurer reserving and settlement practices may significantly impact the data used to estimate an insurer's expected cost, the author does not expect that they will significantly impact the time lag between NPL site listing and insurer notification.

We have now completed the discussion on estimating an insurer's potential Superfund-related liability at current NPL sites. The following section addresses how an actuary might estimate an insurer's potential liability stemming from future NPL sites.

Incurred but not Remediated: Estimating the Cost of Future Sites

To estimate an insurer's liability stemming from future Superfund sites, the model assumes that an estimate of the total, ultimate number of NPL sites is available to the actuary. For reference, some estimates of the total number of NPL sites from different sources have been compiled in [11]. Then, the number of future sites can be calculated directly as the estimated, total Superfund site count, less the number of current NPL sites.

While there are several approaches to estimating true IBNR, one approach the author has seen is to multiply the total estimated cost to the insurer of current sites by the ratio of IBNR sites to current sites. This approach assumes that the percentage of current Superfund sites with no currently identified PRPs (referred to as "orphan sites") is similar to the percentage of future Superfund sites with no PRPs. It also assumes—among other things—a relatively stable average NPL site cost over time. On a present value basis, the shift over time from relatively expensive, shorter-term remedies (i.e., treatment) to relatively less expensive, longer term remedies (like containment and the more recent, "other" remedies) yields an overall downward cost trend. But does the duration of a typical, thirty-year (or longer) containment remedy applied against relatively low—but inflating—annual costs outweigh the high, up-front cost of treatment on an *undiscounted* basis? This would be a good area for future research.

The author's preferred approach is to estimate the total claim cost on future sites using a four step procedure:

1. estimate the percentage distribution of future sites by site type (e.g., chemical plants, landfills, etc.) based on recently listed sites and sites currently proposed for listing on the NPL,

2. estimate the future number of sites for each site type by applying the percentage distribution above to the up-front estimate of the total number of future Superfund sites,

3. multiply the future site counts for each site type calculated above by its respective future average site cost (which might be based on the cost of recently-listed, similar types of NPL sites), and

4. assume the insurer's percentage of future site costs for each site type is proportional to the insurer's percentage of current site costs for that site type.

Clearly, actuarial judgment may be applied at any step along the way, as desired.

Finally, some comments on the theory of "barrel scraping" are in order. According to [12], barrel scraping is "the theory that a disproportionate number of the worst problems were discovered and listed in the early years because of their obviousness, and that the (Superfund) program will increasingly be 'scraping the bottom of the barrel' as additional sites are listed." However, when evaluating how the average cleanup cost for NPL sites has changed (and will change) over time, the actuary should consider four additional items:

1. In addition to the few, ultra-costly "megasites," many more sites listed in the early to mid-1980s were subsequently de-listed with minimal if any remedial activities necessary. (The smaller costs associated with these non-remediated sites may have stemmed from short-term removal actions, RI/FS activities, monitoring costs, etc.) Like the megasites, these "microsites" were predominantly listed on the NPL between 1983 and 1986, and contributed to the average cleanup cost for sites listed during those years. As a result, the average cleanup cost of sites listed on the NPL from 1983 to 1986 is lower than it would otherwise be, were it not for the presence of these microsites.

2. Improved site-screening technology over time, as well as a revised hazard ranking scoring approach (discussed earlier in this paper), has led to a significant reduction in (and possible elimination of) the number of microsites listed on the NPL during the late 1980s to mid-1990s. The removal of low-cost sites from the list of potential

NPL sites yields an average site cost for this time period that is higher than it would otherwise be, were it not for the changes in site-screening technology and the HRS scoring approach.

3. During the mid-1990s, the EPA initiated an effort to take advantage of more cost-effective technology by issuing RoD amendments that superceded the more costly remedies selected in earlier RoDs (in those instances where the remedies had not yet been implemented). As a result, the improvements in the cost-effectiveness of cleanup efforts that are expected to benefit currently listed sites are also benefiting previously listed sites (in the form of these RoD amendments). The impact of these RoD amendments, therefore, is to bring the average cost of currently and previously listed sites closer together than they would otherwise be, were it not for these RoD amendments promoting currently available technology on older Superfund sites.

4. Governors' Concurrence legislation enacted in 1995 (as noted earlier in this paper) required the EPA to receive approval from a state before listing a site located there on the NPL. As of this writing, it remains the EPA's policy to determine a state's position on the listing of a particular site before proposing it for inclusion on the NPL. This is important because, according to a GAO study [13], "Officials of 26 (60 percent) of the 44 states (surveyed) told us that they are more likely to support listing sites with cleanup costs that are very high compared to those for other types of sites." This implies that the cost reduction benefits discussed in the previous item may actually result in fewer future site listings, since the majority of states would be looking to list sites with higher cleanup costs. It would also likely result in an increase in the average cost of future Superfund sites, relative to the average future site cost that would otherwise have

been expected (since the sites *not* listed would be those that are less costly).

Another consideration that might imply a possible *downward* shift in historical site costs over time is the shift from EPA-led efforts to PRP-led efforts. The theory is that a PRP spending its own money may have greater incentive for cost control than the EPA, which may be spending money it hopes to collect later from PRPs. In conjunction with item 4 above, however, the author believes that the expected impact of this issue is more of a decrease in the number of future Superfund sites than a change in the average cost of future Superfund sites, since these future sites where the costs could be lowered might no longer be listed.

In summary, based on all of the above, it is the author's opinion that the average undiscounted Superfund site cleanup cost may not have changed very much over time, and that the average cleanup cost of future Superfund sites might, in fact, be larger than the average cost of currently listed sites (depending on the extent of the impact of item 4 above)—or at the very least, not necessarily be lower than the average cost of currently listed sites, as is implied by the barrel scraping theory.[22]

Does the barrel scraping theory apply to non-NPL sites? The author's opinion about this is similar to his opinion about barrel scraping at NPL sites, though for different reasons:

- The GAO survey noted above implies that the majority of states favor supporting the most costly sites for NPL listing

[22]It would be interesting to test the impact of the barrel scraping theory on sites listed to date using actual cost data (or at least estimated costs from RoDs). However, as of this writing, less than half the sites listed since January of 1991 (after the change in the HRS approach) appear to have had even a single RoD issued for them, per CERCLIS. For sites listed since January of 1995 (the year Governors' Concurrence legislation and some of the SACM initiatives were introduced), less than one-third of the sites listed appear to have had any RoDs issued so far. Further complicating this study is the fact that estimating the number and cost of future RoDs needed on these sites (both where some RoDs have been issued as well as where none have yet been issued) requires assumptions about what the number and costs of those RoDs will likely be—which in a sense puts the cart before the horse, requiring one to answer the barrel scraping question by first assuming it to be true or false.

on a going-forward basis. Shifting the other potential NPL sites into state Superfund programs (which, as will be discussed further later in this paper, are generally considered to have a lower average cleanup cost) will tend to raise the average cleanup cost of non-NPL sites in recent years and into the future. And, while there may be some administrative cost reductions stemming from the "transplanting" of NPL sites from the EPA to the states' jurisdictions, the author believes it unlikely that this jurisdictional shift alone would bring the cost of an otherwise Superfund-worthy site down from the average NPL site level to the average non-NPL site level.

- With the EPA's introduction of the Brownfields initiative in the mid-1990s (which promotes cleanup efforts through financial rewards, rather than enforcement-related penalties), many potential hazardous waste sites that might have otherwise been addressed through state or federal enforcement are now being addressed with the voluntary cooperation of the responsible parties. Many states have since instituted similar programs.

A potentially responsible party's decision whether or not to voluntarily clean a site under these programs is likely based on that site's expected cleanup cost, relative to the benefits derived from performing the cleanup (e.g., tax benefits, improved public perception). The author believes that the non-NPL sites cleaned under these initiatives are likely the less costly ones, since the other sites' cleanup costs may be more likely to outweigh the benefits of performing those cleanups (which may partially explain why few if any expensive Superfund site cleanup efforts are voluntary). As a result, if it is believed that *voluntary* cleanup efforts are not likely subject to insurance recoveries, then the removal of these smaller, less costly sites from the potentially insurable universe of non-NPL sites also yields an increase in the average non-NPL site cleanup cost relevant to insurers.

Based on the above, the author believes that increased state Superfund capacity for larger cleanup and enforcement-related efforts over time, in conjunction with more recent federal and state initiatives centered on achieving voluntary cooperation from responsible parties for the smaller cleanup efforts, may have resulted in an increase in the average non-NPL site cleanup cost over time *for those sites potentially relevant to insurers*—or at the very least, not necessarily a decrease, as would be implied by the barrel scraping theory.

In summary, then, the author believes that the future average cost for both NPL and non-NPL sites may be larger than historical levels. In the case of Superfund, this is due largely to a reduction in the number of expected future sites with smaller associated costs. In the case of non-NPL sites, this is due to an increase in the number of higher cost sites (e.g., the “dropping down” of some otherwise Superfund-worthy sites) in addition to the removal of some of the less costly sites (e.g., the voluntary cleanups).

It is important to stress that many of the reasons the author questions the barrel scraping theory stem from political changes (e.g., the Brownfields initiative, Governors’ Concurrence legislation) and technological changes (e.g., improvements in site-screening technology) that—in the author’s opinion—mitigate (if not eliminate) the likely impact of the barrel scraping theory. Were it not for these issues, the author would probably support the barrel scraping theory as well.

8. RUMMAGE SALE: KNOWN SITE SETTLEMENTS AND POLICY BUYBACKS

A policy buyback represents an agreement between an insurer and an insured whereby the insurer pays money to the insured in exchange for which the insured provides a full or partial release from any future liability relating to a policy or set of policies. In the event of a full policy buyback, the insurer is relieved of all responsibility for both case reserves and IBNR. In the event of a partial policy buyback, the insurer is typically relieved of responsibility for both case reserves and IBNR relating to specific causes of loss only.

A known site settlement represents an agreement between an insurer and an insured whereby the insurer pays money to the insured in exchange for which the insured provides a release from any future liability relating to known sites only. This relieves the insurer of responsibility for case reserves on the claims relating to those sites. However, the insurer may remain potentially liable for claims relating to other current sites, if claims relating to them were not included in the settlement. Also, since the insurer remains potentially liable for that insured's claims relating to future sites, a known site settlement does not eliminate IBNR.

While these are significant issues, a detailed discussion of them is outside the scope of this paper. In general, however, the reader should note the following:

1. Adjustments for historical policy buybacks can be made by running the model excluding them, and then adding to the model results the costs paid by the insurer to achieve them.
2. Adjustments for historical known site settlements can be made in the same way as described for policy buybacks. Alternatively, adjustments for these site settlements may be made by subtracting from the model's results the difference between the estimated and actual amount relating to settled claims. For example, if a ten claim site settlement was estimated to cost a total of $5 million in the model but actually settled for $3 million, then $2 million should be subtracted from the model's results.[23]

[23]The reader should note that a likely reason for the $2 million difference is the timing of the insurer's payments. The $5 million output from the model assumes that the insurer may be liable for costs as the policyholder incurs them over a long period of time. If the insurer settles the claim when the insured still has future payments to make (as is

Determining which of these two approaches to use may depend on whether or not the actuary finds it easier to search for and remove linkages between insureds and sites up-front (i.e., before running the model), or to review the model's results and adjust for any relevant linkages it identified (i.e., after running the model).

3. Adjustments to reflect future site settlements and policy buybacks may be made by reviewing trends in historical site settlement and buyback activity. Relevant issues include trends in the number, timing and average cost of buybacks and known site settlements.

4. When estimating known site settlement and policy buyback adjustments, the actuary should be mindful of the possibility that they could yield increases in the results, rather than reductions. This typically occurs in connection with policy buybacks where the policyholders—each linked to a large number of sites—have policies with high attachment points. In these cases, insurers are sometimes willing to buy their way out of possible future coverage, even though the expectation is that none of those insureds' claims would penetrate the covered layers. While this is a legitimate thing for an insurer to do, the result is still a situation where the actual cost may be greater than the expected.

9. GARBAGE IN, GARBAGE OUT: REMOVAL ACTION COSTS

As Exhibit 3 shows, removal actions are typically restricted to a one-year duration and a $2 million cost limit. There have been many instances where removal costs have exceeded this figure significantly, however, like the Summitville Mine site in

---

frequently the case), the insurer will presumably only pay the costs incurred by the insured to date plus the present value of the insured's expected future costs at the time of that settlement (though the discount rate used would likely also reflect the transfer of uncertainty from the insurer back to the insured).

Colorado, where more than $70 million has been obligated for removal actions alone. These costs are not included in the RoDs, and may produce enough variability in severity to have a material impact on the total cost at a particular site. To the extent that an insured (or insurer) may become liable for these removal costs, it may be worthwhile to consider modeling both remedial and removal costs. In addition, the actuary should try to stay abreast of the continuing stream of environmental liability-related rulings over time, to determine if any other environmental activities (beyond removal and remedial actions) may need to be included in this type of analysis.

10. CONSTRUCTION COMPLETE? (A FEW THOUGHTS ON NON-NPL SITE CLEANUP COSTS)

There are several important differences between Superfund and non-Superfund sites that should be considered when adapting this Superfund-based approach to non-Superfund sites, including the following:

- RoDs are only issued for Superfund sites. RoD-like cost information is not readily available for non-Superfund sites, though it has been generally accepted that cleaning an average (less hazardous) non-NPL site will be significantly less costly than cleaning the average (more hazardous) NPL site. However, within the context of comparing particular types of NPL and non-NPL sites (e.g., landfills listed on the NPL vs. landfills being addressed through state enforcement activities), this is a debatable point. Many actuaries have postulated that the level of hazard and cost at a particular site are directly related,[24] but it is more likely that the EPA's selected remedy for a site based on its relative hazard level (i.e., treat the worst sites and contain the rest) drives the cost. This is important, because for non-NPL sites—where the EPA may not be

[24]The author's negative view of this argument—and the rationale for it—are detailed in Appendix C.

involved—if a particular state does not share the EPA's philosophy, the possible relationship between hazard and cost might not hold. Three more arguments in favor of higher than expected non-NPL site enforcement-based cleanup costs include: (1) some states may not have reported to the EPA all of their hazardous waste sites—many of which may be Superfund-worthy—simply to avoid the perceived delays in cleanups, (2) many states in the past have not considered as many alternative remedies as the EPA prior to determining the selected remedy, which may have caused more cost-effective and equally viable remedies to be excluded from the non-NPL site cleanup alternatives, (3) non-NPL sites requiring no cleanup actions will not produce claims, and sites requiring small-scale efforts will likely be dealt with through voluntary cleanup programs, which might not be considered insurable. Clearly, the removal of these smaller claims from the insurable non-NPL universe will tend to raise the relevant average non-NPL enforcement-based cleanup cost.

- While an estimate of the ultimate number of Superfund sites may be based on the current number of sites already on the NPL and those still in CERCLIS awaiting their NPL-status determination, there is no single, generally accepted estimate of either the current or total number of non-NPL sites that will require cleanup through enforcement (non-voluntary program) actions.
- Estimating an insured's potential liability at a Superfund site frequently includes an estimate based on the number and names of other PRPs at that site. Neither the number nor the names of potentially responsible parties is readily available at most non-NPL sites, however, though it is generally accepted that non-NPL sites have far fewer PRPs than NPL sites (frequently as few as one!). And, similar to NPL sites, even if the number and names of all PRPs for a given non-NPL site were available, a PRP's share of liability might not correlate well with the number of PRPs potentially sharing the cleanup cost

at that site. An additional problem is that not all states apply "retroactive, strict, and joint and several" liability standards. As noted earlier, estimating a given PRP's expected liability share is one of the most difficult aspects of estimating an insurer's environmental liabilities.

- Relevant characteristics applicable to non-Superfund sites may differ from those of Superfund sites, even if RoD-like cost data were available, due to (among other things) differences in state-by-state cleanup requirements and the types of site in each category. For example, Superfund will rarely include leaking underground storage tank (LUST) sites, since these are almost always filled with petroleum—not a substance to which Superfund moneys are intended to respond. (These are addressed under RCRA; see Appendix D.) As a result, when LUST cleanup efforts are required, they will almost always be addressed as non-NPL sites. Small fuel leaks and drycleaner sites will also typically be addressed as non-NPL sites, usually too small and not hazardous enough to warrant NPL listing. It is worth noting that the types of non-NPL sites discussed here (i.e., small fuel leaks, drycleaner sites and LUSTs) tend to be less costly on average than the types of sites typically found on the NPL (e.g., manufacturing and chemical plants) resulting in a lower overall average cost for non-NPL sites than for NPL sites. However, for sites that appear both on and off the NPL (such as landfills), comparisons between NPL and non-NPL site costs may be reasonable.

11. DISCOUNTING THE PROBLEM: WHAT'S IT WORTH TO YOU?

While this topic is clearly deserving of a paper in its own right, a brief introduction to some relevant concepts is included here. In most discounting analyses, three items are required: an estimate of undiscounted total cost, a payout pattern and a discount rate. To discount Superfund liabilities, three additional values are useful: a Superfund cost incurral pattern (indicating the timing of costs incurred by those actually cleaning up the Superfund site, regardless of any cost-sharing agreements or future reimbursements which may apply), a probability of payment (based on the idea that the insurer may or may not be successful in denying liability for the claim altogether), and an estimate of the insured's share of liability for site cleanup costs.

The Superfund cost incurral pattern is necessary because the insurer's potential cost burden relates to future costs associated with Superfund cleanup in addition to those previously incurred. In a car collision claim, an insurer's payment is typically made after the car is repaired and the cost to fix the car is known. In Superfund liability claims, however, cleanup costs are incurred before, during and after an insurer may be found liable for site cleanup costs. Once found liable, an insurer may reimburse the insured for past costs incurred to date in connection with that site's cleanup efforts, but may be reluctant to pre-pay future annual cleanup costs which the insured will incur over the next several years at that site. As a result, the payout pattern for an insurer found liable for site cleanup costs at a given site would be comprised of (1) a first payment, based on cleanup costs incurred to date by its insureds at that site, and (2) annual payments beginning the following year, equal to the cleanup costs to be incurred by the insureds in each subsequent year in which cleanup efforts are required.[25] If the insurer is attempting to deny liability for this claim, however, an additional lag may be necessary to reflect the time between when the insured first notified the insurer of the cleanup claim and when the determination is later

[25] In practice, once liability has been determined, the insurer may instead offer to simply reimburse the insured's past costs and offer the insured the net present value of the future costs to be incurred in connection with the site's cleanup efforts. This present value concept should not be confused with the idea of discounting reserves for statutory reporting purposes. As an example, suppose that in three years, an insurer will extinguish its liabilities to an insured for a particular site by paying the present value (at that time) of costs to be incurred after that date. For simplicity's sake, also assume that the insured will have spent nothing on site cleanup up to that point, and that the payment amount will be $133,100. This $133,100 represents the insurer's current, undiscounted liability to that insured at that site. Assuming, for example, a discount rate of 10% applies, the discounted value of that claim would be calculated as $\$133,100/(1.10)^3$, or $100,000.

made regarding whether or not coverage applies. If it is felt that a determination of liability would take three more years, for example, item (1) above would be the sum of the incurred to date costs, plus the next three years of annual payments, and would be presumed payable (pending determination of liability) three years from today. Item (2) would, therefore, begin with the fourth year of annual payments, and would be assumed to begin one year thereafter. This translation of the Superfund incurral pattern to the insurer's payout pattern is referred to in this paper as the "litigation lag." The litigation lag may be estimated from numerous sources, including the information underlying the selection of the probability of payment at a particular site, and allocated loss adjustment expense (ALAE) development (if there is sufficient history to produce a reasonable and reliable pattern).

The probability of payment represents the fact that, unlike more traditional claims, there is a chance that the insurer will not become obligated to pay for site cleanup costs. This value should differ at least by state, based on relevant court decisions in each state. Similarly, the estimated share of liability reflects the fact that an insured might be held responsible only for a portion of the total cleanup costs at a site, limiting the insurer's liability at that site to its insured's share of liability at that site. This is an important consideration, which, as noted above, is beyond the scope of this paper.

With these issues in mind, one approach that might be used to estimate the discounted Superfund liabilities of an insurer is to (1) estimate the amount and timing of the Superfund cleanup costs incurred at each site (regardless of who will ultimately bear liability for them) using site costs and site cost incurral patterns based on the adjusted RoD costs described earlier in this paper, (2) multiply each of the annual cleanup costs by the estimated share of responsibility borne by the insured, (3) reallocate the insured's Superfund site costs at each point in time based on each site's estimated litigation lag, (4) remove from the litigation lag-adjusted cost incurral pattern the costs incurred before the attachment point is reached and the costs incurred after the policy limit is exhausted, (5) multiply each of the remaining annual cleanup costs by the probability that coverage applies, (6) add together the reallocated costs for all Superfund sites within each calendar year to estimate the costs to be paid by the insurer relating to all Superfund claims in that year, and then (7) discount the Superfund claim payment stream using the selected discount rate.

An additional issue, of course, is the discount rate that should be applied. One approach might be to tie the discount rate in some way to the U. S. Treasury Bond rate in effect at the appropriate point in time (e.g., year-end for statutory reporting purposes), with a duration closest to the estimated RoD cleanup duration for the OU(s) in question. Alternatively (and depending upon the reason for discounting the costs), an insurer could consider the discount rate underlying previous coverage buybacks. The author suggests consulting [14] prior to selecting a discount rate.

12. A PRELIMINARY ASSESSMENT: SOME CONCLUDING THOUGHTS

The author hopes that this paper will serve as a stepping stone for future research into several areas noted throughout this paper, as well as other areas of environmental liability analyses. There is certainly enough that still needs to be done, including:

- research into non-NPL site counts and costs (including what drives them, and how they differ from NPL site cost and count drivers),
- research into other current and future environmental liability issues that should have an impact on our environmental analyses,
- development of alternate environmental liability models, and

- development of environmental (Superfund and non-Superfund) reserve discounting models (with an eye toward acceptability to regulators).

What might be said of the Superfund program in recent years could also apply to actuaries estimating its costs—much has been done, but plenty of work still remains.

REFERENCES

[1] Federal Register, Part VII, Environmental Protection Agency, *Final and Proposed Amendments to National Oil and Hazardous Substances Contingency Plan; National Priorities List*, 8, September 1983, p. 40659.

[2] Federal Register, Part V, Environmental Protection Agency, *National Priorities List for Uncontrolled Hazardous Waste Sites; Rule*, 25, April 1995, p. 20331.

[3] Federal Register, Environmental Protection Agency, *The National Priorities List for Uncontrolled Hazardous Waste Sites; Deletion Policy for Resource Conservation and Recovery Act Facilities*, 20, March 1995, p. 14642.

[4] United States Code, Title 42, Chapter 103, Section 9604(c) (3).

[5] United States General Accounting Office, *Superfund Program Management*, GAO/HR-97-14, February 1997, pp. 9–10.

[6] United States Environmental Protection Agency, *Estimated O&M Costs for RODs: Historical Trends and Projected Costs Through Fiscal Year 2040*, CH2M Hill, 31, May 1995, pp. 3–5.

[7] United States Environmental Protection Agency, *Remedial Action Costing Procedures Manual*, EPA/600/8-87/049, October 1987, pp. 3–21.

[8] Russell, Milton, and Kimberly L. Davis with Ingrid Koehler, *Resource Requirements for NPL Sites*, University of Tennessee, Joint Institute for Energy and Environment, Knoxville, TN, 1996, Appendix A, p. A-10.

[9] United States General Accounting Office, *Superfund: Operations and Maintenance Activities Will Require Billions of Dollars*, GAO/RCED-95-259, September 1995, pp. 1–2.

[10] United States Environmental Protection Agency, *Statement of Carol M. Browner, Administrator, U.S. Environmental Protection Agency Before the Committee on Environment and Public Works—U.S. Senate*, 5, March 1997.

[11] American Academy of Actuaries, *Costs Under Superfund: A Summary of Recent Studies and Comments on Reform*, Public Policy Monograph, August 1995, p. 3.

[12] Congressional Budget Office, *The Total Costs of Cleaning Up Nonfederal Superfund Sites*, January 1994, p. 20.

[13] United States General Accounting Office, *Hazardous Waste: Unaddressed Risks at Many Potential Superfund Sites*, GAO/RCED-99-8, November 1998, p. 26.

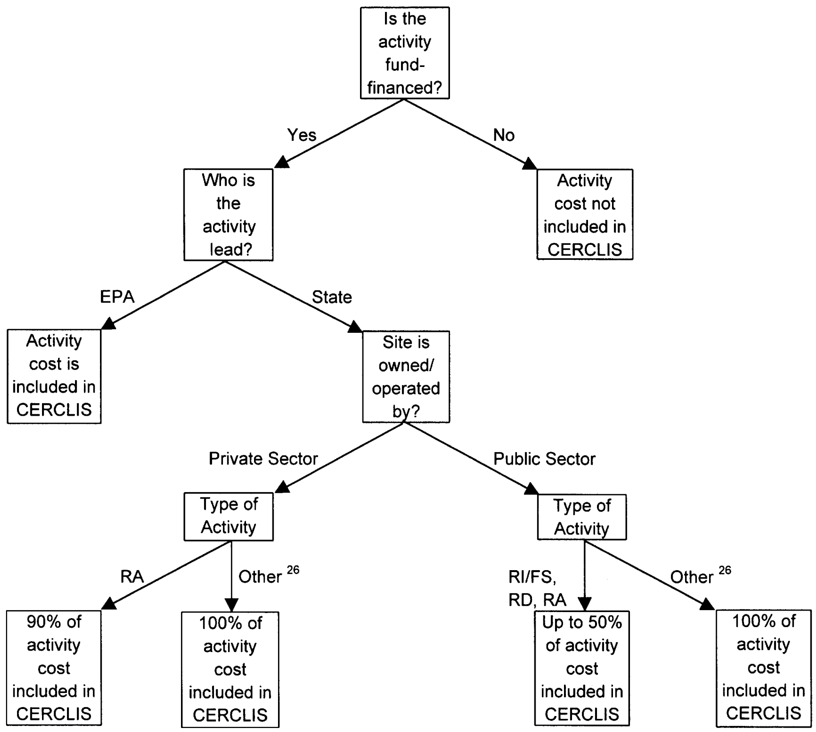
[14] Actuarial Standards Board, Actuarial Standard of Practice No. 20, "Discounting of Property and Casualty Loss and Loss Adjustment Expense Reserves," April 1992.

[15] United States Environmental Protection Agency, Office of Solid Waste, *Draft Regulatory Impact Analysis for the Final Rulemaking on Corrective Action for Solid Waste Management Units Proposed Methodology for Analysis*, March 1993.

[16] United States General Accounting Office, *Hazardous Waste: Remediation Waste Requirements Can Increase the Time and Cost of Cleanups*, GAO/RCED-98-4, October 1997.

EXHIBIT 1

INTERPRETING CERCLIS COST DATA



[26] Excluding those O&M costs not considered eligible for Superfund funding.

EXHIBIT 2

CERCLIS DATA VS. ROD DATA

	CERCLIS	RoDs	Comments
Timeframe	Contains actual, historical incurred to date costs	Contain estimated, prospective costs	CERCLIS also includes information on planned activities
Whose Expenditures are Included?	EPA only	Anyone who will be required to perform the relevant activities	If EPA partially funds an activity, adjustments must be made to derive the total cost from CERCLIS. (See Exhibit 1.)
Cost of Remedy Construction	Included in Remedial Action	Included in Capital Cost	
Cost of Remedy Implementation	At Least Partially Included in Remedial Action	Included in O&M Cost	Percentage of cost included in CERCLIS varies by site ownership, activity lead (i.e., EPA, state, or PRP) and type of activity
Cost of Performing O&M Activities	Not Included in CERCLIS	Included in O&M Cost	
Oversight of Remedial Action, Where Necessary	Included in Remedial Action Cost	Not Included in RoDs	
Oversight of O&M Activities, Where Necessary	Included in O&M Cost	Not Included in RoDs	
Cost Level of Dollar Values	Nominal (Undiscounted)	Discounted	

EXHIBIT 3

Steps in The Process of Listing a Site on the NPL

Activity	Estimated Average Duration[27]	Estimated Average Site Cost[27]	Comments
Site Discovery			Not all potentially hazardous waste sites are reported to the EPA to be included in CERCLIS; many non-NPL sites have never been on CERCLIS.
Preliminary Assessment (PA)			Characterizes threat based on off-site analysis of readily available data: is it imminent (removal action needed), serious (site inspection needed) or not serious (archive the site)?
Removal Action	Up to 1 Year	Up to $2,000,000	Generally targets immediate threats only; remaining serious threats dealt with in Site Inspection (SI) phase. Time and spending limits may be increased if immediate risk to public health or the environment remains.
Site Inspection (SI)			Preliminary and cursory on-site qualitative evaluation of hazard posed
Hazard Ranking			Numerical quantification of the degree of hazard posed at the site based on data collected during PA and SI phases
Proposed for Placement on NPL			Proposal to NPL of sites with Hazard Ranking Scores of 28.5 or greater, State Priority Sites and sites to be listed at the ATSDR's request.

Final Placement on NPL			
Remedial Investigation/Feasibility Study (RI/FS)[28]	18–30 Months	$1,250,000	Actual cleanup cost should be between 50% and 200% of the RI estimate, and then 70% and 150% of the Feasibility Study (RoD) estimate
Remedial Design (RD)	12–18 Months	$1,260,000	Actual cleanup cost should be between 95% and 115% of the RD estimate
Remedial Action (RA)	12–36 Months	$22,500,000	Remedy-related construction and implementation activities. Average duration displayed is per RA; average cost per site.
Construction Completion			Occurs when no further physical construction is required, whether or not remedy has been implemented (i.e., may be before RA completion).
Ongoing Operations & Maintenance (O&M)	22 Years	$5,360,000 on a PV Basis; $20,795,000 Undiscounted[29]	EPA guidance limits duration to 30 years for costing purposes, regardless of actual expectations. Average duration displayed relates to O&M activities for a single RoD; average cost displayed is per site.
Deletion from NPL			Typically occurs when EPA determines that no further response is required to protect human health or the environment.

[27]Duration and cost information from various sources, including [2], [4], CERCLIS and conversations with the EPA.

[28]This results in the issuance of a Record of Decision (RoD) which details the problem, the selected remedy, and its expected costs. The scope at this phase is still based more on assumptions than hard data.

[29]Undiscounted cost assumes $400,000 first year O&M costs are expended, and a 30 year duration, per Federal Register. Also assumes annual inflation of 3%, with O&M activities beginning 3 years after RoD issuance. Discounted cost assumes no inflation and a 5.8% discount rate, which was applied against the first year of O&M costs, as well as each year thereafter.

EXHIBIT 4

DISCOUNT RATE GUIDANCE

Publication Date	Publication Title	Discount Rate[30]
Jun-93	Revisions to OMB Circular A-94 on Guidelines and Discount Rates for Benefit-Cost Analysis (OSWER Directive 9355.3-20)[31]	7%
Oct-88	Guidance for Conducting Remedial Investigations and Feasibility Studies Under CERCLA	5%
Mar-84	Remedial Action Costing Procedures Manual	10%

[30] Since the annual O&M costs included in RoDs are not increased for inflation over time, the discount rate used to calculate their present value also excludes a provision for inflation. For this reason, the discount rates shown here reflect pre-tax, after inflation discount rates.
[31] The referenced OMB circular is available through the internet, at http://www.whitehouse.gov/WH/EOP/OMB/html/circulars/a094/a094.html#7

EXHIBIT 5

Records of Decision (RoDs) Per Operable Unit (OU) with at Least One RoD

Year Listed on NPL	Cumulative Records of Decision/Cumulative OUs with at Least One Record of Decision Issued by Number of Years Since NPL Listing: 1	2	3	4	5	6	7	8	9	10	11	12	13	14
83	1.05	1.02	1.02	1.03	1.03	1.03	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.05
84	1.00	1.00	1.00	1.03	1.03	1.03	1.03	1.03	1.05	1.05	1.05	1.05	1.06	
85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00		
86	1.00	1.00	1.00	1.01	1.02	1.04	1.04	1.05	1.05	1.06	1.06			
87		1.00	1.00	1.00	1.00	1.00	1.02	1.02	1.03	1.03				
88														
89	1.00	1.00	1.00	1.01	1.01	1.01	1.04	1.04						
90	1.10	1.04	1.02	1.02	1.01	1.02	1.03							
91	1.00	1.00	1.00	1.00	1.00	1.00								
92	1.00	1.00	1.00	1.00	1.08									
93														
94	1.00	1.00	1.00											
95	1.00	1.00												
96	1.00													

EXHIBIT 5

Records of Decision (RoDs) Per Operable Unit (OU) with at Least One RoD

(Continued)

	Age to Age Factors													
	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12	12–13	13–14	
83	0.972	0.999	1.016	0.997	1.002	1.005	0.997	1.001	1.002	1.001	0.999	1.006	1.004	
84	1.000	1.000	1.026	1.007	0.992	1.004	1.004	1.012	1.003	1.004	0.998	1.004		
85	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000			
86	1.000	1.000	1.011	1.014	1.016	1.002	1.006	0.996	1.012	1.003				
87		1.000	1.000	1.000	1.000	1.018	0.997	1.012	0.998					
88														
89	1.000	1.000	1.010	1.005	0.998	1.027	1.003							
90	0.942	0.987	0.994	0.997	1.008	1.007								
91	1.000	1.000	1.000	1.000	1.000									
92	1.000	1.000	1.000	1.083										
93														
94	1.000	1.000												
95	1.000													
Age-to-Age	1.000	1.000	1.000	1.021	1.004	1.009	1.001	1.004	1.003	1.002	0.999	1.005	1.004	
Age-to-Ultimate	1.075	1.075	1.075	1.075	1.053	1.048	1.039	1.038	1.033	1.030	1.028	1.029	1.024	1.020
RoDs/OU	1.00	1.00	1.00		1.08	1.00	1.03	1.04		1.03	1.06	1.00	1.06	1.05
Ult RoD/OU Ratio	1.08	1.08	1.08		1.08	1.05	1.07	1.08		1.06	1.09	1.03	1.08	1.07

Overall Average RoDs/OU =	1.07
Average, 1983–1986 =	1.07
Average, 1987–1996 =	1.07

Notes: The value of this ratio for the most recent years and the overall average is a consistent 1.07.

The value for the most recent diagonal for site listing year 1992 represents 7 RoDs issued on 7 OUs through calendar year-end 1995, and 6 more RoDs issued on 5 more OUs during calendar year 1996, yielding 13 RoDs on 12 OUs, and the 1.0833 ratio. Due to the limited number of RoDs and OUs for this year, the average of the most recent three years' results was used for this value.

Approaches to estimating the tail factor for this type of analysis is left as a subject of further research.

EXHIBIT 6

Operable Units (OUs) with at Least One RoD Per Site with at Least One RoD

Year Listed on NPL	Cumulative OUs with at Least One RoD Issued/Cumulative Sites with at Least One RoD Issued, by Number of Years Since NPL Listing: 1	2	3	4	5	6	7	8	9	10	11	12	13	14
83	1.06	1.06	1.15	1.15	1.19	1.27	1.32	1.36	1.40	1.41	1.44	1.48	1.51	1.54
84	1.00	1.00	1.05	1.06	1.12	1.15	1.15	1.22	1.26	1.33	1.33	1.38	1.41	
85	1.00	2.00	2.00	2.00	1.33	2.00	2.00	2.00	2.00	2.00	2.00	2.00		
86	1.00	1.05	1.14	1.19	1.28	1.28	1.34	1.39	1.40	1.42	1.45			
87		1.00	1.00	1.05	1.06	1.07	1.17	1.18	1.21	1.21				
88														
89	1.11	1.17	1.26	1.24	1.25	1.26	1.28	1.30						
90	1.00	1.08	1.07	1.08	1.11	1.13	1.17							
91	1.00	1.00	1.00	1.00	1.00	1.00								
92	1.00	1.00	1.00	1.00	1.09									
93														
94	2.50	2.00	1.44											
95	1.33	1.40												
96	1.00													

EXHIBIT 6

Operable Units (OUs) with at Least One RoD Per Site with at Least One RoD

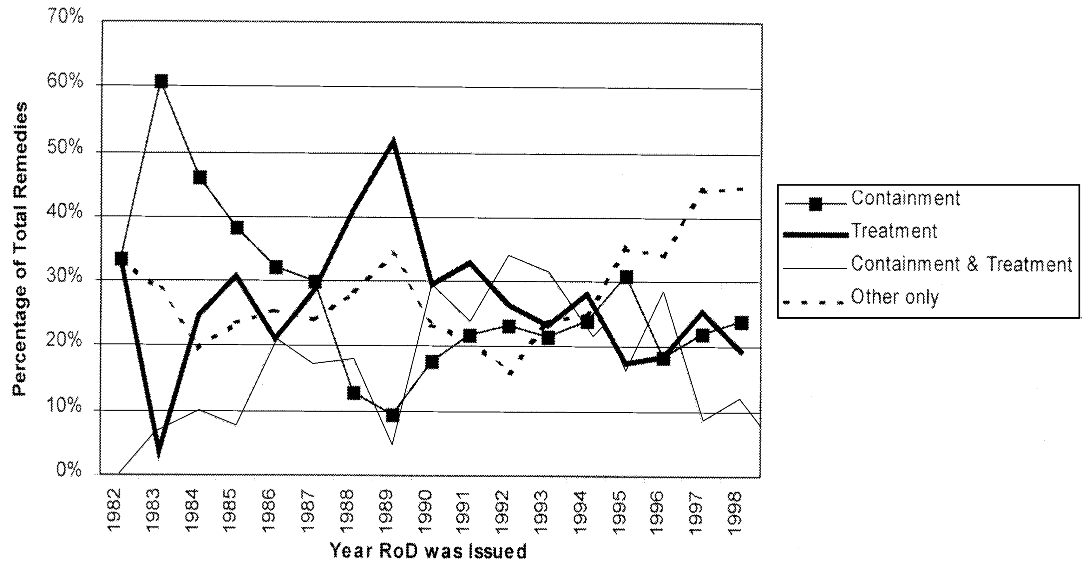
(Continued)

	Age to Age Factors													
	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11	11–12	12–13	13–14	
83	1.00	1.08	1.00	1.04	1.06	1.04	1.03	1.03	1.01	1.02	1.02	1.02	1.02	
84	1.00	1.05	1.01	1.06	1.03	1.00	1.06	1.03	1.06	1.00	1.03	1.03		
85	2.00	1.00	1.00	0.67	1.50	1.00	1.00	1.00	1.00	1.00	1.00			
86	1.05	1.08	1.04	1.08	1.01	1.04	1.04	1.01	1.01	1.02				
87		1.00	1.05	1.01	1.01	1.09	1.01	1.02	1.00					
88														
89	1.06	1.07	0.99	1.01	1.01	1.01	1.02							
90	1.08	0.99	1.01	1.02	1.02	1.04								
91	1.00	1.00	1.00	1.00	1.00									
92	1.00	1.00	1.00	1.09										
93														
94	0.80	0.72												
95	1.05													
Age-to-Age	1.04	1.02	1.00	1.03	1.01	1.05	1.03	1.02	1.02	1.01	1.02	1.02	1.02	
Age-to-Ultimate	1.40	1.35	1.32	1.31	1.27	1.26	1.20	1.17	1.15	1.13	1.12	1.10	1.07	1.05
OUs/Site	1.00	1.40	1.44	—	1.09	1.00	1.17	1.30	—	1.21	1.45	2.00	1.41	1.54
Ult OUs/Site	1.40	1.89	1.90	—	1.38	1.26	1.41	1.53	—	1.37	1.62	2.20	1.52	1.62
Ratio														

Overall Average OUS/Site =	1.59
Average, 1982–1986 =	1.66
Average, 1987–1996 =	1.52
Average, 1987–1994 =	1.47

EXHIBIT 7

NON-FEDERAL SUPERFUND SITE REMEDY SELECTION TRENDS



Notes: This data relates to the remedies for each of the more than 2,800 contaminated media addressed in approximately 1,400 RoDs.
"Containment" includes both containment-only remedies, and RoDs where both "Containment" and "Other" remedies were selected.
"Treatment" includes both treatment-only remedies, and RoDs where both "Treatment" and "Other" remedies were selected.
"Other" includes RoDs where only "Other" remedies were selected.

APPENDIX A

SAMPLE RECORD OF DECISION (RoD) ABSTRACT

General Site Information

Site Name: MOTOR WHEEL
EPA ID: MID980702989 EPA Region: 05
Metro Statistical Area: 4040
Street: 2401 N HIGH ST (REAR)
City: LANSING TWP State: MI Zip: 48909
Congressional District: 08
County Code: 065 County Name: INGHAM
National Priority List (NPL) Status: F
Proposed NPL Update Number: Final NPL Update Number:
Ownership Indicator: OH
Federal Facility Flag: N Federal Facility Docket: F
Latitude: 4245390 Longitude: 08432060
LL Source: E LL Accuracy:
Incident Type: Incident Category: P
Resource and Recovery Act Facility: FMS SS ID: 05S5
Dioxin Tier: USGS Hydro Unit: 04050004
Site Description:

Remediation Information (Records of Decision)

Site Name: MOTOR WHEEL
EPA ID: MID980702989
Operable Unit:
ROD ID: EPA/ROD/R05-91/172 ROD Date: 09/30/91
Contaminant: VOCS
BENZENE
PCE
TCE
TOLUENE
XYLENES
ORGANICS
PAHS

PCBS
PESTICIDES
METALS
ARSENIC
CHROMIUM
LEAD
O&M Costs: Estimated Costs:
Keys: NONE

Abstract:

THE 24-ACRE MOTOR WHEEL SITE IS AN INACTIVE INDUSTRIAL WASTE DISPOSAL SITE IN LANSING, INGHAM COUNTY, MICHIGAN. LAND USE IN THE AREA IS PREDOMINANTLY INDUSTRIAL. THE SITE OVERLIES A GLACIAL TILL AND A GLACIAL AQUIFER. FROM 1938 TO 1978, THE MOTOR WHEEL CORPORATION USED THE SITE FOR THE DISPOSAL OF SOLID AND LIQUID INDUSTRIAL WASTES INCLUDING PAINTS, SOLVENTS, LIQUID ACIDS AND CAUSTICS, AND SLUDGE. WASTES WERE DISPOSED OF IN TANKS, BARRELS, SEEPAGE PONDS, AND OPEN FILL OPERATIONS. AN ESTIMATED 210,000 CUBIC YARDS OF WASTE FILL IS IN PLACE ONSITE. AS A RESULT OF DISPOSAL PRACTICES, CONTAMINANTS HAVE LEACHED THROUGH THE SOIL AND INTO THE UNDERLYING GLACIAL AQUIFER AND PERCHED ZONE. BETWEEN 1970 AND 1982, AT LEAST THREE ONSITE CLEAN-UP ACTIONS WERE INITIATED. IN 1970, THE STATE REQUIRED THE REMOVAL AND OFFSITE DISPOSAL OF SOLID WASTES, PAINT SLUDGE, AND OILS FROM SEEPAGE PONDS AND BACKFILLING OF EXCAVATED POND AREAS. IN 1978, INDUSTRIAL WASTES AND DEGRADED SOIL WERE EXCAVATED AND STOCKPILED ONSITE UNDER A CLAY COVER.

IN 1982, THE SITE OWNERS REMOVED THREE 10,000-GALLON TANKS, THEIR CONTENTS, AND SURROUND-

ING CONTAMINATED SOIL, ALONG WITH CONTAMINATED FILL MATERIAL CONTAINING AN UNKNOWN QUANTITY OF DRUMS. THIS RECORD OF DECISION (ROD) ADDRESSES THE WASTE MASS AND GROUND WATER CONTAMINATION IN THE PERCHED ZONE AND THE GLACIAL AQUIFER. THE PRIMARY CONTAMINANTS OF CONCERN AFFECTING THE SOIL, DEBRIS, AND GROUND WATER ARE VOCS INCLUDING BENZENE, PCE, TCE, TOLUENE, AND XYLENES; ORGANICS INCLUDING PAHS, PCBS, AND PESTICIDES; AND METALS INCLUDING ARSENIC, CHROMIUM, AND LEAD.

THE SELECTED REMEDIAL ACTION FOR THIS SITE INCLUDES BACKFILLING THE NORTHERN PORTION OF THE FILL AREA WITH 125,000 CUBIC YARDS OF FILL; CAPPING THE DISPOSAL AREA WITH A 14.9-ACRE MULTI-MEDIA CAP; INSTALLING A SLURRY WALL AT THE WESTERN AND SOUTHERN BOUNDARY OF THE DISPOSAL AREA; INSTALLING GROUND WATER RECOVERY WELLS OR TRENCHES DOWNGRADIENT, AND A COLLECTION TRANSFER SYSTEM TO DELIVER WATER TO AN ONSITE TREATMENT FACILITY; PRETREATING GROUND WATER ONSITE TO REMOVE IRON AND MANGANESE USING AERATION, CLARIFICATION, AND FILTRATION IF NEEDED, FOLLOWED BY ONSITE TREATMENT USING AIR STRIPPING AND CARBON ADSORPTION; USING ACTIVATED ALUMINA TO REMOVE FLUORIDE FROM GROUND WATER, FOLLOWED BY OFFSITE DISCHARGE OF THE TREATED WATER TO A PUBLICLY OWNED TREATMENT WORKS (POTW); MONITORING GROUND WATER; AND IMPLEMENTING INSTITUTIONAL CONTROLS INCLUDING DEED AND GROUND WATER USE RESTRICTIONS, AND SITE ACCESS RESTRICTIONS SUCH AS FENCING. THE ESTIMATED PRESENT WORTH COST FOR THIS REMEDIAL ACTION IS $30,720,300, WHICH INCLUDES A CAPITAL COST OF

$11,083,300 AND AN ANNUAL O&M COST OF $1,277,400 FOR 30 YEARS. PERFORMANCE STANDARDS OR GOALS; GROUND WATER CLEAN-UP GOALS ARE BASED ON STATE HEALTH-BASED STANDARDS OR METHOD DETECTION LIMITS (MDL), WHICHEVER IS HIGHER. CHEMICAL-SPECIFIC GOALS INCLUDE BENZENE 1 UG/L (STATE), PCE 1 UG/L (MDL), TCE 3 UG/L (STATE), TOLUENE 800 UG/L (STATE), XYLENES 300 UG/L (STATE), AND LEAD 5 UG/L (STATE).

Remedy:

THIS OPERABLE UNIT ADDRESSES REMEDIATION OF GROUNDWATER AND SOURCE CONTROL BY REDUCING THE POTENTIAL FOR CONTINUING GROUNDWATER CONTAMINATION FROM THE ON-SITE WASTE MASS AND REDUCING THE THREAT FROM CONTAMINATED GROUNDWATER THROUGH TREATMENT. THE MAJOR ELEMENTS OF THE SELECTED REMEDY INCLUDE;

* INSTALLATION OF AN APPROXIMATELY 11.3 ACRE MICHIGAN ACT 64 CAP OVER THE DISPOSAL AREA;

* BACK-FILLING TO COVER EXPOSED FILL AREAS AND TO ESTABLISH AN ACCEPTABLE SLOPE IN THE EXCAVATED AREA OF THE SITE FOR EXTENSION OF THE CAP;

* EXTRACTION OF CONTAMINATED GROUNDWATER FROM THE PERCHED ZONE AND THE GLACIAL AQUIFER AND TREATMENT OF THE GROUNDWATER BY AIR STRIPPING, GRANULAR ACTIVATED CARBON, AND ALUMINA REACTION ON-SITE AND TREATMENT OF THE OFF GASES;

* SITE DEED RESTRICTIONS TO LIMIT DEVELOPMENT AND LAND USE AND TO PREVENT INSTALLATION OF DRINKING WATER WELLS OR OTHER INTRUSIVE ACTIVITY AT THE SITE; AND

* GROUNDWATER MONITORING TO ASSESS THE STATE OF THE REMEDIATION.

* A SLURRY WALL WILL BE INSTALLED TO FACILITATE THE DEWATERING OF THE PERCHED ZONE AQUIFER.

APPENDIX A

EXHIBIT 1

DERIVING AN UNDISCOUNTED ESTIMATE OF REMEDIATION COSTS USING ROD DATA

		Amount	Source
(1)	Total Present Value	30,720,300	RoD
(2)	Total Capital Costs	11,083,300	RoD
(3)	Implied Present Value of O&M Costs	19,637,000	(1)–(2)
(4)	Annual O&M Cost	1,277,400	RoD
(5)	Years of O&M Cost	30	RoD
(6)	Assumed Discount Rate	5%	Assumption based on RoD date and Exhibit 4
(7)	Calculated O&M Present Value	19,636,769	Present value of (4) per year for (5) years discounted at a rate of (6).[32]
(8)	Assumed Inflation Rate	3.0%	Selected by actuary
(9)	Initial Estimate of Undiscounted O&M Cost	60,772,836	Future value of (4) per year for (5) years compounded at a rate of (8).[33]
(10)	Assumed Delay from RoD Issuance to Start of Cleanup Effort (in Years)	1.5	Selected by actuary[34]
(11)	Lag-Adjusted Capital Costs	11,585,771	(2)∗[1+(8)]^(10)
(12)	Assumed Duration of Construction Effort (in Years)	2.0	Selected by actuary[34]
(13)	Lag and Duration-Adjusted Capital Costs	11,759,557	(11)/2+(11)/2∗[1+(8)][35]
(14)	Assumed Delay from Construction Completion to O&M Start-up (in Years)	1.0	Selected by actuary[34]
(15)	Lag-Adjusted O&M Costs	69,418,786	(9)∗[1+(8)]^[(10)+(12)+(14)]
(16)	Total Estimated Undiscounted RoD Cost	81,178,343	(13)+(15)
(17)	Ratio: Total Estimated Undiscounted RoD Cost to Present Value Estimate in RoD	264%	(16)/(1)

[32] Note the similarity to the implied O&M cost from item (3).

[33] This is approximately 60% greater than the value derived by simply multiplying the annual O&M cost by the number of years applicable.

[34] Information useful to help estimate this lag may be obtained from CERCLIS.

[35] Assumes that, inflation aside, the two year total cost in current dollars is equally allocable between the two years of construction activities. It may be more realistic to assume a larger percentage applies to the first year, during which some large initial costs are incurred (i.e., equipment), as opposed to subsequent years, which may require predominantly materials and labor.

APPENDIX B

DIGGING UP MORE DIRT: AN APPROACH TO ESTIMATING FUTURE PRP COUNTS ON CURRENT SUPERFUND SITES

This appendix documents the approach outlined in the accompanying exhibits. Note that although this data has received a limited "scrubbing," due to various data quality issues outside the scope of this paper, *the reader should not rely on its quality or accuracy for use in analyses.* One adjustment made to the data is the removal of those PRPs that may relate to sites that are either still under review (i.e., they may eventually, but have not yet become Superfund sites) or sites that have been removed from CERCLIS and placed on NFRAP (i.e., they are expected to receive no further attention from the EPA). In addition, exact duplicate PRP entries at a given site were also removed, though in some cases, due to differences in the name for that PRP (e.g., General Electric Co. vs. GE), they may remain in the data.

Exhibit 1 of Appendix B displays PRP counts by year of NPL site listing and PRP notification, based on CERCLIS and PRP data at year-end 1995. The reader can see that, for sites listed on the NPL in 1983, 1,632 PRPs received notification of their potential liability at that site in 1982. In addition, 2,096 more PRPs received notification of their potential liability in 1983 on these sites.

Exhibit 2 restates the information on Exhibit 1 in "parallelogram" format. The column headings now reflect the difference in time between a PRP's notification of potential liability at a site and that site's placement on the NPL. On Page 2 of Exhibit 2, we can see that, for sites listed on the NPL in 1983, there were 1,632 PRPs notified of their potential liability at those sites one year earlier (in 1982). Another 2,096 PRPs were notified of their potential liability at sites listed in 1983 during 1983, and yet another 1,097 PRPs were notified of their potential links to sites listed in 1983 one year after those sites were listed (in 1984).

Exhibit 3 restates the incremental information in Exhibit 2 on a cumulative basis. Continuing our example, Page 2 of Exhibit 3 shows us that 1,742 PRPs received notice of potential liability at NPL sites listed in 1983 by the end of the year before those sites were listed (1982), and 3,838 PRPs were notified of their potential liability at those sites by the end of the year those sites were listed (1983). At the end of the year after these sites were listed (1984), 4,935 PRPs had been notified of potential links to those sites.

Exhibit 4 is simply "parallelogram" age-to-age factors, based on Exhibit 3. Page 2 shows us a development factor indicating that, for NPL sites listed in 1983, the growth in the number of PRPs notified of their potential liability at those sites between one and two years after those sites were listed is 33.6% $(6,592/4,935 = 1.336)$. Pages 2 and 3 also include the selection of age-to-age factors, as shown below the diagonal line. (It is worth repeating here that the development factors selections included here are for explanatory purposes only, and should not be relied on as "industry PRP development factors." Many additional adjustments to the PRP data should be made prior to evaluating the factors for that purpose.)

Exhibit 5 displays the age-to-ultimate factors corresponding to the age-to-age factors in Exhibit 4. Using our example, the selected factors imply a belief that, for sites listed on the NPL in 1983, no additional PRP notifications will be sent out (i.e., the age-to-ultimate development factor is 1.000). For sites listed in 1995, however, the expected number of PRPs yet to be notified of their links to these sites is expected to be 63.2% of the number of PRPs already linked to those sites (since the age-to-ultimate factor selected is 1.632). The author stresses again that the tail factor of 1.000 is displayed here for explanatory purposes only. It may be too early to truly expect no additional PRP development. Considerations and approaches which may be used to estimate PRP development tail factors may be a worthwhile area of future research.

Exhibit 6 summarizes our results and completes this explanation. The exhibit implies that, under the assumptions used here, 91.1% of PRPs have already been notified of their potential liability at current Superfund sites by year-end 1995. As a result, an estimate of the total number of claims relating to Superfund sites listed on the NPL as of year-end 1995 might be estimated by multiplying the current claim count on current Superfund sites by 1.10 (= 1/91.1%), further adjusted as necessary for any applicable collateral suit defendant and claim report lags. Then, subtracting the number of claims reported to date from the total number of expected claims yields an estimate of the number of future claims on current sites.

APPENDIX B

EXHIBIT 1

PAGE 1

PRP DATA AT YEAR-END 1995 (QUASI-SCRUBBED)
RAW DATA FORMAT: INCREMENTAL COUNTS

Year Listed	Year PRP Received Notice of Potential Liability at Site							
on NPL	1980	1981	1982	1983	1984	1985	1986	1987
1982	0	0	0	0	0	0	0	0
1983	0	110	1,632	2,096	1,097	1,657	1,690	306
1984	0	0	19	123	29	497	32	48
1985	0	0	0	0	0	0	0	0
1986	0	0	0	2	59	147	289	587
1987	0	0	0	5	4	0	99	88
1988	0	0	0	0	0	0	0	0
1989	0	0	0	18	25	27	71	551
1990	0	0	0	0	0	2	8	61
1991	0	0	0	0	0	1	3	0
1992	0	0	0	0	0	0	0	0
1993	0	0	0	0	0	0	0	0
1994	0	0	0	0	0	0	0	8
1995	0	0	0	0	0	0	0	0

APPENDIX B

Exhibit 1

Page 2

PRP Data At Year-End 1995 (Quasi-Scrubbed) Raw Data Format: Incremental Counts

Year Listed	Year PRP Received Notice of Potential Liability at Site							
on NPL	1988	1989	1990	1991	1992	1993	1994	1995
1982	0	0	0	0	0	0	0	0
1983	673	854	1,306	430	1,916	110	98	3
1984	105	79	919	203	31	298	332	1
1985	0	0	0	0	0	0	0	0
1986	315	398	580	490	778	393	9	0
1987	18	90	132	327	243	32	6	0
1988	0	0	0	0	0	0	0	0
1989	131	1,100	385	493	278	108	470	2
1990	383	300	301	238	190	28	31	2
1991	6	19	0	10	5	0	0	0
1992	1	3	1	0	558	17	3	1
1993	0	0	0	0	0	0	0	0
1994	4	646	1	98	73	57	10	0
1995	0	0	0	0	11	2	14	7

APPENDIX B

EXHIBIT 2

PAGE 1

PRP DATA AT YEAR-END 1995 (QUASI-SCRUBBED) ''PARALLELOGRAM'' DATA FORMAT: INCREMENTAL COUNTS

Year Listed	Year PRP Received Notice of Potential Liability at Site, Relative to NPL Listing Year							
on NPL	−12	−11	−10	−9	−8	−7	−6	−5
1982								
1983								
1984								
1985								
1986								
1987								
1988								
1989							18	25
1990								2
1991							1	3
1992								
1993								
1994						8	4	646
1995								

APPENDIX B

EXHIBIT 2

PAGE 2

PRP DATA AT YEAR-END 1995 (QUASI-SCRUBBED) "PARALLELOGRAM" DATA FORMAT: INCREMENTAL COUNTS

Year Listed	Year PRP Received Notice of Potential Liability at Site, Relative to NPL Listing Year								
on NPL	−4	−3	−2	−1	0	1	2	3	4
1982									
1983			110	1,632	2,096	1,097	1,657	1,690	306
1984			19	123	29	497	32	48	105
1985									
1986		2	59	147	289	587	315	398	580
1987	5	4	0	99	88	18	90	132	327
1988									
1989	27	71	551	131	1,100	385	493	278	108
1990	8	61	383	300	301	238	190	28	31
1991	0	6	19	0	10	5	0	0	0
1992	1	3	1	0	558	17	3	1	
1993									
1994	1	98	73	57	10	0			
1995		11	2	14	7				

APPENDIX B

Exhibit 2

Page 3

PRP Data At Year-End 1995 (Quasi-Scrubbed) ''Parallelogram'' Data Format: Incremental Counts

Year Listed on NPL	Year PRP Received Notice of Potential Liability at Site, Relative to NPL Listing Year: 5	6	7	8	9	10	11	12
1982								
1983	673	854	1,306	430	1,916	110	98	3
1984	79	919	203	31	298	332	1	
1985								
1986	490	778	393	9	0			
1987	243	32	6	0				
1988								
1989	470	2						
1990	2							
1991								
1992								
1993								
1994								
1995								

APPENDIX B

EXHIBIT 3

PAGE 1

PRP DATA AT YEAR-END 1995 (QUASI-SCRUBBED) "PARALLELOGRAM" DATA FORMAT: CUMULATIVE COUNTS

Year Listed	Year PRP Received Notice of Potential Liability at Site, Relative to NPL Listing Year							
on NPL	−12	−11	−10	−9	−8	−7	−6	−5
1982								
1983								
1984								
1985								
1986								
1987								
1988								
1989							18	43
1990								2
1991							1	4
1992								
1993								
1994						8	12	658
1995								

APPENDIX B

EXHIBIT 3

PAGE 2

PRP DATA AT YEAR-END 1995 (QUASI-SCRUBBED) ''PARALLELOGRAM'' DATA FORMAT: CUMULATIVE COUNTS

Year Listed	Year PRP Received Notice of Potential Liability at Site, Relative to NPL Listing Year								
on NPL	–4	–3	–2	–1	0	1	2	3	4
1982									
1983			110	1,742	3,838	4,935	6,592	8,282	8,588
1984			19	142	171	668	700	748	853
1985									
1986		2	61	208	497	1,084	1,399	1,797	2,377
1987	5	9	9	108	196	214	304	436	763
1988									
1989	70	141	692	823	1,923	2,308	2,801	3,079	3,187
1990	10	71	454	754	1,055	1,293	1,483	1,511	1,542
1991	4	10	29	29	39	44	44	44	44
1992	1	4	5	5	563	580	583	584	
1993									
1994	659	757	830	887	897	897			
1995		11	13	27	34				

APPENDIX B

Exhibit 3

Page 3

PRP Data At Year-End 1995 (Quasi-Scrubbed) ''Parallelogram'' Data Format: Cumulative Counts

Year Listed on NPL	Year PRP Received Notice of Potential Liability at Site, Relative to NPL Listing Year: 5	6	7	8	9	10	11	12
1982								
1983	9,261	10,115	11,421	11,851	13,767	13,877	13,975	13,978
1984	932	1,851	2,054	2,085	2,383	2,715	2,716	
1985								
1986	2,867	3,645	4,038	4,047	4,047			
1987	1,006	1,038	1,044	1,044				
1988								
1989	3,657	3,659						
1990	1,544							
1991								
1992								
1993								
1994								
1995								

APPENDIX B

Exhibit 4

Page 1

PRP Data at Year-End 1995 (Quasi-Scrubbed)
''Parallelogram'' Data Format: Age-to-Age Development Factors

Year Listed	Development Based on PRP Notification Year, Relative to NPL Listing Year							
on NPL	−12 to −11	−11 to −10	−10 to −9	−9 to −8	−8 to −7	−7 to −6	−6 to −5	−5 to −4
1982								
1983								
1984								
1985								
1986								
1987								
1988								
1989							2.389	1.628
1990								5.000
1991							4.000	1.000
1992								
1993								
1994						1.500	54.833	1.002
1995								

APPENDIX B

EXHIBIT 4

PAGE 2

PRP DATA AT YEAR-END 1995 (QUASI-SCRUBBED)
"PARALLELOGRAM" DATA FORMAT: AGE-TO-AGE DEVELOPMENT FACTORS

Year Listed	Development Based on PRP Notification Year, Relative to NPL Listing Year								
on NPL	−4 to −3	−3 to −2	−2 to −1	−1 to 0	0 to 1	1 to 2	2 to 3	3 to 4	4 to 5
1982									
1983					1.286	1.336	1.256	1.037	1.078
1984				1.204	3.906	1.048	1.069	1.140	1.093
1985									
1986		30.500	3.410	2.389	2.181	1.291	1.284	1.323	1.206
1987	1.800	1.000	12.000	1.815	1.092	1.421	1.434	1.750	1.318
1988									
1989	2.014	4.908	1.189	2.337	1.200	1.214	1.099	1.035	1.147
1990	7.100	6.394	1.661	1.399	1.226	1.147	1.019	1.021	1.001
1991	2.500	2.900	1.000	1.345	1.128	1.000	1.000	1.000	1.050
1992	4.000	1.250	1.000	112.600	1.030	1.005	1.002	1.014	1.050
1993							1.007	1.014	1.050
1994	1.149	1.096	1.069	1.011	1.000	1.051	1.007	1.014	1.050
1995		1.182	2.077	1.259	1.053	1.051	1.007	1.014	1.050

APPENDIX B

EXHIBIT 4

PAGE 3

PRP DATA AT YEAR-END 1995 (QUASI-SCRUBBED)
''PARALLELOGRAM'' DATA FORMAT: AGE-TO-AGE DEVELOPMENT FACTORS

Year Listed	Development Based on PRP Notification Year, Relative to NPL Listing Year							
on NPL	5 to 6	6 to 7	7 to 8	8 to 9	9 to 10	10 to 11	11 to 12	Tail Factor
1982								
1983	1.092	1.129	1.038	1.162	1.008	1.007	1.000	1.000
1984	1.986	1.110	1.015	1.143	1.139	1.000	1.000	1.000
1985						1.003	1.000	1.000
1986	1.271	1.108	1.002	1.000	1.074	1.003	1.000	1.000
1987	1.032	1.006	1.000	1.076	1.074	1.003	1.000	1.000
1988			1.006	1.076	1.074	1.003	1.000	1.000
1989	1.001	1.074	1.006	1.076	1.074	1.003	1.000	1.000
1990	1.099	1.074	1.006	1.076	1.074	1.003	1.000	1.000
1991	1.099	1.074	1.006	1.076	1.074	1.003	1.000	1.000
1992	1.099	1.074	1.006	1.076	1.074	1.003	1.000	1.000
1993	1.099	1.074	1.006	1.076	1.074	1.003	1.000	1.000
1994	1.099	1.074	1.006	1.076	1.074	1.003	1.000	1.000
1995	1.099	1.074	1.006	1.076	1.074	1.003	1.000	1.000

APPENDIX B

Exhibit 5

Page 1

PRP Data at Year-End 1995 (Quasi-Scrubbed)
''Parallelogram'' Data Format: Age-to-Ultimate Development Factors

Year Listed	Development Based on PRP Notification Year, Relative to NPL Listing Year							
on NPL	−12 to ult	−11 to ult	−10 to ult	−9 to ult	−8 to ult	−7 to ult	−6 to ult	−5 to ult
1980								
1981								
1982								
1983								
1984								
1985								
1986								
1987								
1988								
1989							254.521	106.544
1990								1,062.289
1991							63.574	15.893
1992								
1993								
1994						173.769	115.846	2.113
1995								

APPENDIX B

EXHIBIT 5

PAGE 2

PRP DATA AT YEAR-END 1995 (QUASI-SCRUBBED)
''PARALLELOGRAM'' DATA FORMAT: AGE-TO-ULTIMATE DEVELOPMENT FACTORS

Year Listed	Development Based on PRP Notification Year, Relative to NPL Listing Year								
on NPL	−4 to ult	−3 to ult	−2 to ult	−1 to ult	0 to ult	1 to ult	2 to ult	3 to ult	4 to ult
1980									
1981									
1982									
1983									
1984				19.131	15.886	4.067	3.881	3.632	3.185
1985									
1986		2,178.656	71.431	20.949	8.767	4.020	3.115	2.425	1.833
1987	241.929	134.405	134.405	11.200	6.172	5.653	3.979	2.774	1.585
1988									
1989	65.448	32.492	6.620	5.567	2.382	1.985	1.636	1.488	1.438
1990	212.458	29.924	4.680	2.818	2.014	1.643	1.433	1.406	1.378
1991	15.893	6.357	2.192	2.192	1.630	1.445	1.445	1.445	1.445
1992	855.522	213.880	171.104	171.104	1.520	1.475	1.467	1.465	1.445
1993							1.475	1.465	1.445
1994	2.109	1.836	1.675	1.567	1.550	1.550	1.475	1.465	1.445
1995		5.043	4.267	2.055	1.632	1.550	1.475	1.465	1.445

APPENDIX B

Exhibit 5

Page 3

PRP Data at Year-End 1995 (Quasi-Scrubbed)
‘‘Parallelogram’’ Data Format: Age-to-Ultimate Development Factors

Year Listed on NPL	Development Based on Year Listed on NPL Relative to PRP Notification Year							
	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult	11 to ult	12–ult
1980								
1981								
1982								
1983								1.000
1984	2.915	1.468	1.323	1.303	1.140	1.001	1.000	1.000
1985						1.003	1.000	1.000
1986	1.520	1.195	1.079	1.077	1.077	1.003	1.000	1.000
1987	1.202	1.165	1.159	1.159	1.077	1.003	1.000	1.000
1988			1.165	1.159	1.077	1.003	1.000	1.000
1989	1.253	1.252	1.165	1.159	1.077	1.003	1.000	1.000
1990	1.376	1.252	1.165	1.159	1.077	1.003	1.000	1.000
1991	1.376	1.252	1.165	1.159	1.077	1.003	1.000	1.000
1992	1.376	1.252	1.165	1.159	1.077	1.003	1.000	1.000
1993	1.376	1.252	1.165	1.159	1.077	1.003	1.000	1.000
1994	1.376	1.252	1.165	1.159	1.077	1.003	1.000	1.000
1995	1.376	1.252	1.165	1.159	1.077	1.003	1.000	1.000

APPENDIX B

EXHIBIT 6

DATA AT YEAR-END 1995 (QUASI-SCRUBBED)
DEVELOPMENT ANALYSIS SUMMARY

Year Listed on NPL	(1) Selected PRP Dvlpmnt Factor	(2) 1/(1) Probability: Current PRP on Current NPL Site	(3) 1–(2) Probability: Future PRP on Current NPL Site	(4) Count of Current PRPs on Current NPL Sites	(5) (1)*(4) Estimate of Ultimate PRPs on Current NPL Sites
1983	1.000	100.0%	0.0%	13,978	13,978
1984	1.000	100.0%	0.0%	2,716	2,717
1985	1.003	99.7%	0.3%	0	0
1986	1.077	92.9%	7.1%	4,047	4,357
1987	1.159	86.3%	13.7%	1,044	1,210
1988	1.165	85.8%	14.2%	0	0
1989	1.252	79.9%	20.1%	3,659	4,581
1990	1.376	72.7%	27.3%	1,544	2,125
1991	1.445	69.2%	30.8%	44	64
1992	1.465	68.3%	31.7%	584	856
1993	1.475	67.8%	32.2%	0	0
1994	1.550	64.5%	35.5%	897	1,390
1995	1.632	61.3%	38.7%	34	55
				28,547	31,332

Estimated Probability of Current PRP on Current Site:
Total(4)/Total(5) = 91.1%

Estimated Probability of Future PRP on Current Site:
[Total(5) − Total(4)]/Total(5) = 8.9%

Estimated PRP Development, All Years Combined:
Total(5)/Total(4) = 1.10

APPENDIX C

COMING CLEAN: THE RELATIONSHIP BETWEEN HAZARD, TIME AND COST

Similar to the note preceding the main text, the author would like to emphasize that the opinions expressed in this Appendix represent the views of the author, and do not necessarily represent the views of the Casualty Actuarial Society, Ernst & Young LLP, or anyone else.

Many have stipulated a relationship among these three quantities, based on the following argument:

- The Superfund was created to address the country's super-hazardous inactive waste sites; as a result, the most hazardous of Superfund sites would have been those first put on the national priorities list (NPL).
- These super-hazardous sites will also tend to be the largest, most complex sites, making them also the most costly.
- If the earliest, most hazardous sites tend to be the most costly, it follows that the later sites, which should be less hazardous, would be less costly.

A test of this hypothesis is displayed in Exhibits 1 and 2 of Appendix C, which test the specific relationship between the year a site was listed on the NPL and the site's Public Health Hazard Category (PHH) by the Agency of Toxic Substances and Disease Registry (ATSDR). These exhibits imply that the average site posted to the NPL in the most recent years is, if anything, more hazardous than the average site posted to the NPL in the program's earliest years.

Before discussing the possible reasons behind this, a few notes about the exhibits are in order. The ATSDR ranking was used in lieu of the Environmental Protection Agency's (EPA's) hazard ranking system (HRS) score for at least five reasons:

1. The EPA only uses the HRS score to separate potential NPL sites from non-NPL sites; it is not the primary tool used to subsequently prioritize which NPL sites are the most hazardous and require the earliest attention. Thus, the EPA itself does not consider the HRS sufficient for differentiating the degree of differences in hazard among NPL sites. The PHH, however, is designed to differentiate hazard levels at any location (NPL or otherwise).

2. As noted in the main text, since the HRS score only needs to reach a value of 28.5 for possible proposal to the NPL, once sufficient exposure pathways have been scored to achieve this, the remainder might not be scored at all, further diminishing the usefulness of the HRS score as a measure of each NPL site's relative hazard level. Again, this shortcut would not present a problem for the EPA's prioritizing of Superfund sites, since the HRS score is not the primary tool used for that purpose.

3. Part of the HRS scoring approach considers the size of the population near the site being scored. As a result, two sites with identical problems and required remedies may have different HRS scores. This does not imply that such differentiation is improper; only that the EPA's HRS score is really a measure of both hazard and the extent of population exposure to that hazard. The PHH, by contrast, does not consider the extent of population exposure, only whether or not there is *any* potential population exposure.

4. While the potential for future spreading of current contamination at a site is clearly considered by both the HRS and the PHH, the HRS score may be more conservative in that the PHH tries to consider the "likely" future spread of contamination, while the EPA's HRS score has historically considered a broader definition.

This is analogous to estimating "likely" vs. "conservative" IBNR amounts.

5. The HRS was updated in December of 1990, which might limit its usefulness as a consistent estimator of hazard over time. In contrast, the PHHs have been relatively consistent since inception.

Despite the above, however, there are some drawbacks to using the ATSDR data as well, including the following:

1. There are seven PHH categories in the ATSDR scoring system: 1 (urgent public health hazard), 2 (public health hazard), 3 (indeterminate public health hazard), 4 (no apparent public health hazard), 5 (no public health hazard), 6 (no hazard conclusion required) and 12 (posed public health hazard only in the past). Since the rankings of the ATSDR are not actually relative (e.g., a ranking of a 5 is not one-fifth as hazardous as a ranking of 1), the average PHH category for a given site listing year is not meaningful. As a result, the median value was used here, as displayed in Exhibit 1 of Appendix C. The percentage of sites posted to the NPL in each year that represent public hazards as evaluated by the ATSDR is also displayed, in Exhibit 2 of Appendix C.

2. There has been a preponderance of sites with a PHH of 3 (indeterminate hazard), largely because the ATSDR felt that the necessary data to reasonably evaluate the "likely" hazard level at many sites was not available. This analysis focused on differentiating the higher hazard levels (PHH categories 1 and 2) from the lower hazard levels (PHH categories 4 and 5) by excluding sites with a PHH of 3 from the review (Scenario 1 of Appendix C, Exhibits 1 and 2). For sensitivity testing purposes, Scenario 2 in these exhibits includes sites with a PHH of 3, and scenarios 3–8 display the impact that these PHH Level 3 sites

would have had on Scenario 1 if they could all have been allocated among the higher and lower hazard levels (1, 2, 4 and 5). For example, Scenario 3 assumes that 25% of the sites with a PHH of 3 are really higher hazard level sites (i.e., would have been a 1 or 2 if sufficient data were available), and 75% are really lower hazard level sites (i.e., would have been a 4 or 5). Scenario 8 assumes all of these sites would have been categorized as higher hazard level sites, and scenarios 4–7 run other scenarios between those two extremes. The author believes that Scenario 4, displaying a 60%/40% split between low and high hazard levels, respectively, is the most likely. This is because, consistent with a conservative tendency stemming from the EPA's need to protect human health, the last thing an EPA site evaluator would want to do is to remove a site from consideration for the NPL, only to later find out that the site was, in fact, Superfund-worthy. As a result, sites with an indeterminate hazard, though plausibly hazardous, are likely not.[36]

3. Some sites have been categorized and recategorized, though only one category should be used per site for this type of analysis. The selected category used here for a given site was determined by first removing all PHHs of 6 and 12 from the data. Then, the site's ranking was selected as either (1) the most recent PHH determined, if no remedial actions (RAs) have begun at that site yet, or (2) the most recent PHH determined prior to the onset

[36]As possible support for (though far from proof of) this, the author reviewed the 109 non-Federal, non-RCRA sites deleted from the NPL which have received PHHs as outlined earlier in this section. Of the 35 sites with a 4 or 5 PHH categorization (likely not hazardous), 80% were deleted with no need for remedial actions (RAs). In contrast, only three of the seven sites with a PHH of 1 or 2 (i.e., 43% of the likely hazardous sites) were deleted with no RAs required. Of the 67 deleted sites with a PHH of 3 (indeterminate hazard), 50 of them (75%) were deleted with no RAs required—which is much closer to 80% (PHHs 4 and 5) than 43% (PHHs 1 and 2). If we can assume that in general, the more hazardous NPL sites tended to require RAs, then the hazard level of sites with a PHH of 3 is more similar on average to the hazard level of sites with a PHH of 4 or 5 than to sites with a PHH of 1 or 2.

of RA activities which have begun at that site (since any cleanup efforts underway hopefully reduce the hazard level at a site by the time the ATSDR begins its review there). Sites with a PHH of 3 were then pulled out of the data for Scenario 1, included in the data for Scenario 2, and redistributed to the other four categories for Scenarios 3–8, as described in the previous item. Sites with no PHHs at all (there were 21 of these), or PHHs completed only after the onset of RA activities (there were 90 of these) were excluded altogether.

Despite these adjustments, however, Exhibit 1 of Appendix C implies that the recent years' median site hazard levels may be greater than those in the earliest years—or, at the very least, not any less hazardous than those in the earliest years. Exhibit 2 of Appendix C also shows a generally greater percentage of higher hazard level sites in the more recent years than in the early years of the program. The data underlying these exhibits is also included, in Exhibit 3 of Appendix C.

The Fallacy of (De)composition: Possible Explanations for the Apparent Non-decreasing Average Hazard over Time

One possible explanation for this somewhat unlikely result is that, although some ultra-hazardous sites were posted to the NPL early in the Superfund program, that doesn't necessarily mean that all sites posted to the NPL early in the Superfund program were ultra-hazardous. There is some intuitive appeal to this idea as well—it is generally accepted that there were approximately 10–20 "megasites" (i.e., sites which are extremely hazardous and costly) posted to the Superfund in the earliest years of the program. However, this is possibly 20 sites out of more than 400 posted to the Superfund in 1983 alone.

It is also possible that in the early years of the program, political pressure might have been exerted to include on the NPL some sites which would have been addressed through state

Superfund programs, if they existed at the time. With almost all states currently having some form of state Superfund program, these potentially less-hazardous sites might now be addressed as non-NPL sites, leaving only the more hazardous ones to be listed on the NPL currently and into the future. Ironically, political pressure is currently being applied in this, the opposite direction, with the states pressing for a more active role in the Superfund cleanup process.

A third possible explanation stems from the fact that, during the program's infancy, there must have been almost by definition a lack of experience in dealing with Superfund site cleanups. Guidance documents useful to assist in determining what is and is not Superfund-worthy take time and experience to develop—neither of which was likely present by 1983, the year the first 400 Superfund sites were listed. This lack of experience stemming from the newness of the program, in conjunction with a possible conservative desire of the EPA to address plausible (rather than just likely) future public health hazards may have led to some sites with undeterminable or even minimal hazard levels being placed on the NPL as a precautionary measure. However, fifteen or more years of experience with the Superfund program, coupled with the issuance and revisions of guidance documents, a revised HRS score and improved technology no doubt helped to decrease the percentage of sites listed on Superfund with an indeterminate hazard level (as shown in the last column in Exhibit 3 of Appendix C). These same factors may help explain the percentage decrease in sites listed with a PHH of 4 or 5 in the more recent years.

In summary then, the author believes that the average hazard level of Superfund sites has actually *increased* over time, rather than decreased, due to the fact that the sites presenting lower level hazards—which may have been included on the NPL in the past—are perhaps being more effectively screened out during the site review process now, leaving only the most hazardous of sites to be included on the NPL.

"Four Score" and Seven Years Ago: Why the Sudden Drop in NPL Site Listings and Lower Hazard Level Scores between 1990 and 1991?

It is noteworthy that in the most recent seven years, there has been a decrease in the average number of sites posted to the NPL per year, as well as a marked decrease in the percentage of those sites with a 4 or 5 PHH value. This is likely due to the revamping of the HRS score in December of 1990. It is also possible (though purely speculative) that this dramatic decrease in additional NPL postings is partially due to the EPA's desire to complete the cleanup process for those sites already in the Superfund pipeline before starting on new sites, rather than to take every site through the Superfund process simultaneously, one step at a time.[37] Adding more sites to the NPL might only increase the number of Superfund sites which will need to wait for attention, possibly reducing the desire to add sites currently to the NPL. As a result, as current cleanup efforts near completion (and many have been completed in the most recent 2–3 years), a significant increase in the number of sites being posted to the NPL annually may be possible in the near future, depending upon (among other things) the probability that a cap is placed on the number of sites permitted on the NPL (explicitly or implicitly).

Breaking New Ground: A New Theory on the (Non-) Relationship of Hazard and Cost

So what does this imply about the hazard/cost relationship? If it exists, it may imply that current Superfund sites could end up on average more costly than those listed in the earlier years. However, this potential cost increase would be offset by the EPA's recent initiatives discussed in the paper, improved tech-

[37] This actually presents a catch-22 situation. Under the first approach, some sites are cleaned, but many others are forced to wait until any actions can be taken. Under the second approach, all sites are addressed immediately (eliminating the problem using the first approach), but no cleanups would be completed (or perhaps even begun) for many years.

nology, and the experience gained with this type of remediation work over the past fifteen years, which may result in a current average site cost not very different from the average cost of previously listed sites.

The author believes that cost is more likely a function of the selected remedy than the indicated hazard. This is an important distinction, because although the remedy is somewhat dependent on the hazard, it is also dependent on the stringency of cleanup requirements in effect at the onset of remediation activities (i.e., the degree of the preference for treatment over containment) and technology available to implement the selected remedy at the time. This is one reason why it is important to consider records of decision (RoDs) for cost analysis purposes. Over the past couple of years, the EPA has been issuing many new RoDs which supplant remedies selected in the original RoDs for many of the sites posted to the NPL early in the Superfund program, based on new technologies and changes in cleanup requirements. Using this recent RoD information allows these aspects of cleanup costs to be effectively captured in actuarial analyses.

The hazard *is* an important consideration—especially for those sites involving groundwater issues—but it is far from the only consideration. And, as indicated in the main text of the paper, the author also believes the party leading the effort (i.e., the PRP, EPA, or other governmental agency) may also be a significant factor.

APPENDIX C

EXHIBIT 1

An Analysis of the Relationship Between NPL Listing Date and Site Hazard Using The Agency of Toxic Substances and Disease Registry's Public Health Hazard (PHH) Rankings

NPL Listing Year	Scenario 1 Median PHH	Scenario 2 Median PHH	Scenario 3 Median PHH	Scenario 4 Median PHH	Scenario 5 Median PHH	Scenario 6 Median PHH	Scenario 7 Median PHH	Scenario 8 Median PHH
1983	2.0	3.0	4.0	4.0	2.0	2.0	2.0	2.0
1984	2.0	3.0	4.0	4.0	2.0	2.0	2.0	2.0
1985	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1986	2.0	3.0	4.0	4.0	2.0	2.0	2.0	2.0
1987	2.0	3.0	4.0	4.0	2.0	2.0	2.0	2.0
1988								
1989	2.5	3.0	4.0	4.0	2.0	2.0	2.0	2.0
1990	2.0	3.0	4.0	2.0	2.0	2.0	2.0	2.0
1991	2.0	2.5	2.0	2.0	2.0	2.0	2.0	2.0
1992	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
1993								
1994	2.0	3.0	3.0	2.0	2.0	2.0	2.0	2.0
1995	2.0	3.0	2.0	2.0	2.0	2.0	2.0	2.0
1996	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
	2.0	3.0	4.0	4.0	2.0	2.0	2.0	2.0

Public Health Hazard Category Code—

1 = Urgent Public Health Hazard
2 = Public Health Hazard
3 = Indeterminate Public Health Hazard
4 = No Apparent Public Health Hazard
5 = No Public Health Hazard

Scenario Descriptions:

1. Excludes sites with PHH of 3
2. Includes sites with PHH of 3

Scenarios 3–8 include PHHs of 3, reallocated to PHHs 2 and 4 in the following proportions:

3. 25% to Level 2/75% to Level 4
4. 40% to Level 2/60% to Level 4
5. 50% to Level 2/50% to Level 4
6. 60% to Level 2/40% to Level 4
7. 75% to Level 2/25% to Level 4
8. 100% to Level 2/0% to Level 4

APPENDIX C

EXHIBIT 2

An Analysis Of The Relationship Between NPL Listing Date And Site Hazard Using The Agency Of Toxic Substances And Disease Registry's Public Health Hazard (PHH) Rankings

Percentage of Sites with Detected Public Health Hazard Levels, by Scenario[38]

NPL Listing Year	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8
1983	52.7%	17.6%	34.3%	44.2%	50.9%	57.6%	67.5%	84.2%
1984	60.5%	19.5%	36.4%	46.6%	53.4%	60.2%	70.3%	87.3%
1985	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
1986	63.3%	23.8%	39.4%	48.8%	55.0%	61.3%	70.6%	86.3%
1987	63.6%	20.9%	37.7%	47.8%	54.5%	61.2%	71.3%	88.1%
1988								
1989	50.0%	15.6%	32.8%	43.1%	50.0%	56.9%	67.2%	84.4%
1990	67.3%	26.7%	41.8%	50.8%	56.9%	62.9%	71.9%	87.0%
1991	100.0%	50.0%	62.5%	70.0%	75.0%	80.0%	87.5%	100.0%
1992	88.9%	61.5%	69.2%	73.8%	76.9%	80.0%	84.6%	92.3%
1993								
1994	64.7%	35.5%	46.8%	53.5%	58.1%	62.6%	69.4%	80.6%
1995	80.0%	44.4%	55.6%	62.2%	66.7%	71.1%	77.8%	88.9%
1996	85.7%	66.7%	72.2%	75.6%	77.8%	80.0%	83.3%	88.9%
	61.1%	22.4%	38.2%	47.7%	54.1%	60.4%	69.9%	85.8%

Public Health Hazard Category Code—

1 = Urgent Public Health Hazard
2 = Public Health Hazard
3 = Indeterminate Public Health Hazard
4 = No Apparent Public Health Hazard
5 = No Public Health Hazard

[38] Calculated as the sum of sites with a PHH of either 1 or 2, divided by all sites included in that scenario. The scenario descriptions are as follows:

1. Excludes sites with PHH of 3
2. Includes sites with PHH of 3

Scenarios 3–8 include PHHs of 3, reallocated to PHHs 2 and 4 in the following proportions:

3. 25% to Level 2/75% to Level 4
4. 40% to Level 2/60% to Level 4
5. 50% to Level 2/50% to Level 4
6. 60% to Level 2/40% to Level 4
7. 75% to Level 2/25% to Level 4
8. 100% to Level 2/0% to Level 4

APPENDIX C

EXHIBIT 3

An Analysis of the Relationship Between NPL Listing Date and Site Hazard Using the Agency of Toxic Substances and Disease Registry's Public Health Hazard (PHH) Rankings Underlying Data

NPL Listing Year	Public Health Hazard (PHH) Category 1	2	3	4	5	Total	PHH 3, as Pct of Total
1983	6	53	223	38	15	335	66.6%
1984	0	23	80	11	4	118	67.8%
1985	2	1	0	0	0	3	0.0%
1986	1	37	100	11	11	160	62.5%
1987	0	14	45	4	4	67	67.2%
1988	0	0	0	0	0	0	
1989	3	24	119	23	4	173	68.8%
1990	0	35	79	14	3	131	60.3%
1991	1	2	3	0	0	6	50.0%
1992	2	14	8	2	0	26	30.8%
1993	0	0	0	0	0	0	
1994	1	10	14	4	2	31	45.2%
1995	1	3	4	1	0	9	44.4%
1996	0	6	2	1	0	9	22.2%
	17	222	677	109	43	1,068	63.4%

21	w/no PHHs 1–5 at site
90	w/PHH completed after onset of RA activities
1	Delisted, then relisted
1,180	Total on NPL

Public Health Hazard Category Code—

1 = Urgent Public Health Hazard
2 = Public Health Hazard
3 = Indeterminate Public Health Hazard
4 = No Apparent Public Health Hazard
5 = No Public Health Hazard
6 = No hazard conclusion (often applies to brief addenda)
12 = Posed Public Health Hazard Only in the Past

Each site may have multiple PHHs. The following approach was used to select one:

PHH values 6 and 12 were excluded from this analysis altogether (2 sites).
If no RAs have begun at that site by 12/31/96, the most recent PHH available was selected.
Otherwise, the most recent PHH prior to onset of RA activities at that site was selected.
21 sites were excluded due to lack of a PHH.
90 sites were excluded because the first PHH review was completed after the onset of RA activities there.

APPENDIX D

WASTE NOT, WANT NOT: REDUCING AND ELIMINATING HAZARDOUS WASTE THROUGH RCRA

Federal solid waste regulation began in 1965 with the Solid Waste Disposal Act, with an emphasis on research and development (R&D) of solid waste disposal practices. This act was amended in 1970 by the Resource Recovery Act, which changed the emphasis from R&D to recycling and waste reduction. The Resource Conservation and Recovery Act (RCRA) was enacted in 1976, and contained regulations on waste management and the prohibition of open dumps. It also required that anyone seeking to operate a hazardous waste Treatment, Storage and Disposal Facility (TSDF) must first receive a permit from the Environmental Protection Agency (EPA) to do so. The Hazardous and Solid Waste Amendments of 1984 significantly expanded the scope of RCRA, adding land disposal restrictions and corrective action requirements addressing the need to clean previous releases of hazardous waste prior to receiving a RCRA permit (under RCRA Subtitle C).

While the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) is overseen by the EPA, RCRA is predominantly state-run (though there are certain minimum Federal requirements). In addition, there is no RCRA-equivalent to CERCLA's Superfund, which the EPA can use to pay for site cleanups if there are no potentially responsible parties (PRPs). RCRA doesn't focus on the concept of PRPs (i.e., on a broad spectrum of possible sources for any necessary corrective action funding), but instead focuses its authority on the current owner/operator of the TSDF. As a result, the cost sharing typically found at National Priorities List (NPL) sites among their many PRPs might not be as prevalent under RCRA. Therefore, even though the average RCRA site cleanup cost is expected to be approximately $15 million [15]—which is less than the frequently-quoted estimates of the average NPL site cleanup cost—there may be a greater financial burden to the entity responsible for corrective action at a RCRA site than to the entity paying only a fraction of the cleanup cost at an NPL site.

Underground storage tanks (USTs) are typically addressed under RCRA, rather than Superfund. This is because most USTs are filled with petroleum, which is not one of the contaminants identified for response actions under the Superfund program.

Despite their differences, RCRA and CERCLA both share the common goal of protecting human health and the environment from adverse contact with hazardous waste. In general, CERCLA approaches this goal retroactively, by requiring clean up of *inactive* hazardous waste sites, while RCRA attempts to address the issue prospectively, through establishment of standards for *active* hazardous waste sites. RCRA standards require tracking hazardous waste from its creation to its ultimate disposition ("cradle-to-grave" monitoring).

CERCLA and RCRA also interact. For example, RCRA cleanup standards may be applied to Superfund cleanups, since CERCLA doesn't actually dictate specific cleanup standards. RCRA sites may become listed on the NPL if a facility requiring cleanup is owned by a bankrupt entity, or an entity who has shown an unwillingness to clean up a particular RCRA site. In this case, the site is eligible for Superfund moneys—and the possibility of response actions by other PRPs, if they can be found. Conversely, Superfund sites may be deferred to the RCRA program under certain circumstances as well, allowing the EPA to focus its efforts (and funding) on other, Superfund-worthy sites.

A recent General Accounting Office (GAO) Study [16] indicated that the cost of cleaning RCRA sites may be higher than it needs to be in several cases, because of three key RCRA requirements:

1. *Land Disposal Restrictions.* According to the GAO Study, the same stringent standards are frequently applied to

both high-risk and relatively low-risk waste targeted for land disposal.

2. *Minimum Technological Requirements.* The GAO study also notes that the same stringent technological requirements may apply to facilities that manage both high-risk waste and facilities managing low-risk waste.

3. *Permit Requirements.* From [16, pp. 8–9], "the administrative cost of obtaining a RCRA permit can range from $80,000 for an on-site treatment unit, such as a tank, to $400,000 for an on-site incinerator, and up to $1 million for a landfill, according to EPA's estimates. In addition to these costs, a party may incur other costs for tasks needed to obtain a permit, such as assessing a site's conditions in order to design a groundwater monitoring system or conducting emissions testing and trial burns from an incinerator. The time required to obtain a permit can also be extensive…getting a permit can take 7 to 9 months for a simple treatment unit, such as a tank, and an additional 5 to 6 years for a more complicated unit, such as a landfill."

The study also discusses how the EPA has attempted to address these issues, and the policy and regulatory alternatives available to entities responsible for RCRA cleanups. However, the report also notes that, both the EPA and GAO believe that "(comprehensive) reform, while necessary, may take some time to implement." [16, p. 18]

Finally, it is worth noting that, due to the significant differences between CERCLA and RCRA noted here, equally significant insurance coverage-related issued may apply. A discussion of these and other coverage-related issues represents yet another potentially fruitful area for additional research.

MODELING LOSSES WITH THE MIXED EXPONENTIAL DISTRIBUTION

CLIVE L. KEATINGE

Abstract

Finding a parametric model that fits loss data well is often difficult. This paper offers an alternative—the semiparametric mixed exponential distribution. The paper gives the reason why this is a good model and explains maximum likelihood estimation for the mixed exponential distribution. The paper also presents an algorithm to find parameter estimates and gives an illustrative example. The paper compares variances of estimates obtained with the mixed exponential distribution with variances obtained with a traditional parametric distribution. Finally, the paper discusses adjustments to the model and other uses of the model.

1. INTRODUCTION

Loss distributions have been a staple of actuarial work for many years. The Casualty Actuarial Society syllabus has included a separate section on the subject since 1985, the year after Hogg and Klugman [5] published *Loss Distributions*. This was the standard actuarial text on the subject until the recent book by Klugman, Panjer, and Willmot [8], *Loss Models: From Data to Decisions*, replaced it. Over the years, numerous authors have published papers dealing with loss distributions. The two books and most papers on the subject emphasize the use of parametric distributions as models for losses. I have found that the set of distributions generally suggested for use is not adequate. Too often, one cannot find a model that fits a data set well. Nonparametric procedures are available, but although they usually produce a good fit to the data, they often do not smooth the data enough.[1] In this paper, I offer an alternative—the semiparametric mixed exponential distribution.

Statisticians have done quite a bit of work with semiparametric mixture models. Lindsay and Lesperance [12] wrote in their 1995 review of semiparametric mixture models, "There has been a surge of interest in semiparametric mixture models in recent years, as statisticians strive to maintain the efficiencies of parametric methods while incorporating minimal assumptions in their models."[2] I will first explain why the mixed exponential distribution is a good model for losses. I will then discuss the theory underlying maximum likelihood estimation with the mixed exponential distribution. Much of this material has been developed in the statistics literature, but I will highlight the relevant parts of it. Next, I will present an algorithm based on Newton's method to find the maximum likelihood parameter estimates. I will follow with an example of the application of this algorithm to a data set from Klugman, Panjer, and Willmot [8] and with a comparison of the variances of estimates obtained from a mixed exponential distribution and a Pareto distribution, which serves as an example of a traditional parametric distribution. I will then address adjustments that may be necessary when using the mixed exponential distribution, with particular emphasis on how to handle the tail. Finally, I will briefly mention that the mixed exponential distribution is useful for more than just modeling losses.

I will not discuss how to account for loss development before fitting a distribution to a set of data. The actuarial literature has not adequately addressed this very important issue, but it is beyond the scope of this paper. Also, I will assume that all data analyzed has received appropriate trending.

---

[1] Although Klugman, Panjer, and Willmot [8] focus primarily on parametric procedures, they do briefly cover nonparametric procedures in Section 2.11.1.

[2] Lindsay [11] has also written a monograph summarizing much of the recent work in mixture models.

2. MOTIVATION

When working with a set of loss data, we usually want to estimate the underlying probability distribution that describes the process that generated the data. It is generally a plausible assumption that this distribution is reasonably smooth. Thus, smoothing out the data should give a better estimate than simply using the empirical distribution itself. To accomplish such smoothing, we may turn to either parametric or nonparametric procedures. However, a parametric procedure often produces a distribution that does not fit the data well, whereas a nonparametric procedure often produces a distribution that is not smooth enough. What we need is something in between a parametric and a nonparametric procedure—a procedure that will provide a distribution that fits the data well, yet still provides an appropriate amount of smoothing.

We can articulate the amount of smoothing we would like by specifying conditions that the derivatives of the survival function, $S(x)$, should satisfy (where x is the loss size).[3] First, note that $S'(x) = -f(x)$, where $f(x)$ is the probability density function. Clearly, $f(x)$ must not be negative, so we should require that $S'(x) \leq 0$. Next, we would like $f(x)$ to be decreasing, so we require that $S''(x) \geq 0$. Beyond that, we would like $f(x)$ to decrease at a decreasing rate, so we require that $S'''(x) \leq 0$. In general, we would like the derivatives of the survival function to change at a slower and slower rate as the loss size x gets larger and larger and to approach zero asymptotically as x approaches infinity.[4] The mathematical formulation of this requirement is that the survival function should possess derivatives of all orders

[3]The survival function equals one minus the cumulative distribution function. Working with the survival function is more convenient than working with the cumulative distribution function.

[4]These conditions are appropriate for most loss distributions encountered in practice, except perhaps where the loss size x is small. In particular, these conditions are not compatible with a probability density function with a nonzero mode. However, we are assuming that we are not particularly interested in the behavior of the survival function where x is small.

such that

$$(-1)^n S^{(n)}(x) \geq 0, \qquad x > 0.$$

Functions with this alternating derivative property are known as completely monotone functions. There is a beautiful theorem due to Bernstein (1928) which states that a function S on $[0,\infty]$ is completely monotone if and only if it is of the form

$$S(x) = \int_0^\infty e^{-\lambda x} w(\lambda)\, d\lambda,$$

where w is nonnegative. Since we are interested in cases where S is a survival function, we will restrict attention to cases where $S(0) = 1$. This requirement forces w to be a probability function (that may be discrete, continuous, or a combination of the two).[5] In other words, any distribution with the alternating derivative property must be a mixture of exponential distributions, and vice versa.[6]

From now on, I will use a discrete formulation of the mixing distribution w, because as will become clear, we usually deal with mixing distributions that are nonzero at a small number of points. Thus, we have

$$S(x) = \sum_{i=1}^{n} w_i e^{-\lambda_i x}, \qquad w_i > 0, \qquad \sum_{i=1}^{n} w_i = 1,$$

where w_i is the mixing weight corresponding to λ_i. Note that the mean of the ith component distribution of the mixture is $1/\lambda_i$.

One of the distinguishing characteristics of the mixed exponential distribution is that it always has a decreasing failure rate. The failure rate is the probability density function divided by the

[5] Another way of stating this is that S is completely monotone with $S(0) = 1$ if and only if it is the Laplace transform of a probability distribution w. See Feller [3, p. 439] for a proof.

[6] I would like to thank Glenn Meyers for pointing out this equivalence relation, with which he had become familiar through the work of Brockett and Golden [2]. They applied this relation to utility functions just as this paper applies it to loss distributions.

survival function.[7] For the mixed exponential distribution, the failure rate is

$$\sum_{i=1}^{n}\left(\frac{w_i e^{-\lambda_i x}}{\sum_{j=1}^{n} w_j e^{-\lambda_j x}}\right)\lambda_i.$$

This is a weighted average of the λ_i's. As x becomes larger, weight moves away from the larger λ_i's and toward the smaller λ_i's, thus decreasing the failure rate.

Most of the parametric distributions traditionally used to model losses have decreasing failure rates, either throughout the entire distribution or at all but small loss sizes. Some are special cases of the mixed exponential distribution. For example, the Pareto distribution is a mixture of exponential distributions with a gamma mixing distribution. See Appendix A for further discussion of this topic. The advantage that the mixed exponential distribution enjoys over parametric distributions is that the mixed exponential distribution is more general and thus likely to provide a better fit to the data while still providing an appropriate amount of smoothing. It is considered semiparametric because no parametric assumption is made about the form of the mixing distribution. We now turn to the problem of estimating the mixing distribution from a given set of data.

3. MAXIMUM LIKELIHOOD THEORY

Maximum likelihood estimation is the only estimation technique I will cover in this paper. Although other techniques are available, the well-known desirable statistical properties of maximum likelihood estimation usually make it the method of choice.

---

[7]See Section 2.7.2 of Klugman, Panjer, and Willmot [8] for a discussion of failure rates. The failure rate is also known as the hazard rate or the force of mortality. In the context of a loss distribution, "failure" means "loss stoppage." A distribution with a decreasing failure rate has an increasing mean residual lifetime (if it exists).

In this section, I will describe the properties underlying maximum likelihood estimation with the mixed exponential distribution. The proofs are in Appendix B.

I will begin by addressing the situation where no grouping, censoring, or truncation is present in the data. The loglikelihood function is

$$\ln L(w_1, w_2, \ldots) = \sum_{k=1}^{m} \ln f(x_k) = \sum_{k=1}^{m} \ln\left(\sum_{i=1}^{\infty} w_i \lambda_i e^{-\lambda_i x_k}\right),$$

where m is the number of observations. We must find the set of w_i's that maximizes the loglikelihood function, subject to the constraints that each of the w_i's must be greater than or equal to zero and the sum of the w_i's must be one. We consider the λ_i's fixed and arbitrarily close together.

This constrained maximum occurs at the unique point at which the following conditions, known as the Karush–Kuhn–Tucker (KKT) conditions, are satisfied:

$$\frac{\partial \ln L}{\partial w_i} = \sum_{k=1}^{m} \frac{\lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{\infty} w_j \lambda_j e^{-\lambda_j x_k}} \leq m, \qquad \text{if} \quad w_i = 0$$

and

$$\frac{\partial \ln L}{\partial w_i} = \sum_{k=1}^{m} \frac{\lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{\infty} w_j \lambda_j e^{-\lambda_j x_k}} = m, \qquad \text{if} \quad w_i > 0.$$

The inequality conditions ensure that we cannot increase the loglikelihood by moving a small amount of weight to a λ_i that has zero weight attached to it. The equality conditions ensure that we cannot increase the loglikelihood by moving weight around among the λ_i's that have positive weight attached to them. The number of positive w_i's at this maximum is at most m. None of the corresponding λ_i's can be less than $1/x_m$, where x_m is the largest observation, and none can be greater than $1/x_1$, where x_1

is the smallest observation. The number of positive w_i's tends to increase with the number of observations, but remains below ten in most practical situations.

For grouped data, the loglikelihood function is

$$\begin{aligned}\ln L(w_1, w_2, \ldots) &= a_1 \ln(1 - S(b_1)) + \sum_{k=2}^{g-1} a_k \ln(S(b_{k-1}) - S(b_k)) \\ &\quad + a_g \ln(S(b_{g-1})) \\ &= a_1 \ln\left(\sum_{i=1}^{\infty} w_i (1 - e^{-\lambda_i b_1})\right) \\ &\quad + \sum_{k=2}^{g-1} a_k \ln\left(\sum_{i=1}^{\infty} w_i (e^{-\lambda_i b_{k-1}} - e^{-\lambda_i b_k})\right) \\ &\quad + a_g \ln\left(\sum_{i=1}^{\infty} w_i (e^{-\lambda_i b_{g-1}})\right),\end{aligned}$$

where g is the number of groups, $a_1, \ldots, a_g$ are the number of observations in each group, and $b_1, \ldots, b_{g-1}$ are the group boundaries. We will assume that any adjacent groups that all have zero observations have been combined into one group.

In this case, the KKT conditions are

$$\begin{aligned}\frac{\partial \ln L}{\partial w_i} &= a_1 \frac{1 - e^{-\lambda_i b_1}}{\sum_{j=1}^{\infty} w_j (1 - e^{-\lambda_j b_1})} + \sum_{k=2}^{g-1} a_k \frac{e^{-\lambda_i b_{k-1}} - e^{-\lambda_i b_k}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})} \\ &\quad + a_g \frac{e^{-\lambda_i b_{g-1}}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{g-1}})} \leq m, \qquad \text{if} \quad w_i = 0\end{aligned}$$

and

$$\frac{\partial \ln L}{\partial w_i} = a_1 \frac{1 - e^{-\lambda_i b_1}}{\sum_{j=1}^{\infty} w_j (1 - e^{-\lambda_j b_1})} + \sum_{k=2}^{g-1} a_k \frac{e^{-\lambda_i b_{k-1}} - e^{-\lambda_i b_k}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})}$$

$$+ a_g \frac{e^{-\lambda_i b_{g-1}}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{g-1}})} = m, \qquad \text{if} \quad w_i > 0.$$

The constrained maximum will occur at a unique point, unless the mixed exponential probabilities for each group are exactly proportional to the number of observations in each group or, in other words, when the data perfectly fits the model. For this situation, we can easily come up with examples where an arbitrarily large number of different mixed exponential distributions, each with an arbitrarily large number of positive w_i's, will maximize the loglikelihood function. However, a perfect fit is highly unlikely unless the number of groups is very small.

When the fit is not perfect, the number of positive w_i's with corresponding λ_i's on $(0,\infty)$ at the maximum is at most $g/2 - 1$ if g is even and $g/2 - 1/2$ if g is odd. In addition to the λ_i's on $(0,\infty)$, there may also be λ_i's at zero or infinity (or both) that have positive w_i's. For an exponential distribution with a λ_i of zero (and thus a mean of infinity), the survival function is a constant function of 1. In actuarial terms, the w_i corresponding to a λ_i of zero would indicate the probability that a loss will completely exhaust all layers of coverage, no matter how high. For an exponential distribution with a λ_i of infinity (and thus a mean of zero), the survival function is a constant function of 0. The w_i corresponding to a λ_i of infinity would indicate the probability that a loss will be zero. The number of positive w_i's tends to increase with the number of groups, but remains below ten in most practical situations.

The development for grouped data applies also to censored grouped data, since the censored data is simply in the last group with an upper bound of infinity. For other situations, such as censored ungrouped data (thus partially grouped and partially ungrouped) or data censored at various points or grouped with various boundaries, the logic is similar to that used above, since we can simply sum the appropriate loglikelihood functions.

With ungrouped data truncated (but not shifted) by a deductible d, the loglikelihood function is

$$\ln L(w_1, w_2, \ldots) = \sum_{k=1}^{m} \ln\left(\frac{f(x_k)}{S(d)}\right) = \sum_{k=1}^{m} \ln\left(\frac{\sum_{i=1}^{\infty} w_i \lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{\infty} w_j e^{-\lambda_j d}}\right)$$

$$= \sum_{k=1}^{m} \ln\left(\sum_{i=1}^{\infty} w_i^* \lambda_i e^{-\lambda_i (x_k - d)}\right),$$

where

$$w_i^* = \frac{w_i e^{-\lambda_i d}}{\sum_{j=1}^{\infty} w_j e^{-\lambda_j d}}.$$

We can thus convert the problem to a problem without a deductible by subtracting d from each observation. We can then recover the w_i's using the formula

$$w_i = \frac{w_i^* e^{\lambda_i d}}{\sum_{j=1}^{\infty} w_j^* e^{\lambda_j d}}.$$

The same process applies for grouped data with d subtracted from each of the group boundaries instead of the observations. However, the formula to recover the w_i's breaks down if one of the λ_i's with a positive w_i^* is infinity, as quite often occurs with grouped data. Using the fitted mixed exponential distribution to extrapolate below a deductible is not a good idea.

If a set of data contains several different deductibles, we can subtract the smallest deductible for which a credible amount of data exists from each observation and the higher deductibles. We would have to subtract additional terms from the loglikelihood function to account for these higher deductibles.[8]

4. A MAXIMUM LIKELIHOOD ALGORITHM

I will now present an algorithm that we can use to find the maximum likelihood estimates of the parameters of a mixed exponential distribution. I have based the algorithm on Newton's method, the details of which are in any textbook on numerical analysis. After I present the algorithm, I will comment on alternatives to it. The steps of the algorithm are:

1. Begin with an initial set of positive w_i's and the λ_i's associated with them. The closer these are to the final estimated values, the faster the convergence will be. However, the algorithm will converge regardless of what the initial values are.

2. Assume that the number of parameters is fixed and use Newton's method to find the indicated change in the parameters. I will call this the Newton step. Each λ_i is a parameter, and all but one of the w_i's are parameters. We must set the remaining w_i equal to one minus the sum of the others. Appendix C shows the derivatives needed to find the Newton step.

3. Adjust the parameters by the amount of the Newton step. If all the λ_i's remain positive, if all the w_i's remain between zero and one, and if the loglikelihood function

[8] See Section 2.10 of Klugman, Panjer, and Willmot [8] for a discussion of estimation with incomplete data.

increases, then go to step 4. If the result does not satisfy all these conditions, then try a backward Newton step, then half a forward step, then half a backward step, then a quarter of a forward step, and so on until the result satisfies all the conditions.

4. If one of the λ_i's is approaching zero or infinity (which can happen only with grouped or censored data), go to step 5. If one of the w_i's is approaching zero, go to step 6. If the Newton step is very small, thus indicating convergence, go to step 7. Otherwise, go back to step 2.
5. If one of the λ_i's is approaching zero, then fix that λ_i at a very small value, so it is effectively zero. If one of the λ_i's is approaching infinity, then fix that λ_i at a very large value, so it is effectively infinity. Remove the fixed λ_i, but not its associated w_i, from the Newton iterative process. Go back to step 2.
6. If one of the w_i's is approaching zero, then adjust the parameters by the proportion of the Newton step that makes this w_i exactly zero. Remove it and its associated λ_i as parameters. Often, this λ_i will be approaching one of the other λ_i's. If the eliminated w_i was close enough to zero, its removal should result in an increase in the loglikelihood function. Go back to step 2.
7. If convergence has occurred, then check to see if the result satisfies the KKT conditions. To do this, check the conditions for λ_i's close enough together so that it is clear that if the result satisfies the conditions at the checked λ_i's, the result will also satisfy the conditions at all others in between. If the result satisfies the KKT conditions, then the loglikelihood function has reached its maximum. If the result does not satisfy the conditions, go to step 8.
8. If the result does not satisfy the KKT conditions, then add an additional λ_i and associated w_i as parameters. The new λ_i should be in the vicinity of where the KKT

function is the largest (and thus where a new λ_i is most needed). Give the new w_i a small positive value and proportionately decrease the other w_i's so the sum of the w_i's remains at 1. The value assigned to the new w_i should be small enough so that the loglikelihood function increases from its previous value. (The algorithm will work regardless of the values of the new λ_i and w_i as long as the loglikelihood function increases from its previous value. If it does not increase, the algorithm may lead right back to the point where it was before the new λ_i and w_i were added.) Go back to step 2.

This algorithm will always converge to the maximum likelihood estimates of the parameters, because the loglikelihood function is concave and its value is increasing with each step of the algorithm. The points where Newton's method converges but the result does not satisfy the KKT conditions correspond to local maxima with the number of λ_i's fixed at a specified number. When the result satisfies the KKT conditions, we have reached the global maximum, with no restriction on the number of λ_i's.

With ungrouped data, the fitted mixed exponential mean will always equal the sample mean. This applies at both the global maximum and local maxima with a fixed number of λ_i's. Also, with ungrouped data, the fitted mixed exponential variance will not be less than the sample variance. This applies only at the global maximum. Appendix C gives the proofs of these statements. With grouped data, these relationships cannot hold, because the values of the individual observations are not available.

The variance relationship for ungrouped data results from the smoothing effect of the mixed exponential distribution. Probability from the sample values is effectively spread to surrounding values where no data was observed, thus increasing the variance. Though this produces an upward bias in the variance of the fitted distribution, it reduces the variance of the estimates of the survival probabilities produced by the fitted distribution, as we will see in Section 6.

This variance relationship also holds for nonparametric smoothing procedures. For parametric distributions, the fitted variance can be either larger or smaller than the sample variance, depending on the particular sample. For both the mixed exponential distribution and parametric distributions, as long as the variance of the actual distribution is finite, the ratio of the fitted variance to the sample variance will approach 1 as the sample size goes to infinity, since both will converge to the actual variance of the distribution. If the variance of the actual distribution is infinite, this will be true for the distribution censored at any point.

The given algorithm is certainly not the only one that can be used to maximize the loglikelihood function. I presented it because Newton's method is well-known and it converges very fast once the parameters are in the vicinity of the solution. Step 3 of the algorithm, trying successively smaller forward and backward Newton steps until the loglikelihood increases, is not elegant, but it does work. One could certainly improve the efficiency of the algorithm, but with the ample computing power now available, any improvements would probably be of marginal benefit in most cases.[9]

One could use a "canned" optimization program (which may use Newton's method with approximations of the derivatives) to maximize the loglikelihood function. Such programs can work well, but one must take care to ensure that the program does not stop before reaching the solution. Also, since the λ_i's are generally of very different magnitudes, a scaling adjustment may be helpful.

5. AN EXAMPLE

I will now illustrate how the algorithm works. I will use some grouped general liability data taken from Table 2.27 of Klugman,

[9]Bohning [1] reviewed several maximum likelihood algorithms that have been proposed for use with semiparametric mixture models.

Panjer, and Willmot [8]. The first three columns of Table 1 show the data. The loss amounts shown are the group boundaries.

We begin by fixing the number of means at one (though we need not begin with one). Instead of referring to the λ_i's associated with a mixed exponential distribution, throughout this example I will refer to the means (the reciprocals of the λ_i's). Regardless of the initial value we select, we will obtain rapid convergence to a mean of 51,190. The second column of Table 2 shows this result. The third column shows the value of the KKT function

$$h(\lambda) = a_1 \frac{1 - e^{-\lambda b_1}}{\sum_{j=1}^{\infty} w_j (1 - e^{-\lambda_j b_1})} + \sum_{k=2}^{g-1} a_k \frac{e^{-\lambda b_{k-1}} - e^{-\lambda b_k}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})} + a_g \frac{e^{-\lambda b_{g-1}}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{g-1}})}$$

for a number of means. As it must, $h(\lambda)$ has a value of 336 (the number of observations) at 51,190, but the function is larger than this everywhere else. Thus, we have not reached the maximum.

Since $h(\lambda)$ is largest at large means, we move a small amount of weight to a large mean. The actual value of this mean or the amount of weight we place on it is not important as long as the loglikelihood increases. With two means, the algorithm converges to means of 13,570 and 176,638 with weights of 0.7566 and 0.2434, respectively. From Table 2, we see that we still have not reached the maximum.

Since $h(\lambda)$ is again largest at large means, we move a small amount of weight to a large mean and proportionately scale back the weights on the existing two means (checking to be sure that the loglikelihood increases). With three means, the algorithm converges to means of 10,598, 73,440, and 686,632 with weights

TABLE 1

COMPARISON OF FITTED DISTRIBUTIONS

Empirical			Mixed Exponential			Transformed Beta			Pareto			Lognormal		
				Mean	Weight									
				0	0.0526	θ		21,239						
				12,336	0.5999	α		0.9102						
				77,922	0.3102	γ		1.1998	θ		14,679	μ		9.4812
				712,302	0.0373	τ		0.6427	α		1.0758	σ		1.7162
				Loglikelihood	−818.26	Loglikelihood		−820.16	Loglikelihood		−820.78	Loglikelihood		−821.33
Loss Amt	No.> Loss Amt	Survival Prob	No.> Loss Amt	Survival Prob	Diff From Empirical	No.> Loss Amt	Survival Prob	Diff From Empirical	No.> Loss Amt	Survival Prob	Diff From Empirical	No.> Loss Amt	Survival Prob	Diff From Empirical
0	336	1.0000	336.00	1.0000	0.00%	336.00	1.0000	0.00%	336.00	1.0000	0.00%	336.00	1.0000	0.00%
2,500	278	0.8274	278.00	0.8274	0.00%	278.46	0.8288	0.17%	283.70	0.8443	2.05%	279.84	0.8329	0.66%
7,500	217	0.6458	216.79	0.6452	−0.10%	215.59	0.6416	−0.65%	215.53	0.6415	−0.68%	210.86	0.6276	−2.83%
12,500	180	0.5357	174.25	0.5186	−3.20%	175.36	0.5219	−2.58%	173.19	0.5155	−3.78%	171.72	0.5111	−4.60%
17,500	144	0.4286	144.26	0.4293	0.18%	147.11	0.4378	2.16%	144.42	0.4298	0.29%	145.55	0.4332	1.08%
22,500	122	0.3631	122.74	0.3653	0.61%	126.25	0.3757	3.48%	123.64	0.3680	1.34%	126.50	0.3765	3.69%
32,500	92	0.2738	95.10	0.2830	3.37%	97.66	0.2907	6.16%	95.69	0.2848	4.01%	100.27	0.2984	8.99%
47,500	73	0.2173	72.65	0.2162	−0.48%	72.17	0.2148	−1.13%	71.10	0.2116	−2.60%	76.14	0.2266	4.30%
67,500	58	0.1726	56.06	0.1668	−3.34%	53.00	0.1577	−8.62%	52.67	0.1568	−9.19%	57.06	0.1698	−1.61%
87,500	47	0.1399	45.15	0.1344	−3.95%	41.60	0.1238	−11.49%	41.67	0.1240	−11.34%	45.14	0.1344	−3.95%
125,000	29	0.0863	31.47	0.0937	8.51%	29.37	0.0874	1.27%	29.77	0.0886	2.65%	31.73	0.0944	9.43%
175,000	22	0.0655	20.82	0.0620	−5.34%	20.89	0.0622	−5.04%	21.42	0.0637	−2.64%	22.02	0.0655	0.09%
225,000	15	0.0446	14.94	0.0445	−0.42%	16.11	0.0479	7.38%	16.65	0.0496	11.01%	16.41	0.0488	9.37%
325,000	9	0.0268	9.54	0.0284	6.04%	10.94	0.0326	21.57%	11.44	0.0341	27.15%	10.32	0.0307	14.61%
475,000	7	0.0208	6.66	0.0198	−4.82%	7.30	0.0217	4.31%	7.72	0.0230	10.30%	6.13	0.0182	−12.49%
675,000	5	0.0149	4.87	0.0145	−2.55%	5.00	0.0149	0.09%	5.34	0.0159	6.83%	3.64	0.0108	−27.26%
1,000,000	3	0.0089	3.08	0.0092	2.54%	3.27	0.0097	9.07%	3.53	0.0105	17.53%	1.94	0.0058	−35.30%

TABLE 2

KARUSH–KUHN–TUCKER FUNCTION

	One Mean		Two Means		Three Means		Four Means	
Mean $= \frac{1}{\lambda}$	$\frac{1}{\lambda}$	$h(\lambda)$	$\frac{1}{\lambda}$	$h(\lambda)$	$\frac{1}{\lambda}$	$h(\lambda)$	$\frac{1}{\lambda}$	$h(\lambda)$
							0	**336.000**
1,000		1173.337		432.190		396.167		335.881
2,000		1060.099		402.882		373.030		334.870
3,000		968.445		381.032		356.735		333.555
4,000		898.142		366.497		347.053		333.025
5,000		842.334		356.811		341.618		333.218
6,000		796.360		350.199		338.684		333.773
7,000		757.342		345.567		337.168		334.415
8,000		723.472		342.262		336.435		334.993
9,000		693.574		339.891		336.118		335.442
10,000		666.848		338.213		336.012		335.748
					10,598	**336.000**		
							12,336	**336.000**
			13,570	**336.000**				
20,000		496.647		340.263		336.150		335.188
30,000		410.371		353.183		336.079		334.770
40,000		360.872		363.845		336.007		335.056
50,000		336.444		369.889		335.970		335.455
	51,190	**336.000**						
60,000		389.835		371.973		335.979		335.775
70,000		971.560		371.194		335.998		335.958
					73,440	**336.000**		
							77,922	**336.000**
80,000		3,995		368.495		335.992		335.997
90,000		14,647,843		364.604		335.944		335.915
100,000		43,187,502		360.067		335.856		335.745
			176,638	**336.000**				
200,000		6,191,258		338.944		334.801		333.922
300,000		32,692,464		414.348		334.964		334.227
400,000		75,160,236		558.705		335.441		335.019
500,000		123,867,653		729.785		335.788		335.598
600,000		172,830,341		903.455		335.960		335.903
					686,632	**336.000**		
700,000		219,258,668		1068.732		335.999		335.999
							712,302	**336.000**
800,000		262,097,414		1221.452		335.950		335.956
900,000		301,125,152		1360.665		335.845		335.827
1,000,000		336,492,659		1486.842		335.708		335.645
2,000,000		554,645,524		2264.678		334.206		333.527
3,000,000		655,188,015		2622.692		333.230		332.125
4,000,000		712,101,357		2825.202		332.618		331.242
5,000,000		748,596,059		2955.002		332.205		330.645
6,000,000		773,959,261		3045.185		331.909		330.217
7,000,000		792,600,348		3111.454		331.688		329.896
8,000,000		806,875,259		3162.194		331.516		329.646
9,000,000		818,155,499		3202.285		331.378		329.447
10,000,000		827,293,145		3234.758		331.266		329.284

of 0.6270, 0.3340, and 0.0390, respectively. Again, we have not reached the maximum.

The KKT function is now largest below the first mean of 10,598. We move a small amount of weight to a small mean and proportionately scale back the weights on the existing three means. When we resume iterating, this smallest mean heads toward zero. We then fix it at a small value (for example, 25, 1% of the first group boundary). Effectively, we assign all the probability associated with this mean to the first group. We resume iterating, and the algorithm converges to the values shown at the top of Table 1. The table shows the first mean as zero, because that is its true value. As the last column of Table 2 shows, the KKT function now never exceeds 336. We have thus reached the maximum likelihood estimates of the mixed exponential parameters.

Table 1 shows the fitted survival probabilities. The fitted and empirical probabilities match exactly at the first group boundary. This will always occur when a mean of zero has a positive weight in the final parameter set, since this is the only way the KKT function can be equal to the number of observations when λ_i is infinity. Likewise, anytime a mean of infinity has a positive weight in the final parameter set, the survival probabilities will match exactly at the last group boundary.

If the data includes various deductibles, attachment points, or policy limits, we can obtain the empirical distribution using the Kaplan–Meier Product-Limit estimator. This estimator provides empirical survival probabilities that take into account the effect of unobserved losses below deductibles and attachment points as well as losses capped by policy limits. Klugman, Panjer, and Willmot [8] cover this estimator briefly. It has historically been used extensively in survival analysis, and Klein and Moeschberger [7] and London [14] cover the subject in more detail.

For comparison, Table 1 also shows the fits for three distributions other than the mixed exponential. The parameterizations of the transformed beta and the Pareto are the same as those that Klugman, Panjer, and Willmot [8] use. See Appendix A for details. The lognormal parameterization is the standard one. The transformed beta provides the best fit, as measured by the log-likelihood, of the distributions used by Klugman, Panjer, and Willmot [8]. The Pareto is a special case of both the mixed exponential and the transformed beta. As expected, the mixed exponential provides the best fit.

We would prefer the mixed exponential distribution if our hypothesis is that the actual distribution has the alternating derivative property, which is a much weaker hypothesis than one that states that the actual distribution follows a particular parametric form. In most situations, I have found little or no justification for a stronger parametric hypothesis.

The usual way to evaluate a hypothesis is to perform a test such as the chi-square goodness-of-fit test. When the parameters are estimated from the data, this test is not appropriate with the mixed exponential distribution, since the mixed exponential does not have a fixed number of parameters. However, with most loss data I have encountered in practice, the appropriateness of the mixed exponential will be evident from a comparison of the fitted and empirical distributions.

For the other three distributions in Table 1, we can perform chi-square goodness-of-fit tests. We will combine the last three groups, and the two groups before the last three, so there are at least five losses in each of the resulting 14 groups. The results are as follows:

Distribution	Chi-square Statistic	Degrees of Freedom	*p*-value
Transformed Beta	9.24	9	0.41
Pareto	10.55	11	0.48
Lognormal	11.12	11	0.43

Another way to evaluate the Pareto hypothesis would be to use a likelihood ratio test. Since the Pareto distribution is a special case of the transformed beta distribution, under the Pareto hypothesis, twice the difference of the maximum values of the loglikelihoods of the Pareto and transformed beta has approximately a chi-square distribution with two degrees of freedom (the difference in the number of parameters). In this case, $2 \times ((-820.16) - (-820.78)) = 1.24$, which yields a p-value of 0.54. Thus, the Pareto distribution would not be rejected in favor of the transformed beta distribution.[10]

In this example, none of the distributions shown in Table 1 would be rejected as possible models for the actual distribution. However, as I mentioned above, hypothesizing a particular parametric distribution is dubious in most cases I have encountered. In general, the larger the data set, the more evident this becomes.

6. VARIANCE

With parametric distributions, we can obtain the asymptotic variances and covariances of the maximum likelihood estimators of the parameters by calculating the covariance matrix. We can then use the covariance matrix to find the asymptotic variances of the maximum likelihood estimators of functions of the parameters that are of interest, such as survival probabilities and limited expected values.[11]

This approach does not work with the semiparametric mixed exponential distribution. Tierney and Lambert [16] obtained a result that implies that the asymptotic variance of the maximum likelihood estimator of a function of mixed exponential parameters is equal to the variance of the empirical estimator for ungrouped data. For a survival probability, the empirical estimator is the sample proportion of observations that exceeds the loss

[10] See Section 2.9 of Klugman, Panjer, and Willmot [8] for a more thorough discussion of these tests.

[11] See Section 2.5 of Klugman, Panjer, and Willmot [8] for a discussion.

amount under consideration. This has a binomial distribution that approaches a normal distribution as the number of observations approaches infinity. This result means that, asymptotically, we do not reduce the variance of our survival probability estimates, or any other estimates based on the mixed exponential parameters, by using the fitted distribution instead of the empirical distribution.

In practice, we do not have infinite samples. To see what happens with finite samples, we must resort to simulation. Tables 3A, 3B, and 3C show the results of simulations using sample sizes of 10, 50, and 250, respectively. In each case, the simulated distribution is the Pareto distribution from Table 1. I used a Pareto distribution to facilitate comparison of the variances of estimates obtained using the mixed exponential distribution with the variances of estimates obtained using the Pareto distribution. The Pareto distribution serves as an example of a parametric distribution with a fixed number of parameters. These tables show estimates of the bias and variance of survival probability estimates based on 10,000 simulations, for a mixed exponential fit without grouping the data, and both a mixed exponential and a Pareto fit with data grouped using the boundaries from Table 1. The tables display bias as a percentage of the actual survival probability, and variance as a ratio to the variance of the empirical estimator. Table 3C also shows the asymptotic variance for the Pareto distribution. I focus on the survival function because any other function of interest can be expressed in terms of the survival function. For example, the limited expected value is simply the integral of the survival function from zero to the limit being considered.

The grouped mixed exponential results are close to the ungrouped results in the middle of the distribution, but are dramatically worse at small loss amounts and in the tail. The reason for this is that the grouped data provides virtually no information about the distribution either below the first group boundary of 2,500 or above the last group boundary of 1,000,000. There-

TABLE 3A

SIMULATION RESULTS—10 OBSERVATIONS

			Ungrouped Mixed Exponential		Grouped Mixed Exponential		Grouped Pareto	
Loss Amt	Survival Probability	10 Times Empirical Variance	Bias	Ratio to Empirical Variance	Bias	Ratio to Empirical Variance	Bias	Ratio to Empirical Variance
10	0.9993	0.00073	−0.25%	1.09	−3.79%	85.35	−0.02%	0.05
100	0.9927	0.00722	−1.20%	0.79	−3.89%	8.34	−0.13%	0.12
1,000	0.9316	0.06376	−3.61%	0.56	−4.58%	0.94	−0.66%	0.37
2,500	0.8443	0.13142	−4.45%	0.56	−4.96%	0.64	−0.72%	0.54
7,500	0.6415	0.22999	−3.99%	0.64	−4.28%	0.65	0.36%	0.72
12,500	0.5155	0.24976	−3.15%	0.69	−3.37%	0.69	1.06%	0.79
17,500	0.4298	0.24508	−2.63%	0.72	−2.79%	0.72	1.14%	0.82
22,500	0.3680	0.23257	−2.40%	0.74	−2.49%	0.74	0.76%	0.83
32,500	0.2848	0.20368	−2.47%	0.75	−2.47%	0.76	−0.70%	0.82
47,500	0.2116	0.16683	−3.13%	0.75	−3.04%	0.75	−3.32%	0.79
67,500	0.1568	0.13219	−4.16%	0.73	−4.03%	0.74	−6.17%	0.74
87,500	0.1240	0.10863	−5.05%	0.71	−4.89%	0.72	−7.94%	0.70
125,000	0.0886	0.08075	−6.23%	0.68	−6.02%	0.68	−8.98%	0.64
175,000	0.0637	0.05968	−7.13%	0.65	−6.72%	0.65	−7.68%	0.61
225,000	0.0496	0.04710	−7.60%	0.63	−6.79%	0.64	−4.98%	0.59
325,000	0.0341	0.03290	−7.90%	0.61	−5.65%	0.63	1.75%	0.58
475,000	0.0230	0.02245	−7.66%	0.58	−1.60%	0.67	12.25%	0.58
675,000	0.0159	0.01564	−7.00%	0.56	6.94%	0.78	25.39%	0.59
1,000,000	0.0105	0.01038	−6.00%	0.54	26.22%	1.02	44.47%	0.61
2,000,000	0.0050	0.00499	−4.31%	0.51	108.32%	1.93	92.25%	0.68
3,000,000	0.0033	0.00324	−3.39%	0.51	205.80%	2.94	131.11%	0.75
5,000,000	0.0019	0.00188	−1.82%	0.51	417.16%	5.05	195.67%	0.85
10,000,000	0.0009	0.00089	2.63%	0.54	982.05%	10.63	321.95%	1.04
20,000,000	0.0004	0.00042	11.91%	0.60	2178.28%	22.38	514.73%	1.31
30,000,000	0.0003	0.00027	19.65%	0.65	3423.15%	34.60	672.44%	1.52
50,000,000	0.0002	0.00016	31.42%	0.69	6002.45%	59.92	937.99%	1.86
100,000,000	0.0001	0.00008	46.04%	0.72	12761.24%	126.28	1469.57%	2.47

10 (Sample Size) Times Empirical Variance

$$= 10 \cdot \frac{\text{Surv Prob} \cdot (1 - \text{Surv Prob})}{10} = \text{Surv Prob} \cdot (1 - \text{Surv Prob})$$

$$\text{Bias} = \frac{\text{Average Simulated Fitted Survival Probability} - \text{Survival Probability}}{\text{Survival Probability}}$$

Ratio to Empirical Variance

$$= \frac{\text{Variance of Simulated Fitted Survival Probabilities}}{\text{Empirical Variance}}$$

TABLE 3B

SIMULATION RESULTS—50 OBSERVATIONS

			Ungrouped Mixed Exponential		Grouped Mixed Exponential		Grouped Pareto	
Loss Amt	Survival Probability	50 Times Empirical Variance	Bias	Ratio to Empirical Variance	Bias	Ratio to Empirical Variance	Bias	Ratio to Empirical Variance
10	0.9993	0.00073	−0.15%	0.87	−1.94%	98.53	0.00%	0.00
100	0.9927	0.00722	−0.59%	0.55	−1.95%	9.50	−0.02%	0.03
1,000	0.9316	0.06376	−1.45%	0.51	−1.98%	0.97	−0.10%	0.27
2,500	0.8443	0.13142	−1.50%	0.59	−1.75%	0.68	−0.09%	0.50
7,500	0.6415	0.22999	−0.63%	0.72	−0.73%	0.74	0.29%	0.75
12,500	0.5155	0.24976	−0.11%	0.77	−0.15%	0.78	0.57%	0.77
17,500	0.4298	0.24508	0.02%	0.79	0.04%	0.80	0.66%	0.75
22,500	0.3680	0.23257	−0.04%	0.80	0.01%	0.80	0.61%	0.73
32,500	0.2848	0.20368	−0.36%	0.80	−0.29%	0.81	0.27%	0.69
47,500	0.2116	0.16683	−0.83%	0.81	−0.75%	0.81	−0.42%	0.68
67,500	0.1568	0.13219	−1.22%	0.81	−1.15%	0.82	−1.22%	0.68
87,500	0.1240	0.10863	−1.46%	0.81	−1.40%	0.82	−1.76%	0.69
125,000	0.0886	0.08075	−1.85%	0.81	−1.85%	0.81	−2.21%	0.69
175,000	0.0637	0.05968	−2.40%	0.80	−2.56%	0.80	−2.09%	0.68
225,000	0.0496	0.04710	−2.97%	0.78	−3.29%	0.78	−1.57%	0.66
325,000	0.0341	0.03290	−3.99%	0.76	−4.27%	0.75	−0.03%	0.62
475,000	0.0230	0.02245	−5.12%	0.72	−3.80%	0.74	2.58%	0.57
675,000	0.0159	0.01564	−6.08%	0.69	0.28%	0.78	5.98%	0.52
1,000,000	0.0105	0.01038	−6.92%	0.65	12.80%	0.93	10.95%	0.47
2,000,000	0.0050	0.00499	−7.59%	0.60	71.75%	1.64	22.91%	0.40
3,000,000	0.0033	0.00324	−7.36%	0.57	143.09%	2.50	32.00%	0.36
5,000,000	0.0019	0.00188	−6.20%	0.56	301.35%	4.36	45.95%	0.31
10,000,000	0.0009	0.00089	−2.70%	0.55	733.42%	9.21	70.11%	0.27
20,000,000	0.0004	0.00042	2.41%	0.55	1653.60%	19.39	101.68%	0.23
30,000,000	0.0003	0.00027	5.82%	0.56	2611.72%	29.98	124.40%	0.21
50,000,000	0.0002	0.00016	10.34%	0.56	4596.97%	51.92	158.47%	0.19
100,000,000	0.0001	0.00008	15.23%	0.57	9799.13%	109.42	216.68%	0.17

50 (Sample Size) Times Empirical Variance

$$= 50 \cdot \frac{\text{Surv Prob} \cdot (1 - \text{Surv Prob})}{50} = \text{Surv Prob} \cdot (1 - \text{Surv Prob})$$

$$\text{Bias} = \frac{\text{Average Simulated Fitted Survival Probability} - \text{Survival Probability}}{\text{Survival Probability}}$$

Ratio to Empirical Variance

$$= \frac{\text{Variance of Simulated Fitted Survival Probabilities}}{\text{Empirical Variance}}$$

TABLE 3C

SIMULATION RESULTS—250 OBSERVATIONS

			Ungrouped Mixed Exponential		Grouped Mixed Exponential		Grouped Pareto		Grouped Pareto Asymptotic Variance
Loss Amt	Survival Probability	250 Times Empirical Variance	Bias	Ratio to Empirical Variance	Bias	Ratio to Empirical Variance	Bias	Ratio to Empirical Variance	Ratio to Empirical Variance
10	0.9993	0.00073	−0.08%	0.59	−1.06%	134.65	0.00%	0.00	0.00
100	0.9927	0.00722	−0.27%	0.40	−1.05%	12.75	0.00%	0.03	0.03
1,000	0.9316	0.06376	−0.58%	0.49	−0.92%	1.10	−0.04%	0.25	0.25
2,500	0.8443	0.13142	−0.48%	0.62	−0.64%	0.73	−0.06%	0.48	0.49
7,500	0.6415	0.22999	−0.03%	0.78	−0.06%	0.80	−0.05%	0.75	0.77
12,500	0.5155	0.24976	0.02%	0.82	0.02%	0.84	−0.03%	0.77	0.78
17,500	0.4298	0.24508	−0.12%	0.84	−0.11%	0.85	−0.05%	0.75	0.75
22,500	0.3680	0.23257	−0.29%	0.86	−0.28%	0.86	−0.08%	0.72	0.72
32,500	0.2848	0.20368	−0.52%	0.89	−0.52%	0.89	−0.18%	0.69	0.68
47,500	0.2116	0.16683	−0.60%	0.90	−0.61%	0.91	−0.34%	0.68	0.66
67,500	0.1568	0.13219	−0.55%	0.90	−0.55%	0.91	−0.51%	0.69	0.67
87,500	0.1240	0.10863	−0.54%	0.90	−0.50%	0.91	−0.63%	0.71	0.68
125,000	0.0886	0.08075	−0.66%	0.89	−0.57%	0.90	−0.74%	0.72	0.70
175,000	0.0637	0.05968	−0.91%	0.88	−0.83%	0.89	−0.75%	0.72	0.70
225,000	0.0496	0.04710	−1.14%	0.87	−1.13%	0.88	−0.67%	0.71	0.69
325,000	0.0341	0.03290	−1.52%	0.87	−1.61%	0.89	−0.39%	0.68	0.66
475,000	0.0230	0.02245	−1.94%	0.86	−1.73%	0.89	0.14%	0.62	0.60
675,000	0.0159	0.01564	−2.39%	0.85	−0.68%	0.88	0.86%	0.56	0.54
1,000,000	0.0105	0.01038	−3.04%	0.83	3.36%	0.96	1.94%	0.48	0.47
2,000,000	0.0050	0.00499	−4.61%	0.75	27.42%	1.78	4.60%	0.36	0.33
3,000,000	0.0033	0.00324	−5.63%	0.71	61.92%	2.85	6.62%	0.30	0.27
5,000,000	0.0019	0.00188	−6.62%	0.66	145.14%	5.12	9.66%	0.23	0.20
10,000,000	0.0009	0.00089	−6.62%	0.60	383.10%	11.02	14.73%	0.16	0.12
20,000,000	0.0004	0.00042	−5.55%	0.55	900.71%	23.32	20.95%	0.11	0.08
30,000,000	0.0003	0.00027	−4.78%	0.53	1442.30%	36.06	25.17%	0.09	0.06
50,000,000	0.0002	0.00016	−3.64%	0.51	2565.35%	62.44	31.15%	0.06	0.04
100,000,000	0.0001	0.00008	−2.43%	0.49	5508.39%	131.56	40.53%	0.04	0.02

250 (Sample Size) Times Empirical Variance

$$= 250 \cdot \frac{\text{Surv Prob} \cdot (1 - \text{Surv Prob})}{250} = \text{Surv Prob} \cdot (1 - \text{Surv Prob})$$

$$\text{Bias} = \frac{\text{Average Simulated Fitted Survival Probability} - \text{Survival Probability}}{\text{Survival Probability}}$$

Ratio to Empirical Variance

$$= \frac{\text{Variance of Simulated Fitted Survival Probabilities}}{\text{Empirical Variance}}$$

fore, the fitted distribution often contains means of either zero or infinity or both.

Because the Pareto is less flexible than the mixed exponential, the Pareto usually provides survival probability estimates with a smaller variance. This effect is most notable at small loss amounts and in the tail. However, this fact illustrates the problem with using the Pareto or other parametric distributions with a fixed number of parameters. If we knew that the actual distribution were a Pareto, we would of course prefer to fit a Pareto instead of a mixed exponential. However, the assumption that the distribution is a Pareto is virtually never valid. If our data set is small, the fit may appear to be good, but the tail is simply a function of the assumption that the distribution is a Pareto. The fitted tail may or may not be anywhere close to the actual tail. If our data set is large, then unless we really do have a Pareto, we will probably observe a poor fit in the tail because the Pareto is not flexible enough. Thus, though the Pareto provides estimates with smaller variance than the mixed exponential, these estimates may be significantly biased if the actual distribution is not a Pareto.

For the ungrouped mixed exponential, as the number of observations increases, the bias gradually disappears, and the ratio of the variance to the empirical variance eventually approaches 1. This process takes longer at small loss amounts and in the tail. For the grouped mixed exponential, the results are similar except that outside the layer boundaries, the estimator remains poor. Note that an empirical estimate of the survival probability is not an option outside the layer boundaries, since an empirical estimator is only available at the layer boundaries. For the Pareto, with 250 observations, the variance is very close to the asymptotic variance, but there is still some significant bias in the tail.

I have displayed results for only one distribution. The most notable feature that differs by distribution is that, generally speaking, for a given number of observations and a given survival probability, the thinner the tail of a distribution, the smaller the variance. Roughly, this is because there is less spread in the mixing distribution of mixed exponential distributions with thinner tails than in those with thicker tails.

7. ADJUSTMENTS AND OTHER USES

In this section, I will first address the issue of estimating the tail of a distribution. Table 1 showed only survival probabilities up to 1,000,000. Table 4 shows survival probabilities up to 100,000,000. The first distribution in the table is the mixed exponential that we fit previously. The second distribution is the mixed exponential that results when we move one claim from the 675,000–1,000,000 group to the 475,000–675,000 group. The survival probabilities are very close to one another except in the tail. When we move one claim, we acquire a mean of infinity with a small positive weight. The survival function now approaches the value of this weight, instead of zero, as the loss amount approaches infinity. For comparison, Table 4 also shows the Pareto and lognormal distributions from Table 1. If we were to move this same claim and then fit a Pareto or lognormal distribution, the tails would be very close to those from Table 1. However, we have no way to tell from the available data whether either of them is anywhere close to the actual tail. The tails of the Pareto and lognormal distributions are between the two mixed exponential tails, and are also very different from one another.

Thus we see that we cannot reliably use the mixed exponential distribution or any parametric distribution to extrapolate beyond the available data. However, an advantage of the mixed exponential is that if other data is available to assist in estimating the tail, or if we simply use judgment, we can find a mixed exponential distribution that both fits the available data and produces the desired tail. For example, suppose we believe that the tail is likely to have a shape like the Pareto tail. We may base this belief on data we have from a similar source or simply judgment. We can

TABLE 4

TAIL COMPARISON

	Mixed Exponential		Mix Exp–1 Clm Moved		Pareto		Lognormal		Mix Exp–Pareto Tail	
	Mean	Weight	Mean	Weight					Mean	Weight
	0	0.0526	0	0.0525					0	0.0526
	12,336	0.5999	12,260	0.5950					12,303	0.5980
	77,922	0.3102	72,792	0.2962	θ	14,679	μ	9.4812	76,185	0.3063
	712,302	0.0373	326,741	0.0497	α	1.0758	σ	1.7162	437,233	0.0340
			Infinity	0.0066					2,216,890	0.0073
									10,459,111	0.0014
									74,727,807	0.0003
Loss Amt	No.> Loss Amt	Survival Probability	No.> Loss Amt	Survival Probability	No.> Loss Amt	Survival Probability	No.> Loss Amt	Survival Probability	No.> Loss Amt	Survival Probability
0	336.00	1.0000	336.00	1.0000	336.00	1.0000	336.00	1.0000	336.00	1.0000
2,500	278.00	0.8274	278.00	0.8274	283.70	0.8443	279.84	0.8329	278.00	0.8274
7,500	216.79	0.6452	216.76	0.6451	215.53	0.6415	210.86	0.6276	216.77	0.6452
12,500	174.25	0.5186	174.24	0.5186	173.19	0.5155	171.72	0.5111	174.24	0.5186
17,500	144.26	0.4293	144.27	0.4294	144.42	0.4298	145.55	0.4332	144.26	0.4293
22,500	122.74	0.3653	122.77	0.3654	123.64	0.3680	126.50	0.3765	122.75	0.3653
32,500	95.10	0.2830	95.13	0.2831	95.69	0.2848	100.27	0.2984	95.12	0.2831
47,500	72.65	0.2162	72.63	0.2162	71.10	0.2116	76.14	0.2266	72.65	0.2162
67,500	56.06	0.1668	55.99	0.1666	52.67	0.1568	57.06	0.1698	56.03	0.1667
87,500	45.15	0.1344	45.07	0.1341	41.67	0.1240	45.14	0.1344	45.10	0.1342
125,000	31.47	0.0937	31.48	0.0937	29.77	0.0886	31.73	0.0944	31.44	0.0936
175,000	20.82	0.0620	20.98	0.0624	21.42	0.0637	22.02	0.0655	20.85	0.0621
225,000	14.94	0.0445	15.13	0.0450	16.65	0.0496	16.41	0.0488	14.99	0.0446
325,000	9.54	0.0284	9.54	0.0284	11.44	0.0341	10.32	0.0307	9.57	0.0285
475,000	6.66	0.0198	6.26	0.0186	7.72	0.0230	6.13	0.0182	6.60	0.0196
675,000	4.87	0.0145	4.34	0.0129	5.34	0.0159	3.64	0.0108	4.81	0.0143
1,000,000	3.08	0.0092	3.00	0.0089	3.53	0.0105	1.94	0.0058	3.26	0.0097
2,000,000	0.76	0.0022	2.25	0.0067	1.69	0.0050	0.57	0.0017	1.60	0.0048
3,000,000	0.19	0.0006	2.22	0.0066	1.09	0.0033	0.26	0.0008	1.10	0.0033
5,000,000	0.01	0.0000	2.22	0.0066	0.63	0.0019	0.09	0.0003	0.65	0.0019
10,000,000	0.00	0.0000	2.22	0.0066	0.30	0.0009	0.02	0.0001	0.30	0.0009
20,000,000	0.00	0.0000	2.22	0.0066	0.14	0.0004	0.00	0.0000	0.14	0.0004
30,000,000	0.00	0.0000	2.22	0.0066	0.09	0.0003	0.00	0.0000	0.09	0.0003
50,000,000	0.00	0.0000	2.22	0.0066	0.05	0.0002	0.00	0.0000	0.05	0.0002
100,000,000	0.00	0.0000	2.22	0.0066	0.03	0.0001	0.00	0.0000	0.03	0.0001

add eight more group boundaries as shown in Table 4 to increase the number of groups to 25. We can then allocate the three claims above 1,000,000 to the nine groups above 1,000,000 so that the empirical survival probabilities above 1,000,000 match those of the Pareto distribution. We can then find a maximum likelihood estimate based on these 25 groups. The last two columns of Table 4 show the resulting distribution. The mixed exponential distribution is flexible enough so that we can append whatever tail we think appropriate while affecting the fit in the lower portion of the distribution very little.

In the example above, we adjusted the data before fitting to produce an appropriate tail. We may need to adjust the data for other reasons. For example, we may have to adjust for loss development. I will not discuss this issue further in this paper. However, such adjustments would change the empirical distribution to which we fit.

Just as we may adjust the data, we may also need to adjust the fitted distribution. The best fitting distribution, which satisfies the KKT conditions, will not, in all cases, be the most appropriate estimate to use. When conditions warrant, we may set any of the means and weights at fixed values before fitting. For example, despite any data adjustments we have made, if the best fitting distribution contains a mean of infinity, we may fix the largest mean and possibly its weight at a value that yields a tail that we feel is more appropriate. As another example, if we are fitting a number of distributions as part of the same project, we may find it convenient to use the same fixed means for each distribution. If the means are not too far apart, the resulting distributions are likely to fit almost as well as if we had not fixed the means. We could also impose constraints on the relationships among the means and weights through the use of Lagrange multipliers. Also, we could, through trial and error, simply select a distribution that visually fits the data well.

We can use the mixed exponential distribution for more than modeling losses. We can use the mixed exponential to model anything where we expect a function with alternating derivatives. For example, I have found it useful in modeling the probability that a claim does not have any allocated loss adjustment expense attached to it as a function of the claim size. This is not a probability function, so we cannot use maximum likelihood estimation. However, we can use a least squares procedure to fit the distribution to the data.

8. CONCLUSION

In this paper, I have tried to provide the background needed for an actuary to begin using the mixed exponential distribution in his or her work. I believe that the combination of flexibility and smoothness that the mixed exponential provides makes it an extremely useful actuarial modeling tool.

REFERENCES

[1] Bohning, Dankmar, "A Review of Reliable Maximum Likelihood Algorithms for Semiparametric Mixture Models," *Journal of Statistical Planning and Inference* 47, 1/2, October 1995, pp. 5–28.

[2] Brockett, Patrick L., and Linda L. Golden, "A Class of Utility Functions Containing All the Common Utility Functions," *Management Science* 33, 8, August 1987, pp. 955–964.

[3] Feller, William, *An Introduction to Probability Theory and Its Applications*, Volume II, Second Edition, New York: John Wiley & Sons, 1971.

[4] Hillier, Frederick S., and Gerald J. Lieberman, *Introduction to Operations Research*, Sixth Edition, New York: McGraw-Hill, 1995.

[5] Hogg, Robert V., and Stuart A. Klugman, *Loss Distributions*, New York: John Wiley & Sons, 1984.

[6] Jewell, Nicholas P., "Mixtures of Exponential Distributions," *The Annals of Statistics* 10, 2, June 1982, pp. 479–484.

[7] Klein, John P., and Melvin L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*, New York: Springer, 1997.

[8] Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot, *Loss Models: From Data to Decisions*, New York: John Wiley & Sons, 1998.

[9] Lindsay, Bruce G., "Properties of the Maximum Likelihood Estimator of a Mixing Distribution," *Statistical Distributions in Scientific Work* 5, editors C. Taillie, G. Patil and B. Baldessari, Boston: D. Reidel, 1981, pp. 95–109.

[10] Lindsay, Bruce G., "The Geometry of Mixture Likelihoods: A General Theory," *The Annals of Statistics* 11, 1, March 1983, pp. 86–94.

[11] Lindsay, Bruce G., *Mixture Models: Theory, Geometry and Applications*, Hayward, California: Institute of Mathematical Statistics, 1995.

[12] Lindsay, Bruce G., and Mary L. Lesperance, "A Review of Semiparametric Mixture Models," *Journal of Statistical Planning and Inference* 47, 1/2, October 1995, pp. 29–39.

[13] Lindsay, Bruce G., and Kathryn Roeder, "Uniqueness of Estimation and Identifiability in Mixture Models," *The Canadian Journal of Statistics* 21, 2, June 1993, pp. 139–147.

[14] London, Dick, *Survival Models and Their Estimation*, Third Edition, Winsted, Connecticut: ACTEX Publications, 1997.

[15] Polyá, George, and Gabor Szegö, *Problems and Theorems in Analysis*, Volume II (revised and enlarged translation by Claude E. Billigheimer of *Aufgaben und Lehrsätze aus der Analysis*, Volume II, Fourth Edition, 1971), Berlin: Springer-Verlag, 1976.

[16] Tierney, Luke, and Diane Lambert, "Asymptotic Efficiency of Estimators of Functionals of Mixed Distributions," *The Annals of Statistics* 12, 4, December 1984, pp. 1380–1387.

APPENDIX A

In this appendix, I will address the issue of which of the parametric distributions generally used to model losses have completely monotone density functions and are thus special cases of the mixed exponential distribution. I will use the same parameterizations that are used in Klugman, Panjer, and Willmot [8].

The transformed beta distribution has probability density function

$$f(x) = \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\gamma(x/\theta)^{\gamma\tau}}{x[1+(x/\theta)^{\gamma}]^{\alpha+\tau}}.$$

If $\gamma\tau > 1$, then $f(x)$ is not completely monotone because it has a nonzero mode.

If $\gamma\tau \leq 1$ and $\gamma \leq 1$, then $f(x)$ is completely monotone. To see this, note that, ignoring factors not involving x, we can write $f(x)$ as the product of $x^{\gamma\tau-1}$ and $[1+(x/\theta)^{\gamma}]^{-\alpha-\tau}$. The first factor is clearly completely monotone. We can use induction with the product rule for differentiation to show that the second factor is completely monotone. Similarly, we can use induction to show that the product of the two factors is also completely monotone. Feller [3, p. 441] gives a short proof of the fact that the product of completely monotone functions is also completely monotone.

Notable special cases of the transformed beta distribution that are also special cases of the mixed exponential distribution are the Pareto (which has γ and τ fixed at 1) and the Burr (which has τ fixed at 1) with $\gamma \leq 1$.

The set of parameters for which $f(x)$ is completely monotone when $\gamma\tau \leq 1$ and $\gamma > 1$ is an open question. If γ is too large, then $f(x)$ will not be completely monotone, but I could not find a proof that would definitively determine the status of all distributions with parameters in this region.

The transformed gamma distribution has probability density function

$$g(x) = \frac{\tau(x/\theta)^{\alpha\tau}e^{-(x/\theta)^{\tau}}}{x\Gamma(\alpha)}.$$

If $\alpha\tau > 1$, then $g(x)$ is not completely monotone because it has a nonzero mode.

If $\tau > 1$, then $g(x)$ is not completely monotone because it has an increasing failure rate in the tail.

If $\alpha\tau \leq 1$ and $\tau \leq 1$, then $g(x)$ is completely monotone. To see this, note that, ignoring factors not involving x, we can write $g(x)$ as the product of $x^{\alpha\tau-1}$ and $e^{-(x/\theta)^{\tau}}$. These are both completely monotone, so their product is completely monotone.

Notable special cases of the transformed gamma distribution that are also special cases of the mixed exponential distribution are the gamma (which has τ fixed at 1) with $\alpha \leq 1$ and the Weibull (which has α fixed at 1) with $\tau \leq 1$.

The inverse transformed gamma, lognormal, and inverse Gaussian distributions are never completely monotone, since they always have nonzero modes.

All of the distributions mentioned, except for the transformed gamma with certain parameters ($\tau > 1$ or $\tau = 1$, $\alpha \geq 1$), have decreasing failure rates in the tail.

APPENDIX B

In this appendix, I will provide proofs of the key properties underlying maximum likelihood estimation with the mixed exponential distribution—first for ungrouped data, then for grouped data.

Ungrouped Data

The loglikelihood function is

$$\ln L(w_1, w_2, \ldots) = \sum_{k=1}^{m} \ln f(x_k) = \sum_{k=1}^{m} \ln \left(\sum_{i=1}^{\infty} w_i \lambda_i e^{-\lambda_i x_k} \right),$$

where m is the number of observations. We must find the set of w_i's that maximizes the loglikelihood function, subject to the constraints that each of the w_i's must be greater than or equal to zero and the sum of the w_i's must be one. From now on, when I refer to maximizing the loglikelihood function, I mean maximizing the loglikelihood function subject to these constraints. We consider the λ_i's fixed and arbitrarily close together. Thus, the only parameters are the w_i's.

The ln function is strictly concave and the sum of strictly concave functions is also strictly concave.[12] This fact allows us to conclude that if more than one set of w_i's maximizes the loglikelihood function, each set must yield identical values of $\sum_{i=1}^{\infty} w_i \lambda_i e^{-\lambda_i x_k}$ for each x_k. If two sets of w_i's yielding different values of $\sum_{i=1}^{\infty} w_i \lambda_i e^{-\lambda_i x_k}$ maximized the loglikelihood function, each set of w_i's on the line segment between them (which would satisfy the constraints) would yield a value of the loglikelihood function greater than the maximum (since $\sum_{i=1}^{\infty} w_i \lambda_i e^{-\lambda_i x_k}$ is a linear function of the w_i's). Clearly, this cannot be.

[12]See Appendix 2 of Hillier and Lieberman [4] for a discussion of concavity and convexity.

We can view maximizing the loglikelihood function as a convex programming problem, since the loglikelihood function is concave and the constraints are linear (and thus convex). The theory of convex programming gives us a set of necessary and sufficient conditions, the Karush–Kuhn–Tucker (KKT) conditions, for the loglikelihood function to be at a maximum. For ungrouped data, these conditions are

$$\frac{\partial \ln L}{\partial w_i} = \sum_{k=1}^{m} \frac{\lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{\infty} w_j \lambda_j e^{-\lambda_j x_k}} \leq u, \qquad \text{if} \quad w_i = 0$$

and

$$\frac{\partial \ln L}{\partial w_i} = \sum_{k=1}^{m} \frac{\lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{\infty} w_j \lambda_j e^{-\lambda_j x_k}} = u, \qquad \text{if} \quad w_i > 0$$

for some number u. If we sum the KKT conditions, giving weight w_i to each element of the sum, we have

$$u = \sum_{i=1}^{\infty} w_i u = \sum_{i=1}^{\infty} w_i \frac{\partial \ln L}{\partial w_i} = \sum_{i=1}^{\infty}\sum_{k=1}^{m} \frac{w_i \lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{\infty} w_j \lambda_j e^{-\lambda_j x_k}}$$

$$= \sum_{k=1}^{m} \frac{\sum_{i=1}^{\infty} w_i \lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{\infty} w_j \lambda_j e^{-\lambda_j x_k}} = m.$$

Thus, we see that u must be equal to m, the number of observations.[13]

[13]See Chapter 13 of Hillier and Lieberman [4] for an introductory treatment of convex programming. Jewell [6] gave a direct derivation of the Karush–Kuhn–Tucker conditions for the mixed exponential case.

We now examine the function

$$h(\lambda) = \sum_{k=1}^{m} \frac{\lambda e^{-\lambda x_k}}{\sum_{j=1}^{\infty} w_j \lambda_j e^{-\lambda_j x_k}}, \qquad 0 \le \lambda \le \infty.$$

To satisfy the KKT conditions, this function must have a maximum of m that occurs at the points corresponding to where w_i is greater than zero. We first note that $h(0) = h(\infty) = 0$, so the w_i's corresponding to λ_i's of zero and infinity must be zero. Taking the derivative of $h(\lambda)$ gives

$$\frac{dh}{d\lambda} = \sum_{k=1}^{m} \frac{(-\lambda x_k + 1)e^{-\lambda x_k}}{\sum_{j=1}^{\infty} w_j \lambda_j e^{-\lambda_j x_k}}.$$

Polyá and Szegö [15] showed that an exponential polynomial of the form $\sum_{k=1}^{m} p_k(\lambda)e^{-\lambda x_k}$ that is not zero everywhere, where p_k is a real ordinary polynomial of degree d_k, has at most $\sum_{k=1}^{m}(d_k + 1) - 1$ zeros.[14] Thus $dh/d\lambda$ has at most $2m - 1$ zeros. When the KKT conditions are satisfied, $dh/d\lambda$ must be zero where $h(\lambda)$ assumes the value m on $(0,\infty)$. Since maxima must alternate with minima (where $dh/d\lambda$ must also be zero), $h(\lambda)$ can assume the value m at no more than m points on $(0,\infty)$. Since the w_i's corresponding to λ_i's of zero and infinity are zero, the number of positive w_i's at the point that the loglikelihood function is at its maximum is at most m, the number of observations.[15] We can also see that none of the corresponding λ_i's can be less than $1/x_m$, where x_m is the largest observation, since every term of the expression for $dh/d\lambda$ is positive for λ less than $1/x_m$. Likewise, none of the λ_i's can be greater than $1/x_1$, where x_1 is the smallest observation, since every term of the expression for $dh/d\lambda$ is negative for λ greater than $1/x_1$.

[14]See Part Five, Problem 75 of Polyá and Szegö [15].

[15]Using a more general technique, Lindsay [10] showed that this is true for mixtures of any type of distribution.

We will now determine whether the loglikelihood can attain its maximum at more than one set of w_i's. We do know that if more than one set yielded the maximum, each set would have to give the same value of $\sum_{i=1}^{n} w_i\lambda_i e^{-\lambda_i x_k}$ for each x_k. Let $\lambda_1,\ldots,\lambda_n$ be the points at which the w_i's are positive where the loglikelihood is at its maximum. If more than one set of w_i's gave the same value of $\sum_{i=1}^{n} w_i\lambda_i e^{-\lambda_i x_k}$ for each x_k, then the function $\sum_{i=1}^{n}(w_i - w_i^*)\lambda_i e^{-\lambda_i x}$ would have at least m zeros, one for each x_k. From Polyá and Szegö's result, this function can have no more than $n-1$ zeros. Since we have already determined that $n \le m$, we have a contradiction. We thus conclude that the loglikelihood attains its maximum at a unique set of w_i's.[16]

Grouped Data

The loglikelihood function is

$$\begin{aligned}
\ln L(w_1, w_2, \ldots) &= a_1 \ln(1 - S(b_1)) + \sum_{k=2}^{g-1} a_k \ln(S(b_{k-1}) - S(b_k)) \\
&\quad + a_g \ln(S(b_{g-1})) \\
&= a_1 \ln\left(\sum_{i=1}^{\infty} w_i(1 - e^{-\lambda_i b_1})\right) \\
&\quad + \sum_{k=2}^{g-1} a_k \ln\left(\sum_{i=1}^{\infty} w_i(e^{-\lambda_i b_{k-1}} - e^{-\lambda_i b_k})\right) \\
&\quad + a_g \ln\left(\sum_{i=1}^{\infty} w_i(e^{-\lambda_i b_{g-1}})\right),
\end{aligned}$$

where g is the number of groups, $a_1,\ldots,a_g$ are the number of observations in each group, and $b_1,\ldots,b_{g-1}$ are the group boundaries. We will assume that any adjacent groups that all have zero observations have been combined into one group. The development is analogous to that for ungrouped data down to where we

[16]The reasoning in this and the previous paragraph is taken from Jewell [6].

examine the function

$$h(\lambda) = a_1 \frac{1 - e^{-\lambda b_1}}{\sum_{j=1}^{\infty} w_j (1 - e^{-\lambda_j b_1})} + \sum_{k=2}^{g-1} a_k \frac{e^{-\lambda b_{k-1}} - e^{-\lambda b_k}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})}$$
$$+ a_g \frac{e^{-\lambda b_{g-1}}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{g-1}})}, \qquad 0 \leq \lambda \leq \infty.$$

We note that $h(0)$ and $h(\infty)$ are not necessarily equal to zero, so the w_i's corresponding to λ_i's of zero and infinity are not necessarily equal to zero. Taking the derivative of $h(\lambda)$ gives

$$\frac{dh}{d\lambda} = a_1 \frac{b_1 e^{-\lambda b_1}}{\sum_{j=1}^{\infty} w_j (1 - e^{-\lambda_j b_1})} + \sum_{k=2}^{g-1} a_k \frac{-b_{k-1} e^{-\lambda b_{k-1}} + b_k e^{-\lambda b_k}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})}$$
$$+ a_g \frac{-b_{g-1} e^{-\lambda b_{g-1}}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{g-1}})}$$
$$= \sum_{k=1}^{g-1} \left[\frac{a_k}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})} - \frac{a_{k+1}}{\sum_{j=1}^{\infty} w_j (e^{-\lambda_j b_k} - e^{-\lambda_j b_{k+1}})} \right]$$
$$\times b_k e^{-\lambda b_k},$$

where $b_0 = 0$ and $b_g = \infty$.

We may now apply Polyá and Szegö's result, except if all of the $g-1$ coefficients in the above equation are zero. This will occur only when the mixed exponential probabilities for each group are exactly proportional to the number of observations in each group or, in other words, when the data perfectly fits the model. For this situation, we can easily come up with examples where an arbitrarily large number of different mixed exponential distributions, each with an arbitrarily large number of positive w_i's, will maximize the loglikelihood function. However, a perfect fit is highly unlikely unless the number of groups is very small.

When the fit is not perfect, Polyá and Szegö's result ensures that $dh/d\lambda$ has at most $g-2$ zeros. Thus, when the KKT conditions are satisfied, $h(\lambda)$ can assume the value m on $(0,\infty)$ at no more than $g/2-1$ points if g is even and no more than $g/2-1/2$ points if g is odd. This places a bound on the number of positive w_i's with corresponding λ_i's on $(0,\infty)$ at the point that the loglikelihood function is at its maximum. In addition, it is possible that the w_i's corresponding to λ_i's of zero and infinity may be positive.

We now move to the proof of uniqueness. Let $\lambda_1,\ldots,\lambda_n$ be the points at which the w_i's are positive where the loglikelihood is at its maximum. If more than one set of w_i's maximized the loglikelihood, each would have to give the same value of $\sum_{i=1}^{n} w_i(e^{-\lambda_i b_{k-1}} - e^{-\lambda_i b_k})$ for each group with a nonzero number of observations (where b_{k-1} and b_k are the group boundaries). Since adjacent groups with zero observations have been combined, the minimum number of such groups will be $g/2$ if g is even and $g/2-1/2$ if g is odd. Therefore, $\sum_{i=1}^{n}(w_i - w_i^*)(e^{-\lambda_i b_{k-1}} - e^{-\lambda_i b_k})$ has to be zero for each of these groups. This implies that, for each group, the function $\sum_{i=1}^{n}(w_i - w_i^*)e^{-\lambda_i x}$ has the same value at both b_{k-1} and b_k. Thus the derivative of this function must be zero somewhere between b_{k-1} and b_k. Therefore, the function $\sum_{i=1}^{n}(w_i - w_i^*)\lambda_i e^{-\lambda_i x}$ must have at least $g/2$ zeros if g is even and at least $g/2-1/2$ zeros if g is odd. From Polyá and Szegö's result, this function can have no more than n^*-1 zeros, where n^* is the number of λ_i's at which the w_i's are positive, excluding λ_i's of zero and infinity (since these terms drop out of the function). Since we have already determined that $n^* \leq g/2 - 1$ if g is even and $n^* \leq g/2 - 1/2$ if g is odd, we have a contradiction. We thus conclude that the loglikelihood attains its maximum at a unique set of w_i's.[17]

[17]Using a more general technique, Lindsay and Roeder [13] derived similar results to those for grouped data shown here. Those results apply to mixtures of a broader class of distributions.

APPENDIX C

Use of Newton's method requires calculation of the gradient vector of first partial derivatives and the Hessian matrix of second partial derivatives of the loglikelihood function.

In the derivatives that follow, w_1 is not a real parameter, but we set w_1 equal to one minus the sum of the other w_i's.[18]

$$\left(\frac{\partial \ln L}{\partial \lambda_i}\right)_k \qquad \text{and} \qquad \left(\frac{\partial \ln L}{\partial w_i}\right)_k$$

refer to the terms of the first partial derivatives corresponding to the kth observation (for ungrouped data) or kth group (for grouped data).

For ungrouped data, the required derivatives are

$$\frac{\partial \ln L}{\partial \lambda_i} = \sum_{k=1}^{m}\left(\frac{\partial \ln L}{\partial \lambda_i}\right)_k = \sum_{k=1}^{m}\frac{w_i(1-\lambda_i x_k)e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j\lambda_j e^{-\lambda_j x_k}},$$

$$i = 1,\ldots,n,$$

$$\frac{\partial \ln L}{\partial w_i} = \sum_{k=1}^{m}\left(\frac{\partial \ln L}{\partial w_i}\right)_k = \sum_{k=1}^{m}\frac{\lambda_i e^{-\lambda_i x_k} - \lambda_1 e^{-\lambda_1 x_k}}{\sum_{j=1}^{n} w_j\lambda_j e^{-\lambda_j x_k}},$$

$$i = 2,\ldots,n,$$

$$\frac{\partial^2 \ln L}{\partial \lambda_i^2} = \sum_{k=1}^{m}\left[\frac{w_i x_k(\lambda_i x_k - 2)e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j\lambda_j e^{-\lambda_j x_k}} - \left(\left(\frac{\partial \ln L}{\partial \lambda_i}\right)_k\right)^2\right],$$

$$i = 1,\ldots,n,$$

[18] An alternative way to formulate the problem would be to keep w_1 as a parameter and use a Lagrange multiplier to ensure that the sum of the w_i's is one.

$$\frac{\partial^2 \ln L}{\partial\lambda_i \partial\lambda_l} = \sum_{k=1}^{m} \left[-\left(\frac{\partial \ln L}{\partial\lambda_i}\right)_k \left(\frac{\partial \ln L}{\partial\lambda_l}\right)_k \right],$$

$$i = 1,\ldots,n, \quad l = 1,\ldots,n, \quad i \neq l,$$

$$\frac{\partial^2 \ln L}{\partial w_i \partial w_l} = \sum_{k=1}^{m} \left[-\left(\frac{\partial \ln L}{\partial w_i}\right)_k \left(\frac{\partial \ln L}{\partial w_l}\right)_k \right],$$

$$i = 2,\ldots,n, \quad l = 2,\ldots,n,$$

$$\frac{\partial^2 \ln L}{\partial\lambda_1 \partial w_i} = \sum_{k=1}^{m} \left[\frac{-(1-\lambda_1 x_k)e^{-\lambda_1 x_k}}{\sum_{j=1}^{n} w_j \lambda_j e^{-\lambda_j x_k}} - \left(\frac{\partial \ln L}{\partial\lambda_1}\right)_k \left(\frac{\partial \ln L}{\partial w_i}\right)_k \right],$$

$$i = 2,\ldots,n,$$

$$\frac{\partial^2 \ln L}{\partial\lambda_i \partial w_i} = \sum_{k=1}^{m} \left[\frac{(1-\lambda_i x_k)e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j \lambda_j e^{-\lambda_j x_k}} - \left(\frac{\partial \ln L}{\partial\lambda_i}\right)_k \left(\frac{\partial \ln L}{\partial w_i}\right)_k \right],$$

$$i = 2,\ldots,n,$$

and
$$\frac{\partial^2 \ln L}{\partial\lambda_i \partial w_l} = \sum_{k=1}^{m} \left[-\left(\frac{\partial \ln L}{\partial\lambda_i}\right)_k \left(\frac{\partial \ln L}{\partial w_l}\right)_k \right],$$

$$i = 2,\ldots,n, \quad l = 2,\ldots,n, \quad i \neq l.$$

For grouped data, the required derivatives are

$$\frac{\partial \ln L}{\partial\lambda_i} = \sum_{k=1}^{g} a_k \left(\frac{\partial \ln L}{\partial\lambda_i}\right)_k = \sum_{k=1}^{g} a_k \frac{w_i(-b_{k-1}e^{-\lambda_i b_{k-1}} + b_k e^{-\lambda_i b_k})}{\sum_{j=1}^{n} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})},$$

$$i = 1,\ldots,n,$$

$$\frac{\partial \ln L}{\partial w_i} = \sum_{k=1}^{g} a_k \left(\frac{\partial \ln L}{\partial w_i}\right)_k = \sum_{k=1}^{g} a_k \frac{(e^{-\lambda_i b_{k-1}} - e^{-\lambda_i b_k}) - (e^{-\lambda_1 b_{k-1}} - e^{-\lambda_1 b_k})}{\sum_{j=1}^{n} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})},$$

$$i = 2,\ldots,n,$$

$$\frac{\partial^2 \ln L}{\partial \lambda_i^2} = \sum_{k=1}^{g} a_k \left[\frac{w_i (b_{k-1}^2 e^{-\lambda_i b_{k-1}} - b_k^2 e^{-\lambda_i b_k})}{\sum_{j=1}^{n} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})} - \left(\left(\frac{\partial \ln L}{\partial \lambda_i}\right)_k\right)^2\right],$$

$$i = 1,\ldots,n,$$

$$\frac{\partial^2 \ln L}{\partial \lambda_i \partial \lambda_l} = \sum_{k=1}^{g} a_k \left[-\left(\frac{\partial \ln L}{\partial \lambda_i}\right)_k \left(\frac{\partial \ln L}{\partial \lambda_l}\right)_k\right],$$

$$i = 1,\ldots,n, \quad l = 1,\ldots,n, \quad i \neq l,$$

$$\frac{\partial^2 \ln L}{\partial w_i \partial w_l} = \sum_{k=1}^{g} a_k \left[-\left(\frac{\partial \ln L}{\partial w_i}\right)_k \left(\frac{\partial \ln L}{\partial w_l}\right)_k\right],$$

$$i = 2,\ldots,n, \quad l = 2,\ldots,n,$$

$$\frac{\partial^2 \ln L}{\partial \lambda_1 \partial w_i} = \sum_{k=1}^{g} a_k \left[\frac{(b_{k-1} e^{-\lambda_1 b_{k-1}} - b_k e^{-\lambda_1 b_k})}{\sum_{j=1}^{n} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})} - \left(\frac{\partial \ln L}{\partial \lambda_1}\right)_k \left(\frac{\partial \ln L}{\partial w_i}\right)_k\right],$$

$$i = 2,\ldots,n,$$

$$\frac{\partial^2 \ln L}{\partial \lambda_i \partial w_i} = \sum_{k=1}^{g} a_k \left[\frac{(-b_{k-1} e^{-\lambda_i b_{k-1}} + b_k e^{-\lambda_i b_k})}{\sum_{j=1}^{n} w_j (e^{-\lambda_j b_{k-1}} - e^{-\lambda_j b_k})} - \left(\frac{\partial \ln L}{\partial \lambda_i}\right)_k \left(\frac{\partial \ln L}{\partial w_i}\right)_k\right],$$

$$i = 2,\ldots,n,$$

$$\text{and} \qquad \frac{\partial^2 \ln L}{\partial \lambda_i \partial w_l} = \sum_{k=1}^{g} a_k \left[-\left(\frac{\partial \ln L}{\partial \lambda_i}\right)_k \left(\frac{\partial \ln L}{\partial w_l}\right)_k \right],$$

$$i = 2,\ldots,n, \quad l = 2,\ldots,n, \quad i \neq l.$$

The Newton step is the inverse of the Hessian matrix multiplied by the negative of the gradient vector. To remove one of the parameters from the iterative process without reconstructing the entire gradient and Hessian, set that parameter's component of the gradient to zero, its diagonal element of the Hessian matrix to one, and the off-diagonal elements of its row and column of the Hessian matrix to zero.

With ungrouped data, the fitted mixed exponential mean will always equal the sample mean at both the global maximum and at local maxima. To see this, first note that each of the $\partial \ln L/\partial w_i$ values must be zero, so the KKT equalities are satisfied. We have seen that this implies that

$$\sum_{k=1}^{m} \frac{\lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j \lambda_j e^{-\lambda_j x_k}} = m, \qquad i = 1,\ldots,n.$$

Since each of the $\partial \ln L/\partial \lambda_i$ values must be zero, we may sum over them to obtain

$$\sum_{i=1}^{n}\sum_{k=1}^{m} \frac{w_i(1-\lambda_i x_k)e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j \lambda_j e^{-\lambda_j x_k}} = \sum_{i=1}^{n}\left[w_i \frac{1}{\lambda_i} \sum_{k=1}^{m} \frac{\lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j \lambda_j e^{-\lambda_j x_k}} \right]$$

$$- \sum_{k=1}^{m}\left[x_k \sum_{i=1}^{n} \frac{w_i \lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j \lambda_j e^{-\lambda_j x_k}} \right] = m \sum_{i=1}^{n} w_i \frac{1}{\lambda_i} - \sum_{k=1}^{m} x_k = 0.$$

Since $1/\lambda_i$ is the mean of the ith exponential distribution in the mixture, we can see that the mixed exponential mean must indeed be equal to the sample mean.

Also, with ungrouped data, the fitted mixed exponential variance will not be less than the sample variance at the global maximum. To see this, first note that at each of the λ_i's with positive weight attached, $d^2h/d\lambda^2$ must be less than or equal to zero. We may sum over these second derivatives, giving weight w_i to each element of the sum, to obtain

$$\sum_{i=1}^{n}\sum_{k=1}^{m}\frac{w_i(\lambda_i x_k^2 - 2x_k)e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j\lambda_j e^{-\lambda_j x_k}} = \sum_{k=1}^{m}\left[x_k^2\sum_{i=1}^{n}\frac{w_i\lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j\lambda_j e^{-\lambda_j x_k}}\right]$$

$$-\sum_{i=1}^{n}\left[w_i\frac{2}{\lambda_i}\sum_{k=1}^{m}\frac{\lambda_i x_k e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j\lambda_j e^{-\lambda_j x_k}}\right]$$

$$=\sum_{k=1}^{m}x_k^2 - \sum_{i=1}^{n}\left[w_i\frac{2}{\lambda_i}\sum_{k=1}^{m}\frac{e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j\lambda_j e^{-\lambda_j x_k}}\right]$$

$$=\sum_{k=1}^{m}x_k^2 - \sum_{i=1}^{n}\left[w_i\frac{2}{\lambda_i^2}\sum_{k=1}^{m}\frac{\lambda_i e^{-\lambda_i x_k}}{\sum_{j=1}^{n} w_j\lambda_j e^{-\lambda_j x_k}}\right]$$

$$=\sum_{k=1}^{m}x_k^2 - m\sum_{i=1}^{n}w_i\frac{2}{\lambda_i^2} \leq 0.$$

To get from the term in the second line above to the second term in the third line, we use the fact that each of the $\partial \ln L/\partial \lambda_i$ values must be zero. Since $2/\lambda_i^2$ is the second moment of the *i*th exponential distribution in the mixture, and since we know that the mixed exponential mean must be equal to the sample mean, we can see that the mixed exponential variance cannot be less than the sample variance.[19]

[19]Lindsay [9] showed that these moment relationships hold for mixtures of a broader class of distributions.

DOWNWARD BIAS OF USING HIGH-LOW AVERAGES FOR LOSS DEVELOPMENT FACTORS

CHENG-SHENG PETER WU

Abstract

This paper extends previous research that studied the downward bias associated with high-low averages, which occurs when high-low averages are applied to data that exhibits a long-tailed property. The current study conducted a comprehensive review of insurance industry data when three-of-five averages are used to determine the age-to-age development factors in setting reserves. The downward bias was analyzed by line of business, premium size, development age, paid and incurred loss development methods, for one hundred and forty paid and incurred loss triangles from seventy insurance companies/groups compiled from the A.M. Best database. The study assumes that the age-to-age development factors are lognormally distributed. The three-of-five average was selected as the representative high-low average because it is commonly used by property/casualty actuaries. The results for this average can be generalized to other types of high-low averages. The results given in the paper are based on a bias formula for a large volume of data. Since the real-world loss development data is limited in volume, the study used large scale simulations to review the effect of limited volume data on the bias.

1. INTRODUCTION

1.A. Downward Bias of Using High-Low Averages for Age-to-Age Factors

Property/casualty actuaries often employ an averaging technique that excludes the same number of observations, split equally between the lowest and highest ranking observations. These averages will be called the high-low averages in this paper. One common application of the averages is the selection of loss development factors.

There are many types of high-low averages, for example, the middle three of the latest five years (three-of-five averages) and the middle six of the latest eight quarters (six-of-eight averages).

The purpose of using high-low averages is to exclude outliers and their disproportional influence on the results. Exclusion of observations requires a great deal of caution, however. According to Neter, Wasserman, and Kutner [8]:

> "… an outlying influential case should not be automatically discarded, because it may be entirely correct and simply represents an unlikely event. Discarding of such an outlying case could lead to the undesirable consequences of increased variances of some of the estimated regression coefficients."

In other words, systematic exclusion of high and low data points would lead to less statistically significant and, hence, less credible estimators.

Moreover, the distribution of insurance loss data exhibits unsymmetrical behavior of skewing toward the right (higher values). This is called the *long-tailed* property. Most typical insurance claims are small amount claims, probably less than a few thousand dollars. However, the remaining small number of claims can have very large losses. For example, automobile large loss claims will reach a few hundred thousand dollars, while medical malpractice or environmental claims can even be multimillion-dollar claims in today's legal climate. Therefore, long-tailed distributions such as lognormal, Pareto, and gamma distributions are better in describing the loss data than the symmetric normal distribution because they reflect the large loss probability. Exhibit 1 shows graphically a lognormal distribution and its long-tailed property of skewing to the right.

Applying high-low averages to loss development factors will result in a systematic downward bias when the loss development data exhibits a *long-tailed* property. This can be illustrated through the following example based on a lognormal assumption.

First, assume that:

- At development age i, the aggregate reported loss or paid loss is equal to L_i.
- From age i to $i+1$, a total loss of l_{i+1} is reported or paid.
- Since insurance losses have a long-tailed property, both L_i and l_{i+1} can be represented by lognormal distributions. If this is the case, then both $\ln(L_i)$ and $\ln(l_{i+1})$ are normally distributed. For the use of lognormal distributions to approximate insurance losses, please see Bowers, et al. [2], Finger [3], and Hogg and Klugman [5].

Based on these assumptions, the age-to-age development factor from age i to $i+1$ can be expressed as follows:

$$D_{i,i+1} = (L_i + l_{i+1})/L_i = 1 + l_{i+1}/L_i.$$

Since the multiplication or division result of two lognormally distributed variables also has a lognormal distribution, $1 + l_{i+1}/L_i$ and $D_{i,i+1}$ are lognormally distributed and should have a long-tailed property:

$$\ln(D_{i,i+1}) \sim N(\mu_i, \sigma_i^2),$$

where μ_i is the mean and σ_i^2 is the variance of the normal distribution for $\ln(D_{i,i+1})$.

One advantage of assuming lognormal distributions for the age-to-age development factors is that the age-to-ultimate factors and, consequently, the ultimate loss estimates are also lognormally distributed:

$$UD_i = D_{i,i+1} \times D_{i+1,i+2} \times D_{i+2,i+3} \times \cdots,$$

where

$$\ln(UD_i) = \ln(D_{i,i+1}) + \ln(D_{i+1,i+2}) + \ln(D_{i,i+1}) + \cdots$$

and

$$\ln(UD_i) \sim N(\mu_i + \mu_{i+1} + \mu_{i+2} + \cdots, \sigma_i^2 + \sigma_{i+1}^2 + \sigma_{i+2}^2 + \cdots).$$

The fact that age-to-age development factors may have a long tail has been noted previously. Hayne's study [4], in quantifying the variability of loss reserves, assumes that age-to-age development factors are lognormally distributed. Kelly [6] and McNichols [7] also conclude that a lognormal assumption is better in describing age-to-age development factors than a normal assumption, based on the fact that lognormal distributions can take only positive values and their long-tailed property reflects the distinct possibility of large development factors.

However, if $D_{i,i+1}$ is lognormally distributed, using high-low averages to estimate $D_{i,i+1}$ will result in a downward bias. Bias is defined as the percentage difference between the mean and the conditional mean, given that the data lie between a specified lower and upper pair of percentile points. The bias is expressed in the following formula whose detailed derivations can be found in the Appendix:

$$\begin{aligned} \text{Bias} &= \frac{\mathrm{E}(D_{i,i+1})'}{\mathrm{E}(D_{i,i+1})} - 1 \\ &= \frac{1}{(1-2p)}[\Phi(\Phi^{-1}(1-p) - \sigma_i) - \Phi(\Phi^{-1}(p) - \sigma_i)] - 1, \end{aligned} \tag{1.1}$$

where:

$E(D_{i,i+1})$ is the expected value of $D_{i,i+1}$,

$E(D_{i,i+1})'$ is the expected value of $D_{i,i+1}$, given that $D_{i,i+1}$ lies between its upper and lower p percentile points

$$\left(\text{i.e.,}\quad \frac{1}{1-2p}\int_{d1}^{d2} t \times f(t)\,dt\right),$$

$f(d)$ is the probability distribution function for $D_{i,i+1}$,

$F(d)$ is the cumulative distribution function for $D_{i,i+1}$,

p represents percentile,

d_1 is the value of $D_{i,i+1}$ when $F(d) = p$,

d_2 is the value of $D_{i,i+1}$ when $F(d) = 1-p$,

and

$\Phi(X)$ is the standard normal distribution function,

$$\int_{-\infty}^{X} \frac{\exp(\frac{1}{2}t^2)}{\sqrt{2\pi}}\,dt.$$

Equation (1.1) indicates that the degree of bias depends only on p and σ_i, the percentage of data being excluded and the shape parameter, but not on μ_i, the location parameter. This suggests that the more data excluded or the more skewed and volatile the distribution, the higher the downward bias is. Exhibit 1 illustrates the downward bias graphically.

Note that we are not limited to only the lognormal assumption. For example, one other commonly used long-tailed distribution is the Pareto distribution. The bias formula similar to Equation (1.1) for the Pareto distribution is also derived in the Appendix. Further analysis indicates that for the age-to-age development factors reviewed in this study, there is no significant difference in the bias result between the lognormal distribution and the Pareto distribution.

1.B. Modified High-Low Averages for the Correction of Downward Bias

Results from Equation (1.1) can be extended to the high-low averages used by property/casualty actuaries. For example, a three-of-five average also excludes the upper and lower 20% of the data. The only difference is that the high-low average is based on a limited volume of data (five data points) and a sample distribution function, while Equation (1.1) is based on a very large volume of data and a cumulative distribution function.

Equation (1.1) provides a basis to correct the bias for the sample high-low average:

$$\text{Modified High-Low Average} = \text{Sample High-Low Average}/(1+\text{Bias}), \qquad (1.2)$$

where the bias is given in Equation (1.1).

Exhibits 2 to 5 display how to correct the downward bias for the three-of-five averages based on Equations (1.1) and (1.2). This example uses product liability paid loss data for a sample company from the A. M. Best database [1].

Exhibit 2 shows two types of averages: five-year straight averages and three-of-five averages. These are factor averages, not volume-weighted averages. Because the data has 10 years of experience, the three-of-five averages can be applied to only the first five development ages. After the fifth development age, all-years averages are used.

The tail factor of 1.0261 selected in Exhibit 2 should be noted. This factor is the ratio of incurred loss to paid loss for the earliest year in the triangle. No further tail development is assumed. The choice of the tail factor will not affect the relative bias level because it is a constant that will be multiplied by the age-to-age development factors.

Results from Exhibit 2 clearly indicate that the five-year averages result in higher estimates than the three-of-five averages. This is consistent with the assumption that age-to-age loss development factors have a long-tailed property.

Fitting lognormal distributions to the age-to-age development factors in Exhibit 2 produces the parameter estimates in Exhibit 3. First, μ_i and σ_i^2 are estimated for each development period. All of the data in each development period are used to estimate these sample parameters, although only the latest five data points are used to select the age-to-age development factors. This approach is used to increase the credibility of the sample parameters. Then, the parameters for the age-to-ultimate development factors for a development age are the sum of all the parameters of the age-to-age factors from that age to ultimate.

Given these lognormal parameter estimates, the three-of-five averages in Exhibit 2 can be modified to correct the downward bias for the averages. The modified three-of-five factors are given in Exhibit 4. For example, the lognormal parameters for the 12-to-24 development factors are: $\mu_1 = 1.9221$, and $\sigma_1^2 = 0.3057$. With $p = 20\%$, a bias of -11.33% is indicated for the three-of-five average based on Equation (1.1).

Exhibit 4 shows the indicated bias for each development period and the modified three-of-five averages. Exhibit 5 compares the estimated ultimate losses and reserves between the five-year averages, the three-of-five averages, and the modified three-of-five averages. For example, the total reserve for the three-of-five averages is approximately 12.0% lower than the reserve for the five-year averages, and is 8.9% lower than the reserve for the modified three-of-five averages. Exhibit 5 does not show the results for the oldest five accident years since there is no difference among methods for these five accident years.

This specific example is for product liability paid loss data. The results of the comprehensive review, testing the biases with differing data volumes, differing lines of business, and paid and incurred loss data will be shown in later sections.

1.C. Limited Volume Data

As mentioned previously, the bias formula given in Equation (1.1) is based on a very large volume of data and a cumulative distribution function, while the real-world data is limited in volume.

Two issues in dealing with a limited volume of data should be noted. First, additional parameter variation is introduced because sample parameters are assumed in place of true parameters. Therefore, when Equation (1.1) is used to estimate the level of bias of real-world data, sample parameters, not the true parameters, are generally used. For example, in Exhibits 3 and 4, the lognormal parameters, $\mu_1 = 1.9221$ and $\sigma_1^2 = 0.3057$, for the 12-to-24 development factors, distribution are based on the nine sample data points in the 12-to-24 development period. We assumed these parameters were the true parameters when the -11.33% of downward bias was indicated by Equation (1.1).

Second, even if the true parameters are known, the indicated bias when sample size is small will not be the same as the indicated bias when sample size is large. For example, Equation (1.1) provides an accurate estimate of bias if 20% of high and low data are excluded from a data set of, for example, a million data points. However, when a three-of-five average is used to estimate the loss development factors, 20% of the high and low data are excluded from a data set of only five data points.

Resolving these limited volume data issues through statistical methods is very difficult, if not impossible, and is beyond the scope of this study. Instead, large scale simulations have been conducted and the simulation results will be presented in the later sections.

2. CURRENT STUDY

2.A. Purposes

The previous section illustrates the potential bias of using high-low averages for loss development factors, and more details can be found in Wu [9]. In light of these results, however, many outstanding questions remain to be answered:

- Do the real-world loss development factors really exhibit a long-tailed property?
- What is the level of the downward bias when the high-low averages are used in setting reserves?
- How does the downward bias vary by line of business, data volume, development age, and between paid and incurred loss development methods?
- What is the effect of limited volume data on the bias?

This study attempts to answer these questions through a comprehensive review of industry data and large scale simulations.

2.B. Data

Data from the A.M. Best database [1] were gathered for the following seven major liability lines:

- workers compensation;
- private passenger automobile liability;
- commercial automobile liability;
- medical malpractice, occurrence;
- medical malpractice, claims-made;
- product liability; and
- other liability.

For each line of business, paid loss and incurred loss triangles on an annual basis were compiled from ten randomly selected insurance companies/groups. In general, the same ten companies were not used for each line of business, but a few companies were repeatedly selected. A total of one hundred and forty triangles were collected. The loss triangles have ten years of experience and cover the period from 1986 to 1995.

The collected data were further broken down into two groups based on the volume of the data. One group, Group A, contains large multi-line and multi-state companies, while the other group, Group B, contains small local and regional companies. Exhibit 6 shows the range of the annual earned premium for the companies within each groups.

2.C. Review Approach

The loss development procedures used to review the A. M. Best data are the same as the procedures given in Exhibits 2 to 5. The following list summarizes the important assumptions in the approach:

- The three-of-five average was selected as the representative high-low average. The results for that average can be extended to other types of high-low averages.
- Due to the fact that the collected loss triangle data have only ten years of history, the three-of-five averages can be applied to only the first five development ages. For the development ages after 72 months, all-years averages were used.
- There is no tail development assumed for the incurred loss method. For the paid tail, the ratio of incurred to paid loss for the oldest accident year in the triangle was used.
- All data points in each development period were used to calculate the lognormal parameters. This was done to increase the credibility of the sample parameters. However, only the lat-

est five points were used to select the age-to-age development factors.

- Large scale simulations were conducted to study the effect of a limited of volume data on the bias when sample parameters are assumed as the true parameters. The simulations also measure the differences between the simulated bias and the bias based on Equation [1].

3. RESULTS AND DISCUSSION

3.A. Long-Tailed Property for Age-to-Age Development Factors

First, the reserve indications for the five-year averages and the three-of-five are compared. Exhibit 6 gives the comparison results by line of business, company size, and paid versus incurred methods.

Exhibit 6 indicates that approximately 70% of the data reviewed show lower reserve indications for the three-of-five averages. This is consistent with the assumption that the age-to-age development factors may have a long tail and the use of high-low averages will result in a downward bias.

Exhibit 6 further indicates that the long tail assumption is more valid for the more volatile lines such as medical malpractice and product liability. On the other hand, the assumption is equally valid for both large and small groups, and for both incurred and paid methods.

3.B. Results by Line of Business

Exhibits 7 to 13 give two types of downward bias by line of business: the bias for the age-to-age development factors and the bias for the reserve indications. The tests were conducted on both the total reserve and the incurred but not reported reserve (IBNR). In each exhibit, the downward bias is indicated by company size and paid versus incurred methods.

The indicated bias given in these exhibits is based on Equation (1.1). For example, Exhibit 11 shows that for the malpractice claims-made data of the large companies in Group A, the indicated minimum, maximum, and average downward biases associated with the three-of-five averages for the 12–24 paid factors are 0.86%, 2.88%, and 2.06%, respectively.

The bias for the reserve indications is the difference between indications based on the three-of-five averages and the modified three-of-five averages. For example, Exhibit 11 shows that for the malpractice claims-made data of the large companies in Group A, the indicated minimum, maximum, and average downward bias for the total reserves for the paid method are 0.61%, 2.86%, and 1.87%, respectively.

From Exhibits 7 to 13, the following observations can be made:

- The indicated bias for the age-to-age factors decreases as the loss data become mature. For workers compensation, private passenger automobile liability, and commercial automobile liability, the bias appears to be insignificant after 72 months of development. On the other hand, the bias is still noticeable after 72 months for medical malpractice, product liability, and other liability.
- The indicated bias for the reserve indications can be substantial, especially for the highly volatile lines such as medical malpractice, product liability, and other liability. The use of high-low averages can easily lead to a downward bias of over 10% for these lines of business.
- In general, the data of small companies shows higher downward bias than the data of large companies. This is because the age-to-age factors become more volatile as the volume of the data decreases.
- There is no systematic difference in the bias level between the paid and incurred factors. At a first glance, this result is some-

what surprising and counterintuitive, because paid loss development factors are larger and more leveraged than incurred loss development factors. However, most internal and external factors, such as claim processing, late reported claims, inflation, underwriting cycles, and economic cycles, affect both paid and incurred loss development factors. As indicated in Equation (1.1), the bias depends on the skewness and volatility of the data, as represented by σ_i, but not on the level or the magnitude of the data, as represented by μ_i. Further research indicates that the sample paid loss factors and incurred loss factors used in the study have similar degrees of skewness. For example, the averages of the sample σ for 12–24 paid and incurred factors for product liability data are not very different, 0.518 and 0.563, respectively.

3.C. Large Scale Simulations for the Limited Volume Data

As mentioned before, in theory, we need to have an infinitely large amount of loss development data in order to apply Equation (1.1) in calculating the downward bias of high-low averages. The real-world data is limited and, therefore, will deviate somewhat from the asymptotic assumptions underlying Equation (1.1). As a result, there are two issues when Equation (1.1) is used with a limited volume of data. First, true means and variances are usually unknown, and sample means and variances from the data need to be used. Second, Equation (1.1) calculates the bias when one assumes that the data volume is very large, while the three-of-five average, for example, uses only five data points.

In order to study the limited volume data effect, we designed a large scale simulation test. The simulation procedures and results are as follows:

1. A set of μ_i and σ_i are selected. The range for μ_i is between 0.1 and 2.0 and the range for σ_i is between 0.002 to 1.2. These ranges are based on the A. M. Best data reviewed in the study. See Exhibits 14 and 15 for the

selected combinations of μ_i and σ_i. These selected combinations of μ_i and σ_i represent the true parameters of the underlying distribution for the simulations.

2. 4,000 lognormal observations based on the selected μ_i and σ_i are generated. Each observation contains five random data points.

3. For each observation, the sample parameters from the five random data points are calculated. The bias using Equation (1.1) with the sample parameters is calculated. The bias result is compared to the bias based on the true parameters of μ_i and σ_i. Since the sample parameters are different from the true parameters of μ_i and σ_i, the bias based on the sample parameters may be higher or lower than the bias based on the true parameters. This is the effect of the use of the sample parameters. Exhibit 14 shows the comparison based on the overall 4,000 generated observations. The result indicates that the bias based on the sample parameters on average will be lower than the bias based on the true parameters. For example, when $\sigma_i = 1.2$ and $\mu_i = 1.0$, the bias on average will be understated by 8.5% for the sample parameters.

4. Finally, for each observation, the three-of-five average is calculated by excluding the lowest and highest data points. The three-of-five average is compared to the expected average of the lognormal distribution with the selected σ_i and μ_i to obtain the downward bias. The downward bias for the observation is compared to the expected downward bias based on Equation (1.1) with the selected σ_i, μ_i, and $p = 20\%$. This is the effect of the limited volume of data since the bias for each of the observations is based on only five data points, while the bias based on Equation (1.1) is based on a large volume of data. Exhibit 15 shows that the bias is tempered somewhat for the limited volume data. For example, when $\sigma_i = 1.2$ and

$\mu_i = 1.0$, the simulated bias for the three-of-five on average is approximately 67.5% of the bias calculated by Equation (1.1) for a large volume of data.

Exhibits 14 and 15 also show that the effects of the limited volume of data on the bias depend primarily on σ_i, not on μ_i. The effects diminish quickly as σ_i decreases.

Please note that the two effects in Exhibits 14 and 15 are separately studied because, in theory, the effect of sample parameters may not exist. This occurs when there is prior knowledge of the true values for μ and σ. With known μ and σ, there still exists the effect for limited sample size as given in Exhibit 15 when only five data points are used to calculate the three-of-five averages.

3.D. Summary of the Results

The current study presents strong evidence, through a comprehensive review of property and casualty insurance industry data, that downward bias will occur when high-low averages are used to determine age-to-age development factors. The review results show the level of the bias by line of business, development age, premium size, and paid versus incurred methods. The results indicate that the downward bias can be substantial, especially for small companies and highly volatile lines.

Equations (1.1) and (1.2) provide a basis to quantify and correct the bias. Equation (1.1) is based on a large volume of data, while only a limited volume of data is available for most real-world applications. The simulation results show that the bias for the limited volume of data, on average, is somewhat lower than what is indicated by Equation (1.1).

4. CONCLUSIONS

Many property and casualty actuaries are undoubtedly aware of the downward bias associated with the high-low averages. While this study focuses on the loss development application, the results and implications should go beyond that application, and can be extended to many other actuarial applications if the underlying data shows a long-tail property.

Also, the real-world data that actuaries deal with daily may have even higher levels of bias than indicated in this study for the following reasons:

- The bias will increase if less mature data or quarterly and semi-annual data are used.
- Due to the data limitation, the results given in this study only include the bias for the first five development periods and real-world data would allow a more thorough bias analysis beyond the fifth development age.
- The bias is demonstrated and quantified through the lognormal assumption in this study. The assumption may understate the thickness of the tail for insurance data (see Hogg and Klugman [5]). If the tail of the loss development factors distribution is more skewed than what is suggested by the lognormal distribution, the bias will be higher than indicated by Equation (1.1).

As usual, many assumptions used in the current study are ideal. Attempts to study the bias under more complicated assumptions are beyond the scope of the current study because they require advanced statistical knowledge. They can be topics for future research, however. For example, nonparametric methods may be used to explain the effects of limited volume. Another interesting topic would be to study the bias when loss development factors are highly correlated between development periods.

Finally, it should be noted that this paper does not attempt to suggest the high-low averaging approach be completely excluded from consideration by actuaries. The paper does attempt to indicate the potential bias if the approach is applied to insurance data on a comprehensive basis without an in-depth understanding of the data. The principle that no one arithmetic approach is superior to or inferior to all others will not and should not be altered by the results given in the paper. Perhaps, the key message delivered by the paper is the need for even more substantial professional judgment by actuaries in promulgating reserving and pricing estimates.

REFERENCES

[1] A. M. Best, *Best's Casualty Loss Reserve Development Series*., Best Database Services, 1996.

[2] Bowers, N. L., et al., *Actuarial Mathematics*, Society of Actuaries, 1986.

[3] Finger, R. J., "Estimating Pure Premiums by Layer—An Approach," *PCAS* LXIII, 1976, pp. 34–52.

[4] Hayne, R. M., "An Estimate of Statistical Variation in Development Factor Methods," *PCAS* LXXII, 1985, pp. 25–43.

[5] Hogg, R. V., and S. A. Klugman, *Loss Distributions*, John Wiley & Sons, Inc., 1984.

[6] Kelly, M. V., "Practical Loss Reserving Method with Stochastic Development Factors," *Insurer Financial Solvency*, Casualty Actuarial Society Discussion Paper Program, 1992, pp. 355–381.

[7] McNichols, J. P., "Simplified Confidence Boundaries Associated with Calendar Year Projections," *Insurer Financial Solvency*, Casualty Actuarial Society Discussion Paper Program, 1992, pp. 465–509.

[8] Neter, J., W. Wasserman, and M. H. Kutner, *Applied Liner Regression Models* (2nd edition), Richard D. Irwin, Inc., 1989.

[9] Wu, C. P., "Bias of Excluding High and Low Data for Long-Tailed Distributions," *Journal of Actuarial Practice* 4, 1996, pp. 143–158.

EXHIBIT 1

SHAPE OF A TYPICAL LOGNORMAL DISTRIBUTION WITH DOWNWARD BIAS OF HIGH-LOW AVERAGE FOR A LOGNORMAL DISTRIBUTION

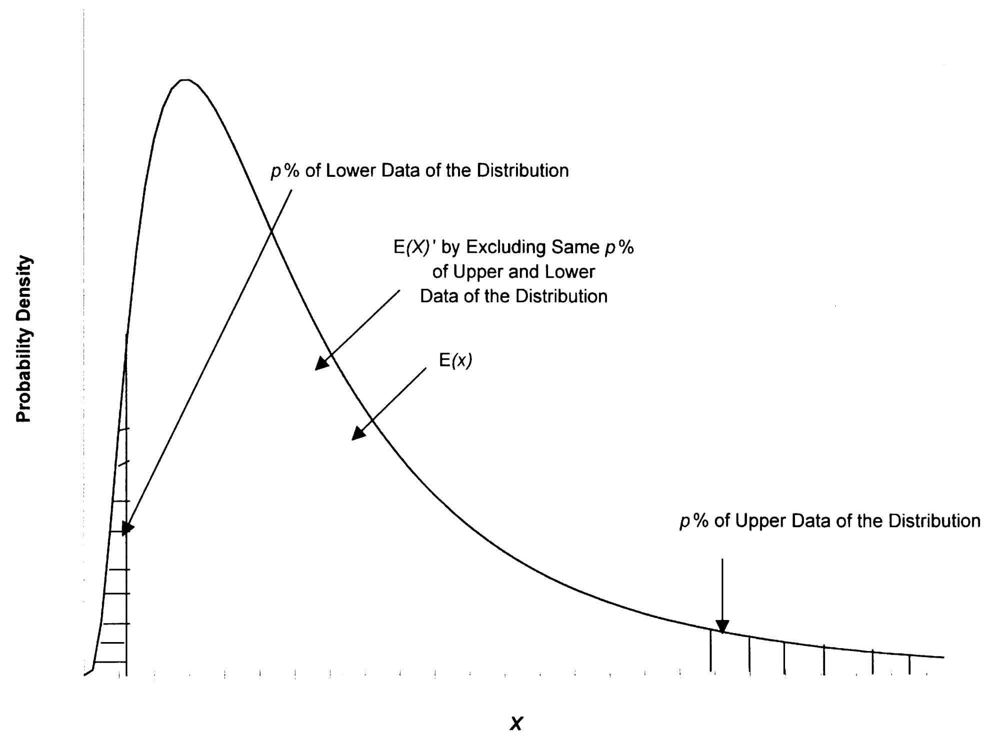


EXHIBIT 2

PRODUCT LIABILITY PAID LOSS AND LOSS DEVELOPMENT FACTOR TRIANGLES*

Paid Losses: (in Thousands)

Accident Year	Earned Premium	Development Period, Months 12	24	36	48	60	72	84	96	108	120
1986	$ 55,779	$ 446	$ 1,618	$ 4,685	$ 7,809	$13,722	$17,849	$18,240	$18,742	$19,076	$19,244
1987	$ 60,737	$ 61	$ 1,336	$ 3,341	$ 6,377	$ 9,596	$11,662	$12,876	$13,301	$13,909	
1988	$ 75,602	$ 302	$ 3,326	$ 6,804	$14,516	$16,254	$17,918	$21,169	$22,000		
1989	$ 82,764	$ 414	$ 3,228	$ 6,125	$12,994	$17,298	$23,505	$25,491			
1990	$103,688	$ 415	$ 3,111	$11,406	$18,249	$20,738	$23,537				
1991	$116,481	$1,747	$ 8,037	$23,063	$33,663	$41,467					
1992	$128,505	$2,956	$14,907	$30,199	$43,563						
1993	$137,629	$2,064	$12,249	$25,737							
1994	$153,565	$2,764	$12,746								
1995	$170,085	$3,232									

Age-to-Age Factors:

Accident Year	Earned Premium	Development Period, Months 12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	Tail**
1986	$ 55,779	3.6250	2.8966	1.6667	1.7571	1.3008	1.0219	1.0275	1.0179	1.0088	
1987	$ 60,737	22.0000	2.5000	1.9091	1.5048	1.2152	1.1042	1.0330	1.0457		
1988	$ 75,602	11.0000	2.0455	2.1333	1.1198	1.1023	1.1814	1.0393			
1989	$ 82,764	7.8000	1.8974	2.1216	1.3312	1.3589	1.0845				
1990	$103,688	7.5000	3.6667	1.6000	1.1364	1.1350					
1991	$116,481	4.6000	2.8696	1.4596	1.2318						
1992	$128,505	5.0435	2.0259	1.4426							
1993	$137,629	5.9333	2.1011								
1994	$153,565	4.6111									
1995	$170,085										
Age-to-Age Development Factors:											
5-Year Average***		5.5376	2.5121	1.7514	1.2648	1.2224	1.0980	1.0333	1.0318	1.0088	1.0261
3-of-5 Average***		5.1960	2.3322	1.7271	1.2331	1.2170	1.0980	1.0333	1.0318	1.0088	1.0261
Age-to-Ultimate Development Factors:											
5-Year Average***		45.6427	8.2423	3.2810	1.8733	1.4811	1.2116	1.1035	1.0680	1.0351	1.0261
3-of-5 Average***		38.0553	7.3240	3.1404	1.8183	1.4746	1.2116	1.1035	1.0680	1.0351	1.0261

*This is product liability paid loss data from a sample company in the A.M. Best database [1].
**The tail factor of 1.0261 is the ratio of incurred to paid loss for 1986.
***For the last four development periods, straight averages are used in place of the specified averages.

EXHIBIT 3

LOGNORMAL PARAMETERS FOR LOSS DEVELOPMENT FACTORS

Natural Logarithm Transformation of the Age-to-Age Factors in Exhibit 2: Accident Year	Development Period, Months 12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120
1986	1.2879	1.0635	0.5108	0.5637	0.2630	0.0216	0.0272	0.0177	0.0087
1987	3.0910	0.9163	0.6466	0.4086	0.1949	0.0991	0.0325	0.0447	
1988	2.3979	0.7156	0.7577	0.1131	0.0974	0.1667	0.0385		
1989	2.0541	0.6405	0.7522	0.2861	0.3066	0.0811			
1990	2.0149	1.2993	0.4700	0.1278	0.1266				
1991	1.5261	1.0542	0.3782	0.2085					
1992	1.6181	0.7060	0.3664						
1993	1.7806	0.7425							
1994	1.5285								
1995									
Age-to-Age Development Factors:									
Estimated Mu	1.9221	0.8922	0.5546	0.2846	0.1977	0.0921	0.0327	0.0312	0.0087
Estimated Sigma Square	0.3057	0.0534	0.0274	0.0306	0.0078	0.0036	0.0000		
Age-to-Ultimate Development Factors:									
Estimated Mu	4.0160	2.0939	1.2017	0.6471	0.3625	0.1648	0.0726	0.0399	0.0087
Estimated Sigma Square	0.4284	0.1228	0.0694	0.0420	0.0114	0.0036	0.0000	0.0000	0.0000

EXHIBIT 4

MODIFIED HIGH-LOW AVERAGES FOR LOSS DEVELOPMENT FACTORS

Age-to-Age Factors in Exhibit 2: Accident Year	12–24	24–36	36–48	48–60	60–72	72–84	84–96	96–108	108–120	
1986	3.6250	2.8966	1.6667	1.7571	1.3008	1.0219	1.0275	1.0179	1.0088	
1987	22.0000	2.5000	1.9091	1.5048	1.2152	1.1042	1.0330	1.0457		
1988	11.0000	2.0455	2.1333	1.1198	1.1023	1.1814	1.0393			
1989	7.8000	1.8974	2.1216	1.3312	1.3589	1.0845				
1990	7.5000	3.6667	1.6000	1.1364	1.1350					
1991	4.6000	2.8696	1.4596	1.2318						
1992	5.0435	2.0259	1.4426							
1993	5.9333	2.1011								
1994	4.6111									
1995										
Age-to-Age Development Factors:										Tail
5-Year Average	5.5376	2.5121	1.7514	1.2648	1.2224	1.0980	1.0333	1.0318	1.0088	1.0261
Lognormal Parameters from Exhibit 3:										
Estimated Mu	1.9221	0.8922	0.5546	0.2846	0.1977	0.0921	0.0327	0.0312	0.0087	
Estimated Sigma Square	0.3057	0.0534	0.0274	0.0306	0.0078	0.0036	0.0000	0.0000	0.0000	
3-of-5 Average	5.1960	2.3322	1.7271	1.2331	1.2170	1.0980	1.0333	1.0318	1.0088	1.0261
% of High and Low Data Excluded	20.0%	20.0%	20.0%	20.0%	20.0%					
Indicated Downward Bias	−11.33%	−2.07%	−1.07%	−1.19%	−0.31%					
Modified 3-of-5 Average	5.8598	2.3816	1.7458	1.2480	1.2207	1.0980	1.0333	1.0318	1.0088	1.0261
Age-to-Ultimate Development Factors:										
5-Year Average	45.6427	8.2423	3.2810	1.8733	1.4811	1.2116	1.1035	1.0680	1.0351	1.0261
3-of-5 Average	38.0553	7.3240	3.1404	1.8183	1.4746	1.2116	1.1035	1.0680	1.0351	1.0261
Modified 3-of-5 Average	44.9738	7.6750	3.2226	1.8460	1.4791	1.2116	1.1035	1.0680	1.0351	1.0261

Development Period, Months (spanning 12–24 through 108–120)

EXHIBIT 5

COMPARISON OF ULTIMATE LOSSES AND RESERVES ACROSS DIFFERENT AVERAGING TECHNIQUES

Age-to-Ult Loss Development Factors

Accident Year	Undeveloped Paid Losses	5-Year Average	3-of-5 Average	Modified 3-of-5 Average
1991	$ 41,467	1.4811	1.4746	1.4791
1992	$ 43,563	1.8733	1.8183	1.8460
1993	$ 25,737	3.2810	3.1404	3.2226
1994	$ 12,746	8.2423	7.3240	7.6750
1995	$ 3,232	45.6427	38.0553	44.9738
Total:	$126,745			

Ultimate Losses

Accident Year	5-Year Average	3-of-5 Average	Modified 3-of-5 Average	Difference 3-of-5 and 5-Year	Difference 3-of-5 and Mod 3-of-5
1991	$ 61,419	$ 61,146	$ 61,334	−0.4%	−0.3%
1992	$ 81,609	$ 79,213	$ 80,417	−2.9%	−1.5%
1993	$ 84,442	$ 80,823	$ 82,940	−4.3%	−2.6%
1994	$105,056	$ 93,351	$ 97,824	−11.1%	−4.6%
1995	$147,500	$122,980	$145,338	−16.6%	−15.4%
Total:	$480,026	$437,513	$467,853	−8.9%	−6.5%

Total Reserves

1991	$ 19,952	$ 19,679	$ 19,867	−1.4%	−0.9%
1992	$ 38,046	$ 35,649	$ 36,854	−6.3%	−3.3%
1993	$ 58,706	$ 55,087	$ 57,203	−6.2%	−3.7%
1994	$ 92,310	$ 80,605	$ 85,078	−12.7%	−5.3%
1995	$144,268	$119,748	$142,106	−17.0%	−15.7%
Total:	$353,281	$310,768	$341,108	−12.0%	−8.9%

EXHIBIT 6

A.M. BEST DATA

	Number of Companies Sampled	Annual Earned Premium from 1986 to 1995 (in Millions) Minimum	Maximum	Average	Data with Lower Reserve Indications for 3-of-5 Averages* Paid Loss Method	Incurred Loss Method
Group A: Multistate, Multiline Insurance Companies/Groups:						
Workers Compensation	5	$426	$ 1,823	$1,029	3	3
Private Passenger Automobile Liability	5	$543	$14,126	$3,651	2	3
Commercial Automobile Liability	5	$151	$ 682	$ 354	2	2
Medical Malpractice-Occurrence	5	$ 14	$ 270	$ 71	5	5
Medical Malpractice-Claims-Made	5	$ 44	$ 700	$ 186	4	3
Product Liability	5	$ 43	$ 218	$ 115	5	5
Other Liability	5	$199	$ 1,221	$ 611	3	3
Total	35				24	24
Group B: Regional or Single State Insurance Companies:						
Workers Compensation	5	$ 14	$ 137	$ 60	2	3
Private Passenger Automobile Liability	5	$ 26	$ 122	$ 62	3	3
Commercial Automobile Liability	5	$ 19	$ 99	$ 47	3	2
Medical Malpractice-Occurrence	5	$ 2	$ 53	$ 17	5	5
Medical Malpractice-Claims-Made	5	$ 20	$ 64	$ 39	5	3
Product Liability	5	$ 5	$ 50	$ 29	5	5
Other Liability	5	$ 12	$ 98	$ 54	5	3
Total	35				28	24
Group A and Group B Combined:						
Workers Compensation	10	$ 14	$ 1,823	—	5	6
Private Passenger Automobile Liability	10	$ 26	$14,126	—	5	6
Commercial Automobile Liability	10	$ 19	$ 682	—	5	4
Medical Malpractice-Occurrence	10	$ 2	$ 270	—	10	10
Medical Malpractice-Claims-Made	10	$ 20	$ 700	—	9	6
Product Liability	10	$ 5	$ 218	—	10	10
Other Liability	10	$ 12	$ 1,221	—	8	6
Total	70				52	48

*Reserve indications were compared between 5-year averages and 3-of-5 averages. This is the data where 3-of-5 averages have a lower reserve indication.

EXHIBIT 7

REVIEW RESULTS OF A.M. BEST WORKERS COMPENSATION DATA

Indicated Downward Bias for 3-of-5 Age-to-Age Factors*:

Paid 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.25%	0.00%	0.00%	0.00%	0.00%
	Maximum	−2.68%	−0.10%	−0.02%	−0.02%	0.00%
	Average	−0.80%	−0.05%	−0.01%	−0.01%	0.00%
Group B-Small to Medium Companies	Miminum	−0.03%	−0.02%	−0.01%	0.00%	0.00%
	Maximum	−0.72%	−0.22%	−0.20%	−0.09%	−0.09%
	Average	−0.25%	−0.07%	−0.05%	−0.02%	−0.02%

Incurred 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.10%	−0.01%	−0.01%	−0.01%	−0.01%
	Maximum	−0.78%	−0.06%	−0.05%	−0.03%	−0.02%
	Average	−0.37%	−0.03%	−0.02%	−0.02%	−0.01%
Group B-Small to Medium Companies	Miminum	−0.07%	−0.02%	−0.01%	0.00%	0.00%
	Maximum	−1.07%	−0.16%	−0.13%	−0.05%	−0.02%
	Average	−0.57%	−0.10%	−0.05%	−0.02%	−0.01%

Indicated Downward Bias for 3-of-5 Reserve Indications:**

		Paid Loss Development Method		Incurred Loss Development Method	
		Total Reserves	IBNR Reserves	Total Reserves	IBNR Reserves
Group A-Large Companies	Miminum	−0.05%	−0.11%	−0.11%	−0.32%
	Maximum	−1.37%	−2.92%	−0.30%	−0.77%
	Average	−0.37%	−0.85%	−0.22%	−0.54%
Group B-Small to Medium Companies	Miminum	−0.06%	−0.15%	−0.09%	−0.32%
	Maximum	−1.37%	−3.63%	−0.73%	−1.73%
	Average	−0.38%	−0.96%	−0.45%	−1.04%

*The indicated downward bias for 3-of-5 factors is based on Equation (1.1).

**The indicated downward bias for reserves is the difference in reserve indications between 3-of-5 averages and modified 3-of-5 averages.

EXHIBIT 8

REVIEW RESULTS OF A.M. BEST PRIVATE PASSENGER AUTOMOBILE LIABILITY DATA

Indicated Downward Bias for 3-of-5 Age-to-Age Factors*:

Paid 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.04%	0.00%	0.00%	0.00%	0.00%
	Maximum	−0.22%	−0.01%	−0.02%	0.00%	0.00%
	Average	−0.08%	−0.01%	−0.01%	0.00%	0.00%
Group B-Small to Medium Companies	Miminum	−0.09%	−0.01%	0.00%	0.00%	0.00%
	Maximum	−0.39%	−0.14%	−0.04%	−0.03%	−0.02%
	Average	−0.16%	−0.04%	−0.02%	−0.01%	0.00%

Incurred 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.01%	0.00%	0.00%	0.00%	0.00%
	Maximum	−0.14%	−0.02%	−0.01%	0.00%	0.00%
	Average	−0.06%	−0.01%	0.00%	0.00%	0.00%
Group B-Small to Medium Companies	Miminum	−0.03%	−0.02%	−0.01%	0.00%	0.00%
	Maximum	−0.20%	−0.07%	−0.04%	−0.03%	−0.01%
	Average	−0.13%	−0.04%	−0.02%	−0.01%	0.00%

Indicated Downward Bias for 3-of-5 Reserve Indications:**

		Paid Loss Development Method		Incurred Loss Development Method	
		Total Reserves	IBNR Reserves	Total Reserves	IBNR Reserves
Group A-Large Companies	Miminum	−0.04%	−0.08%	−0.03%	−0.08%
	Maximum	−0.17%	−0.38%	−0.13%	−1.93%
	Average	−0.08%	−0.21%	−0.08%	−0.56%
Group B-Small to Medium Companies	Miminum	−0.11%	−0.59%	−0.08%	−0.23%
	Maximum	−0.59%	−1.55%	−2.31%	−7.39%
	Average	−0.27%	−0.98%	−0.60%	−2.14%

*The indicated downward bias for 3-of-5 factors is based on Equation (1.1).

**The indicated downward bias for reserves is the difference in reserve indications between 3-of-5 averages and modified 3-of-5 averages.

EXHIBIT 9

Review Results of A.M. Best Commercial Automobile Liability Data

Indicated Downward Bias for 3-of-5 Age-to-Age Factors*:

Paid 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A–Large Companies	Miminum	−0.03%	−0.01%	0.00%	0.00%	0.00%
	Maximum	−1.77%	−0.16%	−0.03%	−0.04%	−0.03%
	Average	−0.56%	−0.07%	−0.02%	−0.02%	−0.01%
Group B-Small to Medium Companies	Miminum	−0.10%	−0.15%	−0.02%	−0.01%	0.00%
	Maximum	−1.21%	−0.43%	−0.13%	−0.07%	−0.07%
	Average	−0.46%	−0.22%	−0.07%	−0.03%	−0.02%

Incurred 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.04%	−0.02%	−0.01%	0.00%	0.00%
	Maximum	−0.91%	−0.18%	−0.11%	−0.06%	−0.02%
	Average	−0.35%	−0.07%	−0.04%	−0.02%	−0.01%
Group B-Small to Medium Companies	Miminum	−0.18%	−0.04%	−0.01%	−0.01%	0.00%
	Maximum	−0.48%	−0.21%	−0.06%	−0.06%	−0.01%
	Average	−0.31%	−0.09%	−0.02%	−0.03%	−0.01%

Indicated Downward Bias for 3-of-5 Reserve Indications:**

		Paid Loss Development Method		Incurred Loss Development Method	
		Total Reserves	IBNR Reserves	Total Reserves	IBNR Reserves
Group A-Large Companies	Miminum	−0.06%	−0.09%	−0.07%	−0.12%
	Maximum	−1.31%	−2.17%	−0.92%	−1.85%
	Average	−0.50%	−0.98%	−0.29%	−0.66%
Group B-Small to Medium Companies	Miminum	−0.39%	−0.77%	−0.22%	−0.59%
	Maximum	−1.32%	−7.32%	−3.63%	−10.37%
	Average	−0.78%	−2.66%	−1.11%	−4.57%

*The indicated downward bias for 3-of-5 factors is based on Equation (1.1).

**The indicated downward bias for reserves is the difference in reserve indications between 3-of-5 averages and modified 3-of-5 averages.

EXHIBIT 10

REVIEW RESULTS OF A.M. BEST MEDICAL MALPRACTICE—OCCURRENCE DATA

Indicated Downward Bias for 3-of-5 Age-to-Age Factors*:

Paid 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−3.84%	−2.02%	−0.34%	−0.12%	−0.09%
	Maximum	−23.00%	−14.79%	−6.24%	−2.01%	−1.55%
	Average	−14.39%	−5.46%	−1.96%	−0.87%	−0.52%
Group B-Small to Medium Companies	Miminum	−10.30%	−1.22%	−0.51%	−0.10%	−0.14%
	Maximum	−22.99%	−9.79%	−2.25%	−2.73%	−0.99%
	Average	−15.75%	−5.75%	−1.31%	−0.92%	−0.37%
Incurred 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.68%	−0.57%	−0.28%	−0.12%	−0.07%
	Maximum	−30.02%	−21.60%	−2.70%	−1.33%	−0.69%
	Average	−12.84%	−7.67%	−0.96%	−0.52%	−0.30%
Group B-Small to Medium Companies	Miminum	−0.23%	−0.32%	−0.27%	−0.14%	−0.07%
	Maximum	−7.88%	−6.08%	−1.33%	−3.61%	−4.88%
	Average	−4.69%	−2.09%	−0.93%	−1.13%	−1.52%

Indicated Downward Bias for 3-of-5 Reserve Indications:**

		Paid Loss Development Method		Incurred Loss Development Method	
		Total Reserves	IBNR Reserves	Total Reserves	IBNR Reserves
Group A-Large Companies	Miminum	−3.19%	−9.40%	−1.14%	−5.43%
	Maximum	−13.49%	−24.56%	−60.72%	−68.37%
	Average	−9.92%	−15.81%	−17.50%	−22.11%
Group B-Small to Medium Companies	Miminum	−4.19%	−8.20%	−0.76%	−1.35%
	Maximum	−18.89%	−39.64%	−43.22%	−283.92%
	Average	−13.58%	−23.67%	−16.65%	−91.94%

*The indicated downward bias for 3-of-5 factors is based on Equation (1.1).
**The indicated downward bias for reserves is the difference in reserve indications between 3-of-5 averages and modified 3-of-5 averages.

EXHIBIT 11

REVIEW RESULTS OF A.M. BEST MEDICAL MALPRACTICE—CLAIMS-MADE DATA

Indicated Downward Bias for 3-of-5 Age-to-Age Factors*:

Paid 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.86%	−0.10%	−0.08%	−0.04%	0.00%
	Maximum	−2.88%	−0.44%	−0.63%	−0.60%	−0.22%
	Average	−2.06%	−0.28%	−0.26%	−0.21%	−0.12%
Group B-Small to Medium Companies	Miminum	−1.45%	−0.39%	−0.11%	−0.05%	−0.01%
	Maximum	−6.95%	−2.31%	−1.04%	−0.24%	−0.78%
	Average	−4.49%	−1.30%	−0.39%	−0.10%	−0.21%

Incurred 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.27%	−0.19%	−0.12%	−0.03%	0.00%
	Maximum	−2.33%	−0.94%	−0.44%	−0.24%	−0.06%
	Average	−0.95%	−0.44%	−0.27%	−0.11%	−0.03%
Group B-Small to Medium Companies	Miminum	−0.49%	−0.07%	−0.07%	−0.04%	−0.03%
	Maximum	−1.45%	−0.36%	−0.32%	−0.26%	−0.54%
	Average	−0.98%	−0.26%	−0.17%	−0.12%	−0.16%

Indicated Downward Bias for 3-of-5 Reserve Indications:**

		Paid Loss Development Method		Incurred Loss Development Method	
		Total Reserves	IBNR Reserves	Total Reserves	IBNR Reserves
Group A-Large Companies	Miminum	−0.61%	−3.10%	−0.52%	−1.27%
	Maximum	−2.86%	−13.79%	−3.64%	−8.05%
	Average	−1.87%	−6.82%	−1.41%	−4.99%
Group B-Small to Medium Companies	Miminum	−3.05%	−3.90%	−1.41%	−0.94%
	Maximum	−4.28%	−68.72%	−7.36%	−51.12%
	Average	−3.89%	−20.40%	−3.01%	−13.71%

*The indicated downward bias for 3-of-5 factors is based on Equation (1.1).

**The indicated downward bias for reserves is the difference in reserve indications between 3-of-5 averages and modified 3-of-5 averages.

EXHIBIT 12

REVIEW RESULTS OF A.M. BEST PRODUCT LIABILITY DATA

Indicated Downward Bias for 3-of-5 Age-to-Age Factors*:

Paid 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−2.44%	−1.45%	−1.02%	−0.30%	−0.16%
	Maximum	−42.19%	−35.08%	−10.36%	−2.04%	−7.65%
	Average	−17.40%	−9.39%	−2.93%	−1.00%	−1.73%
Group B-Small to Medium Companies	Miminum	−1.44%	−0.70%	−0.13%	−0.16%	−0.03%
	Maximum	−13.52%	−5.34%	−3.33%	−1.72%	−0.90%
	Average	−7.08%	−2.59%	−1.19%	−0.62%	−0.26%

Incurred 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−1.42%	−1.00%	−0.17%	−0.18%	−0.09%
	Maximum	−27.35%	−17.13%	−2.49%	−3.51%	−4.15%
	Average	−18.17%	−7.00%	−1.34%	−1.02%	−1.15%
Group B-Small to Medium Companies	Miminum	−4.23%	−0.85%	−0.50%	−0.34%	−0.06%
	Maximum	−21.73%	−6.71%	−4.27%	−3.70%	−1.83%
	Average	−9.84%	−3.34%	−2.64%	−1.70%	−0.73%

Indicated Downward Bias for 3-of-5 Reserve Indications:**

		Paid Loss Development Method		Incurred Loss Development Method	
		Total Reserves	IBNR Reserves	Total Reserves	IBNR Reserves
Group A-Large Companies	Miminum	−3.04%	−6.11%	−1.94%	−4.63%
	Maximum	−68.50%	−77.61%	−39.88%	−45.01%
	Average	−22.20%	−27.14%	−22.00%	−28.70%
Group B-Small to Medium Companies	Miminum	−1.59%	10.47%	−1.55%	−5.61%
	Maximum	−5.82%	−15.54%	−12.89%	−35.88%
	Average	−3.26%	−1.52%	−5.48%	−22.19%

*The indicated downward bias for 3-of-5 factors is based on Equation (1.1).
**The indicated downward bias for reserves is the difference in reserve indications between 3-of-5 averages and modified 3-of-5 averages.

EXHIBIT 13

REVIEW RESULTS OF A.M. BEST OTHER LIABILITY DATA

Indicated Downward Bias for 3-of-5 Age-to-Age Factors*:

Paid 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.30%	−0.12%	−0.04%	−0.05%	−0.02%
	Maximum	−21.90%	−2.04%	−0.41%	−0.21%	−0.23%
	Average	−7.16%	−0.63%	−0.17%	−0.12%	−0.09%
Group B-Small to Medium Companies	Miminum	−1.03%	−0.40%	−0.17%	−0.03%	−0.02%
	Maximum	−8.18%	−3.97%	−4.41%	−0.67%	−0.24%
	Average	−2.98%	−2.28%	−1.29%	−0.33%	−0.10%

Incurred 3-of-5 Averages		12–24 Months	24–36 Months	36–48 Months	48–60 Months	60–72 Months
Group A-Large Companies	Miminum	−0.12%	−0.09%	−0.03%	−0.02%	−0.01%
	Maximum	−3.31%	−0.59%	−0.16%	−0.11%	−0.10%
	Average	−1.23%	−0.29%	−0.09%	−0.07%	−0.05%
Group B-Small to Medium Companies	Miminum	−0.42%	−0.38%	−0.07%	−0.05%	−0.02%
	Maximum	−21.96%	−2.24%	−1.53%	−0.50%	−0.32%
	Average	−8.06%	−0.87%	−0.47%	−0.20%	−0.12%

Indicated Downward Bias for 3-of-5 Reserve Indications:**

		Paid Loss Development Method		Incurred Loss Development Method	
		Total Reserves	IBNR Reserves	Total Reserves	IBNR Reserves
Group A-Large Companies	Miminum	−0.70%	−0.91%	−0.46%	−0.80%
	Maximum	−11.64%	−27.59%	−1.99%	−2.85%
	Average	−3.90%	−8.33%	−1.01%	−1.76%
Group B-Small to Medium Companies	Miminum	−1.47%	−3.91%	−1.14%	−2.08%
	Maximum	−14.28%	−21.24%	−9.29%	−18.19%
	Average	−5.27%	−8.29%	−4.45%	−9.50%

*The indicated downward bias for 3-of-5 factors is based on Equation (1.1).
**The indicated downward bias for reserves is the difference in reserve indications between 3-of-5 averages and modified 3-of-5 averages.

EXHIBIT 14

EFFECT OF SAMPLE PARAMETERS
RATIO OF AVERAGE BIAS
BASED ON SIMULATED SAMPLE PARAMETERS VS. TRUE PARAMETERS

	μ			
σ	2.000	1.000	0.500	0.100
1.200	90.6%	91.5%	91.2%	91.8%
0.900	93.2%	93.2%	94.9%	94.1%
0.500	97.5%	97.7%	97.3%	97.9%
0.100	99.5%	99.9%	99.5%	99.6%
0.050	100.2%	98.8%	100.4%	100.9%
0.002	99.4%	100.6%	100.9%	97.9%

EXHIBIT 15

EFFECT OF LIMITED SAMPLE SIZE
RATIO OF SIMULATED BIAS TO BIAS BASED ON EQUATION (1.1) FOR THREE-OF-FIVE AVERAGES

	μ			
σ	2.000	1.000	0.500	0.100
1.200	68.3%	67.5%	67.4%	67.1%
0.900	80.7%	80.2%	80.6%	80.6%
0.500	93.1%	92.8%	93.6%	93.8%
0.100	99.8%	99.8%	99.9%	99.7%
0.050	99.9%	99.9%	99.9%	99.9%
0.002	100.0%	100.0%	100.0%	100.0%

APPENDIX

DOWNWARD BIAS FOR TWO LONG-TAILED DISTRIBUTIONS

This Appendix shows the derivations of the downward bias based on the cumulative distribution functions for two long-tailed distributions, lognormal and Pareto. Many of the details of these two distributions can be found in Hogg and Klugman [5] or other statistical texts.

First, the following list specifies the global notations for the two distributions:

$E(X)$: expected value for random variable X;

$E(X)'$: expected value of X when excluding the upper $p\%$ and lower $p\%$ of data;

$F(x)$: cumulative probability function;

$f(x)$: probability density function;

p: percentile;

x_1: value of X when $F(x) = p$;

x_2: value of X when $F(x) = 1 - p$;

Φ: standard normal distribution function $= \int_{-\infty}^{x} \frac{\exp(\frac{1}{2}x^2)}{\sqrt{2\pi}} dx$;

ϕ: standard normal density function $= \exp(\frac{1}{2}x^2)/\sqrt{2\pi}$.

A.1. Lognormal Distribution

a. Probability Density Function:

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

b. Cumulative Probability Function:

$$F(x) = \int_0^\infty \frac{\exp\left(\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right)}{x\sigma\sqrt{2\pi}} dx.$$

Let

$$x = e^{\sigma y + \mu}, \qquad \text{then} \quad y = \frac{\ln x - \mu}{\sigma}, \quad \text{and} \quad dx = e^{\sigma y + \mu}\sigma\, dy.$$

$$F(x) = \int_{-\infty}^{\ln x - \mu/\sigma} \frac{e^{-y^2/2} e^{\sigma y + \mu}\sigma}{e^{\sigma y + \mu}\sigma\sqrt{2\pi}} dy = \Phi\left(\frac{\ln x - \mu}{\sigma}\right).$$

$$F(x_1) = \Phi\left(\frac{\ln x_1 - \mu}{\sigma}\right) = p, \qquad x_1 = e^{(\Phi^{-1}(p)\sigma + \mu)}.$$

$$F(x_2) = \Phi\left(\frac{\ln x_2 - \mu}{\sigma}\right) = 1 - p, \qquad x_2 = e^{(\Phi^{-1}(1-p)\sigma + \mu)}.$$

c. Expected Value of X:

$$\mathrm{E}(X) = \int_0^\infty x \frac{e^{-1/2(\ln x - \mu/\sigma)^2}}{x\sigma\sqrt{2\pi}} dx = \int_0^\infty \frac{e^{-1/2(\ln x - \mu/\sigma)^2}}{\sigma\sqrt{2\pi}} dx.$$

Let

$$y = \frac{\ln x - \mu - \sigma^2}{\sigma}, \qquad \text{then} \quad x = e^{\sigma y + \mu + \sigma^2}, \quad \text{and}$$

$$dx = e^{\sigma y + \mu + \sigma^2}\sigma\, dy.$$

$$\mathrm{E}(X) = \int_0^\infty \frac{e^{-1/2(y+\sigma)^2} e^{\sigma y + \mu + \sigma^2}\sigma}{\sigma\sqrt{2\pi}} dx$$

$$= e^{(\mu + (1/2)\sigma^2)} \int_0^\infty \frac{e^{-(1/2)y^2}}{\sqrt{2\pi}} dx = e^{(\mu + (1/2)\sigma^2)}.$$

d. Expected Value of X when Excluding Upper $p\%$ and Lower $p\%$ of Data:

$$\mathrm{E}(X)' = \int_{x_1}^{x_2} x \frac{e^{-1/2(\ln x - \mu/\sigma)^2}}{(1-2p)x\sigma\sqrt{2\pi}} dx = \int_{x_1}^{x_2} \frac{e^{-1/2(\ln x - \mu/\sigma)^2}}{(1-2p)\sigma\sqrt{2\pi}} dx.$$

Let

$$y = \frac{\ln x - \mu - \sigma^2}{\sigma}, \qquad \text{then} \quad x = e^{\sigma y + \mu + \sigma^2}, \quad \text{and}$$

$$dx = e^{\sigma y + \mu + \sigma^2} \sigma \, dy.$$

$$\mathrm{E}(X)' = \frac{e^{(\mu + (1/2)\sigma^2)}}{(1-2p)} \int_{(\ln x_1 - \mu - \sigma^2)/\sigma}^{(\ln x_2 \mu - \sigma^2/\sigma} \frac{e^{-1/2y^2}}{\sqrt{2\pi}} dx$$

$$= \frac{e^{(\mu + (1/2)\sigma^2)}}{(1-2p)} \left(\Phi\left(\frac{\ln x_2 - \mu - \sigma^2)}{\sigma} \right) - \Phi\left(\frac{\ln x_1 - \mu - \sigma^2}{\sigma} \right) \right).$$

$$x_1 = e^{(\Phi^{-1}(p)\sigma + \mu)} \qquad \text{and} \qquad x_2 = e^{(\Phi^{-1}(1-p)\sigma + \mu)}, \quad \text{then}$$

$$\mathrm{E}(X)' = \frac{e^{(\mu + (1/2)\sigma^2)}}{(1-2p)} [\Phi(\Phi^{-1}(1-p) - \sigma)) - \Phi(\Phi^{-1}(p) - \sigma)].$$

e. Downward Bias for Excluding Upper $p\%$ and Lower $p\%$ of Data:

$$\text{Bias} = \frac{\mathrm{E}(x)'}{\mathrm{E}(x)} - 1$$

$$= \frac{1}{(1-2p)} [\Phi(\Phi^{-1}(1-p) - \sigma)) - \Phi(\Phi^{-1}(p) - \sigma)] - 1.$$

The above result indicates that the degree of bias depends on p, the percentage of data being excluded, and σ, the shape factor, only. The bias does not depend on μ, the location parameter.

A.2. Pareto Distribution

a. Probability Density Function:

$$f(x) = \alpha \lambda^{\alpha} (\lambda + x)^{-\alpha - 1}, \qquad x > 0.$$

b. Cumulative Probability Function:

$$F(x) = \int_0^x \alpha\lambda^\alpha(\lambda + x)^{-\alpha-1}\,dx = -\left(\frac{\lambda}{\lambda + x}\right)^\alpha\Bigg|_0^x = 1 - \left(\frac{\lambda}{\lambda + x}\right)^\alpha.$$

$$F(x_1) = p, \qquad \text{then} \quad x_1 = \lambda \times \left(\frac{1}{(1-p)^{1/\alpha}} - 1\right).$$

$$F(x_2) = 1 - p, \qquad \text{then} \quad x_2 = \lambda \times \left(\frac{1}{p^{1/\alpha}} - 1\right).$$

c. Expected Value of X:

$$\begin{aligned}
\mathrm{E}(X) &= \int_0^\infty x\alpha\lambda^\alpha(\lambda + x)^{-\alpha-1}\,dx = -\left(\frac{\lambda}{\lambda + x}\right)^\alpha x\Bigg|_0^\infty \\
&\quad + \int_0^\infty \lambda^\alpha(\lambda + x)^{-\alpha}\,dx \\
&= \int_0^\infty \lambda^\alpha(\lambda + x)^{-\alpha}\,dx = -\frac{\lambda}{\alpha - 1}\left(\frac{\lambda}{\lambda + x}\right)^{-(\alpha-1)}\Bigg|_0^\infty \\
&= \frac{\lambda}{\alpha - 1}.
\end{aligned}$$

d. Expected Value of X when Excluding Upper $p\%$ and Lower $p\%$ of Data:

$$\begin{aligned}
\mathrm{E}(X)' &= \int_{x_1}^{x_2} x\frac{\alpha\lambda^\alpha(\lambda + x)^{-\alpha-1}}{1 - 2p}dx = -x\,\frac{\left(\frac{\lambda}{\lambda + x}\right)^\alpha}{1 - 2p}\Bigg|_{x_1}^{x_2} \\
&\quad + \int_{x_1}^{x_2} \frac{\lambda^\alpha(\lambda + x)^{-\alpha}}{1 - 2p}dx \\
&= -x\,\frac{\left(\frac{\lambda}{\lambda + x}\right)^\alpha}{1 - 2p}\Bigg|_{x_1}^{x_2} - \frac{\lambda\left(\frac{\lambda}{\lambda + x}\right)^{(\alpha-1)}}{(\alpha - 1)(1 - 2p)}\Bigg|_{x_1}^{x_2}.
\end{aligned}$$

Since

$$\frac{\lambda}{\lambda + x_1} = \frac{\lambda}{\lambda + \lambda\left(\frac{1}{(1-p)^{1/\alpha}} - 1\right)} = (1-p)^{1/\alpha}, \qquad \text{and}$$

$$\frac{\lambda}{\lambda + x_2} = \frac{\lambda}{\lambda + \lambda\left(\frac{1}{p^{1/\alpha}} - 1\right)} = p^{1/\alpha},$$

then,

$$\begin{aligned} \mathrm{E}(X)' &= \frac{\lambda}{1-2p}\Bigg[-p^{\alpha-1/\alpha}(1-p^{1/\alpha}) + (1-p)^{\alpha-1/\alpha}(1-(1-p)^{1/\alpha}) \\ &\qquad - \frac{p^{\alpha-1/\alpha}}{\alpha-1} + \frac{(1-p)^{\alpha-1/\alpha}}{\alpha-1}\Bigg] \\ &= \frac{\lambda}{(\alpha-1)(1-2p)}[\alpha(-p^{\alpha-1/\alpha} + (1-p)^{\alpha-1/\alpha}) - (\alpha-1)(1-2p)]. \end{aligned}$$

e. Downward Bias for Excluding Upper $p\%$ and Lower $p\%$ of Data:

$$\begin{aligned} \text{Bias} &= \frac{\mathrm{E}(X)'}{\mathrm{E}(X)} - 1 \\ &= \frac{\alpha}{(1-2p)}[-p^{\alpha-1/\alpha} + (1-p)^{\alpha-1/\alpha} - (1-2p)]. \end{aligned}$$

Again, the degree of bias for Pareto distribution depends on p and α only, the percentage of excluded data and the shape factor, but not on λ, the location parameter.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXXIII

LOSS PREDICTION BY GENERALIZED LEAST SQUARES

LEIGH J. HALLIWELL

DISCUSSION BY KLAUS D. SCHMIDT

Abstract

In a recent paper on loss reserving, Halliwell suggests predicting outstanding claims by the method of generalized least squares applied to a linear model. An example is the linear model given by

$$E[Z_{i,k}] = \mu + \alpha_i + \gamma_k,$$

where $Z_{i,k}$ is the total claim amount of all claims which occur in year i and are settled in year $i + k$. The predictor proposed by Halliwell is known in econometrics but it is perhaps not well-known to actuaries. The present discussion completes and simplifies the argument used by Halliwell to justify the predictor; in particular, it is shown that there is no need to consider conditional distributions.

1. LOSS RESERVING

For $i,k \in \{0,1,\ldots,n\}$, let $Z_{i,k}$ denote the total claim amount of all claims which occur in year i and are settled in year $i+k$. We assume that the *incremental claims* $Z_{i,k}$ are observable for $i + k \leq n$ and that they are non-observable for $i + k > n$. The observable incremental claims are represented by the *run-off triangle* (Table 1).

The non-observable incremental claims are to be predicted from the observable ones. Whether or not certain predictors are

TABLE 1

Occurrence Year	Development Year 0	1	…	k	…	$n-i$	…	$n-1$	n
0	$Z_{0,0}$	$Z_{0,1}$	…	$Z_{0,k}$	…	$Z_{0,n-i}$	…	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1,0}$	$Z_{1,1}$	…	$Z_{1,k}$	…	$Z_{1,n-i}$	…	$Z_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$Z_{i,0}$	$Z_{i,1}$	…	$Z_{i,k}$	…	$Z_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$Z_{n-k,0}$	$Z_{n-k,1}$	…	$Z_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$Z_{n-1,0}$	$Z_{n-1,1}$							
n	$Z_{n,0}$								

preferable to others depends on the stochastic mechanism generating the data. It is thus necessary to first formulate a stochastic model and to fix the properties the predictors should have.

For example, we may assume that the incremental claims satisfy the *linear model* given by

$$E[Z_{i,k}] = \mu + \alpha_i + \gamma_k,$$

with real parameters $\mu, \alpha_0, \alpha_1, \ldots, \alpha_n, \gamma_0, \gamma_1, \ldots, \gamma_n$ such that $\sum_{i=0}^{n} \alpha_i = 0 = \sum_{k=0}^{n} \gamma_k$. This means that the expected incremental claims are determined by an overall mean μ and corrections α_i and γ_k depending on the *occurrence year* i and the *development year* k, respectively.

2. THE LINEAR MODEL WITH MISSING OBSERVATIONS

The model considered in the previous section is a special case of the linear model considered by Halliwell [2]:

Let $\mathbf{Y}$ be an $(m \times 1)$ random vector satisfying

$$E[\mathbf{Y}] = \mathbf{X}\boldsymbol{\beta}$$

and

$$\mathrm{Var}[\mathbf{Y}] = \mathbf{S}$$

for some known $(m \times k)$ design matrix $\mathbf{X}$, some unknown $(k \times 1)$ parameter vector $\boldsymbol{\beta}$, and some known $(m \times m)$ matrix $\mathbf{S}$ which is assumed to be positive definite.

We assume that some but not all coordinates of $\mathbf{Y}$ are observable. Without loss of generality, we may and do assume that the first p coordinates of $\mathbf{Y}$ are observable while the last $q := m - p$ coordinates of $\mathbf{Y}$ are non-observable. We may thus write

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix},$$

where $\mathbf{Y}_1$ consists of the observable coordinates of $\mathbf{Y}$, and $\mathbf{Y}_2$ consists of the non-observable coordinates of $\mathbf{Y}$. Accordingly, we partition the design matrix $\mathbf{X}$ into

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}.$$

We assume that

$$\mathrm{Rank}(\mathbf{X}_1) = k \leq p.$$

Then $\mathbf{X}$ has full rank and $\mathbf{X}'\mathbf{X}$ is invertible.

Following Halliwell, we partition $\mathbf{S}$ into

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix},$$

where

$$\begin{aligned} \mathbf{S}_{11} &:= \mathrm{Cov}[\mathbf{Y}_1, \mathbf{Y}_1] = \mathrm{Var}[\mathbf{Y}_1] \\ \mathbf{S}_{12} &:= \mathrm{Cov}[\mathbf{Y}_1, \mathbf{Y}_2] \\ \mathbf{S}_{21} &:= \mathrm{Cov}[\mathbf{Y}_2, \mathbf{Y}_1] \\ \mathbf{S}_{22} &:= \mathrm{Cov}[\mathbf{Y}_2, \mathbf{Y}_2] = \mathrm{Var}[\mathbf{Y}_2]. \end{aligned}$$

Then $\mathbf{S}_{11}$ and $\mathbf{S}_{22}$ are positive definite, and we also have $\mathbf{S}'_{21} = \mathbf{S}_{12}$. Moreover, $\mathbf{S}_{22} - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}$ is positive definite. Then $\mathbf{S}_{11}$ and $\mathbf{S}_{22} - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}$ are invertible, and there exist invertible matrices $\mathbf{A}$ and $\mathbf{D}$ satisfying

$$\mathbf{A}'\mathbf{A} = \mathbf{S}_{11}^{-1}$$

and

$$\mathbf{D}'\mathbf{D} = (\mathbf{S}_{22} - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12})^{-1}.$$

Define

$$\mathbf{C} := -\mathbf{D}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}$$

and let

$$\mathbf{W} := \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}.$$

Then we have

$$\mathbf{W}'\mathbf{W} = \mathbf{S}^{-1}.$$

In the following sections, we study the problem of estimating $\boldsymbol{\beta}$ and of predicting $\mathbf{Y}_2$ by estimators or predictors based on $\mathbf{Y}_1$.

3. ESTIMATION

Let us first consider the problem of estimating $\boldsymbol{\beta}$.

A random vector $\hat{\boldsymbol{\beta}}$ with values in $\mathbf{R}^k$ is

– a *linear estimator* (of $\boldsymbol{\beta}$) if it satisfies $\hat{\boldsymbol{\beta}} = \mathbf{B}\mathbf{Y}_1$ for some matrix $\mathbf{B}$,

– an *unbiased estimator* (of $\boldsymbol{\beta}$) if it satisfies $E[\hat{\boldsymbol{\beta}}] = \boldsymbol{\beta}$, and

– an *admissible estimator* (of $\boldsymbol{\beta}$) if it is linear and unbiased.

A linear estimator $\hat{\boldsymbol{\beta}} = \mathbf{B}\mathbf{Y}_1$ of $\boldsymbol{\beta}$ is unbiased if and only if $\mathbf{B}\mathbf{X}_1 = \mathbf{I}_k$.

A particular admissible estimator of $\boldsymbol{\beta}$ is the *Gauss–Markov estimator* $\boldsymbol{\beta}^*$, which is defined as

$$\boldsymbol{\beta}^* := (\mathbf{X}_1'\mathbf{S}_{11}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{S}_{11}^{-1}\mathbf{Y}_1.$$

Among all admissible estimators of $\boldsymbol{\beta}$, the Gauss–Markov estimator is distinguished due to the *Gauss–Markov Theorem*:

THEOREM 3.1 *The Gauss–Markov estimator $\boldsymbol{\beta}^*$ satisfies*

$$\operatorname{Var}[\boldsymbol{\beta}^*] = (\mathbf{X}_1'\mathbf{S}_{11}^{-1}\mathbf{X}_1)^{-1}.$$

Moreover, for each admissible estimator $\hat{\boldsymbol{\beta}}$, the matrix

$$\operatorname{Var}[\hat{\boldsymbol{\beta}}] - \operatorname{Var}[\boldsymbol{\beta}^*]$$

is positive semidefinite.

In a sense, the Gauss–Markov Theorem asserts that the Gauss–Markov estimator has minimal variance among all admissible estimators of $\boldsymbol{\beta}$. Since

$$\begin{aligned} E[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})] &= E[\operatorname{tr}((\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}))] \\ &= E[\operatorname{tr}((\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})')] \\ &= \operatorname{tr}(E[(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})']) \\ &= \operatorname{tr}(\operatorname{Var}[\hat{\boldsymbol{\beta}}]). \end{aligned}$$

we see that the Gauss–Markov estimator also minimizes the *expected quadratic estimation error* over all admissible estimators of $\boldsymbol{\beta}$.

4. PREDICTION

Let us now turn to the problem of predicting $\mathbf{Y}_2$.

A random vector $\hat{\mathbf{Y}}_2$ with values in $\mathbf{R}^q$ is

– a *linear predictor* (of $\mathbf{Y}_2$) if it satisfies $\hat{\mathbf{Y}}_2 = \mathbf{Q}\mathbf{Y}_1$ for some matrix $\mathbf{Q}$,

– an *unbiased predictor* (of $\mathbf{Y}_2$) if it satisfies $E[\hat{\mathbf{Y}}_2] = E[\mathbf{Y}_2]$, and

– an *admissible predictor* (of $\mathbf{Y}_2$) if it is linear and unbiased.

A linear predictor $\hat{\mathbf{Y}}_2 = \mathbf{Q}\mathbf{Y}_1$ of $\mathbf{Y}_2$ is unbiased if and only if $\mathbf{Q}\mathbf{X}_1 = \mathbf{X}_2$.

For an admissible estimator $\hat{\boldsymbol{\beta}}$, define

$$\mathbf{Y}_2(\hat{\boldsymbol{\beta}}) := \mathbf{X}_2\hat{\boldsymbol{\beta}} - \mathbf{D}^{-1}\mathbf{C}(\mathbf{Y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}})$$

and

$$\mathbf{h}(\hat{\boldsymbol{\beta}}) := -(\mathbf{C}\mathbf{X}_1 + \mathbf{D}\mathbf{X}_2)(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + (\mathbf{C}\mathbf{e}_1 + \mathbf{D}\mathbf{e}_2),$$

where $\mathbf{e}_1 := \mathbf{Y}_1 - \mathbf{X}_1\boldsymbol{\beta}$ and $\mathbf{e}_2 := \mathbf{Y}_2 - \mathbf{X}_2\boldsymbol{\beta}$. Then $\mathbf{Y}_2(\hat{\boldsymbol{\beta}})$ is an admissible predictor of $\mathbf{Y}_2$.

Following Halliwell, we have the following

LEMMA 4.1 *The identities*

$$\mathbf{Y}_2 = \mathbf{Y}_2(\hat{\boldsymbol{\beta}}) + \mathbf{D}^{-1}\mathbf{h}(\hat{\boldsymbol{\beta}})$$

as well as

$$E[\mathbf{h}(\hat{\boldsymbol{\beta}})] = \mathbf{0}$$

and

$$\mathrm{Var}[\mathbf{h}(\hat{\boldsymbol{\beta}})] = (\mathbf{C}\mathbf{X}_1 + \mathbf{D}\mathbf{X}_2)\mathrm{Var}[\hat{\boldsymbol{\beta}}](\mathbf{C}\mathbf{X}_1 + \mathbf{D}\mathbf{X}_2)' + \mathbf{I}_q$$

hold for each admissible estimator $\hat{\boldsymbol{\beta}}$; in particular, the matrix

$$\mathrm{Var}[\mathbf{h}(\hat{\boldsymbol{\beta}})] - \mathrm{Var}[\mathbf{h}(\boldsymbol{\beta}^*)]$$

is positive semidefinite.

From the last assertion of Lemma 4.1, which is a consequence of the Gauss–Markov theorem, Halliwell concludes that the *Gauss–Markov predictor* $\mathbf{Y}_2(\boldsymbol{\beta}^*)$ is the best unbiased linear predictor of $\mathbf{Y}_2$. This conclusion, however, is not justified in his paper. A partial justification is given by the following

LEMMA 4.2 *For each admissible estimator $\hat{\boldsymbol{\beta}}$, the matrix*

$$\mathrm{Var}[\mathbf{Y}_2 - \mathbf{Y}_2(\hat{\boldsymbol{\beta}})] - \mathrm{Var}[\mathbf{Y}_2 - \mathbf{Y}_2(\boldsymbol{\beta}^*)]$$

is positive semidefinite.

The proof of this lemma is that since $\mathbf{Y}_2(\hat{\boldsymbol{\beta}})$ is an unbiased predictor of $\mathbf{Y}_2$, we have

$$\begin{aligned}\mathrm{Var}[\mathbf{Y}_2 - \mathbf{Y}_2(\hat{\boldsymbol{\beta}})] &= E[(\mathbf{Y}_2 - \mathbf{Y}_2(\hat{\boldsymbol{\beta}}))(\mathbf{Y}_2 - \mathbf{Y}_2(\hat{\boldsymbol{\beta}}))'] \\ &= E[(\mathbf{D}^{-1}\mathbf{h}(\hat{\boldsymbol{\beta}}))(\mathbf{D}^{-1}\mathbf{h}(\hat{\boldsymbol{\beta}}))'] \\ &= \mathbf{D}^{-1}E[\mathbf{h}(\hat{\boldsymbol{\beta}})(\mathbf{h}(\hat{\boldsymbol{\beta}}))'](\mathbf{D}^{-1})' \\ &= \mathbf{D}^{-1}\mathrm{Var}[\mathbf{h}(\hat{\boldsymbol{\beta}})](\mathbf{D}^{-1})'.\end{aligned}$$

Now the assertion follows from Lemma 4.1.

We may even push the discussion a bit further: Why should we confine ourselves to predictors which can be written as $\mathbf{Y}_2(\hat{\boldsymbol{\beta}})$ for some admissible estimator $\hat{\boldsymbol{\beta}}$? There may be other unbiased linear predictors $\hat{\mathbf{Y}}_2$ for which

$$\mathrm{Var}[\mathbf{Y}_2 - \mathbf{Y}_2(\boldsymbol{\beta}^*)] - \mathrm{Var}[\mathbf{Y}_2 - \hat{\mathbf{Y}}_2]$$

and hence

$$\mathrm{Var}[\mathbf{Y}_2 - \mathbf{Y}_2(\hat{\boldsymbol{\beta}})] - \mathrm{Var}[\mathbf{Y}_2 - \hat{\mathbf{Y}}_2]$$

is positive semidefinite. The following result improves Lemma 4.2:

THEOREM 4.3 *For each admissible predictor $\hat{\mathbf{Y}}_2$, the matrix*

$$\mathrm{Var}[\mathbf{Y}_2 - \hat{\mathbf{Y}}_2] - \mathrm{Var}[\mathbf{Y}_2 - \mathbf{Y}_2(\boldsymbol{\beta}^*)]$$

is positive semidefinite.

A proof of this theorem can also be presented. Consider a matrix $\mathbf{Q}$ satisfying

$$\hat{\mathbf{Y}}_2 = \mathbf{Q}\mathbf{Y}_1$$

and hence $\mathbf{Q}\mathbf{X}_1 = \mathbf{X}_2$. Letting

$$\mathbf{Q}^* := \mathbf{S}_{21}\mathbf{S}_{11}^{-1} + (\mathbf{X}_2 - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{X}_1)(\mathbf{X}_1'\mathbf{S}_{11}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{S}_{11}^{-1},$$

we obtain

$$\begin{aligned}\mathbf{Y}_2(\boldsymbol{\beta}^*) &= \mathbf{X}_2\boldsymbol{\beta}^* - \mathbf{D}^{-1}\mathbf{C}(\mathbf{Y}_1 - \mathbf{X}_1\boldsymbol{\beta}^*)\\ &= \mathbf{X}_2\boldsymbol{\beta}^* + \mathbf{S}_{21}\mathbf{S}_{11}^{-1}(\mathbf{Y}_1 - \mathbf{X}_1\boldsymbol{\beta}^*)\\ &= \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{Y}_1 + (\mathbf{X}_2 - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{X}_1)\boldsymbol{\beta}^*\\ &= \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{Y}_1 + (\mathbf{X}_2 - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{X}_1)(\mathbf{X}_1'\mathbf{S}_{11}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{S}_{11}^{-1}\mathbf{Y}_1\\ &= \mathbf{Q}^*\mathbf{Y}_1.\end{aligned}$$

Since $\mathbf{Q}^*\mathbf{X}_1 = \mathbf{X}_2 = \mathbf{Q}\mathbf{X}_1$, we have

$$\begin{aligned}&\mathrm{Cov}[\mathbf{Y}_2 - \mathbf{Y}_2(\boldsymbol{\beta}^*), \mathbf{Y}_2(\boldsymbol{\beta}^*) - \hat{\mathbf{Y}}_2]\\ &\quad= \mathrm{Cov}[\mathbf{Y}_2 - \mathbf{Q}^*\mathbf{Y}_1, \mathbf{Q}^*\mathbf{Y}_1 - \mathbf{Q}\mathbf{Y}_1]\\ &\quad= (\mathbf{S}_{21} - \mathbf{Q}^*\mathbf{S}_{11})(\mathbf{Q}^* - \mathbf{Q})'\\ &\quad= -(\mathbf{X}_2 - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{X}_1)(\mathbf{X}_1'\mathbf{S}_{11}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'(\mathbf{Q}^* - \mathbf{Q})'\\ &\quad= -(\mathbf{X}_2 - \mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{X}_1)(\mathbf{X}_1'\mathbf{S}_{11}^{-1}\mathbf{X}_1)^{-1}(\mathbf{Q}^*\mathbf{X}_1 - \mathbf{Q}\mathbf{X}_1)'\\ &\quad= \mathbf{0},\end{aligned}$$

and hence

$$\begin{aligned}\mathrm{Var}[\mathbf{Y}_2 - \hat{\mathbf{Y}}_2] &= \mathrm{Var}[(\mathbf{Y}_2 - \mathbf{Y}_2(\boldsymbol{\beta}^*)) + (\mathbf{Y}_2(\boldsymbol{\beta}^*) - \hat{\mathbf{Y}}_2)]\\ &= \mathrm{Var}[\mathbf{Y}_2 - \mathbf{Y}_2(\boldsymbol{\beta}^*)] + \mathrm{Var}[\mathbf{Y}_2(\boldsymbol{\beta}^*) - \hat{\mathbf{Y}}_2].\end{aligned}$$

The assertion follows.

Theorem 4.3 asserts that the Gauss–Markov predictor minimizes the variance of the prediction error over all admissible predictors of $\mathbf{Y}_2$. Since

$$E[(\mathbf{Y}_2 - \hat{\mathbf{Y}}_2)'(\mathbf{Y}_2 - \hat{\mathbf{Y}}_2)] = \mathrm{tr}(\mathrm{Var}[\mathbf{Y}_2 - \hat{\mathbf{Y}}_2]),$$

we see that the Gauss–Markov predictor also minimizes the *expected quadratic prediction error* over all admissible predictors of $\mathbf{Y}_2$.

5. A RELATED OPTIMIZATION PROBLEM

To complete the discussion of the predictor proposed by Halliwell, we consider the following optimization problem:

Minimize $\quad E[(\mathbf{Y}-\mathbf{X}\hat{\beta})'\mathbf{S}^{-1}(\mathbf{Y}-\mathbf{X}\hat{\beta})]$

over all admissible estimators $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$.

We thus aim at minimizing an objective function in which there is no discrimination between the observable and the non-observable part of **Y**; this distinction, however, is present in the definition of an admissible estimator.

Because of $\mathbf{S}^{-1} = \mathbf{W}'\mathbf{W}$ and the structure of **W**, it is easy to see that the objective function of the optimization problem can be decomposed into an approximation part and a prediction part:

LEMMA 5.1 *The identity*

$$\begin{aligned} &E[(\mathbf{Y}-\mathbf{X}\hat{\beta})'\mathbf{S}^{-1}(\mathbf{Y}-\mathbf{X}\hat{\beta})] \\ &\quad = E[(\mathbf{Y}_1-\mathbf{X}_1\hat{\beta})'\mathbf{S}_{11}^{-1}(\mathbf{Y}_1-\mathbf{X}_1\hat{\beta})] \\ &\qquad + E[(\mathbf{Y}_2-\mathbf{Y}_2(\hat{\beta}))'\mathbf{D}'\mathbf{D}(\mathbf{Y}_2-\mathbf{Y}_2(\hat{\beta}))] \end{aligned}$$

holds for each admissible estimator $\hat{\boldsymbol{\beta}}$.

Moreover, using similar arguments as before, the three expectations occurring in Lemma 5.1 can be represented as follows:

THEOREM 5.2 *The identities*

$$\begin{aligned} &E[(\mathbf{Y}-\mathbf{X}\hat{\beta})'\mathbf{S}^{-1}(\mathbf{Y}-\mathbf{X}\hat{\beta})] \\ &\quad = (p+q)-2k+\mathrm{tr}((\mathbf{WX})\mathrm{Var}[\hat{\beta}](\mathbf{WX})') \end{aligned}$$

as well as

$$\begin{aligned} &E[(\mathbf{Y}_1-\mathbf{X}_1\hat{\beta})'\mathbf{S}_{11}^{-1}(\mathbf{Y}_1-\mathbf{X}_1\hat{\beta})] \\ &\quad = p-2k+\mathrm{tr}((\mathbf{AX}_1)\mathrm{Var}[\hat{\beta}](\mathbf{AX}_1)') \end{aligned}$$

and

$$\begin{aligned} &E[(\mathbf{Y}_2 - \mathbf{Y}_2(\hat{\boldsymbol{\beta}}))'\mathbf{D}'\mathbf{D}(\mathbf{Y}_2 - \mathbf{Y}_2(\hat{\boldsymbol{\beta}}))] \\ &\quad = q + \mathrm{tr}((\mathbf{C}\mathbf{X}_1 + \mathbf{D}\mathbf{X}_2)\mathrm{Var}[\hat{\boldsymbol{\beta}}](\mathbf{C}\mathbf{X}_1 + \mathbf{D}\mathbf{X}_2)') \end{aligned}$$

hold for each admissible estimator $\hat{\boldsymbol{\beta}}$*.*

Because of Theorem 5.2, each of the three expectations occurring in Lemma 5.1 is minimized by the Gauss–Markov estimator $\boldsymbol{\beta}^*$. We have thus again justified the restriction to predictors of $\mathbf{Y}_2$, which can be written as $\mathbf{Y}_2(\hat{\boldsymbol{\beta}})$ for some admissible estimator $\hat{\boldsymbol{\beta}}$.

The technical details concerning the proofs of the results of this section can be found in Schmidt [4].

6. CONDITIONING

Following the example of $\mathbf{Y}$ having a multivariate normal distribution, Halliwell uses arguments related to the conditional distribution of $\mathbf{Y}_2$ with respect to $\mathbf{Y}_1$; in particular, he claims that $\mathbf{Y}_2(\boldsymbol{\beta}^*)$ is the conditional expectation $E(\mathbf{Y}_2 \mid \mathbf{Y}_1)$ of $\mathbf{Y}_2$ with respect to $\mathbf{Y}_1$. This is not true in general; without particular assumptions on the distribution of $\mathbf{Y}$, the conditional expectation $E(\mathbf{Y}_2 \mid \mathbf{Y}_1)$ may fail to be linear in $\mathbf{Y}_1$, and the unbiased linear predictor of $\mathbf{Y}_2$ based on $\mathbf{Y}_1$ minimizing the expected quadratic loss may fail to be the conditional expectation $E(\mathbf{Y}_2 \mid \mathbf{Y}_1)$.

Moreover, since the identities of Lemma 4.1 hold for each admissible estimator $\hat{\boldsymbol{\beta}}$ (and not only for the Gauss–Markov estimator $\boldsymbol{\beta}^*$), Halliwell's arguments [2, p. 482] suggest that each admissible estimator $\hat{\boldsymbol{\beta}}$ satisfies

$$E(\mathbf{Y}_2 \mid \mathbf{Y}_1) = \mathbf{X}_2\hat{\boldsymbol{\beta}} - \mathbf{D}^{-1}\mathbf{C}(\mathbf{Y}_1 - \mathbf{X}_1\hat{\boldsymbol{\beta}})$$

and

$$\mathrm{Var}(\mathbf{Y}_2 \mid \mathbf{Y}_1) = \mathbf{D}^{-1}\mathrm{Var}[\mathbf{h}(\hat{\boldsymbol{\beta}})](\mathbf{D}^{-1})'.$$

Again, this cannot be true since in both cases the left hand side depends only on $\mathbf{Y}_1$, whereas the right hand side also varies with the matrix $\mathbf{B}$ defining the admissible estimator $\hat{\boldsymbol{\beta}} = \mathbf{B}\mathbf{Y}_1$.

More generally, when only unconditional moments of the distribution of the random vector $\mathbf{Y}$ are specified, it is impossible to obtain any conclusions concerning the conditional distribution of its non-observable part $\mathbf{Y}_2$ with respect to its observable part $\mathbf{Y}_1$.

REMARKS

Traditional least squares theory aims at minimizing the *quadratic loss*

$$(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'\mathbf{S}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}),$$

where all coordinates of $\mathbf{Y}$ are observable. It also involves considerations concerning the variance of $\hat{\boldsymbol{\beta}}$, and it usually handles prediction as a separate problem which has to be solved after estimating $\boldsymbol{\beta}$.

In Section 5 of the present paper, we proposed instead to minimize the *expected quadratic loss*

$$E[(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'\mathbf{S}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})],$$

where some but not all of the coordinates of $\mathbf{Y}$ are observable and the admissible estimators of $\boldsymbol{\beta}$ are unbiased and linear in the observable part $\mathbf{Y}_1$ of $\mathbf{Y}$. This approach has several advantages:

– The expected quadratic loss can be expressed in terms of $\text{var}[\hat{\boldsymbol{\beta}}]$ such that minimization of the expected quadratic loss and minimization of $\text{var}[\hat{\boldsymbol{\beta}}]$ turns out to be the same problem (see Theorem 5.2).

– The expected quadratic loss can be decomposed in a canonical way into an approximation part and a prediction part such that the expected quadratic loss and its two components are simultaneously minimized by the Gauss–Markov estimator (see Lemma 5.1).

– Inserting the Gauss–Markov estimator in the prediction part of the expected quadratic loss provides an unbiased linear predictor for the non-observable part $\mathbf{Y}_2$ of $\mathbf{Y}$.

We thus obtain the predictor proposed by Halliwell [2] by a direct approach which avoids conditioning. This predictor was first proposed by Goldberger [1] (see also Rao and Toutenburg [3; Theorem 6.2]).

ACKNOWLEDGEMENT

I would like to thank Michael D. Hamer who provided Theorem 4.3 and its proof.

REFERENCES

[1] Goldberger, Arthur S., "Best Linear Unbiased Prediction in the Generalized Linear Regression Model," *Journal of the American Statistical Association* 57, 1962, pp. 369–375.

[2] Halliwell, Leigh J., "Loss Prediction by Generalized Least Squares," *PCAS* LXXXIII, 1996, pp. 436–489.

[3] Rao, C. R., and H. Toutenburg, *Linear Models: Least Squares and Alternatives*, Berlin–Heidelberg–New York, Springer, 1995.

[4] Schmidt, Klaus D., "Prediction in the Linear Model: A Direct Approach," *Metrika* 48, 1998, pp. 141–147.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXXIII

LOSS PREDICTION BY GENERALIZED LEAST SQUARES

LEIGH J. HALLIWELL

DISCUSSION BY MICHAEL D. HAMER

Abstract

The paper by Halliwell [1] and the Discussion of Halliwell's paper by Dr. Schmidt both consider the form of "best" linear unbiased estimators for unknown quantities based on observable values. This paper proposes a general definition of "best" called Uniformly Best (UB) to distinguish it from previous definitions and provides various equivalent forms for the definition. It shows the existence and uniqueness of such UB linear unbiased estimators under fairly general conditions, provides an alternative formulation of the definition of UB for unbiased estimators, and discusses how Dr. Schmidt's proposed optimization problem relates to the proposed UB definition.

ACKNOWLEDGEMENT

I would like to thank Mr. Halliwell and Dr. Schmidt for their valuable comments on earlier drafts of this paper, and Dr. Schmidt for suggesting a shorter version of the proof to Theorem 6.1. Needless to say, the remaining opinions and errors are mine.

1. THE STRUCTURE OF THE VARIABLES

We follow the notation used by Halliwell and Schmidt. An n-dimensional random vector Y is vertically partitioned into a

p-dimensioned vector Y_1 of observable outcomes and an $n-p$ dimensioned vector Y_2 of unobservable outcomes. It is assumed that Y takes the form

$$Y = X\beta_0 + e,$$

where e is an n-dimensional random vector of "error" terms with zero mean and $(n \times n)$ dimensional non-singular variance-covariance matrix S (thus $\mathrm{E}[ee^T] = S$ where e^T represents the transpose of e, and S is positive definite), X is a given $(n \times r)$ "design" matrix, and β_0 is an unknown parameter vector of dimension r.

The matrix X and vector e can also be partitioned so that we may write

$$Y_1 = X_1\beta_0 + e_1 \qquad \text{and} \qquad Y_2 = X_2\beta_0 + e_2,$$

where X_1 is a $(p \times r)$ matrix and X_2 is a $(n-p \times r)$ matrix, and we assume that X_1 is of full rank r.

2. A PROPOSED DEFINITION OF "BEST"—THE OBJECTIVE FUNCTION

Halliwell provides a definition of "best" in Appendix A of [1], where he considers linear unbiased estimators β for the unknown vector β_0. We use this as a basis for proposing a more general definition of a "best" estimator P of a "target" quantity T. We call this definition Uniformly Best to distinguish it from other definitions of "best" used in [1–3].

Firstly, we provide a definition of a non-negative definite matrix:

DEFINITION 2.1 Non-Negative Definite. *An $(n \times n)$ matrix M is non-negative definite if $\alpha^T M \alpha \geq 0$ for any n-dimensional vector α.*

Halliwell provides an extensive review of non-negative definite matrices in Appendix A of [2]. Perhaps the most relevant characteristic for our purposes is that any non-negative definite matrix M can be written in the form $W^T W$ for some matrix W, and conversely that any matrix of the form $W^T W$ is non-negative definite.

We use the concept of non-negative definite in our proposed definition of "best" as follows:

DEFINITION 2.2 Uniformly "Best" (UB) Estimator. *A estimator P^* of a target quantity T is uniformly "best" (UB) if, for any other admissible estimator P, the matrix $\{\text{Var}(T - P) - \text{Var}(T - P^*)\}$ is non-negative definite.*

For an n-dimensional random vector z, the upper-case $\text{Var}(z)$ is the $(n \times n)$ dimensional variance-covariance matrix of z where

$$\text{Var}(z) = \text{E}[(z - \text{E}[z])(z - \text{E}[z])^T].$$

Elsewhere, we will use the lower-case $\text{var}(x)$ to denote the variance of a scalar random variable x.

To assist in understanding the nature of a UB estimator, we provide the following "equivalence" result:

LEMMA 2.1 *Suppose we consider estimators P that belong to some given admissible set J. The following statements are equivalent:*

(a) *There exists an estimator P^* in J that is the UB estimator of T.*

(b) *For any admissible P belonging to J, the matrix $\{\text{Var}(T - P) - \text{Var}(T - P^*)\}$ is non-negative definite.*

(c) *P^* minimizes $\alpha^T \text{Var}(T - P)\alpha$ over all admissible P for any α of appropriate dimension.*

(d) *P^* minimizes $\text{var}(\alpha^T(T - P))$ over all admissible P for any α of appropriate dimension.*

Proof (a) and (b) are equivalent from Definition 2.2.

From Definition 2.1 and (b), we have

$$\alpha^T[\mathrm{Var}(T-P)-\mathrm{Var}(T-P^*)]\alpha \geq 0$$

for any suitable α and for any P belonging to J. Then

$$\alpha^T\mathrm{Var}(T-P)\alpha \geq \alpha^T\mathrm{Var}(T-P^*)\alpha$$

for any P belonging to J, and so (c) follows. To show (d), we have

$$\begin{aligned}\alpha^T\mathrm{Var}(T-P)\alpha &= \alpha^T\mathrm{E}[(T-P-\mathrm{E}[T-P])(T-P-\mathrm{E}[T-P])^T]\alpha\\ &= \mathrm{E}[\alpha^T(T-P-\mathrm{E}[T-P])(T-P-\mathrm{E}[T-P])^T\alpha]\\ &= \mathrm{var}(\alpha^T(T-P)).\end{aligned}$$

The definition of UB given in (d) provides us with an objective function that we show below is easy to work with, and is perhaps the easiest to conceptualize. $\alpha^T(T-P)$ can be interpreted as the "length" of the projection of the stochastic vector representing the difference between the target T and the estimator P onto any fixed vector α. The UB estimator P^* minimizes the variance of this projection and does so for any given α.

The UB criterion is potentially quite difficult to meet. Expanding out $\mathrm{var}(\alpha^T(T-P))$ we have:

$$\mathrm{var}(\alpha^T(T-P)) = \sum\sum \alpha_i\alpha_j\,\mathrm{cov}(T_i-P_i,T_j-P_j).$$

The UB estimator P^* must minimize this double sum of products for any possible choice of α_i. However, UB estimators do exist for suitable admissible sets and targets, as shown below.

3. CONSTRAINTS ON ADMISSIBLE ESTIMATORS AND TARGETS

The definition of UB does not put any particular constraints on the admissible sets of estimators or on the form of the "target" quantities. However, it may be necessary to do so to ensure the existence of UB estimators.

(a) *Constraints on Admissible Sets Of Estimators.* Following Halliwell and Schmidt, we wish to consider estimators P that are linear in Y_1 and unbiased estimators of their "targets" T, so we define the set J of linear unbiased estimators as follows:

DEFINITION 3.1 The Admissible Set $J = J(Y_1, T)$. *An estimator P belongs to J if it is*

- *linear in Y_1 and hence of the form $P = QY_1$ where Q is a $(n \times p)$ matrix;*
- *unbiased, so that* $\mathrm{E}[P] = \mathrm{E}[T]$.

(b) *Constraints on "Targets".* We also need to define the "target" quantity T that is being estimated. For the Gauss–Markov theorem it is β_0, but elsewhere in [1] and in Schmidt's paper Y_2 and Y are also considered. To encompass all these possibilities, we consider a general form

$$T = F_1 Y_1 + F_2 Y_2 + A\beta_0,$$

where F_1, F_2 and A are variables. Since T is a vector of dimension n, F_1 is an $(n \times p)$ matrix, F_2 is an $(n \times n - p)$ matrix, and A is an $(n \times r)$ matrix.

4. EXISTENCE OF A UB LINEAR UNBIASED ESTIMATOR FOR T

The following theorem shows that there are many situations in which a UB solution not only exists but is unique.

THEOREM 4.1 *If $T = F_1 Y_1 + F_2 Y_2 + A\beta_0$ and P belongs to the admissible set J, a unique UB linear unbiased estimator $P^* = Q^* Y_1$ exists, and*

$$P^* = F_1 Y_1 + F_2 y_2(\beta^*) + A\beta^*,$$

where

$$y_2(\beta^*) = X_2\beta^* + S_{21}S_{11}^{-1}(Y_1 - X_1\beta^*) \qquad \textit{and}$$
$$\beta^* = (X_1^T S_{11}^{-1} X_1)^{-1} X_1^T S_{11}^{-1} Y_1.$$

Proof A proof of this theorem is presented in the Appendix.

Note the appearance of the Gauss–Markov estimator β^* and the predictor $y_2(\beta^*)$ discussed by Halliwell and Schmidt.

Theorem 4.1 has several interesting special cases.

CASE 1 The UB estimator for β_0

We set

$$F_1 = F_2 = 0 \qquad \text{and} \qquad A = \begin{bmatrix} I(r) \\ 0 \end{bmatrix}$$

where $I(r)$ is an $(r \times r)$ identity matrix. Then

$$Q^* = \begin{bmatrix} \beta^* \\ 0 \end{bmatrix}$$

as required by the Gauss–Markov Theorem, and the definition of UB is consistent with the Gauss–Markov notion of "best".

CASE 2 The UB estimator for Y_2

We set

$$F_1 = A = 0 \qquad \text{and} \qquad F_2 = \begin{bmatrix} 0 \\ I(n-p) \end{bmatrix},$$

where $I(n-p)$ is an $(n-p \times n-p)$ identity matrix. Then

$$Q^* = \begin{bmatrix} 0 \\ y_2(\beta^*) \end{bmatrix},$$

the form of the "best" predictor suggested by Halliwell.

CASE 3 The UB estimator for Y_1

We set

$$F_2 = A = 0 \qquad \text{and} \qquad F_1 = \begin{bmatrix} I(p) \\ 0 \end{bmatrix}.$$

Then

$$Q^* = \begin{bmatrix} Y_1 \\ 0 \end{bmatrix}.$$

Case 3 seems trivial, for of course the difference between an estimator Y_1 and target Y_1 will have zero variance. However, this result still "fits" our process, because the estimator Y_1 is certainly linear in Y_1 and unbiased.

CASE 4 The UB estimator for Y

We set

$$A = 0, \qquad F_1 = \begin{bmatrix} I(p) \\ 0 \end{bmatrix} \qquad \text{and} \qquad F_2 = \begin{bmatrix} 0 \\ I(n-p) \end{bmatrix}.$$

Then

$$Q^* = \begin{bmatrix} Y_1 \\ y_2(\beta^*) \end{bmatrix}.$$

The UB estimator for Y is thus a linear combination of the UB estimators for Y_1 and Y_2. This last result will be used in Section 6.

5. A FURTHER CHARACTERIZATION OF UB

In his Discussion, Schmidt proposes a related optimization problem in which the objective function to be minimized is $\mathrm{E}[(Y - X\beta)^T S^{-1}(Y - X\beta)]$.

We generalize Schmidt's objective function by replacing S^{-1} with any non-negative definite matrix H, and use this to define another type of estimator, which we will call Generalized Schmidt Best.

DEFINITION 5.1 Generalized Schmidt Best (GSB) Estimator. *An estimator P^* of a target quantity T is GSB if it minimizes*

$$\mathrm{E}[(T - P)^T H(T - P)]$$

over all admissible estimators P for any $(n \times n)$ non-negative definite matrix H.

How does a GSB estimator relate to a UB estimator? Rather surprisingly, the answer is that when the admissible set consists of unbiased estimators, if one exists, then they both exist and are the same.

THEOREM 5.1 *If the admissible estimators P of a general target T are all unbiased, an estimator P^* is UB if and only if it is GSB.*

Proof From our discussion of non-negative matrices, we know we can write $H = WW^T$ for some $(n \times n)$ matrix W. Now let z_i be a vector whose ith component is 1 and whose other components are all zero.

(i) Suppose a UB estimator P^* exists. For any other unbiased estimator P and any $H = WW^T$,

$$
\begin{aligned}
&\mathrm{E}[(T-P)^T WW^T (T-P)] \\
&\quad = \mathrm{trace}\{\mathrm{E}[W^T (T-P)(T-P)^T W]\} \\
&\quad = \mathrm{trace}\{W^T \,\mathrm{Var}(T-P) W\}, \qquad \text{since} \quad \mathrm{E}[T-P] = 0 \\
&\quad = \Sigma z_i^T W^T \,\mathrm{Var}(T-P) W z_i, \qquad \text{where the sum is over } i \\
&\quad = \Sigma \alpha_i^T \,\mathrm{Var}(T-P)\alpha_i \qquad \text{for} \quad \alpha_i = W z_i \\
&\quad \geq \Sigma \alpha_i^T \,\mathrm{Var}(T-P^*)\alpha_I, \qquad \text{since} \quad P^* \text{ is UB} \\
&\quad = \mathrm{E}[(T-P^*)^T WW^T (T-P^*)].
\end{aligned}
$$

Thus P^* is also GSB.

(ii) Suppose a GSB estimator P^* exists but P^* is not UB. This means, for some $\alpha^{\#}$ and for some admissible P, we must have

$$\alpha^{\#T}\mathrm{Var}(T-P^*)\alpha^{\#} > \alpha^{\#T}\mathrm{Var}(T-P)\alpha^{\#}.$$

We can construct the matrix $W^{\#} = \{\alpha^{\#}, \alpha^{\#} \ldots, \alpha^{\#})$ so that $\alpha^{\#} = W^{\#} z_i$ for any i. Then

$$z_i^T W^{\#T}\mathrm{Var}(T-P^*)W^{\#} z_i > z_i^T W^{\#T}\mathrm{Var}(T-P)W^{\#} z_i \qquad \text{for any } z_i.$$

Thus

$$\begin{aligned} &\mathrm{E}[(T-P^*)^T W^\# W^{\#T}(T-P^*)] \\ &\quad > \mathrm{E}[(T-P)^T W^\# W^{\#T}(T-P)], \end{aligned}$$

which contradicts the assumption that P^* is GSB. Thus P^* is also UB.

This proof does not require that the admissible estimators be linear in Y_1, nor does it impose any constraint on the form of the "target" T. But it is likely that a "best" solution will not always exist unless there are further restrictions on the admissible estimator set and the target because the UB and GSB conditions are so strong. When T is linear in Y_1, Y_2 and β_0 and the set J consists of linear unbiased estimators, Theorem 4.1 tells us that a UB estimator does exist, and then, from Theorem 5.1, the GSB estimator will be the same as a UB estimator.

More generally, we can use Theorem 5.1 to state an extended "equivalence" result.

LEMMA 5.1 *If the admissible set only contains unbiased estimators of a general "target" T, the following statements are equivalent (but not necessarily true):*

(a) *There exists a P^* that is the UB estimator of T for all admissible estimators P.*

(b) *For any unbiased P, the matrix* $\{\mathrm{Var}(T-P)-\mathrm{Var}(T-P^*)\}$ *is non-negative definite.*

(c) *P^* minimizes* $\alpha^T\mathrm{Var}(T-P)\alpha$ *over all admissible P for any α of appropriate dimension.*

(d) *P^* minimizes* $\mathrm{var}(\alpha^T(T-P))$ *over all admissible P for any α of appropriate dimension.*

(e) *P^* minimizes* $\mathrm{E}[(T-P)^T H(T-P)]$ *over all admissible P for any non-negative definite matrix H of appropriate dimension.*

If we expand the objective function in (e), we get

$$E[(T-P)^T H(T-P)] = \sum\sum h_{ij}\,\text{cov}(T_i - P_i, T_j - P_j),$$

and the UB estimator P^* minimizes this double sum over all possible choices of h_{ij} provided the h_{ij} belong to a non-negative definite matrix. This is more general than (d), which corresponds to the case where $h_{ij} = \alpha_i\alpha_j$. (Note: we can think of any non-negative definite matrix as a possible variance-covariance matrix if we allow the possibility that some of the variances may be zero. In this context, (d) corresponds to the case where all correlations are either $+1$ or -1, and (e) generalizes this to correlations in between.)

6. RELATIONSHIP BETWEEN "BEST" AND SCHMIDT'S OPTIMIZATION PROBLEM

In his Discussion and in [3], Schmidt suggests an optimization problem as a way of justifying the form of the "best" estimators for Y_1 and Y_2. Schmidt shows that his optimization problem can be decomposed into two parts, one involving only Y_1 and the other involving only Y_2. Further, he shows that the solution to the initial optimization problem is achieved by $\beta = \beta^*$, the Gauss–Markov estimator for β_0, and β^* minimizes each of the parts separately. In view of this optimization, Schmidt proposes that the solutions to the separate optimization problems of the parts are "best" estimators for Y_1 and Y_2, respectively.

The objective function for his optimization problem is a special case of the GSB objective function when $H = S^{-1}$ and the target $T = Y$. In addition, however, Schmidt's optimization problem requires that the admissible estimators belong to a set K, where

$$K = \{P : P = X\beta \text{ where } \beta = BY_1 \text{ and } BX_1 = I(r)\}.$$

This constraint means that the estimators in K are linear unbiased estimators of Y, but also the estimators BY_1 are also unbiased estimators of β_0.

Although Schmidt's optimization looks like the GSB objective function and K is a subset of J, the solution to Schmidt's optimization is not in general a UB estimator for Y. This is because K does not include all linear unbiased estimators of Y, and in general (except in the special circumstance detailed below) the UB linear unbiased estimator of Y is not in K.

THEOREM 6.1 *Unless X_1 is square, the UB linear unbiased estimator for Y will not belong to K.*

Proof From Theorem 4.1, the UB estimator of Y among all unbiased linear estimators is

$$P^* = \begin{bmatrix} Y_1 \\ y_2(\beta^*) \end{bmatrix}$$

and it is unique. If P^* belonged to K, we would require $X_1 B^* = I(p)$ as well as $B^* X_1 = I(r)$, where $I(p)$ and $I(r)$ are $(p \times p)$ and $(n \times n)$ identity matrices, respectively. However,

$$\begin{aligned} r = \text{trace}(I(r)) &= \text{trace}(B^* X_1) \\ &= \text{trace}(X_1 B^*), \\ &= \text{trace}(I(p)) = p. \end{aligned}$$

since $\text{trace}(AB) = \text{trace}(BA)$ for any matrices A and B,

This is a contradiction unless $r = p$, in which case X_1 and B are square.

The solution to Schmidt's optimization for a "target" Y is the vector

$$\begin{bmatrix} X_1 \beta^* \\ X_2 \beta^* \end{bmatrix}$$

which in general is quite different to the UB estimator

$$P^* = \begin{bmatrix} Y_1 \\ y_2(\beta^*) \end{bmatrix}.$$

Nevertheless, Schmidt's analysis does produce the UB estimator for Y_2. To get the "best" estimator for Y_2, Schmidt minimizes $\mathrm{E}[(Y_2 - y_2(\beta))^T D^T D(Y_2 - y_2(\beta))]$ for a particular matrix D related to S^{-1}, over possible β belonging to the set $L^* = \{\beta : \beta = BY_1$ and $BX_1 = I(r)\}$. In this case, L^* contains β^*, and the corresponding estimator $y_2(\beta^*)$ belongs to J and is UB. Because of this, we know that $y_2(\beta^*)$ will be a solution to Schmidt's optimization for any matrix D.

The "best" estimator for Y_1 derived by Schmidt's analysis is $X_1\beta^*$, which compares to the UB estimator Y_1. Using the arguments of Theorem 6.1, it can be shown that L^* does not contain a β such that $Y_1 = X_1\beta$ unless X_1 is square.

If the above restrictions on the admissible estimators in Schmidt's optimization are removed, we know from Lemma 5.1 that the resulting solution(s) will be UB. In these circumstances, Schmidt's optimization problem may then be generalized by replacing S^{-1} in the objective function with any non-negative definite matrix of appropriate dimension.

7. SUMMARY

We have proposed a general definition of "best" that we have termed Uniform Best (UB) and that is consistent with the Gauss–Markov Theorem. We have also provided a number of equivalent forms of the UB definition. We have then shown that for a "target" T linear in Y_1 and Y_2 there is always a unique UB linear unbiased estimator of the form QY_1. We have also shown that a generalization of the optimization problem proposed by Schmidt provides yet another characterization of UB. Finally, we have shown that the admissibility conditions imposed by Schmidt on the set of estimators in his optimization problem generally prevent the solution to his problem from being UB, although his "best" and the UB linear unbiased estimators for Y_2 are the same.

REFERENCES

[1] Halliwell, Leigh J., "Loss Prediction By Generalized Least Squares," *PCAS* LXXXIII, 1996, pp. 436–489.

[2] Halliwell, Leigh J., "Conjoint Prediction Of Paid And Incurred Losses," Casualty Actuarial Society *Forum* 1, Summer 1997, pp. 241–380.

[3] Schmidt, Klaus D., "Prediction In The Linear Model: A Direct Approach," *Metrika* 48, 1998, pp. 141–147.

APPENDIX

PROOF OF THEOREM 4.1

Consider two linear unbiased estimators P and P^* for T. Then

$$\mathrm{E}[P] = \mathrm{E}[QY_1] = QX_1\beta_0 = \mathrm{E}[T] = \mathrm{E}[P^*] = Q^*X_1\beta_0.$$

Since this must hold for any β_0, we have $(Q^* - Q)X_1 = 0$.

Then, for any α,

$$\begin{aligned}\mathrm{var}(\alpha^T(T-P)) &= \mathrm{var}(\alpha^T(T-P^*) + \alpha^T(P^*-P))\\ &= \mathrm{var}(\alpha^T(T-P^*)) + \mathrm{var}(\alpha^T(P^*-P))\\ &\quad + 2\,\mathrm{cov}(\alpha^T(T-P^*), \alpha^T(P^*-P)).\end{aligned}$$

Now

$$\begin{aligned}&\mathrm{cov}(\alpha^T(T-P^*), \alpha^T(P^*-P))\\ &\quad = \mathrm{E}[\alpha^T(T-P^*)\alpha^T(P^*-P)]\\ &\quad = \mathrm{E}[\alpha^T(T-P^*)(P^*-P)^T\alpha]\\ &\quad = \mathrm{E}[\alpha^T((F_1-Q^*)Y_1 + F_2Y_2)Y_1^T(Q^*-Q)^T\alpha]\\ &\quad = \alpha^T\{(F_1-Q^*)\mathrm{E}[Y_1Y_1^T] + F_2\mathrm{E}[Y_2Y_1^T]\}(Q^*-Q)^T\alpha\\ &\quad = \alpha^T\{(F_1-Q^*)S_{11} + F_2S_{21}\}(Q^*-Q)^T\alpha.\end{aligned}$$

Suppose $(F_1 - Q^*)S_{11} + F_2S_{21}$ is of the form GX_1^T, so that

$$Q^* = F_1 + F_2S_{21}S_{11}^{-1} - GX_1^TS_{11}^{-1}.$$

Then

$$\begin{aligned}\mathrm{cov}(\alpha^T(T-P^*), \alpha^T(P^*-P)) &= \alpha^TGX_1^T(Q^*-Q)^T\alpha\\ &= 0, \qquad \text{since} \quad (Q^*-Q)X_1 = 0.\end{aligned}$$

So, for any admissible P,

$$\begin{aligned}\operatorname{var}(\alpha^T(T-P)) &= \operatorname{var}(\alpha^T(T-P^*)) + \operatorname{var}(\alpha^T(P^*-P))\\ &\geq \operatorname{var}(\alpha^T(T-P^*)).\end{aligned}$$

Since $P^* = Q^*Y_1$ minimizes $\operatorname{var}(\alpha^T(T-P))$, by Lemma 2.1, it is UB.

We now solve for the form of G. The unbiased property of estimators $P = QY_1$ for T requires that

$$\mathrm{E}[T] = F_1X_1\beta_0 + F_2X_2\beta_0 + A\beta_0 = \mathrm{E}[P] = QX_1\beta_0$$

and, since this holds for any β_0, we have

$$F_1X_1 + F_2X_2 + A = QX_1.$$

Then we have

$$Q^* = F_1X_1 + F_2S_{21}S_{11}^{-1}X_1 - GX_1^TS_{11}^{-1}X_1 = F_1X_1 + F_2X_2 + A,$$

and so

$$G = \{F_2(S_{21}S_{11}^{-1}X_1 - X_2) - A\}(X_1^TS_{11}^{-1}X_1)^{-1}.$$

Substituting this back into the expression for Q^* gives

$$Q^* = F_1 + F_2S_{21}S_{11}^{-1} - \{F_2(S_{21}S_{11}^{-1}X_1 - X_2) - A\}B^*,$$

where

$$B^* = (X_1^TS_{11}^{-1}X_1)^{-1}X_1^TS_{11}^{-1}.$$

Rearranging, we get

$$Q^* = F_1 + F_2\{X_2B^* + S_{21}S_{11}^{-1}(I - X_1B^*)\} + AB^*.$$

Finally, multiplying through by Y_1 gives

$$P^* = F_1Y_1 + F_2y_2(\beta^*) + A\beta^*,$$

where

$$y_2(\beta^*) = X_2\beta^* + S_{21}S_{11}^{-1}(Y_1 - X_1\beta^*) \qquad \text{and}$$

$$\beta^* = B^*Y_1 = (X_1^TS_{11}^{-1}X_1)^{-1}X_1^TS_{11}^{-1}Y_1.$$

So far we have shown the existence of a "best" estimator. Consider another admissible estimator $P^{**} = Q^{**}Y_1$. Because P^* minimizes $\text{var}(\alpha^T(T-P))$, we have from above that

$$\text{var}(\alpha^T(T-P^{**})) = \text{var}(\alpha^T(T-P^*)) + \text{var}(\alpha^T(P^*-P^{**})).$$

If P^{**} also minimizes $\text{var}(\alpha^T(T-P))$, then

$$\text{var}(\alpha^T(T-P^{**})) = \text{var}(\alpha^T(T-P^*)),$$

and so

$$\text{var}(\alpha^T(P^*-P^{**})) = 0 \qquad \text{for any } \alpha.$$

Substituting $(P^*-P^{**}) = (Q^*-Q^{**})Y_1$ into this equation gives

$$\begin{aligned}\text{var}(\alpha^T(P^*-P^{**})) &= \text{var}(\alpha^T(Q^*-Q^{**})Y_1)\\ &= \alpha^T(Q^*-Q^{**})S_{11}(Q^*-Q^{**})^T\alpha = 0.\end{aligned}$$

S_{11}, the variance-covariance matrix of Y_1, is positive definite, so this implies

$$\alpha^T(Q^*-Q^{**}) = 0 \qquad \text{for any } \alpha.$$

Since Q^* and Q^{**} are independent of α, we must have $Q^* = Q^{**}$, and so the "best" estimator $P^* = Q^*Y_1$ is also unique.

AUTHOR'S RESPONSE TO DISCUSSIONS OF PAPER PUBLISHED IN VOLUME LXXXIII

LOSS PREDICTION BY GENERALIZED LEAST SQUARES

LEIGH J. HALLIWELL

1. INTRODUCTION

Having had the pleasure of seeing my paper in the *Proceedings*, I am even more pleased now that Klaus Schmidt and Michael Hamer have deigned to discuss it. But even with their discussions, most of the subject of statistically modeling loss triangles remains *terra incognita*; and I hope that actuaries and academics will continue to explore it.

2. BACKGROUND

Since I wrote the paper late in 1994, I have learned more about statistical modeling. I recommend for interested readers to examine my 1997 *Forum* paper, "Conjoint Prediction of Paid and Incurred Losses," especially its Appendices A and C. Nevertheless, I stand by the conclusions of the earlier paper:

> This paper will argue that the linear modeling and the least squares estimation found in the literature to date have overlooked an important condition of the linear model. In particular, the models for development factors regress random variables against other random variables. Stochastic regressors violate the standard linear model. Moreover, the model assumes that errors are uncorrelated, but stochastic regressors violate this assumption as well. This paper will show that what actuaries are really seeking is found in a general linear model; i.e., a model with nonstochastic regressors but with an error matrix that allows for correlation. [2, p. 436]

> [The use of stochastic regressors] is the fundamental problem with the CL [Chain Ladder] method. Rather than try to rehabilitate it, this paper introduces a different model that honors all the conditions of the Gauss–Markov theorem. [2, p. 441]
>
> A theory becomes very attractive when it unifies partial explanations. Such is the case with loss covariance. CL, prior hypothesis, or BF [Bornhuetter–Ferguson]—which to choose? The answer will lie on a continuum dependent on the variance matrix of the incremental losses. [2, p. 447]
>
> Generalized least squares is a better method of loss prediction than the chain ladder and the other loss development methods. Even when linear models are imposed on loss development methods, they incorporate stochastic regressors, and the estimates are not guaranteed to be either best or unbiased. The confidence intervals derived therefrom are not trustworthy. The fault lies in trying to make the level on one variable affect the level of the next, whereas the statistical idea is that the departure of one variable from its mean affects the departure of the next from its mean. This is the idea of covariance, and it is accommodated in the general linear model and generalized least squares estimation. [2, p. 456]

The problem of stochastic regressors quells my enthusiasm for empirically testing chain-ladder statistical models (as, for example, Gary Venter [6] recommends). The technique of instrumental variables [4, p. 577 and 5, p. 198] solves this problem; but the obvious instrument for a lagged loss is its exposure. And when exposure becomes a regressor, the lagged loss often lacks significance, as Glen Barnett and Ben Zehnwirth have discovered [1, p. 10]. So I am hopeful that actuaries will find their way back to the no-frills "additive model" [2, pp. 442, 449] and thence begin to consider non-trivial covariance structures.[1]

3. AUTHOR'S COMMENTS ON ORIGINAL PAPER

Before responding to the discussions I will point out two flaws of the paper. The first flaw concerns pages 450f. and Exhibit 3. I derived an estimate of β, reweighted the observations, and derived a second estimate of β. I remarked, "The estimate for β changes negligibly (no change within the first ten decimal places)." [2, p. 451] Such a negligible change should have clued me that the estimates of β were identical, the difference owing to computational precision. If one regresses Y against X with error variances σ, the estimate is:

$$\frac{\sum_i \frac{x_i y_i}{\sigma_{ii}}}{\sum_i \frac{x_i x_i}{\sigma_{ii}}}.$$

Therefore, the estimate is invariant to a scale change of the variances. Now the second model applied scale factors according to age. But each element of $\hat{\beta}$ depends on observations of the same age, which have been affected by the same scale factor. Thus the estimate is unchanged.[2]

The second flaw concerns the degrees of freedom in the estimate of σ^2. There were thirty-six observations, eight parameters in β, and two parameters in the variance matrix. I claimed there to be $36 - 8 - 2 = 26$ degrees of freedom [2, p. 453]. But the two parameters that had been estimated in the variance matrix are not like those of β. There is no theoretically right way of accounting for the variance parameters, and twenty-eight degrees of freedom is just as acceptable as twenty-six.

[1]My session "Regression Models and Loss Reserving" at the 1999 Casualty Loss Reserve Seminar presents this broad subject with theory and examples.

[2]I am grateful to William A. Niemczyk for pointing this out to me.

4. AUTHOR'S COMMENTS ON DISCUSSIONS

Drs. Schmidt and Hamer have confined their discussions to the Gauss–Markov theorem and to the best linear unbiased predictor. This is natural, since the Gauss–Markov theorem is the most mathematical topic of the paper and is new material to most actuaries (at least in its matrix form). In several of my papers I have complained that we actuaries know too little about statistical modeling and the matrix algebra that it utilizes. I myself learned what little I know by a time-consuming study of materials outside the actuarial syllabus, particularly [4]. And I believe that even the new actuarial syllabus does not adequately cover this topic. However, I wish that these discussions had gotten beyond the Gauss–Markov theorem and treated the undesirability of stochastic regressors and the distinction between loss covariance and loss development.

Dr. Schmidt's finish, "We thus obtain the predictor proposed by Halliwell by a direct approach which avoids conditioning," provides the basis for my two-fold response. First, as to conditioning, my treatment of the predictor in Appendix C does not depend on Bayes' theorem and a loss distribution. In fact, I wrote that **e** is "not necessarily normal" [2, pp. 480, 473]. However, perhaps I invited Dr. Schmidt's criticism when I used conditional-expectation notation [2, pp. 445, 482f] and said that the unknown elements "are affected by the known elements in a Bayesian sense, through the variance matrix." [2, p. 444] My Appendix B demonstrated that if **e** is multivariate normal, the predictor can be derived by Bayes' theorem; but I did not say that conditional probability was the rationale of the predictor.

And second, Drs. Schmidt and Hamer have made my argument rigorous, and shown that one can bypass the estimation of β on the way to estimating $\mathbf{Y}_2$ (the "direct approach"). I concur with their assertions that the proof in my Appendix C was not strict, and that it confined itself "to predictors which can be written as $y_2(\hat{\beta})$ for some admissible estimator $\hat{\beta}$." I had realized these things when I wrote my paper on conjoint prediction [3]. There I formulated the partitioned model (*p* observations and *q* predictions):

$$\begin{bmatrix} \mathbf{Y}_{1(p\times 1)} \\ \cdots\cdots \\ \mathbf{Y}_{2(q\times 1)} \end{bmatrix} = \begin{bmatrix} X_{1(p\times k)} \\ \cdots\cdots \\ X_{2(q\times k)} \end{bmatrix} \beta_{(k\times 1)} + \begin{bmatrix} \mathbf{e}_1 \\ .. \\ \mathbf{e}_2 \end{bmatrix}, \qquad \text{where}$$

$$\mathrm{Var}\begin{bmatrix} \mathbf{e}_1 \\ .. \\ \mathbf{e}_2 \end{bmatrix} = \begin{bmatrix} S_{11(p\times p)} & \vdots & S_{12(p\times q)} \\ \cdots\cdots & \vdots & \cdots\cdots \\ S_{21(q\times p)} & \vdots & S_{22(q\times q)} \end{bmatrix}$$

And I showed [3, p. 328] that the best linear unbiased predictor of $\mathbf{Y}_2$ is:

$$\hat{\mathbf{Y}}_2 = (S_{21}S_{11}^{-1} + (X_2 - S_{21}S_{11}^{-1}X_1)(X_1'S_{11}^{-1}X_1)^{-1}X_1'S_{11}^{-1})\mathbf{Y}_1$$

This agrees with Dr. Schmidt's Theorem 4.3, whose proof Dr. Hamer has provided. This formulation is direct because the estimator $\hat{\mathbf{Y}}_2$ does not involve $\hat{\beta}$. However, if $X_2 = \mathbf{I}_k$ and $\mathbf{e}_2$ is a zero matrix (and hence S_{21} and S_{22} are zero matrices), then $\mathbf{Y}_2 = \beta$, and:

$$\begin{aligned} \hat{\beta} = \hat{\mathbf{Y}}_2 &= (S_{21}S_{11}^{-1} + (X_2 - S_{21}S_{11}^{-1}X_1)(X_1'S_{11}^{-1}X_1)^{-1}X_1'S_{11}^{-1})\mathbf{Y}_1 \\ &= (0S_{11}^{-1} + (I_k - 0S_{11}^{-1}X_1)(X_1'S_{11}^{-1}X_1)^{-1}X_1'S_{11}^{-1})\mathbf{Y}_1 \\ &= (X_1'S_{11}^{-1}X_1)^{-1}X_1'S_{11}^{-1}\mathbf{Y}_1 \end{aligned}$$

So the estimation of β is a special case of the estimation of $\mathbf{Y}_2$ [3, p. 331], which Dr. Hamer calls Case 1 of his Theorem 4.1.[3]

That really is all that I need to say about the Gauss–Markov theorem and best linear unbiased prediction. The task now, as

[3]Dr. Hamer devotes his Appendix to deriving the best linear unbiased estimator (BLUE) of $F_1\mathbf{Y}_1 + F_2\mathbf{Y}_2 + A\beta$. Though correct, the form of this derivation is overly complex. I have shown [3, p. 335f] that the estimator is a linear operator; hence, the BLUE of this expression is $F_1\hat{\mathbf{Y}}_1 + F_2\hat{\mathbf{Y}}_2 + A\hat{\beta} = F_1\mathbf{Y}_1 + F_2\hat{\mathbf{Y}}_2 + A\hat{\beta}$.

I see it, is to get actuaries to understand that this theory is not just a mathematical nicety. Though perhaps not a Copernican revolution, it is revolutionary nonetheless. As it makes inroads, we will see less of development factors and loss adjustments and more of modeling and exposure adjustments.

REFERENCES

[1] Barnett, Glen, and Ben Zehnwirth, "Best Estimates for Reserves," Casualty Actuarial Society *Forum*, Fall 1998, pp. 1–54.

[2] Halliwell, Leigh J., "Loss Prediction by Generalized Least Squares," *PCAS* LXXXIII, 1996, pp. 436–489.

[3] Halliwell, Leigh J., "Conjoint Prediction of Paid and Incurred Losses," Casualty Actuarial Society *Forum*, Summer 1997, pp. 241–380.

[4] Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lütkepohl, and T.-C. Lee, *Introduction to the Theory and Practice of Econometrics*, Second Edition, New York, John Wiley, 1988.

[5] Pindyck, Robert S., and Daniel L. Rubinfeld, *Econometric Models and Economic Forecasts*, Fourth Edition, Boston, Irwin/McGraw-Hill, 1998.

[6] Venter, Gary G., "Testing the Assumptions of Age-to-Age Factors," *PCAS* LXXXV, 1998, pp. 807–847.

APPENDIX A

As an appendix, I wish to comment on the optimization problem of Dr. Schmidt's fifth section, and on Dr. Hamer's generalization of it. Though this problem has occasioned some interesting mathematics, I see the problem as a sidelight, as only loosely related to the Gauss–Markov theorem.

Dr. Schmidt wishes to find the admissible estimator $\hat{\beta}$ that minimizes:

$$E[(\mathbf{Y} - \mathrm{X}\hat{\beta})'\mathrm{S}^{-1}(\mathbf{Y} - \mathrm{X}\hat{\beta})].$$

Estimator $\hat{\beta}$ is admissible if and only if it is a linear function of $\mathbf{Y}_1$ and it is unbiased. In his third section he shows that admissible estimators are of the form $\mathrm{B}_{(k\times p)}\mathbf{Y}_1$ for $\mathrm{BX}_1 = \mathrm{I}_k$, and $\mathrm{Var}[\hat{\beta}] = \mathrm{Var}[\mathrm{B}\mathbf{Y}_1] = \mathrm{B}\,\mathrm{Var}[\mathbf{Y}_1]\mathrm{B}' = \mathrm{BS}_{11}\mathrm{B}'$.

As I had done [2, p. 480f], he factors S^{-1} as $\mathrm{W}'\mathrm{W}$, where:

$$\mathrm{W} = \begin{bmatrix} \mathrm{A}_{(p\times p)} & 0_{(p\times q)} \\ \mathrm{C}_{(q\times p)} & \mathrm{D}_{(q\times q)} \end{bmatrix},$$

such that

$$\begin{aligned} \mathrm{A}'\mathrm{A} &= \mathrm{S}_{11}^{-1}, \\ \mathrm{D}'\mathrm{D} &= (\mathrm{S}_{22} - \mathrm{S}_{21}\mathrm{S}_{11}^{-1}\mathrm{S}_{12})^{-1}, \qquad \text{and} \\ \mathrm{C} &= -\mathrm{DS}_{21}\mathrm{S}_{11}^{-1}. \end{aligned}$$

Now:

$$\begin{aligned} \mathrm{W}(\mathbf{Y} - \mathrm{X}\hat{\beta}) &= \begin{bmatrix} \mathrm{A}_{(p\times p)} & 0_{(p\times q)} \\ \mathrm{C}_{(q\times p)} & \mathrm{D}_{(q\times q)} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_1 - \mathrm{X}_1\hat{\beta} \\ \mathbf{Y}_2 - \mathrm{X}_2\hat{\beta} \end{bmatrix} \\ &= \begin{bmatrix} \mathrm{A}(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta}) \\ \mathrm{C}(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta}) + \mathrm{D}(\mathbf{Y}_2 - \mathrm{X}_2\hat{\beta}) \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} A(\mathbf{Y}_1 - X_1\hat{\beta}) \\ D\mathbf{Y}_2 - DX_2\hat{\beta} + DD^{-1}C(\mathbf{Y}_1 - X_1\hat{\beta}) \end{bmatrix}$$

$$= \begin{bmatrix} A(\mathbf{Y}_1 - X_1\hat{\beta}) \\ D(\mathbf{Y}_2 - X_2\hat{\beta} + D^{-1}C(\mathbf{Y}_1 - X_1\hat{\beta})) \end{bmatrix}$$

$$= \begin{bmatrix} A(\mathbf{Y}_1 - X_1\hat{\beta}) \\ D(\mathbf{Y}_2 - y_2(\hat{\beta})) \end{bmatrix}$$

Therefore:

$$\begin{aligned}
&(\mathbf{Y} - X\hat{\beta})'S^{-1}(\mathbf{Y} - X\hat{\beta}) \\
&\qquad = (\mathbf{Y} - X\hat{\beta})'W'W(\mathbf{Y} - X\hat{\beta}) \\
&\qquad = (W(\mathbf{Y} - X\hat{\beta}))'(W(\mathbf{Y} - X\hat{\beta})) \\
&\qquad = [(\mathbf{Y}_1 - X_1\hat{\beta})'A' \quad (\mathbf{Y}_2 - y_2(\hat{\beta}))'D'] \begin{bmatrix} A(\mathbf{Y}_1 - X_1\hat{\beta}) \\ D(\mathbf{Y}_2 - y_2(\hat{\beta})) \end{bmatrix} \\
&\qquad = (\mathbf{Y}_1 - X_1\hat{\beta})'A'A(\mathbf{Y}_1 - X_1\hat{\beta}) \\
&\qquad\quad + (\mathbf{Y}_2 - y_2(\hat{\beta}))'D'D(\mathbf{Y}_2 - y_2(\hat{\beta})) \\
&\qquad = (\mathbf{Y}_1 - X_1\hat{\beta})'S_{11}^{-1}(\mathbf{Y}_1 - X_1\hat{\beta}) \\
&\qquad\quad + (\mathbf{Y}_2 - y_2(\hat{\beta}))'D'D(\mathbf{Y}_2 - y_2(\hat{\beta}))
\end{aligned}$$

And we have Dr. Schmidt's Lemma 5.1:

$$\begin{aligned}
E[(\mathbf{Y} - X\hat{\beta})'S^{-1}(\mathbf{Y} - X\hat{\beta})] &= E[(\mathbf{Y}_1 - X_1\hat{\beta})'S_{11}^{-1}(\mathbf{Y}_1 - X_1\hat{\beta})] \\
&\quad + E[(\mathbf{Y}_2 - y_2(\hat{\beta}))'D'D(\mathbf{Y}_2 - y_2(\hat{\beta}))]
\end{aligned}$$

To prove his Theorem 5.2 we have to review the trace function. The trace of a square matrix Q is defined as the sum of its diagonal elements: $\mathrm{tr}(Q_{(n\times n)}) = \sum_{i=1}^{n} q_{ii}$. Some theorems that should be obvious are:

$$\begin{aligned}\text{tr}(\alpha\text{Q}) &= \alpha\,\text{tr}(\text{Q})\\ \text{tr}(\text{Q}') &= \text{tr}(\text{Q})\\ \text{tr}(\text{Q}_1 + \text{Q}_2) &= \text{tr}(\text{Q}_1) + \text{tr}(\text{Q}_2)\\ \text{tr}(\text{I}_n) &= n\end{aligned}$$

If Q is (1×1), then $\text{tr}(\text{Q}) = q_{11} = \text{Q}$. (For our purposes we may ignore the distinction between a scalar and a one-element matrix.) And if **Q** is a random matrix:

$$\begin{aligned}\text{tr}(E[\mathbf{Q}]) &= \sum_{i=1}^{n} E[\mathbf{q}_{ii}]\\ &= E\left[\sum_{i=1}^{n} \mathbf{q}_{ii}\right]\\ &= E[\text{tr}(\mathbf{Q})]\end{aligned}$$

But a theorem that is not obvious is that if A is $(m \times n)$ and B is $(n \times m)$, then the traces of AB and BA are equal. The proof is:

$$\begin{aligned}\text{tr}(\text{AB}) &= \sum_{i=1}^{m} [\text{AB}]_{ii}\\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} b_{ji}\right)\\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} b_{ji} a_{ij}\right)\\ &= \sum_{j=1}^{n} [\text{BA}]_{jj} = \text{tr}(\text{BA})\end{aligned}$$

With this knowledge of the trace we can prove Theorem 5.2.

We reduce the first term on the right side of Lemma 5.1, mindful of the fact that the expressions within the expectation operators are (1×1) matrices:

$$\begin{aligned}
&E[(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})'\mathrm{S}_{11}^{-1}(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})] \\
&\quad = E[\mathrm{tr}((\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})'\mathrm{S}_{11}^{-1}(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta}))] \\
&\quad = E[\mathrm{tr}(\mathrm{S}_{11}^{-1}(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})')] \\
&\quad = \mathrm{tr}(E[\mathrm{S}_{11}^{-1}(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})']) \\
&\quad = \mathrm{tr}(\mathrm{S}_{11}^{-1}E[(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})']) \\
&\quad = \mathrm{tr}(\mathrm{S}_{11}^{-1}E[(\mathbf{Y}_1 - \mathrm{X}_1\mathrm{B}\mathbf{Y}_1)(\mathbf{Y}_1 - \mathrm{X}_1\mathrm{B}\mathbf{Y}_1)']) \\
&\quad = \mathrm{tr}(\mathrm{S}_{11}^{-1}E[((\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathbf{Y}_1)((\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathbf{Y}_1)'])
\end{aligned}$$

But because $\hat{\beta}$ is admissible, $\mathrm{BX}_1 = \mathrm{I}_k$ and:

$$\begin{aligned}
E[(\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathbf{Y}_1] &= (\mathrm{I}_p - \mathrm{X}_1\mathrm{B})E[\mathbf{Y}_1] \\
&= (\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathrm{X}_1\beta \\
&= \mathrm{X}_1\beta - \mathrm{X}_1\mathrm{B}\mathrm{X}_1\beta \\
&= \mathrm{X}_1\beta - \mathrm{X}_1\mathrm{I}_k\beta \\
&= 0
\end{aligned}$$

So:

$$\begin{aligned}
&E[((\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathbf{Y}_1)((\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathbf{Y}_1)'] \\
&\quad = \mathrm{Var}[(\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathbf{Y}_1] \\
&\quad = (\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathrm{Var}[\mathbf{Y}_1](\mathrm{I}_p - \mathrm{X}_1\mathrm{B})' \\
&\quad = (\mathrm{I}_p - \mathrm{X}_1\mathrm{B})\mathrm{S}_{11}(\mathrm{I}_p - \mathrm{X}_1\mathrm{B})'
\end{aligned}$$

Therefore:

$$
\begin{aligned}
&E[(\mathbf{Y}_1 - X_1\hat{\beta})' S_{11}^{-1}(\mathbf{Y}_1 - X_1\hat{\beta})] \\
&\quad = \operatorname{tr}(S_{11}^{-1}E[((I_p - X_1B)\mathbf{Y}_1)((I_p - X_1B)\mathbf{Y}_1)'])\\
&\quad = \operatorname{tr}(S_{11}^{-1}(I_p - X_1B)S_{11}(I_p - X_1B)')\\
&\quad = \operatorname{tr}(S_{11}^{-1}(S_{11} - S_{11}B'X_1' - X_1BS_{11} + X_1BS_{11}B'X_1'))\\
&\quad = \operatorname{tr}(I_p - B'X_1' - S_{11}^{-1}X_1BS_{11} + S_{11}^{-1}X_1\operatorname{Var}[\hat{\beta}]X')\\
&\quad = \operatorname{tr}(I_p) - \operatorname{tr}(B'X_1') - \operatorname{tr}(S_{11}^{-1}X_1BS_{11}) + \operatorname{tr}(S_{11}^{-1}X_1\operatorname{Var}[\hat{\beta}]X')
\end{aligned}
$$

But

$$
\begin{aligned}
\operatorname{tr}(I_p) &= p,\\
\operatorname{tr}(B'X_1') &= \operatorname{tr}(X_1B),\\
\operatorname{tr}(S_{11}^{-1}X_1BS_{11}) &= \operatorname{tr}(X_1BS_{11}S_{11}^{-1})\\
&= \operatorname{tr}(X_1B) = \operatorname{tr}(BX_1) = \operatorname{tr}(I_k) = k, \qquad \text{and}\\
\operatorname{tr}(S_{11}^{-1}X_1\operatorname{Var}[\hat{\beta}]X_1') &= \operatorname{tr}(A'AX_1\operatorname{Var}[\hat{\beta}]X_1')\\
&= \operatorname{tr}(AX_1\operatorname{Var}[\hat{\beta}]X_1'A')\\
&= \operatorname{tr}((AX_1)\operatorname{Var}[\hat{\beta}](AX_1)').
\end{aligned}
$$

So we arrive at the second equation of Theorem 5.2:

$$
\begin{aligned}
&E[(\mathbf{Y}_1 - X_1\hat{\beta})' S_{11}^{-1}(\mathbf{Y}_1 - X_1\hat{\beta})]\\
&\quad = \operatorname{tr}(I_p) - \operatorname{tr}(X_1B) - \operatorname{tr}(S_{11}^{-1}X_1BS_{11}) + \operatorname{tr}(S_{11}^{-1}X_1\operatorname{Var}[\hat{\beta}]X')\\
&\quad = p - 2k + \operatorname{tr}((AX_1)\operatorname{Var}[\hat{\beta}](AX_1)')
\end{aligned}
$$

Then we reduce the second term:

$$
\begin{aligned}
&E[(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))'\mathrm{D}'\mathrm{D}(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))] \\
&\quad = E[\mathrm{tr}((\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))'\mathrm{D}'\mathrm{D}(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta})))] \\
&\quad = E[\mathrm{tr}(\mathrm{D}'\mathrm{D}(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))')] \\
&\quad = \mathrm{tr}(E[\mathrm{D}'\mathrm{D}(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))'])\\
&\quad = \mathrm{tr}(\mathrm{D}'\mathrm{D}E[(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))'])\\
&\quad = \mathrm{tr}(\mathrm{D}'\mathrm{D}\,\mathrm{Var}[\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta})])\\
&\quad = \mathrm{tr}(\mathrm{D}\,\mathrm{Var}[\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta})]\mathrm{D}')
\end{aligned}
$$

The next-to-last step follows from the fact that $\mathrm{y}_2(\hat{\beta})$ is an admissible predictor of $\mathbf{Y}_2$ (as Dr. Schmidt states in his fourth section); hence, $E[\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta})] = 0$. But according to Lemma 4.1, $\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}) = \mathrm{D}^{-1}h(\hat{\beta})$ and:

$$\mathrm{Var}[h(\hat{\beta})] = (\mathrm{CX}_1 + \mathrm{DX}_2)\mathrm{Var}[\hat{\beta}](\mathrm{CX}_1 + \mathrm{DX}_2)' + \mathrm{I}_q$$

So by substitution we arrive at the third equation of Theorem 5.2:

$$
\begin{aligned}
&E[(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))'\mathrm{D}'\mathrm{D}(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))] \\
&\quad = \mathrm{tr}(\mathrm{D}\,\mathrm{Var}[\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta})]\mathrm{D}')\\
&\quad = \mathrm{tr}(\mathrm{D}\,\mathrm{Var}[\mathrm{D}^{-1}h(\hat{\beta})]\mathrm{D}')\\
&\quad = \mathrm{tr}(\mathrm{Var}[\mathrm{DD}^{-1}h(\hat{\beta})])\\
&\quad = \mathrm{tr}(\mathrm{Var}[h(\hat{\beta})])\\
&\quad = q + \mathrm{tr}((\mathrm{CX}_1 + \mathrm{DX}_2)\mathrm{Var}[\hat{\beta}](\mathrm{CX}_1 + \mathrm{DX}_2)')
\end{aligned}
$$

Dr. Schmidt denotes the Gauss–Markov estimator

$$(\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathbf{Y}_1$$

as β^*. Adapting my notation to his, I can restate the last formula of my Appendix A [2, p. 474] as:

$$\mathrm{Var}[\hat{\beta}] - \mathrm{Var}[\beta^*] = \{\mathrm{BA}^{-1} - (\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{A}'\} \times \{\mathrm{BA}^{-1} - (\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{A}'\}' \geq 0,$$

where, as above, $\mathrm{A}'\mathrm{A} = \mathrm{S}_{11}^{-1}$ and $\mathrm{BX}_1 = \mathrm{I}_k$. And equality obtains if and only if:

$$\begin{aligned}\mathrm{BA}^{-1} - (\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{A}' &= 0\\ \mathrm{BA}^{-1} &= (\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{A}'\\ \mathrm{B} &= (\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{A}'\mathrm{A}\\ &= (\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{S}_{11}^{-1}\end{aligned}$$

Therefore, $\mathrm{Var}[\hat{\beta}] - \mathrm{Var}[\beta^*]$ is non-negative definite (or, as Dr. Schmidt calls it, positive semidefinite).[4]

Winding up the optimization problem, we have:

$$\begin{aligned}&E[(\mathbf{Y} - \mathrm{X}\hat{\beta})'\mathrm{S}^{-1}(\mathbf{Y} - \mathrm{X}\hat{\beta})] - E[(\mathbf{Y} - \mathrm{X}\beta^*)'\mathrm{S}^{-1}(\mathbf{Y} - \mathrm{X}\beta^*)]\\ &\quad= E[(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})'\mathrm{S}_{11}^{-1}(\mathbf{Y}_1 - \mathrm{X}_1\hat{\beta})]\\ &\qquad - E[(\mathbf{Y}_1 - \mathrm{X}_1\beta^*)'\mathrm{S}_{11}^{-1}(\mathbf{Y}_1 - \mathrm{X}_1\beta^*)]\\ &\qquad + E[(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))'\mathrm{D}'\mathrm{D}(\mathbf{Y}_2 - \mathrm{y}_2(\hat{\beta}))]\\ &\qquad - E[(\mathbf{Y}_2 - \mathrm{y}_2(\beta^*))'\mathrm{D}'\mathrm{D}(\mathbf{Y}_2 - \mathrm{y}_2(\beta^*))]\\ &\quad= \mathrm{tr}((\mathrm{AX}_1)(\mathrm{Var}[\hat{\beta}] - \mathrm{Var}[\beta^*])(\mathrm{AX}_1)')\\ &\qquad + \mathrm{tr}((\mathrm{CX}_1 - \mathrm{DX}_2)(\mathrm{Var}[\hat{\beta}] - \mathrm{Var}[\beta^*])(\mathrm{CX}_1 - \mathrm{DX}_2)')\end{aligned}$$

[4] See [3, pp. 306–309] for an overview of non-negative definite matrices.

The arguments of the trace functions are non-negative definite matrices, whose diagonal elements must be non-negative. Therefore, the traces are non-negative, and:

$$E[(\mathbf{Y}-\mathrm{X}\hat{\beta})'\mathrm{S}^{-1}(\mathbf{Y}-\mathrm{X}\hat{\beta})]-E[(\mathbf{Y}-\mathrm{X}\beta^*)'\mathrm{S}^{-1}(\mathbf{Y}-\mathrm{X}\beta^*)]\geq 0$$

$$E[(\mathbf{Y}-\mathrm{X}\beta^*)'\mathrm{S}^{-1}(\mathbf{Y}-\mathrm{X}\beta^*)]\leq E[(\mathbf{Y}-\mathrm{X}\hat{\beta})'\mathrm{S}^{-1}(\mathbf{Y}-\mathrm{X}\hat{\beta})]$$

β^* minimizes the expected quadratic loss, though it may not be unique among all admissible estimators of β.

This problem has led Dr. Hamer to define the "generalized Schmidt best (GSB)" estimator as the admissible (i.e., linear-in-$\mathbf{Y}_1$ and unbiased) estimator $\mathbf{P}^*$ that minimizes $E[(\mathbf{Y}_2-\mathbf{P})'\mathrm{W}' \cdot \mathrm{W}(\mathbf{Y}_2-\mathbf{P})]$ over all admissible $\mathbf{P}$, regardless of W.[5] He proves in his Theorem 5.1 that $\mathbf{P}^*$ is GSB if and only if it is the best linear unbiased predictor $\hat{Y}_2$. Therefore, GSB and "uniformly best (UB)" are equivalent. Now the set of admissible estimators in Dr. Schmidt's problem is a subset of the set of those in Dr. Hamer's definition; hence, $\hat{Y}$ will dominate $X\beta^*$ in the optimization of $E[(\mathbf{Y}-\mathbf{P})'\mathrm{W}'\mathrm{W}(\mathbf{Y}-\mathbf{P})]$.

In his Section 6 Dr. Hamer proves that $\mathrm{X}\beta^*$ is best if and only if X_1 is square. I wish to present here another proof. The relevant formulas are:

$$\hat{Y}=\begin{bmatrix}\hat{\mathbf{Y}}_1\\ \hat{\mathbf{Y}}_2\end{bmatrix}$$

$$=\begin{bmatrix}\mathbf{Y}_1\\ (\mathrm{S}_{21}\mathrm{S}_{11}^{-1}+(\mathrm{X}_2-\mathrm{S}_{21}\mathrm{S}_{11}^{-1}\mathrm{X}_1)(\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{S}_{11}^{-1})\mathbf{Y}_1\end{bmatrix}$$

$$=\begin{bmatrix}\mathrm{I}_p\\ \mathrm{S}_{21}\mathrm{S}_{11}^{-1}+(\mathrm{X}_2-\mathrm{S}_{21}\mathrm{S}_{11}^{-1}\mathrm{X}_1)(\mathrm{X}_1'\mathrm{S}_{11}^{-1}\mathrm{X}_1)^{-1}\mathrm{X}_1'\mathrm{S}_{11}^{-1}\end{bmatrix}\mathbf{Y}_1$$

[5] I've changed his notation, but not his meaning.

$$X\beta^* = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} (X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} \mathbf{Y}_1$$
$$= \begin{bmatrix} X_1 (X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} \\ X_2 (X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} \end{bmatrix} \mathbf{Y}_1$$

The two estimators are identical (i.e., equal, regardless of the value of $\mathbf{Y}_1$) if and only if $X_1(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} = I_p$ and

$$\begin{aligned} &X_2(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} \\ &\quad = S_{21} S_{11}^{-1} + (X_2 - S_{21} S_{11}^{-1} X_1)(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1}. \end{aligned}$$

However, if $X_1(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} = I_p$:

$$\begin{aligned} &X_2(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} \\ &\quad = S_{21} S_{11}^{-1} + X_2(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} - S_{21} S_{11}^{-1} I_p \\ &\quad = S_{21} S_{11}^{-1} + X_2(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} \\ &\qquad - S_{21} S_{11}^{-1} X_1 (X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} \\ &\quad = S_{21} S_{11}^{-1} + (X_2 - S_{21} S_{11}^{-1} X_1)(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} \end{aligned}$$

Therefore, the two estimators are identical if and only if $X_1(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} = I_p$.

Now if $X_1(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} = I_p$:

$$\begin{aligned} p &= \mathrm{tr}(I_p) \\ &= \mathrm{tr}(X_1 (X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1}) \\ &= \mathrm{tr}((X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} X_1) \\ &= \mathrm{tr}(I_k) \\ &= k \end{aligned}$$

And if $p = k$, then since the rank of X_1 is k (guaranteeing that $X_1' S_{11}^{-1} X_1$ has an inverse), X_1 has an inverse. And:

$$\begin{aligned}
X_1(X_1' S_{11}^{-1} X_1)^{-1} X_1' S_{11}^{-1} &= X_1(X_1)^{-1}(S_{11}^{-1})^{-1}(X_1')^{-1} X_1' S_{11}^{-1} \\
&= I_p(S_{11}^{-1})^{-1} I_p S_{11}^{-1} \\
&= (S_{11}^{-1})^{-1} S_{11}^{-1} \\
&= I_p
\end{aligned}$$

So $X\beta^*$ is best if and only if X_1 is square, in which case the observations constitute a system of simultaneous equations that has the unique solution $\beta^* = X_1^{-1}\mathbf{Y}_1$.

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXXV

AGGREGATION OF CORRELATED RISK PORTFOLIOS: MODELS AND ALGORITHMS

SHAUN S. WANG, PH.D.

DISCUSSION BY GLENN MEYERS

Abstract

In response to a request for proposal from the Committee on the Theory of Risk, Shaun Wang has written a paper that significantly advances, to quote the proposal, "the development of tools and models that improve the accuracy of the estimation of aggregate loss distributions for blocks of insurance risks."

Dr. Wang's charge was to "assume a book of business is the union of disjoint classes of business each of which has an aggregate distribution. …The classes of business are NOT independent. …The problem is how do you calculate the aggregate distribution for the whole book." Dr. Wang's paper covers a variety of dependency models and computational methods.

This discussion of his paper delves more deeply into a particular dependency model—correlation caused by parameter uncertainty—and then shows how his work applies to calculating the aggregate loss distribution for this case with one particular computational method—Fourier Inversion.

1. BACKGROUND

The collective risk model has long been one of the primary tools of actuarial science. One can view that model as a computer simulation where one first picks a random number of claims and then sums the random loss amounts for each claim. Simulating the distribution of losses for the collective risk model can (even today) be time consuming so, over the years, a number of mathematical methods have been developed to shorten the computing time. Klugman, Panjer, and Willmot [6, Ch. 4], provide an excellent description of the current computational methods.

The early uses of the collective risk model were mostly theoretical illustrations of the role of insurer surplus and profit margins. Such illustrations are still common today in insurance educational readings such as Bowers, Gerber, Jones, Hickman and Nesbitt [3, Ch. 13].

By the late 1970s, members of the Casualty Actuarial Society were beginning to use the collective risk model as input for real world insurance decisions. The early applications of the collective risk model included retrospective rating, e.g., Meyers [7], and aggregate stop loss reinsurance, which is described by Patrik [10]. Bear and Nemlick [2] provide further examples of the use of the collective risk model in the pricing of reinsurance contracts.

Some of these early efforts recognized the fact that the parameters of the collective risk model were unknown. Patrik and John [5] introduced parameter uncertainty by treating the parameters of the claim severity and claim count distributions as random variables. Heckman and Meyers [4] followed with an efficient computational algorithm that allows for some particular forms of parameter uncertainty in the collective risk model.

It is easy and instructive to consider the effect of parameter uncertainty on the variance of a distribution. Let X be a random variable that depends on a parameter θ. Then:

$$\mathrm{Var}[X] = \underbrace{\mathrm{E}_\theta[\mathrm{Var}[X \mid \theta]]}_{\text{Process Variance}} + \underbrace{\mathrm{Var}_\theta[\mathrm{E}[X \mid \theta]]}_{\text{Parameter Variance}}. \tag{1.1}$$

If there is no parameter uncertainty, the parameter variance will be zero. Introducing parameter uncertainty will increase the unconditional variance.

Suppose $X_1,\ldots,X_n$ are identically distributed random variables that depend on a parameter θ. Let $\mathrm{E}[X \mid \theta]$ and $\mathrm{Var}[X \mid \theta]$ be their common mean and variance given θ. Assume further that the X_i's are conditionally independent given θ. Then:

$$\mathrm{E}\left[\sum_{i=1}^{n} X_i \mid \theta\right] = n \cdot \mathrm{E}[X \mid \theta] \qquad \text{and}$$

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i \mid \theta\right] = n \cdot \mathrm{Var}[X \mid \theta].$$

Unconditionally:

$$\begin{aligned}\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] &= \mathrm{E}_\theta\left[\mathrm{Var}\left[\sum_{i=1}^{n} X_i \mid \theta\right]\right] + \mathrm{Var}_\theta\left[\mathrm{E}\left[\sum_{i=1}^{n} X_i \mid \theta\right]\right] \\ &= \underbrace{n \cdot \mathrm{E}_\theta[\mathrm{Var}[X \mid \theta]]}_{\text{Process Variance}} + \underbrace{n^2 \cdot \mathrm{Var}_\theta[\mathrm{E}[X \mid \theta]]}_{\text{Parameter Variance}}. \qquad (1.2)\end{aligned}$$

In most insurance situations, $\mathrm{E}_\theta[\mathrm{Var}[X \mid \theta]] \gg \mathrm{Var}_\theta[\mathrm{E}[X \mid \theta]]$, and we should expect the process variance to be dominant for small n. But as n increases, the parameter variance becomes increasingly important. This becomes apparent by looking at the coefficient of variation:

$$\begin{aligned}\mathrm{CV}\left[\sum_{i=1}^{n} X_i\right] &= \frac{\sqrt{n \cdot \mathrm{E}_\theta[\mathrm{Var}[X \mid \theta]] + n^2 \cdot \mathrm{Var}_\theta[\mathrm{E}[X \mid \theta]]}}{n \cdot \mathrm{E}[X]} \\ &\underset{n\to\infty}{\longrightarrow} \frac{\sqrt{\mathrm{Var}_\theta[\mathrm{E}[X \mid \theta]]}}{\mathrm{E}[X]} > 0. \qquad (1.3)\end{aligned}$$

More generally, we expect parameter uncertainty to play a minor role for small insureds and to play a major role for large insureds or for a reasonably sized insurance company.

In situations where parameter uncertainty affects several lines of insurance simultaneously, we expect high losses in one line to be associated with high losses in another line. Thus parameter uncertainty generates correlation. There are, of course, other generators of correlation. One example is in property insurance, where natural disasters cause damage to properties in close proximity.

Meyers and Schenker [9] provided some statistical methods of quantifying parameter uncertainty using observations spanning a period of years. However, any statistical method for quantifying parameter uncertainty requires considerable judgment because:

1. Data is scarce. You get one observation per insured per year.
2. The source of the historical variability in the parameters is often identifiable (at least after the fact). The user might not expect that source of variability to be present in future years. However, other sources of variability may arise.

2. DYNAMIC FINANCIAL ANALYSIS

The Casualty Actuarial Society coined the term "Dynamic Financial Analysis" (DFA) in the wake of the efforts to create a risk-based capital formula for insurers. To do DFA, one must often create an aggregate loss distribution for an entire insurance company. Now, for an insurance company, the primary source of parameter uncertainty is change over time. Thus parameter uncertainty will be a very important component in any collective risk model when it is applied to an entire insurance company.

As mentioned above, quantifying parameter uncertainty involves a fair amount of judgment. For example:

- Uncertain inflation will affect all claims simultaneously.

- Changes in the general economy can affect various lines of insurance in special ways. For example, directors and officers liability claims are more likely in a recession.
- Insurance companies write liability insurance at several different policy limits. We expect uncertainty in the claim frequency to affect policy limits in the same way.

The ultimate goal of DFA is to make financial decisions based on controlling the risk of an entire insurance company. DFA necessarily involves the more general concept of covariance, which can be driven by mechanisms other than parameter uncertainty. Practitioners familiar with the collective risk model should make the effort to express their knowledge in financial language. On the other hand, as we shall show, the collective risk model—with parameter uncertainty—can enrich the financial models.

3. PARAMETER UNCERTAINTY AND CORRELATION

For the h^{th} line of insurance let:

μ_h = Expected claim severity;

σ_h^2 = Variance of the claim severity distribution;

λ_h = Expected claim count; and

$\lambda_h + c_h \cdot \lambda_h^2$ = Variance of the claim count distribution.

Following Heckman and Meyers [4], we call c_h the contagion parameter. If the claim count distribution is:

Poisson, then $c_h = 0$;

negative binomial, then $c_h > 0$; and

binomial with n trials, then $c_h = -1/n$.

A good way to view the collective risk model is by a Monte Carlo simulation.

Simulation Algorithm #1
The Collective Risk Model Without Parameter Uncertainty

1. For lines of insurance 1 to n, select a random number of claims, K_h, for each line of insurance h.
2. For each line of insurance h, select random claim amounts Z_{hk}, for $k = 1,\ldots,K_h$. Each Z_{hk} has a common distribution $\{Z_h\}$.
3. Set $X_h = \sum_{k=1}^{K_h} Z_{hk}$.
4. Set $X = \sum_{h=1}^{n} X_h$.

The collective risk model describes the distribution of X. In this section we restrict ourselves to calculating the covariance structure of X. In the next section we will show how to calculate the entire distribution of X.

If we assume that K_h is independent of K_g for $g \neq h$, and that Z_h is independent of K_h, we have:

$$\begin{aligned}\mathrm{Var}[X_h] &= \mathrm{E}_{K_h}[\mathrm{Var}[X_h \mid K_h]] + \mathrm{Var}_{K_h}[\mathrm{E}[X_h \mid K_h]]\\ &= \lambda_h \cdot \sigma_h^2 + \mu_h^2 \cdot (\lambda_h + c_h \cdot \lambda_h^2). \end{aligned} \tag{3.1}$$

Also

$$\mathrm{Cov}[X_g, X_h] = 0 \qquad \text{for} \quad g \neq h. \tag{3.2}$$

We now introduce parameter uncertainty that affects the claim count distribution for several lines of insurance simultaneously. We partition the lines of insurance into covariance groups $\{G_i\}$. Our next version of the collective risk model is defined as follows.

Simulation Algorithm #2
The Collective Risk Model with Parameter Uncertainty in the Claim Count Distributions

1. For each covariance group i, select $\alpha_i > 0$ from a distribution with:

$$\mathrm{E}[\alpha_i] = 1 \qquad \text{and} \qquad \mathrm{Var}[\alpha_i] = g_i.$$

   g_i is called the covariance generator for the covariance group i.

2. For line of insurance h in covariance group i, select a random number of claims K_{hi} from a distribution with mean $\alpha_i \cdot \lambda_{hi}$.

3. For each line of insurance h in covariance group i, select random claim amounts Z_{hik} for $k = 1,\ldots,K_h$. Each Z_{hik} has a common distribution $\{Z_{hi}\}$.

4. Set $X_{hi} = \sum_{k=1}^{K_{hi}} Z_{hik}$.

5. Set $X_{\bullet i} = \sum_{h \in G_i} X_{hi}$.

6. Set $X = \sum_{i=1}^{n} X_{\bullet i}$.

We have:

$$\begin{aligned}\mathrm{Cov}[X_{di}, X_{hi}] = {} & \mathrm{E}_{\alpha_i}[\mathrm{Cov}[X_{di}, X_{hi} \mid \alpha_i]] \\ & + \mathrm{Cov}_{\alpha_i}[\mathrm{E}[X_{di} \mid \alpha_i], \mathrm{E}[X_{hi} \mid \alpha_i]].\end{aligned}$$

For $d \neq h$, X_{di} and X_{hi} are conditionally independent. Thus $\mathrm{Cov}[X_{di}, X_{hi} \mid \alpha_i] = 0$ and

$$\begin{aligned}\mathrm{Cov}[X_{di}, X_{hi}] &= \mathrm{Cov}_{\alpha_i}[\alpha_i \cdot \lambda_{di} \cdot \mu_{di}, \alpha_i \cdot \lambda_{hi} \cdot \mu_{hi}] \\ &= g_i \cdot \lambda_{di} \cdot \mu_{di} \cdot \lambda_{hi} \cdot \mu_{hi}. \qquad (3.3)\end{aligned}$$

Also,

$$\begin{aligned}
\mathrm{Cov}[X_{hi}, X_{hi}] &= \mathrm{Var}[X_{hi}] \\
&= \mathrm{E}_{\alpha_i}[\mathrm{Var}[X_{hi} \mid \alpha_i]] + \mathrm{Var}_{\alpha_i}[\mathrm{E}[X_{hi} \mid \alpha_i]] \\
&= \mathrm{E}_{\alpha_i}[\alpha_i \cdot \lambda_{hi} \cdot \sigma_{hi}^2 + \mu_{hi}^2 \cdot (\alpha_i \cdot \lambda_{hi} + \alpha_i^2 \cdot c_{hi} \cdot \lambda_{hi}^2)] \\
&\quad + \mathrm{Var}_{\alpha_i}[\alpha_i \cdot \lambda_{hi} \cdot \mu_{hi}] \\
&= \lambda_{hi} \cdot \sigma_{hi}^2 + \mu_{hi}^2 \cdot (\lambda_{hi} + (1 + g_i) \cdot c_{hi} \cdot \lambda_{hi}^2) + g_i \cdot \lambda_{hi}^2 \cdot \mu_{hi}^2.
\end{aligned} \tag{3.4}$$

And:

$$\mathrm{Cov}[X_{di}, X_{hj}] = 0 \qquad \text{for} \quad i \neq j. \tag{3.5}$$

We now introduce parameter uncertainty in the severity distributions. Let β be a positive random variable with $\mathrm{E}[1/\beta] = 1$ and $\mathrm{Var}[1/\beta] = b$. Following Heckman and Meyers [4], we call b the mixing parameter. Let $X_{hi}^{\beta} = X_{hi}/\beta$ *for all h and i*. Then:

$$\begin{aligned}
\mathrm{Cov}[X_{di}^{\beta}, X_{hj}^{\beta}] &= \mathrm{E}_{\beta}[\mathrm{Cov}[X_{di}/\beta, X_{hj}/\beta]] \\
&\quad + \mathrm{Cov}_{\beta}[\mathrm{E}[X_{di}/\beta], \mathrm{E}[X_{hj}/\beta]] \\
&= \mathrm{Cov}[X_{di}, X_{hj}] \cdot (1 + b) + b \cdot \mathrm{E}[X_{di}] \cdot \mathrm{E}[X_{hj}].
\end{aligned} \tag{3.6}$$

From Equations 3.3 to 3.6, we see that the first term of Equation 3.6 will be zero whenever $i \neq j$, and the second term will be positive whenever $b > 0$.

To calculate the coefficient of correlation, ρ_{XY}, between two separate lines of insurance with random losses X and Y, we use Equations 3.3 to 3.6 and the relationship:

$$\rho_{XY} = \frac{\mathrm{Cov}[X, Y]}{\sqrt{\mathrm{Var}[X] \cdot \mathrm{Var}[Y]}}. \tag{3.7}$$

We illustrate the effect of parameter uncertainty on correlation with an example. We use the illustrative claim severity distribu-

TABLE 3.1

CLAIM COUNT DISTRIBUTION PARAMETERS

Covariance Group	Covariance Generator	Line of Insurance	λ	c
#1	.01	GL-$1M	Varies	0.00
		GL-$5M	Varies	0.00
#2	.02	AL-$1M	Varies	0.01
		AL-$5M	Varies	0.01

tions for general liability and automobile liability given in Appendix A. Table 3.1 gives the covariance group and claim count distribution parameters. The examples use $b = 0.01$.

Table 3.2 gives the correlation matrices for the claim count distributions[1] and the aggregate loss distributions for each line of insurance with $\lambda = 10$, 100, and 100,000. Note that as λ increases the coefficients of correlation approach a limiting value. We can calculate that limiting value by dropping the terms with λ_{hi} (small compared with terms with λ_{hi}^2) in Equation 3.4. If $c = 0$, the limiting coefficients of correlation are 1.0.[2]

If we modify the claim severity distribution by a deductible, with p being the probability of exceeding the deductible, we must then change the λ parameter of a negative binomial claim count distribution by replacing λ with $p \cdot \lambda$. The contagion parameter c remains unchanged.[3] We can then apply Equations 3.3 to 3.7 to the modified claim count and claim severity distributions. Table 3.2 gives the resulting correlation matrices.

These examples show the practical utility of having correlation coefficients that are generated by a model. One should not

[1]We calculated claim count covariances from Equations 3.3 to 3.6 using $\mu_{hi} = 1$ and $\sigma_{hi} = 0$.

[2]Holding c as a constant while varying λ uses the interpretation of c as quantifying parameter uncertainty within a single line of insurance. See Heckman and Meyers [4] for details.

[3]This is proven on pp. 266–7 of Klugman et al. [6]. Note that, in our parameterization, $\lambda = r \cdot \beta$ and $c = 1/r$.

TABLE 3.2

ILLUSTRATED CORRELATION MATRICES

Expected Claim Count = 10

	Claim Count Correlations				**Total Loss Correlations**			
	GL-$1M	GL-$5M	AL-$1M	AL-$5M	GL-$1M	GL-$5M	AL-$1M	AL-$5M
GL-$1M	1.00000	0.09091	0.00000	0.00000	1.00000	0.01361	0.00412	0.00354
GL-$5M	0.09091	1.00000	0.00000	0.00000	0.01361	1.00000	0.00355	0.00305
AL-$1M	0.00000	0.00000	1.00000	0.15361	0.00412	0.00355	1.00000	0.00560
AL-$5M	0.00000	0.00000	0.15361	1.00000	0.00354	0.00305	0.00560	1.00000

Expected Claim Count = 1,000

	Claim Count Correlations				**Total Loss Correlations**			
	GL-$1M	GL-$5M	AL-$1M	AL-$5M	GL-$1M	GL-$5M	AL-$1M	AL-$5M
GL-$1M	1.00000	0.90909	0.00000	0.00000	1.00000	0.57819	0.18826	0.17271
GL-$5M	0.90909	1.00000	0.00000	0.00000	0.57819	1.00000	0.17671	0.16212
AL-$1M	0.00000	0.00000	1.00000	0.64103	0.18826	0.17671	1.00000	0.32042
AL-$5M	0.00000	0.00000	0.64103	1.00000	0.17271	0.16212	0.32042	1.00000

Expected Claim Count = 100,000

	Claim Count Correlations				**Total Loss Correlations**			
	GL-$1M	GL-$5M	AL-$1M	AL-$5M	GL-$1M	GL-$5M	AL-$1M	AL-$5M
GL-$1M	1.00000	0.99900	0.00000	0.00000	1.00000	0.99272	0.34743	0.34674
GL-$5M	0.99900	1.00000	0.00000	0.00000	0.99272	1.00000	0.34705	0.34636
AL-$1M	0.00000	0.00000	1.00000	0.66203	0.34743	0.34705	1.00000	0.73582
AL-$5M	0.00000	0.00000	0.66203	1.00000	0.34674	0.34636	0.73582	1.00000

Limiting Correlations as the Expected Claim Count Approaches Infinity

	Claim Count Correlations				**Total Loss Correlations**			
	GL-$1M	GL-$5M	AL-$1M	AL-$5M	GL-$1M	GL-$5M	AL-$1M	AL-$5M
GL-$1M	1.00000	1.00000	0.00000	0.00000	1.00000	1.00000	0.35048	0.35048
GL-$5M	1.00000	1.00000	0.00000	0.00000	1.00000	1.00000	0.35048	0.35048
AL-$1M	0.00000	0.00000	1.00000	0.66225	0.35048	0.35048	1.00000	0.74564
AL-$5M	0.00000	0.00000	0.66225	1.00000	0.35048	0.35048	0.74564	1.00000

Ground Up Expected Count = 1,000 with a $100,000 Deductible

	Claim Count Correlations				**Total Loss Correlations**			
	GL-$1M	GL-$5M	AL-$1M	AL-$5M	GL-$1M	GL-$5M	AL-$1M	AL-$5M
GL-$1M	1.00000	0.43740	0.00000	0.00000	1.00000	0.38533	0.12445	0.11282
GL-$5M	0.43740	1.00000	0.00000	0.00000	0.38533	1.00000	0.11355	0.10294
AL-$1M	0.00000	0.00000	1.00000	0.21918	0.12445	0.11355	1.00000	0.20181
AL-$5M	0.00000	0.00000	0.21918	1.00000	0.11282	0.10294	0.20181	1.00000

use empirical correlation coefficients if they were applied to an insured with a different exposure, or if a deductible were imposed.

4. CALCULATING THE AGGREGATE LOSS DISTRIBUTION BY FOURIER INVERSION

In this section, we show how to use direct Fourier inversion to calculate the aggregate loss distribution described by Simulation Algorithm #2. We begin by summarizing the method of Heckman and Meyers [4] using the more compact notation of Klugman et al. [6, p. 316].[4]

Let Z be a random variable representing claim severity. Define the Fourier transform of Z as:

$$\phi_Z(t) \equiv \mathrm{E}[e^{itZ}].$$

A fundamental property of Fourier transforms is that:

$$\phi_{\underbrace{Z+\cdots+Z}_{\text{K Times}}}(t) = \phi_Z(t)^K,$$

where the Z's are independent.

Let K be a random variable representing claim count. Define the probability generating function (pgf) of a claim count distribution as:

$$P_K(t) \equiv \mathrm{E}[t^K].$$

Define the aggregate loss

$$X = \underbrace{Z + \cdots + Z}_{\text{K Times}}.$$

We then have:

$$\phi_X(t) = \mathrm{E}[(\phi_Z(t))^K] = P_K(\phi_Z(t)). \tag{4.1}$$

[4]Wang describes a similar process using the Fast Fourier Transform.

Let $X_1,\ldots,X_n$ be independent random variables of aggregate losses. Then:

$$\phi_{X_1+\cdots+X_n}(t) = \prod_{i=1}^{n} \phi_{X_i}(t). \tag{4.2}$$

Heckman and Meyers [4] provide a way to obtain the distribution of $X_1 + \cdots + X_n$ and the distribution[5] of $(X_1 + \cdots + X_n)/\beta$ given the Fourier transform $\phi_{X_1+\cdots+X_n}(t)$ and that β has a gamma distribution.

To summarize, Fourier inversion turns the time-consuming process of simulating the sum of random variables into the mathematically complex, but doable, process of multiplying the Fourier transforms of the random variables and then inverting this product. Until now, we have been assuming that the claim count distributions are independent and that the claim severity distribution is independent of the claim count.

To remove the assumption that the claim count distributions are independent, Wang uses the multivariate Fourier transform which is defined by:

$$\phi_{X_1,\ldots,X_n}(t_1,\ldots,t_n) = \mathrm{E}[e^{i(t_1 X_1+\cdots+t_n X_n)}]$$

and has the property that:

$$\phi_{X_1+\cdots+X_n}(t) = \phi_{X_1,\ldots,X_n}(t,\ldots,t). \tag{4.3}$$

When the lines of insurance are correlated, we can then apply the Heckman/Meyers Fourier inversion formula to Equation 4.3 to obtain the aggregate loss distribution.

We now use Equation 4.3 to calculate the Fourier transform for the aggregate loss distribution described by Simulation Algorithm #2—the collective risk model with parameter uncertainty

[5] See Equation 3.6 and the preceding paragraph.

in the claim count distributions.

$$\phi_{X_{\bullet i}}(t) = \phi_{X_{1i},\ldots,X_{n_i i}}(t,\ldots,t)$$

(from Equation 4.3)

$$= \mathrm{E}_{\alpha_i}[\phi_{X_{1i},\ldots,X_{n_i i}}(t,\ldots,t) \mid \alpha_i]$$

$$= \mathrm{E}_{\alpha_i}\left[\prod_{h=1}^{n_i} \phi_{X_{hi}}(t) \mid \alpha_i\right]$$

(Equation 4.2 applies since the X_{hi}'s are *conditionally* independent.)

$$= \mathrm{E}_{\alpha_i}\left[\prod_{h=1}^{n_i} P_{K_{hi}}(\phi_{Z_{hi}}(t)) \mid \alpha_i\right]. \qquad (4.4)$$

(from Equation 4.1)

Since the covariance groups are independent:

$$\phi_X(t) = \prod_{i=1}^{n} \phi_{X_{\bullet i}}(t). \qquad (4.5)$$

To complete the model description, we need to specify:

- the distribution of the a_i's;
- the pgf's $P_{K_{hi}}(t)$; and
- the Fourier transforms of the severity distributions $\phi_{Z_{hi}}(t)$.

We will use a three-point discrete distribution for a_i. Let:

$$\begin{aligned}
\alpha_{i1} &= 1 - \sqrt{3g_i} & \Pr\{\alpha_i = \alpha_{i1}\} &= 1/6 \\
\alpha_{i2} &= 1 & \Pr\{\alpha_i = \alpha_{i2}\} &= 2/3 \\
\alpha_{i3} &= 1 + \sqrt{3g_i} & \Pr\{\alpha_i = \alpha_{i3}\} &= 1/6
\end{aligned} \qquad (4.6)$$

This discrete distribution was motivated by an approximation to Equation 4.4 when a_i has a normal distribution. Equation 4.4

then becomes:

$$\mathrm{E}_{\alpha_i}\left[\prod_{h=1}^{n_i} P_{K_{hi}}(\phi_{Z_{hi}}(t)) \mid \alpha_i\right] = \frac{1}{\sqrt{2\pi g_i}} \int_{-\infty}^{\infty} \left[\prod_{h=1}^{n_i} P_{K_{hi}}(\phi_{Z_{hi}}(t)) \mid \alpha_i\right] \cdot e^{-(\alpha_i - 1)^2/2g_i} d\alpha_i; \tag{4.7}$$

by using the Gauss–Hermite three-point quadrature formula:

$$\int_{-\infty}^{\infty} f(x)\cdot e^{-x^2}dx \approx \frac{\sqrt{\pi}}{6} f\left(-\sqrt{\frac{3}{2}}\right) + \frac{2\sqrt{\pi}}{3} f(0) + \frac{\sqrt{\pi}}{6} f\left(\sqrt{\frac{3}{2}}\right); \tag{4.8}$$

with the change of variables:

$$x = \frac{\alpha_i - 1}{\sqrt{2g_i}}.$$

One can use a higher-order formula, obtainable from many texts on numerical analysis. See, for example, Ralston [11].

Appendix B of Klugman et al. [6] provides the pgf's for a wide variety of claim count distributions. We provide two examples here, translated into this paper's notation.

For the negative binomial claim count distribution:

$$P_{K_{hi}}(t) \mid \alpha_i = (1 - c_{hi}\cdot \lambda_{hi}\cdot \alpha_i \cdot (t-1))^{-1/c_{hi}}.$$

For the Poisson claim count distribution:

$$P_{K_{hi}}(t) \mid \alpha_i = e^{-\lambda_{hi}\cdot\alpha_i\cdot(t-1)}.$$

The Fourier transform of a claim severity distribution with probability density function $f(z)$:

$$\phi_Z(t) = \int_0^{\infty} e^{itx} f(x)\,dx.$$

This integral does not have a closed form for most of the commonly used claim severity distributions. Heckman and Meyers

[4] get around that difficulty by approximating the cumulative distribution function (cdf), $F(z)$, with a piecewise linear cdf, for which the integral does have a closed form.

To summarize this section, we have shown how to calculate the multivariate Fourier transform of the collective risk model with correlations generated by parameter uncertainty. We then used the direct Fourier inversion formulas of Heckman and Meyers to calculate the corresponding aggregate loss distribution.

Note that one could use the Fast Fourier Transform methods discussed by Wang.

5. AN ILLUSTRATIVE EXAMPLE

We now illustrate the effect of covariance on the aggregate loss distribution of the hypothetical XYZ Insurance Company. XYZ writes commercial lines exclusively—workers compensation, general liability, commercial auto and commercial property. Table 5.1 provides summary statistics for XYZ's book of business.

Following are some additional remarks about XYZ's loss distribution.

- We set the mixing parameter $b = 0.01$.
- The claim severity distributions are piecewise linear approximations to mixed exponential distributions. See Appendix A for details. Also, the standard deviations for the claim severity distributions reflect the mixing generated by the mixing parameter, b.
- The claim count distributions are all negative binomial.
- The correlations between the claim count distributions of the coverages in a given line are driven by the covariance generator listed with the first coverage of the line.

TABLE 5.1

XYZ SUMMARY LOSS STATISTICS
LINE/COVERAGE SUMMARY STATISTICS
AGGREGATE SUMMARY STATISTICS

Aggregate Mean	1,004,422,886
Aggregate Standard Deviation	156,034,063
Mixing Parameter	0.010000

Line Name/ Liability Limit	E[Count]	Std[Count]	E[Severity]	Std[Severity]	E[Tot.Loss]	Covariance Generator
WC-$5M Limit	80,000.00	8,005.00	5,339.89	52,927.43	427,191,200	
GL-$5M Limit	200.00	42.61	40,348.87	160,218.51	8,069,774	0.020000
GL-$2M Limit	800.00	163.27	39,892.11	152,516.66	31,913,688	
GL-$1M Limit	2,200.00	444.68	36,966.16	124,853.59	81,325,552	
GL-$.5M Limit	1,250.00	253.72	31,085.63	87,532.67	38,857,038	
AL-$5M Limit	350.00	53.03	12,809.55	99,730.27	4,483,342	0.010000
AL-$2M Limit	1,350.00	194.89	12,626.84	94,724.36	17,046,234	
AL-$1M Limit	3,700.00	528.08	11,456.65	76,434.03	42,389,605	
AL-$.5M Limit	2,300.00	329.59	9,131.21	50,896.52	21,001,783	
APhD	1,100.00	159.44	4,360.00	6,331.53	4,796,000	
CP-$50M Limit	2,000.00	667.83	10,999.77	224,488.75	21,999,540	0.100000
CP-$10M Limit	8,000.00	2,666.83	6,999.95	45,887.29	55,999,600	
CP-$5M Limit	18,500.00	6,165.08	6,499.98	24,515.84	120,249,630	
CP-$2M Limit	10,000.00	3,333.17	6,199.99	13,467.32	61,999,900	
CP-$1M Limit	11,000.00	3,666.33	6,100.00	11,066.55	67,100,000	

FIGURE 5.1

XYZ AGGREGATE LOSS DISTRIBUTION

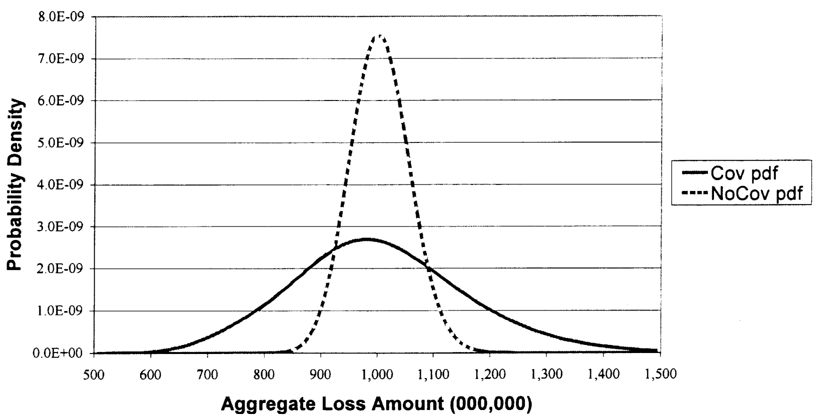


TABLE 5.2

COMPARISON OF AGGREGATE LOSS DISTRIBUTIONS† WITH AND WITHOUT THE COVARIANCE GENERATORS AND THE MIXING PARAMETER

	WO/Covariance	W/Covariance
Aggregate Mean	1,004,422,886	1,004,422,886
Aggregate Std. Dev.	52,698,873	156,034,063

Aggregate	Cumulative Probability		Limited Pure Premium Ratio	
Loss	WO/Covariance	W/Covariance	WO/Covariance	W/Covariance
500,000,000	0.00000	0.00000	0.49780	0.49780
600,000,000	0.00000	0.00070	0.59736	0.59734
700,000,000	0.00000	0.01617	0.69692	0.69634
800,000,000	0.00001	0.08782	0.79648	0.79136
900,000,000	0.01954	0.25528	0.89570	0.87477
1,000,000,000	0.47643	0.51146	0.97685	0.93653
1,100,000,000	0.96097	0.74683	0.99909	0.97282
1,200,000,000	0.99970	0.89181	1.00000	0.99004
1,300,000,000	1.00000	0.96115	1.00000	0.99688
1,400,000,000	1.00000	0.98831	1.00000	0.99916
1,500,000,000	1.00000	0.99703	1.00000	0.99981
1,600,000,000	1.00000	0.99935	1.00000	0.99996
1,700,000,000	1.00000	0.99987	1.00000	0.99999
1,800,000,000	1.00000	0.99998	1.00000	1.00000
1,900,000,000	1.00000	1.00000	1.00000	1.00000
2,000,000,000	1.00000	1.00000	1.00000	1.00000

†The cumulative probability is the probability that the aggregate loss amount is less than the stated loss amount. The limited pure premium is the expected aggregate loss when limited to the stated loss amount. The limited pure premium ratio is the limited pure premium divided by the expected aggregate loss.

Appendix B gives the correlation matrices generated by mixing the claim count and claim severity distributions.

Table 5.2 and Figure 5.1 illustrate the significant effect that correlations have on the aggregate loss distribution of XYZ Insurance Company.

6. CONCLUSION

We congratulate Dr. Wang for his fine work in introducing dependency into the collective risk model. This discussion has attempted to expand the applicability of his work and illustrate its importance in Dynamic Financial Analysis.

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, pp. 105–172.

[2] Bear, Robert A. and Kenneth J. Nemlick, "Pricing the Impact of Adjustable Features and Loss Sharing Provisions of Reinsurance Treaties," *PCAS* LXXVII, 1990, pp. 60–123.

[3] Bowers, N. L., H. U. Gerber, J. C. Hickman, D. A. Jones, and C. J. Nesbitt, *Actuarial Mathematics*, Second Edition, Society of Actuaries, 1997.

[4] Heckman, Philip E. and Glenn G. Meyers, "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *PCAS* LXX, 1983, pp. 22–61.

[5] John, Russell T. and Gary S. Patrik, "Pricing Excess-of-Loss Casualty Working Cover Reinsurance Treaties," *Pricing Property and Casualty Insurance Products*, Casualty Actuarial Society Discussion Paper Program, 1980, pp. 399–474.

[6] Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot, *Loss Models: From Data to Decisions*, John Wiley & Sons, 1998.

[7] Meyers, Glenn G., "An Analysis of Retrospective Rating," *PCAS* LXVII, 1980, pp. 110–143.

[8] Meyers, Glenn G., "An Analysis of the Capital Structure of an Insurance Company," *PCAS* LXXVI, 1989, pp. 147–170.

[9] Meyers, Glenn G. and Nathaniel A. Schenker, "Parameter Uncertainty in the Collective Risk Model," *PCAS* LXX, 1983, pp. 111–143.

[10] Patrik, Gary S., "Reinsurance," *Foundations of Casualty Actuarial Science*, Third Edition, Casualty Actuarial Society, 1996, Chapter 6.

[11] Ralston, Anthony A., *A First Course in Numerical Analysis*, McGraw-Hill Inc., 1965.

APPENDIX A

THE CLAIM SEVERITY DISTRIBUTIONS

The Heckman/Meyers algorithm requires that the cumulative distribution functions for the claim severity distributions be piecewise linear. Users of the algorithm usually have an analytic model for claim severity, so some approximation is necessary. This appendix gives the analytic models used in this paper and their piecewise linear approximations. The claim severity distributions are merely illustrative and the reader should note that we did not derive the claim severity distributions from any proprietary data available to us.

This paper uses the mixed exponential claim severity model for all lines of insurance. The cumulative distribution function (cdf) is given by:

$$F(x) = 1 - \sum_{i=1}^{4} w_i \cdot e^{-x/b_i}. \tag{A.1}$$

The limited average severity (LAS) is given by:

$$L(x) = \sum_{i=1}^{4} w_i \cdot b_i \cdot (1 - e^{-x/b_i}). \tag{A.2}$$

A piecewise linear cdf approximates each mixed exponential cdf. For the specified values $x_0, x_2, \ldots, x_{2n}$, the piecewise linear cdf has the same value as its corresponding mixed exponential cdf, and the piecewise linear LAS has the same value as its corresponding mixed exponential LAS. We accomplish this matching of the LAS values by setting:

$$x_{2n-1} = \frac{L(x_{2n}) - L(x_{2n-2}) - x_{2n} \cdot (1 - F(x_{2n})) + x_{2n-2} \cdot (1 - F(x_{2n-2}))}{F(x_{2n}) - F(x_{2n-2})} \tag{A.3}$$

TABLE A.1

MIXED EXPONENTIAL PARAMETERS

Line Names	b_1	b_2	b_3	b_4	w_1	w_2	w_3	w_4
WC	1,000	10,000	100,000	500,000	0.940	0.040	0.015	0.005
GL	1,000	10,000	100,000	500,000	0.350	0.500	0.100	0.050
AL	1,000	2,500	10,000	500,000	0.360	0.500	0.120	0.020
APhD	1,000	5,000	10,000	15,000	0.360	0.500	0.120	0.020
CP-$50M Limit	2,000	5,000	20,000	5,000,000	0.360	0.500	0.139	0.001
CP-$10M Limit	2,000	5,000	20,000	1,000,000	0.360	0.500	0.139	0.001
CP-$5M Limit	2,000	5,000	20,000	500,000	0.360	0.500	0.139	0.001
CP-$2M Limit	2,000	5,000	20,000	200,000	0.360	0.500	0.139	0.001
CP-$1M Limit	2,000	5,000	20,000	100,000	0.360	0.500	0.139	0.001

and

$$F(x_{2n-1}) = F(x_{2n}) - (F(x_{2n}) - F(x_{2n-2}))\frac{x_{2n-1} - x_{2n-2}}{x_{2n} - x_{2n-2}}. \tag{A.4}$$

Table A.1 gives the parameters of the mixed exponential distributions used in this paper. Table A.2 gives the piecewise linear approximations for two of these distributions. The values $x_0, x_2, \ldots$ are the same for all of the piecewise linear distributions used in this paper.

TABLE A.2

PIECEWISE LINEAR APPROXIMATIONS TO MIXED EXPONENTIAL DISTRIBUTIONS

WC-$5M Limit	*w*'s	Means	GL-$5M Limit	*w*'s	Means
Exp #1	0.940	1,000	Exp #1	0.350	1,000
Exp #2	0.040	10,000	Exp #2	0.500	10,000
Exp #3	0.015	100,000	Exp #3	0.100	100,000
Exp #4	0.005	500,000	Exp #4	0.050	500,000
Loss Amount	cdf	LAS	Loss Amount	cdf	LAS
0.00	0.000000	0.00	0.00	0.000000	0.00
49.15	0.045700	48.02	49.21	0.019500	48.73
100.00	0.089867	95.43	100.00	0.038392	98.05
149.19	0.131200	139.18	149.37	0.056200	145.08
200.00	0.171217	182.31	200.00	0.073565	192.43
342.56	0.276533	292.95	343.62	0.120000	322.15
500.00	0.371892	399.35	500.00	0.162648	456.43
729.42	0.494340	529.40	733.42	0.219840	645.21
1,000.00	0.598159	652.18	1,000.00	0.269918	846.51
1,419.20	0.727210	793.58	1,443.94	0.339720	1,155.13
2,000.00	0.820353	924.97	2,000.00	0.395447	1,506.79
2,883.28	0.911960	1,043.19	3,256.69	0.485113	2,210.19
5,000.00	0.950186	1,189.09	5,000.00	0.549751	3,051.45
6,797.29	0.960808	1,269.07	7,275.66	0.618840	3,997.45
10,000.00	0.966769	1,385.05	10,000.00	0.676551	4,957.25
14,264.10	0.972925	1,513.63	14,236.37	0.749097	6,173.83
20,000.00	0.977502	1,655.80	20,000.00	0.802420	7,466.28
30,790.44	0.983013	1,868.83	30,030.69	0.861207	9,153.31
50,000.00	0.986108	2,165.42	50,000.00	0.890736	11,630.07
72,261.57	0.988482	2,448.25	71,743.39	0.908547	13,812.20
100,000.00	0.990386	2,741.34	100,000.00	0.922253	16,202.71
142,933.77	0.992801	3,102.25	143,357.97	0.939641	19,196.72
200,000.00	0.994618	3,461.20	200,000.00	0.952951	22,238.65
306,605.45	0.996837	3,916.67	311,738.74	0.970510	26,514.86
500,000.00	0.998060	4,410.19	500,000.00	0.980932	31,085.63
700,063.34	0.998817	4,722.62	702,893.51	0.988239	34,213.12
1,000,000.00	0.999323	5,001.59	1,000,000.00	0.993229	36,966.16
1,343,154.66	0.999707	5,168.02	1,343,292.63	0.997074	38,630.66
2,000,000.00	0.999908	5,294.21	2,000,000.00	0.999084	39,892.11
2,493,216.63	0.999985	5,320.55	2,492,457.58	0.999848	40,155.10
5,000,000.00	1.000000	5,339.89	5,000,000.00	0.999998	40,348.87

APPENDIX B

Correlation Matrix for Claim Counts

	WC-$5M Limit	GL-$5M Limit	GL-$2M Limit	GL-$1M Limit	GL-$0.5M Limit	AL-$5M Limit	AL-$2M Limit	AL-$1M Limit	AL-$0.5M Limit	APhD	CP-$50M Limit	CP-$10M Limit	CP-$5M Limit	CP-$2M Limit	CP-$1M Limit
WC-$5M Limit	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GL-$5M Limit	0.0000	1.0000	0.4599	0.4644	0.4624	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GL-$2M Limit	0.0000	0.4599	1.0000	0.4848	0.4828	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GL-$1M Limit	0.0000	0.4644	0.4848	1.0000	0.4875	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GL-$0.5M Limit	0.0000	0.4624	0.4828	0.4875	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AL-$5M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.4572	0.4624	0.4606	0.4553	0.0000	0.0000	0.0000	0.0000	0.0000
AL-$2M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	0.4572	1.0000	0.4853	0.4834	0.4779	0.0000	0.0000	0.0000	0.0000	0.0000
AL-$1M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	0.4624	0.4853	1.0000	0.4889	0.4834	0.0000	0.0000	0.0000	0.0000	0.0000
AL-$0.5M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	0.4606	0.4834	0.4889	1.0000	0.4815	0.0000	0.0000	0.0000	0.0000	0.0000
APhD	0.0000	0.0000	0.0000	0.0000	0.0000	0.4553	0.4779	0.4834	0.4815	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CP-$50M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.8984	0.8987	0.8985	0.8985
CP-$10M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8984	1.0000	0.9002	0.9000	0.9000
CP-$5M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8987	0.9002	1.0000	0.9003	0.9003
CP-$2M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8985	0.9000	0.9003	1.0000	0.9001
CP-$1M Limit	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.8985	0.9000	0.9003	0.9001	1.0000

APPENDIX B

CORRELATION MATRIX FOR AGGREGATE LOSSES

	WC-$5M Limit	GL-$5M Limit	GL-$2M Limit	GL-$1M Limit	GL-$0.5M Limit	AL-$5M Limit	AL-$2M Limit	AL-$1M Limit	AL-$0.5M Limit	APhD	CP-$50M Limit	CP-$10M Limit	CP-$5M Limit	CP-$2M Limit	CP-$1M Limit
WC-$5M Limit	1.0000	0.1859	0.2577	0.2879	0.2841	0.1500	0.2530	0.3311	0.3248	0.3758	0.1186	0.1915	0.1952	0.1954	0.1955
GL-$5M Limit	0.1859	1.0000	0.3090	0.3452	0.3406	0.0596	0.1004	0.1314	0.1290	0.1492	0.0471	0.0760	0.0775	0.0776	0.0776
GL-$2M Limit	0.2577	0.3090	1.0000	0.4784	0.4721	0.0825	0.1392	0.1822	0.1787	0.2068	0.0653	0.1054	0.1074	0.1075	0.1076
GL-$1M Limit	0.2879	0.3452	0.4784	1.0000	0.5275	0.0922	0.1555	0.2035	0.1997	0.2310	0.0729	0.1177	0.1200	0.1201	0.1202
GL-$0.5M Limit	0.2841	0.3406	0.4721	0.5275	1.0000	0.0910	0.1535	0.2009	0.1971	0.2280	0.0719	0.1162	0.1184	0.1185	0.1186
AL-$5M Limit	0.1500	0.0596	0.0825	0.0922	0.0910	1.0000	0.1629	0.2132	0.2091	0.2420	0.0380	0.0614	0.0625	0.0626	0.0626
AL-$2M Limit	0.2530	0.1004	0.1392	0.1555	0.1535	0.1629	1.0000	0.3595	0.3528	0.4081	0.0641	0.1035	0.1054	0.1056	0.1056
AL-$1M Limit	0.3311	0.1314	0.1822	0.2035	0.2009	0.2132	0.3595	1.0000	0.4616	0.5341	0.0838	0.1354	0.1380	0.1381	0.1382
AL-$0.5M Limit	0.3248	0.1290	0.1787	0.1997	0.1971	0.2091	0.3528	0.4616	1.0000	0.5240	0.0823	0.1329	0.1354	0.1355	0.1356
APhD	0.3758	0.1492	0.2068	0.2310	0.2280	0.2420	0.4081	0.5341	0.5240	1.0000	0.0952	0.1537	0.1566	0.1568	0.1569
CP-$50M Limit	0.1186	0.0471	0.0653	0.0729	0.0719	0.0380	0.0641	0.0838	0.0823	0.0952	1.0000	0.5384	0.5486	0.5492	0.5496
CP-$10M Limit	0.1915	0.0760	0.1054	0.1177	0.1162	0.0614	0.1035	0.1354	0.1329	0.1537	0.5384	1.0000	0.8860	0.8869	0.8876
CP-$5M Limit	0.1952	0.0775	0.1074	0.1200	0.1184	0.0625	0.1054	0.1380	0.1354	0.1566	0.5486	0.8860	1.0000	0.9038	0.9045
CP-$2M Limit	0.1954	0.0776	0.1075	0.1201	0.1185	0.0626	0.1056	0.1381	0.1355	0.1568	0.5492	0.8869	0.9038	1.0000	0.9054
CP-$1M Limit	0.1955	0.0776	0.1076	0.1202	0.1186	0.0626	0.1056	0.1382	0.1356	0.1569	0.5496	0.8876	0.9045	0.9054	1.0000

APPENDIX B

Covariance Matrix for Aggregate Losses

	WC-$5M Limit	GL-$5M Limit	GL-$2M Limit	GL-$1M Limit	GL-$0.5M Limit	AL-$5M Limit	AL-$2M Limit	AL-$1M Limit	AL-$0.5M Limit	APhD	CP-$50M Limit	CP-$10M Limit	CP-$5M Limit	CP-$2M Limit	CP-$1M Limit
WC-$5M Limit	3.897E+15	3.447E+13	1.363E+14	3.474E+14	1.660E+14	1.915E+13	7.282E+13	1.811E+14	8.972E+13	2.049E+13	9.398E+13	2.392E+14	5.137E+14	2.649E+14	2.866E+14
GL-$5M Limit	3.447E+13	8.823E+12	7.778E+12	1.982E+13	9.470E+12	3.618E+11	1.376E+12	3.421E+12	1.695E+12	3.870E+11	1.775E+12	4.519E+12	9.704E+12	5.003E+12	5.415E+12
GL-$2M Limit	1.363E+14	7.778E+12	7.182E+13	7.838E+13	3.745E+13	1.431E+12	5.440E+12	1.353E+13	6.702E+12	1.531E+12	7.021E+12	1.787E+13	3.838E+13	1.979E+13	2.141E+13
GL-$1M Limit	3.474E+14	1.982E+13	7.838E+13	3.737E+14	9.543E+13	3.646E+12	1.386E+13	3.447E+13	1.708E+13	3.900E+12	1.789E+13	4.554E+13	9.779E+13	5.042E+13	5.457E+13
GL-$0.5M Limit	1.660E+14	9.470E+12	3.745E+13	9.543E+13	8.760E+13	1.742E+12	6.624E+12	1.647E+13	8.161E+12	1.864E+12	8.548E+12	2.176E+13	4.673E+13	2.409E+13	2.607E+13
AL-$5M Limit	1.915E+13	3.618E+11	1.431E+12	3.646E+12	1.742E+12	4.183E+12	1.536E+12	3.820E+12	1.893E+12	4.322E+11	9.863E+11	2.511E+12	5.391E+12	2.780E+12	3.008E+12
AL-$2M Limit	7.282E+13	1.376E+12	5.440E+12	1.386E+13	6.624E+12	1.536E+12	2.126E+13	1.452E+13	7.196E+12	1.643E+12	3.750E+12	9.546E+12	2.050E+13	1.057E+13	1.144E+13
AL-$1M Limit	1.811E+14	3.421E+12	1.353E+13	3.447E+13	1.647E+13	3.820E+12	1.452E+13	7.677E+13	1.789E+13	4.086E+12	9.326E+12	2.374E+13	5.097E+13	2.628E+13	2.844E+13
AL-$0.5M Limit	8.972E+13	1.695E+12	6.702E+12	1.708E+13	8.161E+12	1.893E+12	7.196E+12	1.789E+13	1.958E+13	2.025E+12	4.620E+12	1.176E+13	2.525E+13	1.302E+13	1.409E+13
APhD	2.049E+13	3.870E+11	1.531E+12	3.900E+12	1.864E+12	4.322E+11	1.643E+12	4.086E+12	2.025E+12	7.626E+11	1.055E+12	2.686E+12	5.767E+12	2.974E+12	3.218E+12
CP-$50M Limit	9.398E+13	1.775E+12	7.021E+12	1.789E+13	8.548E+12	9.863E+11	3.750E+12	9.326E+12	4.620E+12	1.055E+12	1.611E+14	1.367E+14	2.936E+14	1.514E+14	1.639E+14
CP-$10M Limit	2.392E+14	4.519E+12	1.787E+13	4.554E+13	2.176E+13	2.511E+12	9.546E+12	2.374E+13	1.176E+13	2.686E+12	1.367E+14	4.003E+14	7.475E+14	3.854E+14	4.171E+14
CP-$5M Limit	5.137E+14	9.704E+12	3.838E+13	9.779E+13	4.673E+13	5.391E+12	2.050E+13	5.097E+13	2.525E+13	5.767E+12	2.936E+14	7.475E+14	1.778E+15	8.276E+14	8.956E+14
CP-$2M Limit	2.649E+14	5.003E+12	1.979E+13	5.042E+13	2.409E+13	2.780E+12	1.057E+13	2.628E+13	1.302E+13	2.974E+12	1.514E+14	3.854E+14	8.276E+14	4.716E+14	4.618E+14
CP-$1M Limit	2.866E+14	5.415E+12	2.141E+13	5.457E+13	2.607E+13	3.008E+12	1.144E+13	2.844E+13	1.409E+13	3.218E+12	1.639E+14	4.171E+14	8.956E+14	4.618E+14	5.516E+14

ADDRESS TO NEW MEMBERS—NOVEMBER 15, 1999

IT IS EASIER TO BECOME AN ACTUARY

LEROY J. SIMON

Heartiest congratulations—first to those accompanying persons who sacrificed and put up with so much while this morning's new Fellows and Associates struggled to reach this great day. All those who have been through it before you know and understand how difficult it is and of the appreciation you deserve. And, of course, a very warm welcome to you new Fellows and Associates on this milestone day. I hate to be the one that has to tell you this but—it is easier to become an actuary *than to be one*. More on that in a moment, but, as a corollary, learn to be good at accepting criticism, you'll get a lot of practice. The basic nature of our work is such that we must at times deliver messages that others do not want to hear; one of their defensive reactions is to blast the messenger. That does not change the quality of the message, however, so just be right in the first place, learn to accept criticism, and have faith in yourself.

When you entered the room this morning you stepped into an environment that was *created* for you. I am speaking in a broad sense of the environment of traditions, spirit, morals, ethics, and the knowledge base…all that has been built to create this Casualty Actuarial Society. You now stand on the shoulders of those who preceded you. What will you do with this opportunity? Thirty or forty years from now when you retire from active business pursuits, whether you like it or not, you will leave a legacy to those who follow—make it the best legacy you possibly can. You owe that to the CAS, you owe it to those who supported you on this path, but even more so, you owe it to yourself. Yes, it is easier to become an actuary than to be one.

You will have many successes over your actuarial career, so you must remind yourself that the greatest enemy of future progress is past success. We are all comfortable with things that we know and understand. It's easy to *apply* a familiar tool to a new problem—much easier than analyzing the problem to see what tools would best apply and then designing a workable technique and carrying out the solution. Experience in a field makes you comfortable—you know the tool to use even before the problem is completely formulated. On the contrary, you must be willing to turn things on their head and look at them in a new light. You must strive to make change a partner, not an enemy; new input an accomplice, not a rival. And above all, you must force yourself to completely, carefully, thoroughly define the problem without jumping to the method of solution before you have a full grasp of the situation. The tools you have learned through the education process have brought you to today and will guide you through your early years as professionals. Be ready to discard them when improvements come along. Yes, *it is* far easier to become an actuary than to be one.

Forty-five years ago today I became a Fellow and the papers presented to the Annual Meeting of the CAS included two on Work*men*'s Compensation, two on health insurance, one on the Boiler and Machinery experience rating plan, and an elementary one on fire insurance ratemaking. Now look at the program for this meeting: securitization of catastrophe exposures, computer technology, complex models, financial services, discounted cash flow.

There is no secret about how we got from the papers of 45 years ago to the presentations of today—CHANGE. And the only way to cope with such dramatic change over the course of your actuarial career is continuing education and continuous adjustment to the new environment. You have to go to a museum today to see a punched card, which was the standard for data processing in 1954. When you get back to the office, look around and you'll see the museum pieces of the future and they'll be in those museums before you retire. Just make sure your actuarial expertise is not at a 1999 level, because it is far easier to become an actuary than to be one.

Once upon a time…we knew that when the insurance policy said that, to be covered, a loss had to be "sudden and accidental" *meant* that the event had to be sudden and it had to be accidental. Of course that is no longer true today. Once upon a time…your product had to have caused the loss in order for you to be liable for damages. No longer; now you only need to be a member of a class that manufactured products something like the one deemed guilty and you are liable. And now…we have a challenge in the courts alleging that the normal operating costs of running an enterprise are covered under the property insurance policy when they involve the Y2K bug. Stay tuned for the outcome of that one.

You must be prepared for equally bizarre attempts to twist and distort the intent of insurance policies to provide funds for some worthwhile social purpose—"worthwhile" that is for others but life-threatening for our industry. Over the course of your careers don't be surprised to encounter something as strange as this: a tornado has struck a devastating blow to a major city and *heavily damaged* a large residential area…70% of the homes in this area have been flattened but 30% have escaped damage. The insurance industry is ordered to pay up the face amount of all fire and homeowners policies within an area described by the authorities and approved by the court. No, it makes no difference whether your insured's property was only partially damaged or not damaged at all; that was an act of God. The act of the courts is to mobilize the resources available and one of the handiest resources is the insurance companies' funds. Impossible? Too far out? Maybe so, but then there was a time when we thought we knew what sudden and accidental meant, and a time when we thought we were covering the liability of a manufacturer for damage done by *his* product, and a time…and a time…and a…well, maybe it's not so impossible after all.

Yes, it is far, far easier to become an actuary than to be one. But then, that's why we have actuaries like you in the CAS. Your median age is 31. You will be in the forefront at the 2014 celebration of the 100th anniversary of the founding of the Casualty Actuarial Society. You're young, bright, responsible individuals, ready to challenge the world and proud to be members of the Casualty Actuarial Society. Keep your pride of today throughout your entire career. Always remember, there are two broad groups of actuaries—casualty and non-casualty.

You are the last CAS graduating class of the 20th century—but let's keep it in perspective—50 years and 50 days from today, you will still be actuaries but you will be closer to the start of the 22nd century than you will be to the year 2000. Good luck. Now that you've done the easier part and *become* an actuary, get out there and do the rest of the job—*be* one. I'll be watching you because, in some small way, you're my class of '99.

PRESIDENTIAL ADDRESS—NOVEMBER 15, 1999

THE CAS IN THE NEW MILLENNIUM

STEVEN G. LEHMANN

In the field of observation, chance favors the prepared mind.

—Louis Pasteur

This will be the last CAS meeting of the old millennium, at least by the way most people count it. It seems an appropriate time to look back at our roots as well as forward to the new millennium.

Eighty-five years ago last Sunday a new actuarial society was born. Led by a Russian immigrant, Dr. Isaac M. Rubinow, the new society was named the Casualty Actuarial and Statistical Society of America. The name was shortened to the Casualty Actuarial Society in 1921.

The founders of the new society, our forefathers, were innovators and pioneers of a new form of insurance called workmen's compensation insurance, certainly a nontraditional area of practice at that time.

It is interesting to note that the initial examination syllabus set in 1915 had six exams, four Associateship exams and two Fellowship exams. Another early priority of our Society was the appointment of a committee to address new methods of reserving for liability and compensation losses (in other words, research). Thus innovation, research, and education have been hallmarks of our Society from its earliest days.

Eighty-five years ago, our roots were formed.

Eighty-five years ago. . . .

Where will the new millennium take us? Let me offer my predictions of what we'll see in the next ten years and beyond.

Globalization

Globalization—an overused word. But it's a fact that we are seeing actuaries from North America relocating to London, Zurich, Hong Kong, and the Far East. U.S. and Canadian companies are becoming global. European and Asian companies are marketing in the U.S. and Canada. My prediction is that in the next ten–twenty years we will see a globalization of business far beyond anything we've seen to date. A truly global world and world economy, where a flight from New York to London or Paris will be as common as a flight from New York to Chicago is today. A world where actuaries move freely from country to country as part of a typical job progression in getting to know their company.

Convergence of Insurance and Financial Services

Secondly, I see a continuation of the blurring of lines between insurance companies and other financial services. Many insurance products are already a mixture of traditional insurance and financial products. Banks and thrift institutions want additional sales opportunities related to their savings and lending activities. Consumers, borrowing money for a car or house, are likely to be in the market for car and house insurance. Banks and thrifts can use their existing facilities to offer these new products with minimal additional capital expenditures for office space and to an existing client base. Insurance companies want additional marketing outlets and access to the established client bases of financial institutions.

It remains to be seen who will come out on top in these mergers. If it's the banks it is critical to our future that the bankers and investment people become familiar with actuaries and what we can do for them.

Mergers and Acquisitions

Not only are we seeing banks and insurance companies merge, we are seeing an ever-increasing number of mergers within the insurance industry and elsewhere. As the number of insurance companies become fewer with these mergers and acquisitions, we will see actuarial jobs eliminated and consolidated. This has not been a big problem for casualty actuaries in the past. It is becoming one today for casualty actuaries and will continue in the future.

Competition from MBAs and Financial Engineers

A major activity of the CAS this year was a series of interviews with CEOs of insurance companies, reinsurers, brokers, and consulting firms to identify the needs of potential employers of actuarial services and to explore how actuaries could meet those needs. You will be reading about the results of these interviews in the coming months, but I want to focus on one aspect of the findings. The general consensus of the CEOs was that actuaries need to develop better general business skills and a broader business perspective.

We are also hearing about a new profession called financial engineers who are finding employment on Wall Street and Bay Street doing things like pricing options, derivatives, and futures.

If actuaries are to become broader-based problem solvers in the field of risk, we will face competition from MBAs and financial engineers. We will also face competition to *recruit* the best and brightest actuarial students from MBA schools and financial engineer programs. If you were a bright, talented math or business student with the opportunity to work on Wall Street now for a large salary versus taking a series of nine very difficult exams, which way would you go? All I can say is, "Thank God for the *Jobs Rated Almanac.*" Compare a two-year MBA program to the five–ten years it takes to achieve Fellowship. Again, thank God for the *Jobs Rated Almanac*. But we can only ride that horse so long.

Technology

Technology today is truly amazing. I could spend this entire address talking about the Internet, hand-held computers and other communication devices, and where technology is going. We are able to optimize class factors using generalized linear modeling and computer techniques that were impossible ten years ago. We can now run out multiple reserve projections at the touch of a button. Who can predict what new technologies will be able to do ten years from now? What I can predict is that technology will continue to advance at a head-splitting rate, and actuaries must be at the forefront of these advances or we face the danger of irrelevance.

The scientist Louis Pasteur once said, “In the field of observation, chance favors the prepared mind.” Chance favors the prepared mind. I think the truth of this pearl of wisdom goes far beyond the observational sciences. I believe this quotation is relevant to the actuarial field generally and to our position at this moment in time, particularly. Far be it from me to suggest that our exam process might be subject to chance. Some of you might say that, but I would never say it. However, I think that most of you would agree with me that chance does favor those whose mind is well-prepared for the actuarial exams. From my experience with actuarial work after the exam process, again chance favors those who prepare well and prepare hard.

Speaking of pearls of wisdom, one of the job requirements for CAS presidents is that they must read all of the past presidential addresses going back to Isaac Rubinow. I dutifully read through them. In fact I read one a day each night, just before bedtime.

Now, how are Pasteur’s words relevant to our actuarial society at this particular juncture in our history? We have just received a very important report from a CAS Task Force on Non-Traditional Areas of Practice. This Task Force report identifies several potential new areas of practice for casualty actuaries and skill sets needed by future actuaries. The opportunity is there but, only if we prepare our actuaries through a determined effort of education and research to move into these areas. We must seize the initiative and begin immediately on both fronts, or the opportunities will surely be lost, perhaps forever.

We must embrace change. As LeRoy Simon said in his address to the new members: "Make change a partner, not an enemy." Change is hard. It is much easier to sit back and say we're doing okay; why should we change what's worked well for casualty actuaries over the years? There are certainly some things we don't want to change. We don't want to lower our standards for admission to the Society. We don't want to change our fundamental principles, standards of practice, and discipline procedures. However, the changes I'm talking about are in the areas of education, research, technology, and development of new areas of practice.

Early in my presidency, I was asked by an actuary, "Why should I care about growth of the CAS and actuaries generally? After all, that will just mean more competition for existing jobs and consulting work." It seems to me that this is shortsighted. While growth may bring some increased competition, I think that growth of the CAS is in all of our interests.

- It gives the profession a louder voice with policy makers and others.
- It brings in new ideas and approaches.
- It opens up opportunities in nontraditional areas of practice, because if the supply of actuaries is not growing, employers will look to others to meet their needs.
- It keeps our organization alive and vital. I say that the day that we quit growing is the day we begin contracting. The day we quit expanding our markets is the day our markets and demand begin to shrink.

So I propose several CAS initiatives to prepare our minds and ourselves for the new millennium:

1. A mobilization of our research and education efforts in the identified priority areas of nontraditional practice as we did so successfully with our DFA effort a few years ago. It will require our CAS committees to make some major changes in our education and research priorities. It will require a major effort of time, commitment, and funds for the CAS.

2. A broadening of our educational process to make our actuaries broader-based business problem solvers. We should make our education more like an MBA program with emphasis on team building, negotiation, and communication skills. And we must find a way to shorten our examination system, particularly in the basic education area. We should rely more on universities, without lowering our standards or giving up examination on key areas of actuarial practice.

3. A major effort by the CAS, perhaps in combination with the SOA, to develop additional strategic planning tools for actuaries that can be applied to the financial services industry.

If we can do these things, I firmly believe the future will be bright indeed for casualty actuaries. It will expand our actuarial horizons and allow actuaries to move into roles of strategic planning and other leadership positions in the insurance and financial services businesses, and it will make our profession more attractive to the best math and business students.

Earlier I poked a little fun at our presidential addresses. But there are indeed some shrewd insights and words of wisdom in those prior addresses, and some common themes. Perhaps the one overriding theme was best expressed by Al Beer. I think he speaks for all of us when he said that he hopes he *made* *a difference*. I hope that each of you here today will endeavor to make a contribution to the CAS and make a difference. The profession will be better for it, and you will be better for it. For the many, many of you who are already making a contribution, I thank you for it.

I previously spoke about the increasing globalization of the actuarial profession. The CAS now has 112 members outside the U.S. and Canada. We are committed to playing a more active role internationally. The CAS recently appointed a new Vice President–International, and we are becoming more active and visible in the IAA, the International Congress of Actuaries, and meetings of the international presidents of actuarial societies in English-speaking countries. We were recently asked for assistance by the Actuarial Society of India to help them set up a general insurance course there. I believe that these and other activities are vital to our long-term success. It provides better service to our members who are overseas and will lead to expanded opportunities for our North American members who would like to work overseas.

Perhaps the most controversial issue I have had to deal with in my year as President was Mutual Recognition. This was a proposal which arose out of the international presidents' meetings. Under the proposal, Fellows of other actuarial societies outside the U.S. and Canada (such as in Australia and the U.K.) who had achieved their Fellowship by examination, who established residency in the U.S. or Canada, and who met certain other requirements would be granted FCAS status. By the same token CAS Fellows who, for example, went to Australia would be automatically granted Fellowship in the Institute of Actuaries of Australia.

During the year, we have had a CAS task force studying this issue, and I have spoken about it to many of you at CAS Regional Affiliate meetings. Many of you expressed sincere concerns about this proposal. After carefully studying and giving full consideration to member concerns, the task force has recommended that automatic Fellowship not be granted due to the high degree of specialization in the CAS in general insurance compared to the other actuarial societies. A factor in the task force's recommendation was that actuaries from other companies can get practice rights in the U.S. via the American Academy of Actuaries. The task force is recommending an increase in our waiver policy, from the current five exams to seven or eight for Fellows of the Institute of Actuaries who have achieved their Fellowship in general insurance under the current syllabus. A similar policy is likely for the Institute of Actuaries of Australia.

Yesterday the CAS Board agreed with this recommendation, subject to additional information by the CAS Education Policy Committee on specific exam waivers.

This has been a difficult issue. With your help, I think we have reached the right conclusion. More than anything, I think it demonstrates the sensitivity of the CAS Board and leadership to membership concerns.

It has been my good fortune to inherit the reins of the CAS from the capable hands of Mavis Walters. Mavis, I'd like to thank you for your efforts on behalf of the CAS and say it was and is a pleasure working with you. I will also be leaving the CAS in the capable hands of Alice Gannon, and Pat Grannan after Alice. I would also like to thank the Executive Council of the CAS—Bob Miccolis, Kevin Thompson, Gary Dean, Dave Chernick, Abbe Bensimon, and Alice Gannon—who have worked very hard this last year and often don't get the recognition they deserve. And to Howard Bolnick, immediate past president of the SOA, for his friendship over the last two years. Also, Tim Tinsley. Tim, I don't know how I could have done it without you. Thank you, and I'll miss working with you. And my wife Judy, who has put up with the travel and long hours that go with the presidency. Thanks for your patience and your support. And to the members of the CAS, thanks for the memories. I've gotten to meet many of you at Regional Affiliate meetings and other meetings of the

CAS. It's truly been the highlight of my professional career as an actuary.

And finally, to my son Todd and the new members of the CAS. I'd like to close with the inspirational words of Stan Hughey, CAS President, 1974:

Keep your roots deep in the CAS fundamentals. Soar with the wings of new developments which provide better solutions.

MINUTES OF THE 1999 CAS ANNUAL MEETING

November 14–17, 1999

SAN FRANCISCO MARRIOTT

SAN FRANCISCO, CALIFORNIA

Sunday, November 14, 1999

The Board of Directors held their regular quarterly meeting from 9:00 a.m. 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 1999 CAS Executive Council members briefly discussed their roles in the Society with the new members. In addition, Robert A. Anker, 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 15, 1999

Registration continued from 7:00 a.m. to 8:00 a.m.

CAS President Steven G. Lehmann opened the business session at 8:00 a.m. and introduced members of the Executive Council and the CAS Board of Directors. Mr. Lehmann also recognized past presidents of the CAS who were in attendance at the meeting, including: Robert A. Anker (1996), Irene K. Bass (1993), Albert J. Beer (1995), Phillip N. Ben-Zvi (1985), Ronald L. Bornhuetter (1975), Charles A. Bryan (1990), Michael Fusco (1989), David G. Hartman (1987), Charles C. Hewitt Jr. (1972), Carlton W. Honebein (1983), Allan M. Kaufman (1994), C.K. "Stan" Khury (1984), W. James MacGinnitie (1979), George D. Morison (1976), Kevin M. Ryan (1988), Jerome A. Scheibl (1980), LeRoy J. Simon (1971), Michael L. Toothman (1991), Mavis A. Walters (1997), and Michael A. Walters (1986).

Mr. Lehmann also recognized special guests in the audience: Howard J. Bolnick, past president of the Society of Actuaries; A. Norman Crowder, president of the Society of Actuaries; Muneo Kawasaki, representative of the president of the Institute of Actuaries of Japan; Lonnie Liu, representative of the chairman of the Actuarial Institute of the Republic of China; David J. Oakden, president-elect of the Canadian Institute of Actuaries; and John P. Ryan, board member of the Institute of Actuaries.

Mr. Lehmann then announced the results of the CAS elections. The next president will be Alice H. Gannon, and the president-elect will be Patrick J. Grannan. Members of the CAS Executive Council for 1999–2000 will be: Curtis Gary Dean, vice president–administration; Mary Frances Miller, vice president–admissions; Abbe Sohne Bensimon, vice president–continuing education; LeRoy A. Boison, vice president–international; David R. Chernick, vice president–programs and communication; and Gary R. Josephson, vice president–research and development. The vice president–international is a new position approved by the Board of Directors in the fall of 1999. New members of the CAS Board of Directors are Amy S. Bouska, Stephen P. D'Arcy, Frederick O. Kist, and Susan E. Witcraft.

Abbe S. Bensimon and Kevin B. Thompson announced the new Associates and Alice H. Gannon announced the new Fellows. The names of these individuals follow.

NEW FELLOWS

Rimma Abian	Lisa A. Bjorkman	Bethany L. Cass
Ethan David Allen	Suzanne E. Black	Jean-François Chalifoux
Mark B. Anderson	Jonathan Everett Blake	Bryan C. Christman
Martin S. Arnold	Ann M. Bok	Darrel W. Chvoy
William P. Ayres	Michael D. Brannon	Gary T. Ciardiello
Richard J. Babel	Anthony E. Cappelletti	Christopher William Cooney
Cynthia A. Bentley	Martin Carrier	

Brian K. Cox
Claudia Barry Cunniff
Karen Barrett Daley
Timothy Andrew Davis
Jean A. DeSantis
Kurt S. Dickmann
Christopher S. Downey
Michael Edward Doyle
Peter F. Drogan
Denis Dubois
Mary Ann Duchna-Savrin
Rachel Dutil
Dawn E. Elzinga
Jean-Pierre Gagnon
Donald M. Gambardella
Gary J. Ganci
Thomas P. Gibbons
John T. Gleba
Matthew E. Golec
Philippe Gosselin
Jay C. Gotelaere
David Thomas Groff
Scott T. Hallworth
Gregory Hansen
Michael B. Hawley
Jodi J. Healy
Noel M. Hehr
Christopher Ross Heim
David E. Heppen
Ronald J. Herrig
Thomas A. Huberty
Brian L. Ingle
James B. Kahn
Chad C. Karls
Mark J. Kaufman
James M. Kelly
Sarah Krutov
James D. Kunce
Jean-Sebastien Lagarde
Yin Lawn
David Leblanc-Simard
Kevin A. Lee
P. Claude Lefebvre
Siu K. Li
Janet G. Lindstrom
Lee C. Lloyd
William R. Maag
David E. Marra
Michael Boyd Masters
Bonnie C. Maxie
Jeffrey F. McCarty
Douglas W. McKenzie
Allison Michelle McManus
James R. Merz
Paul W. Mills
Christopher J. Monsour
David Patrick Moore
François L. Morissette
Matthew C. Mosher
Roosevelt C. Mosley
Donna M. Nadeau
Catherine A. Neufeld
Hiep T. Nguyen
Randall S. Nordquist
Michael A. Nori
James L. Nutting
Christopher Edward Olson
Denise R. Olson
David Anthony Ostrowski
Teresa K. Paffenback
Charles Pare
M. Charles Parsons
Luba O. Pesis
Karen L. Queen
Kathleen Mary Quinn
Yves Raymond
Hany Rifai
John W. Rollins
Seth Andrew Ruff
David L. Ruhm
Tracy A. Ryan
Rajesh V. Sahasrabuddhe
Michael C. Schmitz
Nathan Alexander Schwartz
Bret Charles Shroyer
Matthew Robert Sondag
Jay Matthew South
Angela Kaye Sparks
Brian Tohru Suzuki
Adam M. Swartz
Nitin Talwalkar
Dom M. Tobey
Jeffrey S. Trichon
Kai Lee Tse
Leslie Alan Vernon
Kyle Jay Vrieze
Edward H. Wagner
Benjamin A. Walden
Robert J. Wallace

Patricia Cheryl White
Wendy L. Witmer
Simon Kai-Yip Wong
Vincent F. Yezzi
Sheng H. Yu

NEW ASSOCIATES

Michael D. Adams
Genevieve L. Allen
Saeeda Behbahany
Penelope A. Bierbaum
Tony Francis Bloemer
Caleb M. Bonds
Maureen Ann Boyle
Jeremy James Brigham
Kin Lun (Victor) Choi
Alan R. Clark
Brian Roscoe Coleman
Douglas Lawrence Dee
Jonathan Mark Deutsch
Richard James Engelhuber
Weishu Fan
Kathleen Marie Farrell
Richard A. Fuller
Rainer Germann
Isabelle Gingras
Peter Scott Gordon
Stephanie Ann Gould
Robert Andrew Grocock
David Lee Handschke
Karen Lerner Jiron
Robert C. Kane
Linda S. Klenk
Ravi Kumar
Julie-Linda Laforce
John B. Landkamer
Aaron Michael Larson
Shangjing Li
Joshua Nathan Mandell
Kevin Paul McClanahan
Ian John McCracken
Shawn Allan McKenzie
Christian Menard
Peter Victor Polanskyj
Josephine Teruel Richardson
Marn Rivelle
Tina Shaw
Joseph Allen Smalley
Michael William Starke
David K. Steinhilber
Stephen James Streff
Josephine L. C. Tan
Javanika Patel Weltig
Rosemary Gabriel Wickham
Apryle Oswald Williams
Dean Michael Winters
Jeffrey S. Wood

Mr. Lehmann then introduced LeRoy J. Simon, a past president of the Society, who presented the Address to New Members.

Following the address, David R. Chernick, vice president–programs and communications, briefly highlighted the meeting's programs and thanked the CAS Program Planning Committee. Mr. Chernick then introduced Gary R. Josephson, chairperson of the CAS Committee on Review of Papers. Mr. Josephson announced that the following would be presented: four *Proceedings* papers, two discussions of previous *Proceedings* papers, and one author's response to a discussion of his paper. In addition, one paper by Dr. Klaus D. Schmidt would be published in the 1999 *Proceedings* but would not be presented at this meeting. (Note: The paper, "The 1999 Table of Insurance Charges," by William R. Gillam, was presented at the 1999 CAS Annual Meeting but is not published in the 1999 *Proceedings*.)

Mr. Josephson began the awards program by announcing that the 1999 Woodward-Fondiller Prize was given to Stephen J. Mildenhall for his paper, "A Systematic Relationship Between Minimum Bias Methods and Generalized Linear Models." Mr. Josephson then presented the 1999 CAS Dorweiler Prize to Gary G. Venter for his paper, "Testing the Assumptions of Age-to-Age Factors." Mr. Mildenhall's paper is published in this edition of the *Proceedings*. Mr. Venter's was published in last year's *Proceedings*, Volume LXXXV.

Mr. Lehmann presented the 1999 CAS Matthew S. Rodermund Service Award to John H. Muetterties, who was chosen for his outstanding contributions to the actuarial profession.

Mr. Lehmann then requested a moment of silence in honor of those CAS members who passed away since November 1998. They are: John R. Bevan, Martin Bondy, Robert L. Hurley, Daniel J. Lyons, and Philipp K. Stern.

In a final item of business, Mr. Lehmann acknowledged a donation of $15,000 from D.W. Simpson & Company to the CAS Trust (CAST). The donation was made October 4, 1999.

Mr. Lehmann then concluded the business session of the Annual Meeting and introduced the featured speaker, Gloria Borger. Borger is a political reporter/columnist and contributing editor for *U.S. News and World Report,* and a regular panelist on PBS' *Washington Week in Review*.

After a refreshment break, the first General Session was held from 10:45 a.m. to 12:15 p.m.

"Past Presidents' Perspectives: An Actuarial Career"

Moderators: Albert J. Beer
President
Munich–American RiskPartners

Michael Fusco
Senior Executive Vice President
Insurance Services Office, Inc.

Panelists: Irene K. Bass
Consulting Actuary
Bass & Khury

Ronald L. Bornhuetter
Chairman, Retired
NAC Re Corporation

Carlton W. Honebein
Consultant

C. K. "Stan" Khury
Consulting Actuary
Bass & Khury

W. James MacGinnitie
Consultant

Following the general session, CAS President Steven G. Lehmann gave his Presidential Address at the luncheon. At the luncheon's end, Mr. Lehmann officially passed on the CAS presidential gavel to the new CAS president, Alice H. Gannon.

After the luncheon, the afternoon was devoted to presentations of concurrent sessions, which included presentations of the *Proceedings* papers. The panel presentations from 1:30 p.m. to 3:00 p.m. covered the following topics:

1. Weather Hedge Products

Moderator: Kenneth J. Bock
Managing Director
American Re Financial Products

Panelists: David Molyneux
Assistant Vice President
Zurich Re North America, Inc.

Paul Murray
Director of Marketing and
Business Development
Castlebridge Partners, LLC

2. Report of the CAS Y2K Work Group

Moderator/
Panelist: Raja R. Bhagavatula
Consulting Actuary
Milliman & Robertson, Inc.

Panelists: Philip D. Miller
Consulting Actuary
Tillinghast-Towers Perrin

Paul G. O'Connell
Principal
PricewaterhouseCoopers LLP

3. The Debate on Competitive Auto Replacement Parts

Moderator: John W. Rollins
Actuary
Florida Farm Bureau Insurance
Companies

Panelists: Robert J. Hurns
Associate Counsel
National Association of Independent
Insurers

Pete A. Tagliapietra
Senior Vice President of Strategic
Planning and Business Development
Mitchell International

4. Securitization: An Update

Moderator: Frederick O. Kist
Senior Vice President & Corporate
Actuary
CNA Insurance Companies

Panelists: David A. Lalonde
Vice President–Risk Transfer Services
Applied Insurance Research

Glenn G. Meyers
Assistant Vice President
Insurance Services Office, Inc.

Susan E. Witcraft
Consulting Actuary
Milliman & Robertson, Inc.

5. Commercial Lines Deregulation—Opportunities and Risks

Moderator: William M. Wilt
Vice President/Senior Analyst
Moody's Investor Service

Panelists: Raul R. Allegue
Second Vice President–Government Affairs
Travelers Property and Casualty

Joseph A. DiGiovianni
Senior Vice President–State Affairs
American Insurance Association

Gregory S. Martino
Deputy Insurance Commissioner
Pennsylvania Insurance Department

The following 1999 *Proceedings* Papers were presented:

1. "The 1999 Table of Insurance Charges"*

Author: William R. Gillam
Quality Casualty Consulting

2. "Downward Bias of Using High-Low Averages for Loss Development Factors"

Author: Cheng-Sheng Peter Wu
Deloite & Touche LLP

* This paper is not included in the 1999 edition of the *Proceedings*.

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. Task Force on Complex Models

   Moderator: Karen F. Terry
   Actuary II
   State Farm Fire & Casualty Company

   Panelists: Paul E. Kinson
   Consulting Actuary
   Liscord, Ward & Roy, Inc.

   Ronald T. Kozlowski
   Consulting Actuary
   Tillinghast-Towers Perrin

2. Questions and Answers with the CAS Board of Directors

   Moderator: Alice H. Gannon
   President-Elect
   Casualty Actuarial Society

   Panelists: Paul Braithwaite
   Senior Vice President
   Zurich Re

   Charles A. Bryan
   Senior Vice President-Chief Actuary
   Nationwide Insurance Company

   Jerome A. Degerness
   President
   Degerness Consulting Services, Inc.

   Richard J. Roth Jr.
   Chief Property/Casualty Actuary
   California Department of Insurance

3. Auto-Choice Reform Act

Moderator:	Michael J. Miller Principal and Consulting Actuary Miller, Herbers, Lehmann & Associates, Inc.
Panelists:	Stephen J. Carroll Senior Economist RAND, The Institute for Civil Justice
	David F. Snyder Assistant General Counsel American Insurance Association
	Elizabeth A. Sprinkel Senior Vice President & Chief Research Officer Insurance Research Council

Proceedings papers presented during this time were:

1. Discussion of "Loss Prediction by Generalized Least Squares"
   (by Leigh J. Halliwell, *PCAS* LXXXIII, 1996, p. 436)

Discussion by:	Michael D. Hamer The Zurich Center

2. Author's Response to Discussion of "Loss Prediction by Generalized Least Squares"
   (by Leigh J. Halliwell, *PCAS* LXXXIII, 1996, p. 436)

Author:	Leigh J. Halliwell American Re-Insurance Company

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 16, 1999

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.:

"Reassessing Seismic Hazards"

Moderator:	Ronald T. Kozlowski Consulting Actuary Tillinghast-Towers Perrin
Panelists:	Michael L. Blanpied Associate Chief Scientist for Scientific Programs, Earthquake Hazards Team United States Geological Survey
	Seth Stein Department of Geological Sciences Northwestern University

"Financial Services Reform"

Moderator:	Mavis A. Walters Executive Vice President Insurance Services Office, Inc.
Panelists:	Martin Carus State Insurance Officer American International Group
	Robert Dibblee Senior Vice President, Government Relations National Association of Independent Insurers
	Woody Girion Chief of Financial Analysis Division California Department of Insurance

Robert W. Stein
Partner
Ernst & Young LLP

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 presented were:

1. Privatization of Workers Compensation Funds

Moderator:	Michael C. Dubin Consulting Actuary Milliman & Robertson, Inc.
Panelists:	Spencer M. Gluck Senior Managing Director Gerling Global Financial Products
	G. Kevin Saba President Capstone Technologies

2. Volunteering Within the CAS—Working to Advance the Profession

Moderator:	Roger A. Schultz Member of the CAS Committee on Volunteer Resources
Panelists:	Nancy A. Braithwaite Chairperson, Syllabus Committee
	Kristine E. Plickys Member, CAS Examination Committee
	Gary E. Shook President, Casualty Actuaries of the Mid-Atlantic Region

The following *Proceedings* papers were presented:

1. "Modeling Losses With the Mixed Exponential Distribution"

   Author: Clive L. Keatinge
   Insurance Servies Office, Inc.

2. Discussion of "Aggregation of Correlated Risk Portfolios: Models & Algorithms"
   (by Shaun S. Wang, *PCAS* LXXXV, 1998, Book 2, p. 848)

   Discussion by: Glenn G. Meyers
   Insurance Services Office, Inc.

Various committee meetings were held from 12:00 p.m. to 5:00 p.m. Certain concurrent sessions that had been presented earlier during the meeting were also repeated from 12:30 p.m. to 2:00 p.m. Additional concurrent sessions presented at this time were:

1. Internet and e-Commerce Exposure

   Moderator: Hilary Rowen
   Partner
   Thelen, Reid & Priest

   Panelists: Julie K. Davis
   Executive Vice President
   Aon Risk Services, Inc.

   Kathryn I. Lovaas
   Vice President, Technology
   St. Paul Companies, Inc.

2. The California Workers Compensation Marketplace

   Moderator: David M. Bellusci
   Senior Vice President and Chief Actuary
   Workers Compensation Insurance Rating Bureau of California

Panelists: Robert T. Reville
Economist
RAND, The Institute for Civil Justice

Alex Swedlow
Principal
Applied Outcomes Research

Following the concurrent sessions, a special Actuarial Standards Board Hearing was held from 2:00 p.m. to 5:30 p.m.

Entertainment and a buffet dinner were held from 7:00 p.m. to 10:00 p.m.

Wednesday, November 17, 1999

Certain concurrent sessions were repeated from 8:00 a.m. to 9:30 a.m. Additional concurrent sessions presented at this time were:

1. The Deregulation of Pacific Rim Insurance Markets

   Moderator: Nancy A. Braithwaite
   Assistant Vice President
   Insurance Services Office

   Panelists: Frank J. Karlinski
   Vice President
   American International Underwriters

   Lee R. Steeneck
   Vice President and Actuary
   General Reinsurance Corporation

2. Introduction to the CAS Examination Committee

   Moderator: Thomas G. Myers
   Vice President
   Prudential Property & Casualty Insurance

   Panelists: J. Thomas Downey
   Manager, Admissions
   Casualty Actuarial Society

Larry A. Haefner
Vice President, Strategic Planning
CGU Insurance Companies

Donald D. Palmer
Manager, Actuarial Services
Manitoba Public Insurance Corporation

3. Discounted Cash Flow Models

Moderator: Robert F. Wolf
Consulting Actuary
William M. Mercer, Inc.

Panelists: Russell E. Bingham
Vice President Corporate Research
The Hartford

Philip S. Borba
Economic Consultant
Milliman & Robertson, Inc.

Richard A. Derrig
Senior Vice President
Automobile Insurers Bureau of Massachusetts

The following *Proceedings* paper was presented:

"Residual Market Pricing"

Author: Richard B. Amundson
Minnesota Department of Commerce

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.

"Technology"

Moderator: Stephen P. Lowe
Chief Actuary
Tillinghast-Towers Perrin

Panelists:	Gayle E. Haskell Risk Manager, Senior Vice President Coregis Insurance Group
	Jeffrey O'Dell Executive Director United Services Automobile Association
	Jaimie Pickles Vice President, Consulting and Actuarial Services InsWeb Corporation

Steven G. Lehmann officially adjourned the 1999 CAS Annual Meeting at 11:45 a.m. after closing remarks and an announcement of future CAS meetings.

Attendees of the 1999 CAS Annual Meeting

The 1999 CAS Annual Meeting was attended by 484 Fellows, 185 Associates, and 61 Guests. The names of the Fellows and Associates in attendance follow:

FELLOWS

Rimma Abian	Nolan E. Asch	Linda L. Bell
Barbara J. Addie	Richard V. Atkinson	Gary F. Bellinghausen
Martin Adler	Roger A. Atkinson	David M. Bellusci
Rhonda K. Aikens	William M. Atkinson	Phillip N. Ben-Zvi
Ethan D. Allen	Karen F. Ayres	Abbe Sohne Bensimon
Timothy Paul Aman	William P. Ayres	Cynthia A. Bentley
Richard B. Amundson	Richard J. Babel	Regina M. Berens
Dean R. Anderson	Anthony J. Balchunas	Steven L. Berman
Mark B. Anderson	D. Lee Barclay	Lisa M. Besman
Scott C. Anderson	W. Brian Barnes	Neil A. Bethel
Charles M. Angell	Irene K. Bass	Raja R. Bhagavatula
Robert A. Anker	Todd R. Bault	David R. Bickerstaff
Steven D. Armstrong	Philip A. Baum	Lisa A. Bjorkman
Martin S. Arnold	Andrea C. Bautista	Suzanne E. Black

Jonathan Everett Blake
Cara M. Blank
Barry E. Blodgett
LeRoy A. Boison
Ann M. Bok
Ronald L. Bornhuetter
Charles H. Boucek
Pierre Bourassa
Amy S. Bouska
Roger W. Bovard
Christopher K. Bozman
Nancy A. Braithwaite
Paul Braithwaite
Michael D. Brannon
Malcolm E. Brathwaite
Margaret A. Brinkmann
J. Eric Brosius
Lisa J. Brubaker
Kirsten R. Brumley
Ron Brusky
Charles A. Bryan
Christopher J. Burkhalter
Jeanne H. Camp
Anthony E. Cappelletti
Kenneth E. Carlton
Martin Carrier
Bethany L. Cass
Jean-François Chalifoux
David R. Chernick
Kasing Leonard Chung
Darrel W. Chvoy
Gary T. Ciardiello
Mark M. Cis
Jo Ellen Cockley
Howard L. Cohen
Jeffrey R. Cole
Robert F. Conger
Eugene C. Connell
Christopher William Cooney
Brian C. Cornelison
Francis X. Corr
Gregory L. Cote
Michael D. Covney
Brian K. Cox
Kathleen F. Curran
Ross A. Currie
Daniel J. Czabaj
Ronald A. Dahlquist
Kenneth S. Dailey
Charles Anthony Dal Corobbo
Karen Barrett Daley
Guy Rollin Danielson
Robert N. Darby
Jeffrey W. Davis
Timothy Andrew Davis
Michael L. DeMattei
Jean A. DeSantis
Curtis Gary Dean
Jerome A. Degerness
Marie-Julie Demers
Kurt S. Dickmann
Behram M. Dinshaw
Scott H. Dodge
John P. Donaldson
Christopher S. Downey
Michael Edward Doyle
Peter F. Drogan
Michael C. Dubin
Denis Dubois
Diane Symnoski Duda
Janet E. Duncan
Rachel Dutil
Tammy L. Dye
Richard D. Easton
Bob D. Effinger
Gary J. Egnasko
Valere M. Egnasko
Donald J. Eldridge
John W. Ellingrod
Paula L. Elliott
Dawn E. Elzinga
Charles C. Emma
Martin A. Epstein
Paul E. Ericksen
Dianne L. Estrada
Glenn A. Evans
Doreen S. Faga
Richard J. Fallquist
Randall A. Farwell
Dennis D. Fasking
Richard I. Fein
Russell S. Fisher
William G. Fitzpatrick
James E. Fletcher
Daniel J. Flick
John R. Forney
Russell Frank
Jacqueline Frank Friedland

Michael Fusco
Jean-Pierre Gagnon
Luc Gagnon
John E. Gaines
Cecily A. Gallagher
Donald M. Gambardella
Gary J. Ganci
Alice H. Gannon
Steven A. Gapp
Robert W. Gardner
Roberta J. Garland
Kathy H. Garrigan
James J. Gebhard
Richard J. Gergasko
Margaret Wendy Germani
Thomas P. Gibbons
John F. Gibson
Richard N. Gibson
Bruce R. Gifford
Judy A. Gillam
William R. Gillam
Michael Ambrose Ginnelly
Nicholas P. Giuntini
Olivia Wacker Giuntini
John T. Gleba
Spencer M. Gluck
Steven F. Goldberg
Philippe Gosselin
Jay C. Gotelaere
Patrick J. Grannan
Gary Grant
Anne G. Greenwalt
Daniel Cyrus Greer
Daniel E. Greer
Cynthia M. Grim
Charles Gruber
Denis G. Guenthner
Larry A. Haefner
David N. Hafling
Kyleen Knilans Hale
Allen A. Hall
Leigh Joseph Halliwell
Scott T. Hallworth
George M. Hansen
Gregory Hansen
Robert L. Harnatkiewicz
Steven Thomas Harr
David C. Harrison
David G. Hartman
Gayle E. Haskell
Marcia C. Hayden
David H. Hays
Jodi J. Healy
Noel M. Hehr
Christopher Ross Heim
Suzanne E. Henderson
David E. Heppen
Kirsten Costello Hernan
Ronald J. Herrig
Richard J. Hertling
Charles C. Hewitt
Kathleen A. Hinds
Alan M. Hines
Robert J. Hopper
Ruth A. Howald
George A. Hroziencik
Thomas A. Huberty
David Dennis Hudson
Jeffrey R. Hughes
Stephen Jameson
Christian Jobidon
Eric J. Johnson
Jennifer Polson Johnson
Kurt J. Johnson
Larry D. Johnson
Marvin A. Johnson
Jeffrey R. Jordan
Gary R. Josephson
John J. Joyce
Jeremy M. Jump
James B. Kahn
Frank J. Karlinski
Chad C. Karls
Allan M. Kaufman
Mark J. Kaufman
Clive L. Keatinge
Glenn H. Keatts
James M. Kelly
Rebecca Anne Kennedy
Allan A. Kerin
C.K. “Stan” Khury
Ann L. Kiefer
Gerald S. Kirschner
Frederick O. Kist
Warren A. Klawitter
Michael F. Klein
Joel M. Kleinman
Craig W. Kliethermes
Leon W. Koch
Timothy F. Koester
John J. Kollar

Ronald T. Kozlowski
Israel Krakowski
Gustave A. Krause
Rodney E. Kreps
Jane Jasper Krumrie
Sarah Krutov
Jeffrey L. Kucera
James D. Kunce
Jason Anthony Kundrot
Howard A. Kunst
Edward M. Kuss
Salvatore T. LaDuca
David A. Lalonde
Timothy J. Landick
Dennis L. Lange
Matthew G. Lange
James W. Larkin
Yin Lawn
David Leblanc-Simard
Kevin A. Lee
Thomas C. Lee
Marc-Andre Lefebvre
P. Claude Lefebvre
Merlin R. Lehman
Steven G. Lehmann
Elizabeth Ann Lemaster
Winsome Leong
Andre L'Esperance
Joseph W. Levin
Jennifer McCullough Levine
Siu K. Li
Peter M. Licht
Orin M. Linden
Janet G. Lindstrom
Barry Lipton
Richard Borge Lord
Stephen P. Lowe
William R. Maag
W. James MacGinnitie
Christopher P. Maher
Lawrence F. Marcus
Blaine C. Marles
Leslie R. Marlo
David E. Marra
Michael Boyd Masters
Bonnie C. Maxie
Kevin C. McAllister
Jeffrey F. McCarty
James B. McCreesh
Douglas W. McKenzie
David W. McLaughry
Allison Michelle McManus
Dennis C. Mealy
William T. Mech
Brian James Melas
Stephen V. Merkey
James R. Merz
Glenn G. Meyers
Robert S. Miccolis
David L. Miller
Mary Frances Miller
Michael J. Miller
Philip D. Miller
Ronald R. Miller
William J. Miller
Paul W. Mills
Neil B. Miner
Camille Diane Minogue
John H. Mize
Frederic James Mohl
David Molyneux
Richard B. Moncher
Christopher J. Monsour
Andrew Wakefield Moody
Rebecca A. Moody
Brian C. Moore
Bruce D. Moore
David Patrick Moore
George D. Morison
François L. Morissette
Robert Joseph Moser
Matthew C. Mosher
Roosevelt C. Mosley
John H. Muetterties
Todd B. Munson
Daniel M. Murphy
Giovanni A. Muzzarelli
Nancy R. Myers
Thomas G. Myers
Donna M. Nadeau
Vinay Nadkarni
Allan R. Neis
Hiep T. Nguyen
Mindy Y. Nguyen
Gary V. Nickerson
William A. Niemczyk
Ray E. Niswander
Randall S. Nordquist
Michael A. Nori
James L. Nutting
Paul G. O'Connell

David J. Oakden
Christopher Edward Olson
Denise R. Olson
William L. Oostendorp
David Anthony Ostrowski
Teresa K. Paffenback
Donald D. Palmer
Charles Pare
Curtis M. Parker
M. Charles Parsons
Kathleen M. Pechan
Wende A. Pemrick
Melanie T. Pennington
Luba O. Pesis
Charles I. Petit
Mark W. Phillips
Daniel C. Pickens
Kristine E. Plickys
Brian D. Poole
Dale S. Porfilio
Stuart Powers
Joseph J. Pratt
Ronald D. Pridgeon
Mark Priven
Arlie J. Proctor
Mark R. Proska
Karen L. Queen
Mark S. Quigley
Kathleen Mary Quinn
Richard A. Quintano
Jeffrey C. Raguse
Kara Lee Raiguel
Donald K. Rainey
Scott E. Reddig
Daniel A. Reppert
Hany Rifai
Tracey S. Ritter
Dennis L. Rivenburgh
Douglas S. Rivenburgh
Sharon K. Robinson
John W. Rollins
Deborah M. Rosenberg
Kevin D. Rosenstein
Gail M. Ross
Richard J. Roth
Jean-Denis Roy
Seth Andrew Ruff
Jason L. Russ
James V. Russell
Kevin M. Ryan
Tracy A. Ryan
Rajesh V. Sahasrabuddhe
Manalur S. Sandilya
Donald D. Sandman
Jerome A. Scheibl
Timothy L. Schilling
Michael C. Schmitz
Roger A. Schultz
Mark E. Schultze
Nathan Alexander Schwartz
Susanne Sclafane
Jeffery J. Scott
Kim A. Scott
Mark R. Shapland
Michelle G. Sheng
Margaret Tiller Sherwood
Jeffrey Parviz Shirazi
Gary E. Shook
Edward C. Shoop
LeRoy J. Simon
David Skurnick
Lee M. Smith
M. Kate Smith
Richard A. Smith
Linda D. Snook
Matthew Robert Sondag
Jay Matthew South
Angela Kaye Sparks
Daniel L. Splitt
Barbara A. Stahley
Thomas N. Stanford
Lee R. Steeneck
John A. Stenmark
Michael J. Steward
Richard A. Stock
Brian Tohru Suzuki
Christian Svendsgaard
Scott J. Swanay
Adam M. Swartz
Andrea M. Sweeny
Susan T. Szkoda
Nitin Talwalkar
Catherine Harwood Taylor
Karen F. Terry
Patricia A. Teufel
Kevin B. Thompson
Barbara H. Thurston
Dom M. Tobey
Darlene P. Tom
Michael L. Toothman
Cynthia Traczyk

Jeffrey S. Trichon
Everett J. Truttmann
Kai Lee Tse
Theresa Ann Turnacioglu
Jean Vaillancourt
Peter S. Valentine
John V. Van de Water
Richard L. Vaughan
Gary G. Venter
Leslie Alan Vernon
Kyle Jay Vrieze
Edward (Ted) H. Wagner
Robert H. Wainscott
Benjamin A. Walden
Glenn M. Walker
Robert J. Wallace
Lisa Marie Walsh
Mavis A. Walters
Michael A. Walters
Jeffrey D. White
Jonathan White
Patricia Cheryl White
Gnana K. Wignarajah
William Robert Wilkins
William M. Wilt
John J. Winkleman
Martha A. Winslow
Michael L. Wiseman
Susan E. Witcraft
David A. Withers
Wendy L. Witmer
Richard G. Woll
Simon Kai-Yip Wong
Patrick B. Woods
Walter C. Wright
Cheng-Sheng P. Wu
Vincent F. Yezzi
Jeffery Michael Zacek
Alexander Guangjian Zhu
John D. Zicarelli
Ralph T. Zimmer

ASSOCIATES

Anthony L. Alfieri
Genevieve L. Allen
Nancy S. Allen
Robert C. Anderson
James A. Andler
Anju Arora
Robert D. Bachler
Paul C. Barone
Andrew S. Becker
Saeeda Behbahany
Eric D. Besman
Penelope A. Bierbaum
Tony Francis Bloemer
Thomas S. Boardman
Caleb M. Bonds
John T. Bonsignore
Lesley R. Bosniack
Maureen Ann Boyle
Richard Albert Brassington
Jeremy James Brigham
Hayden Heschel Burrus
Michelle L. Busch
Stephanie T. Carlson
Kin Lun (Victor) Choi
Wei Chuang
Michelle Codere
Brian Roscoe Coleman
Thomas V. Daley
Douglas Lawrence Dee
William Der
Sean R. Devlin
David K. Dineen
Gordon F. Diss
Sharon C. Dubin
François Richard Dumontet
James Robert Elicker
Richard James Engelhuber
Gregory James Engl
Brian A. Evans
Joseph G. Evleth
Charles V. Faerber
Weishu Fan
Kathleen Marie Farrell
William P. Fisanick
Chauncey E. Fleetwood
David Michael Flitman

Charles D. Foley
Kai Y. Fung
Charles E. Gegax
Isabelle Gingras
Theresa Giunta
Todd Bennett Glassman
Terry L. Goldberg
Peter Scott Gordon
Stephanie Ann Gould
Robert Andrew Grocock
Christopher Gerald Gross
Nasser Hadidi
Rebecca N. Hai
David Lee Handschke
Adam D. Hartman
Gary M. Harvey
Philip E. Heckman
Kevin B. Held
Joseph A. Herbers
Thomas Edward Hinds
David D. Hu
Jane W. Hughes
Jeffrey R. Ill
Philip M. Imm
Susan Elizabeth Innes
David H. Isaac
Jean-Claude Joseph Jacob
Karen Lerner Jiron
William Rosco Jones
James W. Jonske
Edwin G. Jordan
Robert C. Kane
Pamela A. Kaplan
David L. Kaufman
Scott A. Kelly
Paul E. Kinson
Linda S. Klenk
Brandelyn C. Klenner
Elina L. Koganski
Andrew M. Koren
Karen Lee Krainz
Richard Scott Krivo
Frank O. Kwon
Robin M. LaPrete
David W. Lacefield
Julie-Linda Laforce
Elaine Lajeunesse
Aaron Michael Larson
Dennis H. Lawton
Bradley R. LeBlond
Stephen E. Lehecka
Todd William Lehmann
Glen Alan Leibowitz
Brendan Michael Leonard
Giuseppe F. Lepera
Shangjing Li
Sharon Xiaoyin Li
James P. Lynch
Joshua Nathan Mandell
Gabriel O. Maravankin
Jason N. Masch
Emma Macasieb McCaffrey
Patrice McCaulley
Kevin Paul McClanahan
Ian John McCracken
Jennifer Ann McCurry
Heather L. McIntosh
Shawn Allan McKenzie
Christian Menard
Richard Ernest Meuret
Karen M. Moritz
John V. Mulhall
Mark Naigles
Henry E. Newman
Lynn Nielsen
Christopher Maurice Norman
Corine Nutting
Mihaela Luminita S. O'Leary
Steven Brian Oakley
Dale F. Ogden
Christy Beth Olson
Rebecca Ruth Orsi
Kerry S. Patsalides
Claude Penland
Amy Ann Pitruzzello
Glen-Roberts Pitruzzello
Peter Victor Polanskyj
Anthony E. Ptasznik
Richard B. Puchalski
Eric K. Rabenold
William Dwayne Rader
Brenda L. Reddick
John Dale Reynolds
Delia E. Roberts
Kim R. Rosen

Richard A. Rosengarten
Brian P. Rucci
George A. Rudduck
John P. Ryan
Michael Sansevero
James C. Santo
Gary Frederick Scherer
Michael L. Scruggs
Tina Shaw
Charles Leo Sizer
Donald P. Skrodenis
David C. Snow
Calvin C. Spence
Benoit St-Aubin
Michael William Starke
Scott T. Stelljes
Carol A. Stevenson
Stephen J. Streff
Chester J. Szczepanski
Josephine L.C. Tan
Richard Glenn Taylor
Laura Little Thorne
Laura M. Turner
Frederick A. Urschel
Scott D. Vandermyde
Claude A. Wagner
Lawrence M. Walder
Gregory S. Wanner
Linda F. Ward
Denise R. Webb
Javanika Patel Weltig
Mark Steven Wenger
David L. Whitley
Rosemary Gabriel Wickham
Apryle Oswald Williams
Jennifer N. Williams
Jerelyn S. Williams
Kendall P. Williams
Robin Davis Williams
Oliver T. Wilson
Dean M. Winters
Brandon L. Wolf
Robert F. Wolf
Jeffrey S. Wood

REPORT OF THE VICE PRESIDENT–ADMINISTRATION

This report provides a summary of CAS activities since the 1998 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 applied to property, casualty, and similar risk exposures;
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* and the *Forum* contribute to the attainment of the first purpose. In addition to the *Proceedings*, three volumes of the *Forum* and the Spring Meeting discussion paper program were published and distributed to members in 1999.

The 1998 *Proceedings* was published in two books for the first time with a total of 1138 pages, the greatest number of pages yet for any *Proceedings*. Included in this volume were sixteen papers and five discussions.

The spring 1999 edition of the *Forum* included six reinsurance call papers plus four additional papers.

The summer 1999 edition of the *Forum* included eight dynamic financial analysis discussion papers as well as an additional paper.

The fall 1999 edition of the *Forum* included thirteen reserving call papers plus two additional papers.

A volume titled *Securitization of Risk* included nine papers from the Spring Meeting discussion paper program.

Note that two of the above volumes focussed on topics that are relatively new to the insurance industry: dynamic financial analysis and risk securitization. The CAS has taken a proactive role in stimulating research and educating its members in developing areas.

In regards to the second purpose, the new syllabus for the revised CAS examination process was released. There will continue to be seven exams required for Associateship, but Fellowship will require nine exams rather than the current ten. The first four exams will be jointly administered with the Society of Actuaries (SOA). The new structure will be effective in the year 2000.

A new class of CAS membership was created in 1998: Affiliate. Affiliate members can 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. In 1999, nine Affiliate members were admitted.

CAS membership continues to grow with 217 new Associates and 137 new Fellows in the last year. The total membership now stands at 3,283. A total of 6,511 candidates registered for 1999 CAS exams.

The CAS Task Force on Mutual Recognition examined whether the CAS should enter into bilateral agreements with other actuarial organizations to grant reciprocal Fellowship status. The task force's report pointed out that the American Academy of Actuaries has a process to allow qualified actuaries to practice in the U.S., that the CAS now offers Affiliate membership, and that some CAS examination waivers are available to actuaries of other exam-giving organizations. The Board resolved not to enter into agreements granting reciprocal Fellowship status.

The third purpose 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	Orlando, FL	787
Annual	San Francisco, CA	727

Seminars:

Topic	*Location*	*Month*	*Registrants*
Ratemaking	Nashville	March	508
Financial Risk Management	Denver	April	157
Reinsurance	Baltimore	June	244
Dynamic Financial Analysis	Chicago	July	209
Casualty Loss Reserves	Scottsdale	September	524
CIA/CAS Appointed Actuary	Montréal	September	300
Health and Managed Care	Hilton Head	October	82
Course on Professionalism	Six locations		217

Limited Attendance Seminars:

Topic	*Location*	*Month*	*Registrants*
Advanced Dynamic Financial Analysis	Boston, MA	July	41
Dynamic Financial Analysis (2)	New York, NY; San Francisco, CA	May; October	32, 35
Managing Asset and Investment Risk	Chicago, IL	April	21
Principles of Finance	Boston, MA	June	23
Practical Applications of Loss Distributions (2)	Washington DC; Los Angeles, CA	January; July	40, 27
Reinsurance	New York, NY	August	63

A new CAS Regional Affiliate, Casualty Actuaries of the Desert States, was recognized. 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 updated. Authors submitted their first drafts for revised chapters, which are being reviewed by the Textbook Rewriting Committee.

To increase the awareness of actuarial science, the fourth purpose, the CAS, jointly with the SOA, sponsors Actuarial Career Information Fairs and other activities. In order to attract more minority students to actuarial science, the Joint CAS/SOA Committee on Minority Recruiting awarded 35 $1,000 scholarships to minority students.

The CAS Web Site, now in its fourth year of existence, supports all four purposes. Following are some highlights from the past year:

1. The home page was redesigned. It loads more quickly, includes more menu items and is scroll free.
2. The Web site search engine was upgraded.
3. Thirty past volumes of the *Proceedings* now can be downloaded from the site.
4. Members are now able to respond to the Participation Survey, Research Survey, and Survey on Nontraditional Practice Areas online.
5. A new section for academics was created.
6. A total of 151 job openings were posted for a fee over the last year in our advertising section, helping to defray the cost of maintaining the Web site.

Also, electronic distribution via e-mail of CAS announcements was initiated in 1999 with 70% of the members participating.

Constitution changes pertaining to the officers of the Society and the composition and duties of the Executive Council were approved by the Fellows on July 31, 1999. Subsequently, the Board of Directors approved the addition of a sixth vice president and elected LeRoy A. Boison to serve in the new position of Vice President–International. The Executive Council then approved three new committees under the Vice President–International: International Oversight, IAA Liaison, and International Issues. These structural changes recognize the need for additional CAS efforts in international activities.

The Research Policy and Management Committee reviewed and evaluated the CAS's research process and its effectiveness. Their Review of CAS Research report was presented to the Board in September. This report concluded that the CAS currently has a significant amount of casualty actuarial research. The challenge is to find ways to make that research more accessible to the members and to expand the research efforts beyond those conducted on a voluntary basis. The report included the results of a 1999 membership survey on CAS research, and made recommendations to increase the value of research to practicing actuaries. These recommendations will be incorporated into the 1999–2000 goals of the Vice President–Research and Development.

The report on the results of the 1998 CAS Membership Survey (conducted every five years) also was presented to the Board in September. A copy of the report was posted on the CAS Web Site. The Executive Council will use the feedback in planning goals for 1999–2000 and after.

The Task Force on Nontraditional Practice Areas presented its report to the Board in November. The task force made recommendations on how the CAS can better serve its members practicing in nontraditional areas, and provide additional opportunities for members interested in working in these areas. Nontraditional areas identified as priorities were asset/liability management and investment policy, valuation of property/casualty insurance companies, enterprise risk management, and securitization/risk financing. It was also recommended that instruction on general business skills be included in the CAS continuing education program. The Board approved recommendations for new CAS initiatives in research and education in nontraditional areas.

The CEO Advisory Task Force also reported its findings to the Board in November. Fourteen property and casualty insurance industry leaders were interviewed to determine how well actuaries are meeting the needs of their organizations. The leaders discussed the skills and talents needed to meet current and future business challenges. The Long Range Planning Committee is reviewing the report and will recommend actions to the Board.

Joint activities with the SOA continue. The CAS is participating on the Joint CAS/CIA/SOA Task Force on Academic Ties, and their report will be distributed to the membership for review and comment. A joint CAS/SOA Board meeting was held on September 16, 1999 for getting to know each other, sharing ideas and discussing topics of common interest.

New members elected to the Board of Directors for next year include Amy S. Bouska, Stephen P. D'Arcy, Frederick O. Kist, and Susan E. Witcraft. The membership elected Patrick J. Grannan to the position of President–Elect, while Alice H. Gannon 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, Mary Frances Miller

Vice President–Continuing Education, Abbe S. Bensimon

Vice President–International, LeRoy A. Boison

Vice President–Programs and Communications, David R. Chernick

Vice President–Research and Development, Gary R. Josephson

The CPA firm of Langan Associates was engaged to examine the CAS books for fiscal year 1999 and its findings will be reported by the Audit Committee to the Board of Directors in February 2000. The fiscal year ended with unaudited net income from operations of $338,255 compared to a budgeted loss of $7,035. This higher than expected net income was primarily the result of exam income from higher than expected exam enrollments in anticipation of the syllabus changes taking effect in the year 2000.

Members' equity now stands at $3,074,859. This represents an increase in equity of $161,898 over the amount reported last year. With rising interest rates in 1999, there was an unrealized loss of $157,000 to adjust the CAS's marketable fixed income investments to market value, which dampened the increase in members' equity.

For 1999–2000, the Board of Directors has approved a budget of approximately $4.3 million, an increase of $400,000 over the prior fiscal year. Members' dues for next year will be $290, an increase of $10, while fees for the Subscriber Program will increase by $10 to $360. A $20 discount is available to members and subscribers who elect to receive the *Forums* and *Discussion Paper Program* in electronic format from the Web site.

Respectfully submitted,
Curtis Gary Dean
Vice President–Administration

FINANCIAL REPORT
FISCAL YEAR ENDED 9/30/99

OPERATING RESULTS BY FUNCTION

FUNCTION	*INCOME*	*EXPENSE*	*DIFFERENCE*
Membership Services	$ 1,148,017	$ 1,349,928 (a)	$ (201,911)
Seminars	1,029,307	897,107	132,200
Meetings	581,529	543,300	38,229
Exams	2,615,075 (b)	2,433,229 (b)	181,846
Publications	42,762	25,844	16,918
TOTAL	$ 5,416,689	$ 5,249,408	$ 167,282

NOTES: (a) Includes loss of $170,973 to adjust marketable securities to market value (SFAS 124).
(b) Includes $1,475,850 of Volunteer Services for income and expense (SFAS 116).

BALANCE SHEET

ASSETS	*9/30/98*	*9/30/99*	*DIFFERENCE*
Checking Accounts	$ 149,088	$ 134,490	$ (14,598)
T-Bills/Notes	3,436,980	3,537,154	100,174
Accrued Interest	49,902	51,708	1,806
Prepaid Expenses	74,072	72,451	(1,621)
Prepaid Insurance	11,184	16,871	5,687
Accounts Receivable	39,461	11,255	(28,206)
Textbook Inventory	12,247	8,174	(4,073)
Computers, Furniture	313,752	386,873	73,121
Less: Accumulated Depreciation	(254,800)	(256,384)	(1,584)
TOTAL ASSETS	$ 3,831,886	$ 3,962,594	$ 130,709

LIABILITIES	*9/30/98*	*9/30/99*	*DIFFERENCE*
Exam Fees Deferred	$ 388,425	$ 500,444	$ 112,019
Annual Meeting Fees Deferred	42,246	29,355	(12,891)
Seminar Fees Deferred	61,440	27,441	(33,999)
Accounts Payable and Accrued Expenses	372,716	263,779	(108,937)
Deferred Rent	15,384	9,018	(6,366)
Unredeemed Vouchers	0	19,800	19,800
Accrued Pension	38,714	37,896	(818)
TOTAL LIABILITIES	$ 918,925	$ 887,735	$ (31,190)

MEMBERS' EQUITY			
Unrestricted	*9/30/98*	*9/30/99*	*DIFFERENCE*
CAS Surplus	$ 2,560,111	$ 2,727,393	$ 167,282
Michelbacher Fund	102,249	105,861	3,612
Dorweiler Fund	2,771	1,911	(860)
CAS Trust	19,765	36,616	16,851
Research Fund	166,207	133,207	(33,000)
ASTIN Fund	43,353	52,046	8,693
Subtotal Unrestricted	$ 2,894,456	$ 3,057,034	$ 162,578
Temporarily Restricted			
Scholarship Fund	$ 6,895	$ 6,738	$ (157)
Rodermund Fund	11,611	11,087	(524)
Subtotal Restricted	18,506	17,825	(681)
TOTAL EQUITY	$ 2,912,962	$ 3,074,859	$ 161,898

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: Paul Braithwaite, Chairperson;
Charles A. Bryan, Anthony J. Grippa, and Richard W. Lo

1999 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 3, 4, 5, 6, and 7, 1999. Examinations for Parts 3B, 4A, 4B, 5A, 5B, 7-United States, 7-Canada, and 9 of the Casualty Actuarial Society were held on November 1, 2, 3, and 4, 1999.

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 1999, and Parts 3A and 3C were given in May and November of 1999. 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 1999 Part 1 CAS Examination, the $200 first prize winner was Jin Li, Wesleyan University. The $100 second prize winners were Karyn Beth Baker, Indiana University; Kevin Neal Bills, Texas A&M; Choongtze Chua, University of Pennsylvania; and Genevieve Couture, University of Laval.

For the Spring 1999 Part 1 CAS Examination, the $200 first prize winner was Eugene Chislenko, Stuyvesant High School. The $100 second prize winners were Tianyang Wang, Nankai University; Qiyu Luo, Peking University; Wei Dong Wang, Peking University; Jianhua Gan, University of Science and Technology of China; and Yasong Yang, Fudan University.

For the Fall 1999 Part 1 CAS Examination, the $200 first prize winners were Zheng Wang, Peking University; and Dan Yue, Renmin University. The $100 second prize winners were Hui Zeng, Peking University; Meng Du, University of Science and Technology of China; and Jiayu Mei, Renmin University.

The following candidates were admitted as Fellows and Associates at the 1999 CAS Spring Meeting in May. By passing Fall 1998 CAS examinations, these candidates successfully fulfilled the Society requirements for Fellowship or Associateship designations.

NEW FELLOWS

Mustafa Bin Ahmad
Betsy A. Branagan
Elliot Ross Burn
Brian Harris Deephouse
Alana C. Farrell
Bruce Daniel Fell
Claudine Helene Kazanecki
Deborah M. King
Eleni Kourou
Dawn M. Lawson
Richard Borge Lord
Michael Shane
Christopher C. Swetonic

NEW ASSOCIATES

Jason R. Abrams
Michael Bryan Adams
Anthony L. Alfieri
Silvia J. Alvarez
Gwendolyn Anderson
Paul D. Anderson
Amy Petea Angell
Anju Arora
Nathalie J. Auger
Amy Lynn Baranek
Patrick Beaudoin
David James Belany
Kristen Maria Bessette
John T. Binder
Mario Binetti
Christopher David Bohn
Mark E. Bohrer
David R. Border
Thomas S. Botsko
Stephane Brisson
Karen Ann Brostrom
Conni Jean Brown
Paul Edward Budde
Julie Burdick
Derek D. Burkhalter
Anthony Robert Bustillo
Allison F. Carp
Daniel George Charbonneau
Nathalie Charbonneau
Todd Douglas Cheema
Yvonne W. Y. Cheng
Julia Feng-Ming Chu
Jeffrey Alan Clements
Jeffrey J. Clinch
Eric John Clymer
Carolyn J. Coe
Steven A. Cohen
Larry Kevin Conlee
Peter J. Cooper
Sean Oswald C. Cooper
Sharon R. Corrigan
David Ernest Corsi
Jose R. Couret
John Edward Daniel
Mujtaba H. Datoo
Catherine L. DePolo
Jean A. DeSantis
Timothy Michael DiLellio
Sophie Duval
James Robert Elicker
Gregory James Engl
Brian Michael Fernandes
Kenneth D. Fikes
Janine Anne Finan

Sean Paul Forbes
Ronnie Samuel Fowler
Mark R. Frank
Serge Gagné
James M. Gallagher
Anne M. Garside
Justin Gordon Gensler
Emily C. Gilde
Theresa Giunta
Todd Bennett Glassman
Paul E. Green Jr.
Joseph Paul Greenwood
Michael S. Harrington
Bryan Hartigan
Jeffery Tim Hay
Qing He
Amy Louise Hicks
Jay T. Hieb
Glenn R. Hiltpold
Glenn Steven Hochler
Brook A. Hoffman
Todd Harrison Hoivik
Terrie Lynn Howard
Paul Jerome Johnson
Bryon Robert Jones
Burt D. Jones
Derek A. Jones
Ung Min Kim
Thomas F. Krause
Isabelle La Palme
Travis J. Lappe
Borwen Lee
Christian Lemay
Brendan Michael Leonard
Karen N. Levine
Sally Margaret Levy
Sharon Xiaoyin Li
Dengxing Lin
James P. Lynch
Kelly A. Lysaght
Kevin M. Madigan
Vahan A. Mahdasian
Atul Malhotra
Albert Maroun
Jason Aaron Martin
Laura Smith McAnena
Timothy L. McCarthy
Rasa Varanka McKean
Sarah Kathryn McNair-Grove
Kirk Francis Menanson
Ain Milner
Michael W. Morro
John-Giang L. Nguyen
Michael Douglas Nielsen
Randall William Oja
Sheri L. Oleshko
Leo Martin Orth Jr.
Gerard J. Palisi
Prabha Pattabiraman
Michael A. Pauletti
Fanny C. Paz-Prizant
Rosemary Catherine Peck
John Michael Pergrossi
Sylvain Perrier
Christopher Kent Perry
Anthony J. Pipia
Jordan J. Pitz
Thomas LeRoy Poklen Jr.
William Dwayne Rader Jr.
Sara Reinmann
Sylvain Renaud
Mario Richard
David C. Riek
Kathleen Frances Robinson
Joseph Francis Rosta Jr.
Janelle Pamela Rotondi
Robert Allan Rowe
Joseph John Sacala
James C. Santo
Frances Ginette Sarrel
Jason Thomas Sash
Jeremy Nelson Scharnick
Jeffery Wayne Scholl
Annmarie Schuster
Peter Abraham Scourtis
David Garrett Shafer
Vladimir Shander
Seth Shenghit
Mark Richard Strona
Jayme P. Stubitz
Stephen James Talley
Jo Dee Thiel-Westbrook
Robert M. Thomas II

Jennifer L. Throm
Gary Steven Traicoff
Andrea Elisabeth Trimble
Brian K. Turner
Jon S. Walters
Douglas M. Warner
David W. Warren
Kevin Earl Weathers
Trevar K. Withers
Meredith Martin Woodcock
Jonathan Stanger Woodruff
Perry Keith Wooley
Yin Zhang
Steven Bradley Zielke

The following candidates successfully completed the following Parts of the Spring 1999 CAS Examinations that were held in May.

Part 3B

Patrick Barbeau
Roger N. Batdorff
Marie-Eve J. Belanger
Jeremy James Brigham
Michael C. Carini
Peggy Chan
Wil Chong
Wai Yip Chow
Benjamin W. Clark
Michael Fong
Richard A. Fuller
Rainer Germann
Guo Harrison
Hans Heldner
Mark D. Heyne
Richard S. Holland
George Joseph
Kathleen L. Koshy
Ravi Kumar
Ting Kwok
Pak-Chuen Li
Ian John McCracken
Edward M. Moore
Michael R. Petrarca
Sean E. Porreca
Stephen D. Riihimaki
Brett A. Roush
Joseph Allen Smalley
Jeffrey S. Wood

Part 4A

Vera E. Afanassieva
Genevieve L. Allen
Stevan S. Baloski
Dan S. Barnett
Alex G. Bedoway
Toby Layne Bennington
Sheila J. Bertelsen
William J. Blatcher
Eli B. Bowman
Jeffrey A. Brueggeman
Randall T. Buda
Douglas J. Busta
Cemal Alp Can
Brian J. Cefola
Sanjeev Chaudhuri
Scott A. Chaussee
Julia Chou
Martin P. Chouinard
Christopher J. Cleveland
Matthew P. Collins
Andrea D. Combs
Costas A. Constantinou
Cameron A. Cook
Sean T. Corbett
John E. Costango
Tighe C. Crovetti
Laura M. Dembiec
Mark R. Desrochers
Christopher P. DiMartino
Pamela G. Doonan
Charles W. Dorman
Dale A. Fethke
William M. Finn

Jill A. Frackenpohl
Louise Frankland
Andre Gagnon
Carol Ann Garney
Alexander R. George
William J. Gerhardt
Christie L. Gilbert
Isabelle Girard
Simon Girard
Jason L. Grove
Eric A. Hatch
Kimberly A. Haza
Arie Haziza
Michael J. Hebenstreit
Brandon L. Heutmaker
Marcy R. Hirner
Kathleen Hobbs
Allen J. Hope
Wendy L. Hopfensperger
Sheng-Fei Huang
Kuo Ming Hung
Christopher W. Hurst
Nathan L. Jones
Julie A. Jordan
Jesse A. Karls
Susan M. Keaveny
Elissa Y. Kim
Jason M. Kingston
Brandon E. Kubitz
Kristine Kuzora
Nathaniel Kwawukume
Jeff A. Lamy
Aaron Michael Larson
Stefan A. Lecher
Michaela Ledlova
Wendy R. Leferson
Ho Shan A. Leung
Julia Leung
Amanda M. Levinson
Carrie L. Lewis
Hayden Anthony Lewis
Jennifer L. Ligon
Lucia A. Lloyd-Kolkin
Winnie Lo
Siew-Won Loh
Daniel A. Lowen
Xiaofeng Lu
Abbe M. Macdonald
Teresa Madariaga
Chaim Markowitz
Susan E. Marra
Michelle C. Martin
Raul G. Martin
Carolyn J. McElroy
Sylwia S. McMichael
Sylvie Menard
Kathleen M. Miller
Richard J. Mills
Kazuko Minagawa
Erica F. Morrone
Joseph J. Muccio
James C. Murphy
Daniel G. Myers
Scott L. Negus
Winnie Ning
Mary A. Noga
Billy J. Onion
Russel W. Oslund
John F. Pagano
Felix Patry
Kristin S. Piltzecker
Etienne Plante-Dube
Stephen R. Prevatt
Elisabeth Prince
Lester Pun
Suzanne M. Reddy
John J. Reid
Erica L. Riggs
Sandra E. Rita
Benoit Robert
Robert C. Roddy
Kevin D. Roll
Charles A. Romberger
Adam J. Rosowicz
Jeffrey N. Roth
Ryan P. Royce
Josef W. Rutkowski
Doris Y. Schirmacher
Bradley J. Schroer
Monica S. Schroeter
Frank W. Shermoen
Walter J. Slobojun
Douglas E. Smith
Jodi L. Smith
Christopher Y. So
Kuixi Karl Song
Brooke S. Spencer
Kyrke O. Stephen
Erik J. Steuernagel
Christina H. Sung
Erica W. Szeto
Josephine L. C. Tan
Robert Bradley Tiger
Phoebe A. Tinney
Michael C. Torre
Jean-Francois Tremblay

Matthew D. Trone
Lawrence A. Vann
Nilesh M. Vasani
Melinda K. Vasecka
Chinatsu H. Vergara
Maxim Viel
Matthew W. Walljasper
Jamie M. Weber
Robert S. Weishaar
Carolyn D. Wettstein
Erica H. Wheeler
Stephen C. Williams
Ian G. Winograd
Jimmy L. Wright
Chung-Shiang Wu
Nien-Chien I. Wu
Run Yan
Chih-Cheng Yang
Lisa Shuk-Han Yeung
Jonathan K. Yu

Part 4B

Jeanene M. del Valle
Christopher B. Abreu
Vera E. Afanassieva
Andrea Ondine Ahern
Sayyed Babar Ali
Afrouz Assadian
Kenneth W. Au
Damian T. Bailey
Stephen M. Balden
Igor Balevich
Stevan S. Baloski
Brent A. Banister
Dana Barlow
Stephanie A. Beach
Van R. Beach
Richard D. Behnke
Jerome C. Bellavance
Jesse A. Beohm
Jean-François Bernard
Timothy P. Bert
Stephen Bertolini
Assia Billig
Timothy S. Bischof
Michael D. Blakeney
Luc Blanchet
Roman G. Blichar
Annabelle Blondeau
Randy D. Blum
Nebojsa Bojer
Lisa Bolduc
Mary A. Borrelli-Margraf
Marie-Andrée C. Boucher
Glenn D. Bowen
Russell H. Brands
Erick A. Brandt
Kevin E. Branson
Ward A. Brigham
Gregor L. Brown
Jason C. Buckholt
Andrew E. Buckley
Suejeudi Buehler
Vanessa N. Butala
Heather M. Byrne
Jun Cai
Caryn C. Carmean
Scott A. Carter
Thomas L. Cawley
Ronald S. Cederburg
Rafael Ignacio Cespedes
Ka Lun Chan
Jung-Chiang Chang
Shao-Chien Chang
Yuan-Yuan Chang
Yves Charbonneau
Kin Shuen Iris Chau
Scott A. Chaussee
Ching-Yi Chen
Hung F. Cheung
Janice Cheung
Sharlean Chiu
Jean S. Choi
Kin Lun (Victor) Choi
Hei Mei Chu
On Lee K. Chu
Yuen Wah (Helen) Chu
Delphina S. M. Chue
Anthony F. Colella
Linda Brant Collins
Christopher L. Cooksey
Gerald D. Cooper
Jean-Pierre Cormier
Thomas Cosenza
Huiying Cui
Aaron T. Cushing
Jacek Czajkowski
Robin S. N. Damm
Smita G. Dave

Rich A. Davey
Christopher P. Davies
David A. DeNicola
Erik L. Donahue
Craig A. Doughty
Shane S. Drew
Alexandre Drouin
Olga Dunaveskaya
Sarah M. Duyos
Cecilia A. Earls
Tomer Eilam
John W. Elbl
Tricia G. English
William H. Erdman
Ross C. Eriksson
Brian C. Evanko
Lauren B. Feldman
Matthew B. Feldman
Matthew D. Fienman
Tim P. Finnegan
Theodore M. Fitzpatrick
Jeffrey R. Fleischer
Robin A. Fleming
Ben Flores
Marc A. Fournier
Teresa M. Fox
Geoffrey A. Fradkin
Dana R. Frantz
Jeffrey J. Fratantaro
Rebecca E. Freitag
Mark Kevin Friedman
Craig D. Fyfe
Patrick P. Gallagher
Brett D. Gardner
Tracey Ann Gardner-Lacy
Roland P. Gatti II
Alexander R. George
Alain C. Georget
Alexis Gerbeau
Karen E. Gibbs
Jean-Philippe Giguere
Christie L. Gilbert
Valerie Gingras
Cary W. Ginter
Robert A. Giulietti
Peter Scott Gordon
Michael J. Gossmann
Stephane Goyer
Aleksey V. Granovsky
Jeffrey S. Grant
Timothy S. Grant
Jean-Pierre Gravel
Daniel Groleau
Jason L. Grove
Xu Gu
Patrik R. Guindon
Richard C. Gunning
Elena Hagi
Kevin J. Halfpenny
John I. Hall
Lynette D. Hamberger
Bradley O. Harris
Eric A. Hatch
Stuart J. Hayes
Sean M. Hayward
Joseph Hebert
James D. Heidt
Gregory L. Helser
David A. Henderson
Deborah L. Herman
Nigel P. Hernandez
Joseph S. Highbarger
Ki Wai Ho
Tuong H. Ho
Guillaume Hodouin
David J. Horn Jr.
Patricia L. Horn
Kaylie Horning
Peter R. Horstman
Steven P. Hoxmeier
Alex B. Huang
Wenjun Huang
Carissa A. Hughes
Edward H. L. Hui
Edward M. Huizinga
David C. Hung
Scott R. Hurt
Windy J. Hutchings
Amy R. Jackson
Karen A. Jackson
Frederic Jacques
Hanna K. Jankowski
Jason T. Jarzynka
Chi-Chung Jen
Bret A. Jensen
Lin Jiang
Charles B. Jin
William P. Jirak
Michael S. Johnson
Shantelle A. Johnson
Jason A. Jones
George Joseph
Sarah Kadlecik

Ronald J. Kalvoda
Hye-Sook Kang
John J. Karwath
Inga Kasatkina
Deborah G. Kasper
Stephen F. Katz
Stacey M. Kidd
Eugene T. Kim
Hyuntae Kim
Sung-Hoon Kim
Beth M. Kirk Malecki
Roman Kizner
Ann E. Klaessy
Linda S. Klenk
Laurie A. Knoke
Joseph G. Korabik
Karen A. Kosiba
Randall M. Koss
Tatiana Kozak
Regina Krasnovsky
Rosanne L. Kropp
Jack D. Krull
Adrian Kryszak
Terry T. Kuruvilla
Faye Kurz
Daniel Y. Kutliroff
Hilary S. N. Kwok
Lucie LaChance
Ying Han Lai
Aaron D. Lambright
Eric S. Lanham
Stephane Lapierre
Jacqueline Win Yu Lau
Kan Yuk A. Lau
Sok Hoon Lau
Yue-Che Lau
Michael L. Laufer
Ross A. Laursen
David L. Lautenschlager
Damon T. Lay
Valerie Lebrun
John H. Lee
Sheung Yuen Lee
Stuart Saiwah Lee
Yee Nin Lee
Christopher J. Lemming
James J. Leonard
Brian P. Levine
Jun Li
Monica Yanhong Li
Monika Lietz
Ching-Yi Lin
Sheng-Lun Lin
Yu-Chu Lin
Wai Tat Ling
Jia Liu
Xin Liu
Ying Liu
Lucia A. Lloyd-Kolkin
Nataliya A. Loboda
Michael J. Lockerman
Robert M. Long Jr.
Wan Li Lu
Todd W. Lueders
PeiQing Luo
JoDee L. Lymburner
Mark W. Malnati
Sarah E. Marr
Danny Martin
Matthew J. Martin
Peter G. Matheos
Susan J. McMains
John D. McMichael
Stephen J. McNamara
Sharad Mehra
Michael E. Mielzynski
Wu Chi Ming
Charles W. Mitchell
Ghada M. Samir Mohamed
Jacqueline M. Mohan
David A. Moore
Jeffrey A. Moore
James C. Murphy
Leonard D. Myers
Neil Narale
James P. Naughton
Jacqueline L. Neal
Scott L. Negus
Hon Fai Ng
Eleasar Ngassa
Jacqueline Nam Phuong Nguyen
Karla J. Nieforth
Stoyko N. Nikolov
Christopher F. Noble
John J. Noel
Mary A. Noga
Janet M. Nowatzki
William S. Ober
Michael A. Onofrietti
Bo Ouyang
Masakazu Ozeki
Kristen J. Pack
Jeremy D. Palmer
Staci P. Palmer

Brenda Papillon
James L. Paprocki
Sandra K. Parsons
David James Pauls
Michel W. Pelletier
Chan H. Phan
Jayne L. Plunkett
Vincent Polis
Annie Pui Ying Poon
Marie-Claude Poulin
Corrie L. Proksa
Jingsu Pu
Jianjun Qian
Guillaume Raymond-Turcotte
Mary E. Reading
Thomas V. Reedy
Erin R. Reid
Kevin J. Reimer
Joe Reschini
Jason C. Richards
Craig A. Roberts
Jeremy C. Roberts
Rebecca D. Robertston
Graham E. Rogers
Randall D. Ross
Paul J. Rostand
Kirk A. Roy
Timothy L. Rozar
Michael M. Rubin
Andrei P. Salomatov
Doris Y. Schirmacher
April Sonia Gale Seixeiro
Mandy M. Y. Seto
Gopi B. Shah
Pui Kei Shek
Wei Sheng
Michelle L. Sheppard
Tai-Ming Shiun
Rene R. Simon
Satbir Singh
Raymond D. Sinnappan
Martine Slight
Stuart N. Slutzky
Todd G. Smith
Ka Ying So
Michael I. Sonin
Alexandra R. St-Onge
Tania E. Staffen
Molly A. Stark
Dominic M. Stephenson
Aaron M. Stoeger
Robert P. Stone
Mark S. Struck
Natalia Borisovna Sullivan
Guohong Sun
Douglas B. Swift
Michael E. Symonds
Sergei A. Syskin
Yuk Lun Szeto
Karl Tanguay
Doyle Adrian Tanner
Alex V. Tartakovski
Lucia Tedesco
Alex M. Terry
Helene Thibault
Michael J. Thomas
Noel J. Thomas
Sterling R. Tiessen
Albert Y. Tiw
Randi H. Topp
Frederic Tremblay
Hubert Tremblay
Maryse Tremblay-Lavoie
Chi-Liang Tsai
Yu-Fang Tseng
Kosei Tsukada
Choi Nai Charlies Tu
Christopher R. Tucek
Stephen H. Underhill
William O. Van Arsdale
Samuel S. Van Blarcom
Shannon C. Vecchiarello
Eric T. Veletzos
Paul A. Vendetti
Kevin K. Vesel
Brian A. Viscusi
Natalie Vishnevsky
Hanny C. Wai
Andrew E. Walinsky
Kate L. Walsh
Qingxian Wang
Simon Lijen Wang
Tianshu Wang
Xiuwen Wang
Yi Wang
David W. Watkins
Bethany R. Webb
Robert S. Weishaar
Thomas E. Weist
Ann Welch
Jean P. West

Amanda M. Westphal
Daniel J. White
Andrew T. Wiest
Dennis D. Wiggill
John W. Wiklund
Shawn A. Wilkin
Duane A. Willis
Andrew J. Witte
Molly B. Witzenburg
Ai-Hua Angela Wong
Chi Kit Wong
Po-Shing Wong
Brian T. Woolfolk
Joshua C. Worsham
Eddie J. Wright
Chi-wai Edwin Wu
Susan A. Wudi
Yu Xiang
Feipeng Xie
Qi Xie
Suixiang Xie
Run Yan
Zhi Kang Yan
Su Yang
Christopher H. Yaure
Kim Fung Yeung
Mark R. Yoest
Stephanie C. Young
Ming-Yeh Yu
Pak Kin Yu
Xiaodong Yu
Peng Zeng
Dong Zhang
Song Zhang
Ji Fang Zhou
Yuhan Zhu

Part 5A

Genevieve L. Allen
Penelope A. Bierbaum
Jonathan E. Bransom
Anthony P. Brown
Jonathan Mark Deutsch
Richard James Engelhuber
Yehoshua Y. Engelsohn
Jieqiu Fan
Weishu Fan
Christine M. Fleming
Stuart G. Gelbwasser
Joseph E. Goldman
Stephanie Ann Gould
Ann E. Green
Stacie R. W. Grindstaff
Kristina S. Heer
Carol I. Humphrey
Shantelle A. Johnson
Linda M. Kane
Brant L. Kizer
Susan L. Klein
John E. Kollar
Aleksandr I. Korb
Ruth M. LeStourgeon
Joshua Nathan Mandell
Paul J. Molinari
Joann C. Ribar
Stephen D. Riihimaki
Jennifer L. Rupprecht
Steven M. Schienvar
Robert E. Schmid
Deniz Selman
Brett M. Shereck
Barry Dov Siegman
Pantelis Tomopoulos
Jennifer L. Vadney
Lisa M. VanDermark
Colleen Ohle Walker
Apryle Oswald Williams
Lianmin Zhou

Part 5B

William J. Albertson
Lara L. Anthony
Paul W. Bauer
Marie-Eve J. Belanger
Melissa L. Borell
Maureen B. Brennan
Melissa L. Brewer
Kevin C. Burke
Michael W. Buttke
David R. Cabana
Hao Chai
Gregory R. Chrin
Millie Chu
Christian J. Coleianne
Matthew P. Collins

Avery F. M. Cook
Christopher L. Cooksey
Jonathan M. Corbett
Sean T. Corbett
Lynn E. Cross
Walter C. Dabrowski
Genine Darrough
Krikor Derderian
Tricia G. English
Weishu Fan
Solomon Carlos Feinberg
Kevin M. Finn
Sean W. Fisher
Beth A. Foremsky
Sylvain Fortier
Chad J. Gambone
Sophie M. L. Georget
Brett A. Gissel
Joseph E. Goldman
John P. Gots
Ruth M. Gregory
Stacie R. W. Grindstaff
Margarita Hambrock
Andrew J. Hazel
Esther Y. Hui
Mark C. Jones
Kelly F. Kahling
David G. Keeton
Robin A. Keeven
Eric J. Kendig
Shenaz Keshwani
Perry A. Klingman
Aleksandr I. Korb
Leland S. Kraemer
Terry T. Kuruvilla
David J. Kuzma
Mai B. Lam
Kan Yuk A. Lau
Stuart Saiwah Lee
Sean M. Leonard
Eric M. Lin
Jing Liu
Erik F. Livingston
Nataliya A. Loboda
Daniel A. Lowen
Hazel J. Luckey
Lynn C. Malloney
Jennifer A. McGrath
Michael P. McKenney
Quynh-Nhu T. Morse
Joseph J. Muccio
David B. Mukerjee
Treva A. Myers
Saeid Nazari
Shannon P. Newman
Matthew P. Nimchek
Lauree J. Nuccio
Ginette Pacansky
Lorie A. Pate
Joy-Ann C. Payne
Shing Chi Poon
Sebastien Portmann
David N. Prario
Michael J. Quigley
Conni A. Rader
David P. Rafferty
Monica L. Ransom
Mary S. Rapp
Michelle L. Reckard
Joann C. Ribar
Renata Ringo
Joseph L. Rizzo
John D. Rosilier
Richard H. Seward
Elizabeth A. Sexauer
Sonja M. Shea
Michelle L. Sheppard
Brett M. Shereck
Keith M. Slonski
Lora L. Smith-Sarfo
Benjamin R. Specht
Matthew D. Trone
Melissa K. Trost
Lien K. Tu
Shannon C. Vecchiarello
Kimberly A. Vogel
Monica S. White
Rosemary Gabriel Wickham
Apryle Oswald Williams
Todd M. Wing
Shing-Ming Wong
Regina E. Wood
Shawn A. Wright
Anthony C. Yoder
Janice M. Young
Megan L. Zack

Part 6

Michael D. Adams
Ariff B. Alidina
Robert E. Allen
Brian M. Ancharski
Kris Bagchi
Brian J. Barth
Saeeda Behbahany .
Nathalie Belanger
Jody J. Bembenek
Darryl R. Benjamin
Jonathan P. Berenbom
Brad D. Birtz
Tony Francis Bloemer
Neil M. Bodoff
Joseph V. Bonanno Jr.
Caleb M. Bonds
Donna Bono-Dowd
Olivier Bouchard
John R. Bower
Maureen Ann Boyle
Maureen B. Brennan
John R. Broadrick
Sara T. Broadrick
Kristin J. Brown
Bruce D. Browning
Elaine K. Brunner
Lisa K. Buege
Angela D. Burgess
Lori L. Burton
Matthew E. Butler
Sandra J. Callanan
James E. Calton
Mary Ellen Cardascia
Samuel C. Cargnel
William Brent Carr
Tracy L. Child
Andrew H. S. Cho
Alan M. Chow
Philip A. Clancey Jr.
Alan R. Clark
Jason T. Clarke
Kevin M. Cleary
Brian Roscoe Coleman
Richard Jason Cook
Hugo Corbeil
Stephen M. Couzens
Brenda K. Cox
Richard R. Crabb
Keith R. Cummings
Kelly K. Cusick
Robert P. Daniel
Mark A. Davenport
Lori Anne Davey
Willie L. Davis
Nicholas J. De Palma
Peter R. DeMallie
Douglas Lawrence Dee
Paul B. Deemer
Krikor Derderian
Timothy M. Devine
Brian S. Donovan
Scott H. Drab
Jeffrey A. Dvinoff
Donna L. Emmerling
Kyle A. Falconbury
Brian A. Fannin
Kathleen Marie Farrell
Junko K. Ferguson
Kristine M. Fitzgerald
Jennifer L. Fitzpatrick
Sharon L. Fochi
Feifei Ford
David Gagnon
Michelle R. Garnock
Genevieve Garon
Dustin W. Gary
Matthew P. Gatsch
Robert W. Geist
Laszlo J. Gere
Gregory Evan Gilbert
Isabelle Gingras
Andrew S. Golfin Jr.
Melanie T. Goodman
Lori A. Gordon
Matthew R. Gorrell
Christopher J. Grasso
Diane E. Grieshop
Donald B. Grimm
Robert Andrew Grocock
Isabelle Groleau
Lisa N. Guglietti
Chantal Guillemette
James C. Guszcza
Serhat Guven
Edward Kofi Gyampo
David B. Hackworth
Barbara Hallock
Marcus R. Hamacher
Faisal O. Hamid
David Lee Handschke
Jason C. Head

Hans Heldner
Mark D. Heyne
David E. Hodges
Suzanne B. Holohan
Margaret M. Hook
Francis J. Houghton Jr.
Derek R. Hoyme
Long-Fong Hsu
Jamison J. Ihrke
Joseph M. Izzo
Jesse T. Jacobs
David R. James
William T. Jarman
Scott R. Jean
Philip J. Jennings
Karen Lerner Jiron
Brian B. Johnson
Erik A. Johnson
Tricia L. Johnson
William B. Johnson
Gregory K. Jones
Theodore A. Jones
Kyewook (Gary) Kang
Barbara L. Kanigowski
Alexander Kastan
Kathryn E. Keehn
Sean M. Kennedy
David R. Kennerud
Susanlisa Kessler
Joseph E. Kirsits
Henry J. Konstanty
James J. Konstanty
Brandon E. Kubitz
Todd J. Kuhl
Darjen D. Kuo
Julie-Linda Laforce
Stephane Lalancette
John B. Landkamer
Frank A. Laterza
Rocky S. Latronica
Anh Tu Le
Jeffrey Leeds
Geraldine Marie Z. Lejano
Twiggy Lemercier
Hayden Anthony Lewis
Shangjing Li
Matthew A. Lillegard
Kenneth Lin
Kathleen T. Logue
William F. Loyd
Yih-Jiuan B. Lu
Eric A. Madia
Alexander P. Maizys
David K. Manski
Timothy J. McCarthy
Kevin Paul McClanahan
John R. McCollough
Jeffrey B. McDonald
Richard J. McElligott
Patrick A. McGoldrick
Shawn Allan McKenzie
Jeffrey S McSweeney
Lawrence J. McTaggart III
Christian Menard
Martin Menard
Ellen E. Mercer
Vadim Y. Mezhebovsky
Paul B. Miles
Kathleen C. Miller
Stephanie A. Miller
Suzanne A. Mills
Jason E. Mitich
Josée Morin
Matthew E. Morin
Rodney S. Morris
Timothy C. Mosler
Gwendolyn D. Moyer
Carole Nader
Jennifer Y. Nei
Brian C. Neitzel
Ronald T. Nelson
Susan K. Nichols
Jill A. Nielsen
James L. Norris
Miodrag Novakovic
Nancy Eugenia O'Dell-Warren
Wade H. Oshiro
Robin V. Padwa
Kelly A. Paluzzi
Phillip J. Panther
Jean-Pierre Paquet
Carolyn Pasquino
Bruce G. Pendergast
Priyantha L. Perera
Matthew J. Perkins
Isabelle Perron
Christopher A. Pett
Faith M. Pipitone
Peter Victor Polanskyj
Gregory T. Preble
Bill D. Premdas

Marie-Josee Racine
John T. Raeihle
Kathleen M. Rahilly-VanBuren
Josephine Teruel Richardson
Marn Rivelle
Ezra J. Robison
Keith A. Rogers
Benjamin G. Rosenblum
Christina B. Rosenzweig
David A. Royce
John C. Ruth
Charles J. Ryherd
Laura B. Sachs
Parr T. Schoolman
Michael F. Schrah
Larry J. Seymour
Tina Shaw
James S. Shoenfelt
Marina Sieh
Annemarie Sinclair
Helen A. Sirois
Lee O. Smith
Steven A. Smith II
Thomas M. Smith
Lisa C. Stanley
Michael William Starke
Amy L. Steburg
David K. Steinhilber
Stephen J. Streff
Mark Sturm
Beth M. Sweeney
Neeza Thandi
Christopher S. Throckmorton
Tamara L. Trawick
Joseph S. Tripodi
Bonnie J. Trueman
Peggy J. Urness
Michael O. Van Dusen
William D. Van Dyke
Susan B. Van Horn
Karen L. VanCleave
Gaetan R. Veilleux
Jennifer A. Vezza
Josephine M. Waldman
Ya-Feng Wang
Chang-Hsien Wei
Javanika Patel Weltig
Joseph C. Wenc
Gary A. Wick
William B. Wilder
Dean M. Winters
Jennifer X. Wu
Mihoko Yamazoe
Mark K. Yasuda
Jacinthe Yelle
Michael G. Young
Stephanie C. Young
Christine Seung Yu
Michael R. Zarember
Xiangfei Zeng
Gene Q. Zhang
Yingjie Zhang
Eric E. Zlochevsky

Part 8-Canada

Suzanne E. Black
Veronique Bouchard
Robert N. Campbell
Jean-François Chalifoux
Louise Chung-Chum-Lam
Steven A. Cohen
Louis Durocher
Hugo Fortin
Philip W. Jeffery
David Leblanc-Simard
P. Claude Lefebvre
Eric Millaire-Morin
François L. Morissette
Charles Pare
Ernest C. Segal

Part 8-United States

Ethan D. Allen
Katherine H. Antonello
Michele S. Arndt
Carl Xavier Ashenbrenner
William P. Ayres
Michael William Barlow
Andrew S. Becker

Ellen A. Berning
Daniel R. Boerboom
Raju Bohra
Sherri Lynn Border
Thomas L. Boyer II
David C. Brueckman
Michelle L. Busch
Victoria J. Carter
Patrick J. Charles
Richard M. Chiarini
Thomas Joseph Chisholm
Wanchin W. Chou
Gary T. Ciardiello
Larry Kevin Conlee
Karen Barrett Daley
Timothy Andrew Davis
Kurt S. Dickmann
Christopher S. Downey
Sara P. Drexler
Stephen C. Dugan
Mark Kelly Edmunds
Laura A. Esboldt
Kristine Marie Firminhac
Tracy Marie Fleck
Michelle L. Freitag
Kevin Jon Fried
Gary J. Ganci
Amy L. Gebauer
Christopher H. Geering
Bernard H. Gilden
Todd Bennett Glassman
Sanjay Godhwani
Francis X. Gribbon
Jacqueline Lewis Gronski
John A. Hagglund
Marc S. Hall
Dawn Marie S. Happ
Michelle Lynne Harnick
Bryan Hartigan
Michael B. Hawley
Jeffery Tim Hay
Qing He
Chad Alan Henemyer
Amy Louise Hicks
Richard M. Holtz
Brian L. Ingle
Craig D. Isaacs
Randall A. Jacobson
Charles B. Jin
Mark J. Kaufman
Scott A. Kelly
Ung Min Kim
Elina L. Koganski
Richard Scott Krivo
Sarah Krutov
Robin M. LaPrete
Travis J. Lappe
Dennis H. Lawton
Ramona C. Lee
James P. Leise
Bradley H. Lemons
Brendan Michael Leonard
Charles Letourneau
John Norman Levy
Sally Margaret Levy
Siu K. Li
Richard P. Lonardo
Jason Aaron Martin
Michael Boyd Masters
David M. Maurer
Douglas W. McKenzie
Sarah K. McNair-Grove
Ain Milner
David Patrick Moore
Lisa J. Moorey
Matthew C. Mosher
Ethan Charles Mowry
Seth Wayne Myers
Michael D. Neubauer
Corine Nutting
James L. Nutting
Steven Brian Oakley
Randall William Oja
Richard D. Olsen
Christopher Edward Olson
David Anthony Ostrowski
Moshe C. Pascher
Lisa Michelle Pawlowski
John R. Pedrick
John M. Pergrossi
Christopher Kent Perry
Daniel B. Perry
Luba O. Pesis
Sean E. Porreca

Christopher David Randall
Sara Reinmann
Scott Reynolds
John W. Rollins
Richard A. Rosengarten
Robert Allan Rowe
David L. Ruhm
Douglas A. Rupp
Joanne E. Russell
James C. Santo
Nathan Alexander Schwartz
William Harold Scully III
Bret Charles Shroyer
Matthew Robert Sondag
Laurence H. Stauffer
Curt A. Stewart
Lisa M. Sukow
Elizabeth Susan Tankersley
Varsha A. Tantri
Michael J. Tempesta
Robert M. Thomas II
Beth S. Thompson
Gary S. Traicoff
Jeffrey S. Trichon
Kai Lee Tse
Therese M. Vaughan
Leslie Alan Vernon
Cameron J. Vogt
Kyle Jay Vrieze
Wade T. Warriner
Dean Allen Westpfahl
William B. Westrate
Kendall P. Williams
Laura Markham Williams
Kah-Leng Wong
Simon Kai-Yip Wong
Jonathan Stanger Woodruff
Vincent F. Yezzi

Part 10

Rimma Abian
Stephen A. Alexander
Mark B. Anderson
Amy Petea Angell
Martin S. Arnold
Peter Attanasio
Richard J. Babel Sr.
Emmanuil Theodore Bardis
Patrick Beaudoin
Nicolas Beaupre
Cynthia A. Bentley
David M. Biewer
Frank J. Bilotti
Lisa A. Bjorkman
Jonathan Everett Blake
Michael J. Bluzer
Mark E. Bohrer
Ann M. Bok
Tobe E. Bradley
Michael D. Brannon
Stephane Brisson
Hayden Heschel Burrus
Anthony E. Cappelletti
Martin Carrier
Sharon C. Carroll
Bethany L. Cass
Joseph G. Cerreta
Michael Joseph Christian
Bryan C. Christman
Andrew K. Chu
Kuei-Hsia Ruth Chu
Darrel W. Chvoy
Christopher William Cooney
Jeffrey Alan Courchene
Brian K. Cox
Claudia Barry Cunniff
Robert E. Davis
Kris D. DeFrain
Jean A. DeSantis
Michael Edward Doyle
Peter F. Drogan
Denis Dubois
Mary Ann Duchna-Savrin
Rachel Dutil
Sophie Duval
Jane Eichmann
Dawn E. Elzinga
Brandon L. Emlen
Gregory James Engl

Kenneth D. Fikes
Jean-Pierre Gagnon
Donald M. Gambardella
Thomas P. Gibbons
John T. Gleba
Matthew E. Golec
Karl Goring
Philippe Gosselin
Jay C. Gotelaere
David Thomas Groff
Rebecca N. Hai
Scott T. Hallworth
Kenneth Jay Hammell
Gregory Hansen
Jonathan B. Hayes
Jodi J. Healy
Noel M. Hehr
Christopher Ross Heim
David E. Heppen
Ronald J. Herrig
Kurt D. Hines
Amy L. Hoffman
Thomas A. Huberty
Ali Ishaq
Philippe Jodin
Burt D. Jones
James B. Kahn
Chad C. Karls
Mark J. Kaufman
James M. Kelly
James D. Kunce
Jean-Sebastien Lagarde
Elaine Lajeunesse
Chingyee Teresa Lam
Yin Lawn
Kevin A. Lee
Dengxing Lin
Shu C. Lin
Janet G. Lindstrom
Diana M. S. Linehan
Lee C. Lloyd
Michelle Luneau
William R. Maag
Atul Malhotra
David E. Marra
Julie Martineau
Bonnie C. Maxie
Jeffrey F. McCarty
Ian John McCracken
Allison Michelle McManus
James R. Merz
Scott A. Miller
Paul W. Mills
Christopher J. Monsour
Roosevelt C. Mosley
Kari S. Mrazek
Donna M. Nadeau
Catherine A. Neufeld
Hiep T. Nguyen
Kari A. Nicholson
Lynn Nielsen
Randall S. Nordquist
Michael A. Nori
Richard A. Olsen
Denise R. Olson
Teresa K. Paffenback
Ajay Pahwa
M. Charles Parsons
Mark Paykin
Julie Perron
Jeffrey J. Pfluger
Anthony George Phillips
Igor Pogrebinsky
Karen L. Queen
Kathleen Mary Quinn
Leonid Rasin
Yves Raymond
Hany Rifai
Seth Andrew Ruff
Tracy A. Ryan
Rajesh V. Sahasrabuddhe
Asif M. Sardar
Gary Frederick Scherer
Michael C. Schmitz
Annmarie Schuster
Stuart A. Schweidel
Meyer Shields
Jay Matthew South
Angela Kaye Sparks
Avivya Simon Stohl
Brian Tohru Suzuki
Adam M. Swartz
Nitin Talwalkar
Jonathan Garrett Taylor
Dom M. Tobey
Jennifer M. Tornquist
Michael J. Toth
Michael C. Tranfaglia
Turgay F. Turnacioglu
Kieh Treavor Ty
Mark A. Verheyen

Martin Vezina
Nathan K. Voorhis
Claude A. Wagner
Edward (Ted) H. Wagner
Benjamin A. Walden
Robert J. Wallace
Shaun S. Wang Ph.D.
Patricia Cheryl White
Jerelyn S. Williams
Wendy L. Witmer
Brandon L. Wolf
Mark L. Woods
Mary C. Woodson
Jeanne Lee Ying
Sheng H. Yu

The following candidates were admitted as Fellows and Associates at the 1999 CAS Annual Meeting in November. By passing Spring 1999 CAS examinations, these candidates successfully fulfilled the Society requirements for Fellowship or Associateship designations.

NEW FELLOWS

Rimma Abian
Ethan David Allen
Mark B. Anderson
Martin S. Arnold
William P. Ayres
Richard J. Babel
Cynthia A. Bentley
Lisa A. Bjorkman
Suzanne E. Black
Jonathan Everett Blake
Ann M. Bok
Michael D. Brannon
Anthony E. Cappelletti
Martin Carrier
Bethany L. Cass
Jean-François Chalifoux
Bryan C. Christman
Darrel W. Chvoy
Gary T. Ciardiello
Christopher William Cooney
Brian K. Cox
Claudia Barry Cunniff
Karen Barrett Daley
Timothy Andrew Davis
Jean A. DeSantis
Kurt S. Dickmann
Christopher S. Downey
Michael Edward Doyle
Peter F. Drogan
Denis Dubois
Mary Ann Duchna-Savrin
Rachel Dutil
Dawn E. Elzinga
Jean-Pierre Gagnon
Donald M. Gambardella
Gary J. Ganci
Thomas P. Gibbons
John T. Gleba
Matthew E. Golec
Philippe Gosselin
Jay C. Gotelaere
David Thomas Groff
Scott T. Hallworth
Gregory Hansen
Michael B. Hawley
Jodi J. Healy
Noel M. Hehr
Christopher Ross Heim
David E. Heppen
Ronald J. Herrig
Thomas A. Huberty
Brian L. Ingle
James B. Kahn
Chad C. Karls
Mark J. Kaufman
James M. Kelly
Sarah Krutov
James D. Kunce
Jean-Sebastien Lagarde
Yin Lawn
David Leblanc-Simard

Kevin A. Lee
P. Claude Lefebvre
Siu K. Li
Janet G. Lindstrom
Lee C. Lloyd
William R. Maag
David E. Marra
Michael Boyd Masters
Bonnie C. Maxie
Jeffrey F. McCarty
Douglas W. McKenzie
Allison Michelle McManus
James R. Merz
Paul W. Mills
Christopher J. Monsour
David Patrick Moore
François L. Morissette
Matthew C. Mosher
Roosevelt C. Mosley
Donna M. Nadeau
Catherine A. Neufeld
Hiep T. Nguyen
Randall S. Nordquist
Michael A. Nori
James L. Nutting
Christopher Edward Olson
Denise R. Olson
David Anthony Ostrowski
Teresa K. Paffenback
Charles Pare
M. Charles Parsons
Luba O. Pesis
Karen L. Queen
Kathleen Mary Quinn
Yves Raymond
Hany Rifai
John W. Rollins
Seth Andrew Ruff
David L. Ruhm
Tracy A. Ryan
Rajesh V. Sahasrabuddhe
Michael C. Schmitz
Nathan Alexander Schwartz
Bret Charles Shroyer
Matthew Robert Sondag
Jay Matthew South
Angela Kaye Sparks
Brian Tohru Suzuki
Adam M. Swartz
Nitin Talwalkar
Dom M. Tobey
Jeffrey S. Trichon
Kai Lee Tse
Leslie Alan Vernon
Kyle Jay Vrieze
Edward H. Wagner
Benjamin A. Walden
Robert J. Wallace
Patricia Cheryl White
Wendy L. Witmer
Simon Kai-Yip Wong
Vincent F. Yezzi
Sheng H. Yu

NEW ASSOCIATES

Michael D. Adams
Genevieve L. Allen
Saeeda Behbahany
Penelope A. Bierbaum
Tony Francis Bloemer
Caleb M. Bonds
Maureen Ann Boyle
Jeremy James Brigham
Kin Lun (Victor) Choi
Alan R. Clark
Brian Roscoe Coleman
Douglas Lawrence Dee
Jonathan Mark Deutsch
Richard James Engelhuber
Weishu Fan
Kathleen Marie Farrell
Richard A. Fuller
Rainer Germann
Isabelle Gingras
Peter Scott Gordon
Stephanie Ann Gould
Robert Andrew Grocock
Rebecca N. Hai
David Lee Handschke
Philip M. Imm
Karen Lerner Jiron
Robert C. Kane

Linda S. Klenk
Ravi Kumar
Julie-Linda Laforce
Chingyee Teresa Lam
John B. Landkamer
Aaron Michael Larson
Shangjing Li
Joshua Nathan Mandell
Kevin Paul McClanahan
Ian John McCracken
Shawn Allan McKenzie
Christian Menard
Peter Victor Polanskyj
Darin L. Rasmussen
Josephine Teruel Richardson
Marn Rivelle
Delia E. Roberts
Tina Shaw
Joseph Allen Smalley
Michael William Starke
David K. Steinhilber
Stephen James Streff
Josephine L. C. Tan
Javanika Patel Weltig
Rosemary Gabriel Wickham
Apryle Oswald Williams
Dean Michael Winters
Jeffrey S. Wood

The following candidates successfully completed the following Parts of the Fall 1999 CAS Examinations that were held in November.

Part 3B

Alan M. Chow
Kelly K. Cusick
Christopher A. Donahue
Kyle A. Falconbury
John S. Flattum
Feifei Ford
Matthew R. Gorrell
Erik A. Johnson
Brian J. Kasper
Kenneth Lin
Timothy C. Mosler
Carole Nader
Bhikhabhai C. Patel
Isabelle Perron
Christopher A. Pett
Robert E. Royer
Benjamin C. Strasser
David B. Thaller
Kieh Treavor Ty
Karen L. VanCleave
Jennifer A. Vezza
William B. Wilder
Xiaodong Yu

Part 4A

Andrea Ondine Ahern
Faisal Ahmed
Jennifer A. Ahner
Muhammad Munawar Ali
Fernando Alberto Alvarado
Brandie J. Andrews
Talal I. Arimah
Jennifer M. K. Arthur
Kevin J. Atinsky
Linda S. Baum
Nicolas Marc Beaudoin
Benjamin Beckman
Nathan L. Bluhm
Alla Bottoni
Jean-Philippe Boucher
John R. Bower
Maureen B. Brennan
John J. Brown
Suejeudi Buehler
Don J. Burbacher
Robert L. Bush

Thomas L. Cawley
Thomas C. Cecil
Rafael Ignacio Cespedes
James Chang
Dionne K. Chisolm
Kevin J. Christy
Stephan Cliche
Aimee B. Cmar
Robert J. Collingwood
Greg E. Conklin
Tina M. Costantino
Stephen M. Couzens
Lynn E. Cross
Karen O'Brien Curtin
Jeannine M. Danner
Rich A. Davey
Chantal Delisle
Brent P. Donaldson
Kevin P. Donnelly
Yvonne M. Duncan
Lisa S. Eichenbaum
Todd A. Ekey
Melissa D. Elliott
Yehoshua Y. Engelsohn
John M. English
Michael D. Ersevim
Michael A. Faria
Kevin M. Finn
March M. Fisher
Jennifer L. Fitzpatrick
Chad J. Gambone
Angela L. Garrett
Sophie M. L. Georget
Lillian Y. Giraldo
Mary T. Glaudell
Jennifer L. Glodowski
Jon H. Gottesfeld
Travis J. Grulkowski
Simon Guenette
Jonathan M. Guy
Benjamin D. Haas
John J. Hageman
Margarita Hambrock
Sunny M. Harrington
Sarah B. Hartung
Dedie C. Holley
Frank E. Horn
Esther Y. Hui
Mohammad A. Hussain
Hsu Hwang-Ming
Elena Ilina
Victoria K. Imperato
Yehuda S. Isenberg
William A. Jaeger
Jennifer L. Janisch
Dana F. Joseph
Eric J. Kendig
Sayeh Khavary
Thomas F. Klem
Aleksandr I. Korb
Bradley S. Kove
Leland S. Kraemer
Vladimir A. Kremerman
Frank K. Kumah
Terry T. Kuruvilla
Bobb J. Lackey
Heather D. Lake
Kan Yuk A. Lau
Eric T. Le
Shannon Rebecca Leckey
Kimi K. Lee
Jeffrey Leeds
Kenneth L. Leonard
Sean M. Leonard
Lorinda A. M. Leshock
Mark A. Lesperance
Frederic Levesque
Nannan Liu
Rachael A. LoBosco
Gwenette K. Lorino
Suzanne S. Luebbe
Keyang Luo
PeiQing Luo
Sally Ann MacFadden
Thomas J. Macintyre
Hilton Mak
Alison L. Matsen
Zinoviy Mazo
Laurence R. McClure II
John D. McMichael
Anne A. McNair
Hernan L. Medina
Paul B. Miles
Yuchun Mu
Loralea A. Mullins
Sureena Binte Mustafa
Natalia Navarova
Jacqueline L. Neal
Tho D. Ngo
Steven A. Nichols
Matthew P. Nimchek
John N. Norman

William S. Ober
Melissa A. Ogden
Brent J. Otto
Ajay Pahwa
Lucia Papa
Neelam P. Patel
Robert Anthony Peterson
Dianne M. Phelps
Genevieve Pigeon
Daniel J. Plasterer
Timothy K. Pollis
Michael J. Quigley
David P. Rafferty
Sheikh M. Rahman
Vijay R. Ramanujan
Lynellen M. Ramirez
Monica L. Ransom
Dong-Jye Rau
Timothy J. Regan
Terri-Beth Reynolds
Michele S. Rosenberg
Marc R. Rothschild
Ray M. Saathoff
Andrei P. Salomatov
Dionne M. Schaaffe
Thomas Schneider
Parr T. Schoolman
Pinchas R. Schreiber
Matthew L. Schutz
Tammy L. Schwartz
Elizabeth A. Sexauer
Clista E. Sheker
Lori A. Sheppard
Brett M. Shereck
Glenn D. Shippey
James M. Smieszkal
Jennifer L. Smith
Todd G. Smith
Molly A. Stark
Jason D. Stubbs
Wei Hua Su
Linda Sun
Adam D. Swope
Sandrine K. Tagni
Dominic A. Tocci
Joseph S. Tripodi
Michael S. Uchiyama
Paul A. Vendetti
Steven R. Waldman
Gary C. Wang
Qingxian Wang
Bethany R. Webb
Jean P. West
Josianne M. Wickham
Joshua C. Worsham
Andrew F. Yashar
Jong H. Yoo
Megan L. Zack
Anna Zieba

Part 4B

Anthony W. Ackley
Karen H. Adams
Jon R. Aerni
Armine Aharonyan
John C. Albrecht
Michael D. Altier
Catherine Ambrozewicz
Dorothy L. Andrews
David H. Anenberg
Ashwin Arora
Yuliya V. Artemov
Richard Audet
John L. Baldan
Cornel Balteanu
Dan S. Barnett
Warren C. Barney
Kim M Basco
Isabelle Belanger
Richard J. Bell III
Sylvain Belley
Andrew W. Bernstein
Maulik Bhansali
William J. Blatcher
Craig J. Blumenfeld
Stephane A. Boileau
Gilbert R. Booher
Nigel B. Branker
Craig R. Bridge
Alma R. Broadbent
Jeffrey A. Brueggeman
Monica M. Buck
Randall T. Buda
Don J. Burbacher
Wei Cai
Francisco Camba
Glenalan C. Cameron
Jonathan H. Camire
Jason A. Campbell
Christina A. Candusso
Stanley R. Caravaggio
Thomas C. Cecil

Raji H. Chadarevian
Man Ho Chan
Phyllis B. Chan
Wai Kit Chan
Yanli Hwang Chan
Margaret A. Chance
Naxine Chang
Jennifer A. Charlonne
Kuo-mei Chen
Qian Chen
Yi Chuan Chen
Yutian Chen
Henry K. Cheng
Marianne Cherkez
Ka-Chun Cheung
Lai Hing Cheung
Xiaolei Chi
Wai Choi
Julia Chou
Ka Ming (Danny) Chow
George E. Christopher
Daisy L. Chu
Wai Yee Salina Chung
Wesley G. Clarke
Sebastien Clement
Aimee B. Cmar
Christian J. Coleianne
John C. Collingwood
Peter D. Collins
Cameron A. Cook
Leanne M. Cornell
John E. Costango
Michael J. Covert
Michael B. Cunningham
Bridget A. Cupp
Richard A. Cuzzone
Walter C. Dabrowski
Ka-Ming Dai
David B. Dalton
James T. Daniels
Ryan E. Daniels
Amy L. DeHart
Sheri Lee de La Boursodiere
Timothy A. DeMars
Craig L. DeSanto
Michael J. Dekker
Laura M. Dembiec
Diana M. Dodu
Brent P. Donaldson
Margaret H. Donavan
Brian S. Donovan
Charles W. Dorman
Kristen S. Dossett
Etienne Dube
Matthew D. Dunscombe
Aaron D. Ekstrom
Malika El Kacemi
Brian Elliott
Troy R. Elliott
Jessica L. Elsinger
Seong-min Eom
Amy R. Eversole
Carlos M. Fajardo
Derek L. Farmer
Mark S. Feldman
Donna K. Ferguson
Anusha M. Fernando
Dale A. Fethke
Lawrence K. Fink
Marten W. Finlator
William M. Finn
Steven M. Fix
Eric P. Fortier
Pierre Fortier
Sebastien Fortin
Robert J. Foskey
Jason L. Franken
Louise Frankland
Gregory A. Frankowiak
Laurence Frappier
Yan Fridman
Eric S. Friedman
Michael C. Fruchter
Paul M. Y. Fung
Joseph Gabriel
Karl Gagnon
Samih S. Geha
Mark X. Geske
Michael P. Gibson
Daniel James Giovannone
Dominique Godin
Noah P. Goldstein
Samantha A. Graber
David S. Graham
Elizabeth A. Grande
Gaelle Gravot
Heather L. Grebe
Christa Green
Veronique Grenon
Stacie R.W. Grindstaff
Stephanie A. Groharing
Isabelle Groleau

Waleed H. Grunden
Serhat Guven
Theodore A. Haard
Marilou I. Halim
Jon E. Hamberg
Sunny M. Harrington
Jason S. Hart
Gary A. Hatfield
Yong Hao He
Joshua E. Hedgecorth
Mindy E. Herzog
Brandon L. Heutmaker
Lauren E. Heyl
Scott P. Higginbotham
Carole K. L. Ho
Tony Yiu Tung Ho
Jeremy A. Hoch
Mitchell H. Hofing
Eric B. Hofman
Michael A. Holderman
James E. Holland Jr.
Jeffrey I. Holm
Melissa S. Holt
Hyunpyo Hong
Wen Cai Hou
Chih-Che Hsiao
Tsu-Yueh Hsueh
Sheng-Fei Huang
Chad A. Hueffmeier
David P. Hughes
Wan Yee Connie Hui
Scott A. Humpert
Jawyih J. Hung
Li-Jiuan Hung
Pui Yuen Hung
Suzette L. Huovinen
Kenneth L. Israelsen
Jesse T. Jacobs
Suzanne Jacques
William A. Jaeger
Steven N. Jankovich
Michael S. Jarmusik
Kurugamega C. Jayawardena
Han Jiang
Lori K. Johnson
Brigitte Joncas
Nathan L. Jones
Julie A. Jordan
Dana F. Joseph
James A. Juillerat
Minas K. Kalachian
Kuei-Hua Kan
Linda M. Kane
Tami J. Karnatz
Jennifer L. Kearon
Susan M. Keaveny
Stephen G. Kelloway
Amy Jieseon Kim
Chung-Hun Kim
Sang W. Kim
Roman Kimelfeld
Melissa J. Kirshenbaum
Linda M. Klaips
Steven T. Knight
Hiu-Wan Ko
John E. Kollar
Karen E. Koop
Aleksandr I. Korb
Alexey P. Kozmin
Julia R. Kraemer
Brandon E. Kubitz
Rohan P. Kumar
Eric M. Kurzrok
Nadya Kuzkina
Kristine Kuzora
Claudel Laguerre
Hooi Lee Lai
Robert Lamarche
Neil A. F. Lamb
Stacey B. Lampkin
James A. Landgrebe
Julie L. Landreville
Andre Landry Jr.
Yuk Yee Lau
Jason A. Lauterbach
Michaela Ledlova
Chengwei Lee
Victor C. Lee
York Hon John Lee
Geraldine Marie Z. Lejano
Twiggy Lemercier
David Sean Leonard
Michael A. Leonberger
Wesley Leong
Charles L. Levine
Jonathan D. Levy
Michael B. Lewis
Bin Li
Hing Keung Li
Jiehui Li
Kin Hing Li
Oi K. Li
Rongmin Li
Jenn Y. Lian
Yong Hua Liang

Jennifer L. Ligon
Ruey Shyan Lin
Yi-Ling Lin
Andy M. Liu
Dong Liu
Guan-Bo Liu
Jianxun Liu
Mei-Chu Liu
Ruixue Liu
Xiaoquing Iris Liu
Zhanzhong Liu
Todd L. Livergood
John Cy Lo
Winnie Lo
Phillip J. Loftus
Michael H. Loretta
Harold E. Luber
Chun-Shuo Ma
Dick Ka Ma
Huixiu Ma
Anna B. Maciejewska
Teresa Madariaga
Lynn C. Malloney
Ratsamy Manoroth
Dan Mao
Roy M. Markham
Rene Martel
Thomas D. Martin
Lora K. Massino
Joseph W. Mawhinney
Michael B. McCarty
James P. McCoy
Joseph N. McDonald
David A. McMahon
Melissa A. McMains
Sylwia S. McMichael
Alexander Medvetsky
Mehul D. Mehta
Andre-Claude Menard
Duane G. Middendorf
Xiaohong Mo
Yi Man Mok
Christopher K. Moore
Anne Morency
Vincent Morin
Donald F. Morrison
Fritzner Mozoul
Yuchun Mu
Sumera Muhammad
Laura M. Murphy
Donald P. Myers
Natalia Navarova
Muhammad H. Nazir
Georgia A. Nelson
Jason G. Neville
Daniel T. Newton
Ka Yee Ng
Kit Wan Ng
May-Yee Ng
Mona Y. Ng
Julie K. Nielsen
Robert Niyazov
Jabran Noor
Russel W. Oslund
Shunli Pan
Hua Ying Pang
Michel Pare
Alexa Patterson
Agnes Paul
Christopher A. Paulus
Brian T. Pedersen
Guanghui Peng
Shu Y. Peng
Robert B. Penwick
Julien Perreault
Michael C. Petersen
Christopher A. Pett
Dianne M. Phelps
David A. Pitts
Ka Lok Po
Sue L. Poduska
Christopher R. Poirier
Flavia H. F. Poon
Daniel P. Post
Stephane Poulin
Stephen R. Prevatt
Marvin R. Puymon
Yubo Qiu
Darryl L. Raines
Heather N. Ramsay-Acosta
Lei Rao-Knight
William C. Reddington
Zia Ur Rehman
John J. Reid
Brent F. Reis
Adam J. Rennison
Danis Rheault
Richard G. Rhode
Wendi L. Richmond
Joseph L. Rizzo
Stanley T. Roberts
Keith A. Rogers
Jeff D. Rohlinger
An Qi Rong
John J. Rosati Jr.
Rebecca B. Rosenbaum
John D. Rosilier

Kelly J. Rosseland
Ryan P. Royce
Katherine I. Russell
Frank A. Santasiero
Janice Pauline D. Santos
Steven J. Savard
Reid M. Schaefer
Andrew F. Schallhorn
Vickie J. Scherr
Ernesto Schirmacher
Thomas W. Schroeder
Ronald J. Schuler
Paul A. Schultz
Brent W. Seiler
Tomasz Serbinowski
Richard H. Seward
Fahad R. Shah
Mayur M. Shah
Heather Shemek
Ye Shen
Brett M. Shereck
Jeremy D. Shoemaker
Andrew P. Shull
Sing Chai Siau
Summer L. Sipes
Robert P. Siwicki
Cory J. Skinner
Amado C. Sleiman
Stephen G. Slocum
Audrey L. Smerchansky
Daren M. Smith
Douglas E. Smith
Wallace G. Smith
Joao M. Soares
Joshua A. Sobol
Eric P. Sock
Marc St-Jacques
Andreas J. Stabno
Amy L. Steburg
Donna B. Steepe
Laura B. Stein
Kyrke O. Stephen
Richard M. Stiens
Kevin H. Strobel
Elizabeth D. Strong
Moshe Stulman
Nicki A. Styka
Louis P. Sugarman
Ju-Young Suh
Bin Sun
Qi Sun
Konrad P. Szatzschneider
Su-Chuan Tai
Takashi Tanemura
Connie W. Tang
Hai Peng Tang
Li Qin Tang
Sebastien Tanguay
Veronique Tanguay
Julie Tanguy
Jeffrey D. Thacher
Deepak Thakor
Sarah E. Theis
Christian A. Thielman
Jonas F. Thisner
Clinton Jay Thompson
Henry K. To
Siu Yin To
Michael C. Torre
Raymond D. Trogdon
Matthew D. Trone
Feng-Hui Tsai
Wen-Tzu Tsai
Jeffery G. Turnbull
Michael S. Uchiyama
Eric R. Ulm
Chris M. Vanden Haak
Jason A. Vary
Nilesh M. Vasani
Sylvain Veilleux
Frederic Venne
Tomas Vezauskas
Maxim Viel
Sebastien Y. Vignola
Remi Villeneuve
John T. Volanski
Benny Wan
Gary C. Wang
Jianbing Wang
Darren M. Welch
Kenneth P. Westman
John J. Whitaker
Gregory A. Whittaker
Timothy P. Wiebe
Andrew P. Wieduwilt
Stephen C. Williams
Rebecca Yang Wilson
Ian G. Winograd
Chun Shan Wong
Kim W. Wong
Laiping Wong
Philip Wong
Shing-Ming Wong
Tak Chi Wong
Yuk Lun Wong

Agnieszka E. Wygladala
Andreas Wyler
Jun Feng Xie
Huan Yao Xu
Xue Mei Xu
Dimitris Xynogalas
Benjamin J. Yang
Yan Yang
Andrew F. Yashar
Manha Yau
Heather M. Yonosh
Janice M. Young
Jiyoung Yue
Raymond R. Y. Yung
Alexandru Zaharia
Ali A. Zaker-Shahrak
Liang Zhang
Xiaoyu Zhang
Wei Dong Zhou

Part 5A

Brian M. Ancharski
Ashaley N. Attoh-Okine
Chris D. Barela
Marie-Eve J. Belanger
Angela D. Burgess
Rachel A. Cills
Brenda K. Cox
Kevin P. Donnelly
Peter M. Doucette
Juan Espadas
Brandon L. Heutmaker
David G. Keeton
Stephen J. Langlois
Sean R. Lawley
Amy E. LeCount
Wendy R. Leferson
Erik F. Livingston
Laurence R. McClure II
Jennifer A. McGrath
Rebecca E. Miller
James C. Murphy
Sureena Binte Mustafa
Loren J. Nickel
Sebastien Portmann
Laura B. Sachs
Anthony D. Salido
Michelle L. Sheppard
Nicki A. Styka
Phoebe A. Tinney
John D. Trauffer
Gaetan R. Veilleux
Kimberly A. Vogel
Tice R. Walker

Part 5B

Shawn C. Adams
Felix F. Aguilar
Aaron D. Albert
Fernando Alberto Alvarado
Daryl S. Atkinson
Nicki C. Austin
Joseph M. Beesack
Richard J. Bell III
Mattthew C. Berasi
Chris M. Bilski
Robin V. Blasberg
Timothy D. Boles
Stephen A. Bowen
Elaine K. Brunner
Lisa K. Buege
Brian P. Bush
Douglas J. Busta
Cemal Alp Can
Rachel A. Cills
Robert J. Collingwood
Kelly K. Cusick
Francis J. Dooley
William E. Doran
Elaine V. Eagle
Jeffrey S. Ernst
Jieqiu Fan
Jennifer L. Fitzpatrick
Tricia D. Floyd
Katherine M. Funk
Timothy S. Grant
Ann E. Green
Diane E. Grieshop
Jeffrey A. Gruel
Deborah J. Gurnon
Jonathan M. Guy
Koichi Hamasaki
Jason C. Harland
Kandace A. Heiser

Keri P. Helgeson
Robert E. Heyen
Carole K. L. Ho
Michael C. Hogan
Elena Ilina
Ronald J. Jankoski
Megan S. Johnson
Madeleine R. Kaestli
Brian M. Karl
Scott M. Klabacha
John E. Kollar
Eric T. Krause
Charles B. Kullmann
Thomas P. Langer
Nancy E. Lanier
Eric N. Laszlo
Xun-Yuan Liang
Steven R. Lindley
Andriy P. Loboda
Michael L. Loritsch
Wing Lowe
PeiQing Luo
Gavin Raj Maistry
Laura S. Marin
Susan E. Marra
Craig L. Merrill
Pantelis N. Messolonghitis
Chad M. Miller
Paul J. Molinari
George C. Moulton
Sureena Binte Mustafa
John J. Myers
Lisa M. Nield
Brent J. Otto
Bruce G. Pendergast
Karen M. Peterka
Robert Anthony Peterson
Terry C. Pfeifer
Timothy K. Pollis
Miriam Polyakov
Terry W. Quakenbush
Benjamin L. Richards
Kevin D. Roll
Randall D. Ross
Farid Sandoghdar
Michael J. Scarborough
Mark W. Schluesche
Robert E. Schmid
Elizabeth M. Scott
Yipei Shen
Ranjit B. Shiralkar
Lance H. Shull
Vijayalakshimi Sridharan
Alexandra R. St-Onge
Christopher J. St. George
Kevin L. Stephenson
Lisa Liqin Sun
Hugh T. Thai
Malgorzata Timberg
Peter R. Vita
Matthew J. Walter
Tom C. Wang
Jennifer X. Wu
Keith Young
Wei Zhang

Part 7-Canada

Patrick Barbeau
Brad D. Birtz
Richard Jason Cook
Jean-François Desrochers
Louis-Christian C. H. Dupuis
John S. Giles
Chantal Guillemette
Patricia A. Hladun
Omar A. Kitchlew
Jean-François Larochelle
W. Scott Lennox
Stephane McGee
Martin Menard
Eric Millaire-Morin
Lambert Morvan
Cosimo Pantaleo
Bill D. Premdas
Nathalie Tremblay
Richard A. Van Dyke

Part 7-United States

Jodie Marie Agan
Brian C. Alvers
Denise M. Ambrogio
Kevin L. Anderson
Peter Attanasio
Maura Curran Baker
Mary P. Bayer
Jody J. Bembenek
Jeremy T. Benson
Jason E. Berkey
Ellen A. Berning
Kofi Boaitey
Mary Denise Boarman
Thomas L. Boyer II
David C. Brueckman
Claude B. Bunick
Fatima E. Cadle
Ronald S. Cederburg
John Celidonio
Hao Chai
Sigen Harry Chen
Brian K. Ciferri
Susan M. Cleaver
Kiera E. Cope
Kevin A. Cormier
Thomas Cosenza
Paul T. Cucchiara
David F. Dahl
Peter R. DeMallie
Patricia A. Deo-Campo Vuong
Mike Devine
Mary Jane B. Donnelly
Kevin G. Donovan
Scott H. Drab
Donna L. Emmerling
Keith A. Engelbrecht
Laura A. Esboldt
Farzad Farzan
Christine M. Fleming
Donia N. Freese
Shina Noel Fritz
Cynthia Galvin
Michael A. Garcia
Dustin W. Gary
Hannah Gee
Laszlo J. Gere
Christie L. Gilbert
Patrick J. Gilhool
Joseph E. Goldman
Andrew S. Golfin Jr.
Olga Golod
Stacey C. Gotham
James C. Guszcza
David B. Hackworth
Dawn Marie S. Happ
Jason C. Head
Pamela B. Heard
Kristina S. Heer
Hans Heldner
Scott E. Henck
Deborah L. Herman
Mark D. Heyne
Robert C. Hill
David E. Hodges
Allen J. Hope
Derek R. Hoyme
Carol I. Humphrey
Rusty A. Husted
Thomas D. Isensee
Michael S. Jarmusik
Gregory O. Jaynes
Philippe Jodin
Steven M. Jokerst
Gregory K. Jones
Lawrence S. Katz
Cheryl R. Kellogg
David R. Kennerud
Susanlisa Kessler
Young Y. Kim
James F. King
Jill E. Kirby
Henry J. Konstanty
Darjen D. Kuo
Christine L. Lacke
Peter H. Latshaw
Doris Lee
Jeffrey Leeds
Joshua Y. Ligosky
Jia Liu
Jing Liu
Rebecca M. Locks
Kathleen T. Logue
Richard P. Lonardo
William F. Loyd
Alexander P. Maizys
Victor Mata
David M. Maurer
Timothy C. McAuliffe
John R. McCollough
Richard J. McElligott

Mitchel Merberg
Vadim Y. Mezhebovsky
Suzanne A. Mills
Matthew K. Moran
Celso M. Moreira
Thomas M. Mount
Joseph J. Muccio
Scott L. Negus
Ronald T. Nelson
Michael D. Neubauer
Stoyko N. Nikolov
Mary A. Noga
Joshua M. Nyros
Rodrick R. Osborn
Carolyn Pasquino
Michael T. Patterson
Wendy W. Peng
Jill E. Peppers
Kevin T. Peterson
Kraig P. Peterson
Kristin S. Piltzecker
Warren T. Printz
Stephen D. Riihimaki
Ezra J. Robison
Scott I. Rosenthal
Bryant E. Russell
Frederick D. Ryan
Laura B. Sachs
Salimah H. Samji
Rachel Samoil
Jennifer A. Scher
Daniel David Schlemmer
Darrel W. Senior
Larry J. Seymour
Paul O. Shupe
Lee O. Smith
Lora L. Smith-Sarfo
Scott G. Sobel
Anthony A. Solak
Christine L. Steele-Koffke
Gary A. Sudbeck
Jonathan L. Summers
Edward Sypher
Neeza Thandi
Tanya K. Thielman
Lien K. Tu
Kieh Treavor Ty
Matthew L. Uhoda
Dennis R. Unver
Justin M. Van Opdorp
Cameron J. Vogt
Josephine M. Waldman
Colleen Ohle Walker
Kristie L. Walker
Janet L. Wang
Shaun S. Wang Ph.D.
Ya-Feng Wang
Petra L. Wegerich
Joseph C. Wenc
Chris J. Westermeyer
Paul D. Wilbert
Amy M. Wixon
Karin H. Wohlgemuth
Terry C. Wolfe
Mihoko Yamazoe
Run Yan
Nora J. Young
Gene Q. Zhang
Lianmin Zhou
Eric E. Zlochevsky

Part 9

Jason R. Abrams
Michael Bryan Adams
Genevieve L. Allen
Amy Petea Angell
Katherine H. Antonello
Wendy Lauren Artecona
Martha E. Ashman
Joel E. Atkins
David Steen Atkinson
Patrick Beaudoin
Kristen Maria Bessette
David M. Biewer
John T. Binder
Linda Jean Bjork
Neil M. Bodoff
Mark E. Bohrer
Veronique Bouchard
Erick A. Brandt
James L. Bresnahan
John R. Broadrick
Sara T. Broadrick
Paul E. Budde
Julie Burdick
John C. Burkett
Robert N. Campbell
Allison F. Carp
Joseph G. Cerreta
Nathalie Charbonneau
Patrick J. Charles

Yvonne W. Y. Cheng
Thomas Joseph Chisholm
Kin Lun (Victor) Choi
Michael Joseph Christian
Julia Feng-Ming Chu
Louise Chung-Chum-Lam
Jason T. Clarke
Jeffrey J. Clinch
Eric John Clymer
Christopher Paul Coelho
Steven A. Cohen
Larry Kevin Conlee
Christopher L. Cooksey
Kathleen T. Cunningham
M. Elizabeth Cunningham
Jonathan Scott Curlee
Kris D. DeFrain
Nancy K. deGelleke
Michael Brad Delvaux
Pamela G. Doonan
Sharon C. Dubin
Tammi B. Dulberger
George T. Dunlap IV
Sophie Duval
Kevin M. Dyke
Mark Kelly Edmunds
Jane Eichmann
Gregory James Engl
Kathleen Marie Farrell
Benedick Fidlow
Kenneth D. Fikes
Ronnie Samuel Fowler
Michelle L. Freitag
Anne M. Garside
Abbe B. Gasparro
Charles E. Gegax
Gregory Evan Gilbert
Isabelle Gingras
Todd Bennett Glassman
Christopher David Goodwin
Francis X. Gribbon
Marvin Harlan Grove
Lisa N. Guglietti
Nasser Hadidi
Rebecca N. Hai
Eric Christian Hassel
Jeffery Tim Hay
Qing He
Amy Louise Hicks
Jay T. Hieb
Glenn R. Hiltpold
Christopher Todd Hochhausler
Todd Harrison Hoivik
Long-Fong Hsu
John F. Huddleston
Craig D. Isaacs
Philip J. Jennings
Weidong Wayne Jiang
Charles B. Jin
Michael S. Johnson
Bryon Robert Jones
William Rosco Jones
Kyewook (Gary) Kang
Sean M. Kennedy
Stacey M. Kidd
Jennifer E. Kish
Brandelyn C. Klenner
Richard Scott Krivo
Scott C. Kurban
Kirk L. Kutch
Isabelle La Palme
Elaine Lajeunesse
William J. Lakins
Chingyee Teresa Lam
Carl Lambert
Travis J. Lappe
Ramona C. Lee
James P. Leise
Christian Lemay
John Norman Levy
Sally Margaret Levy
Shangjing Li
Xiaoying Liang
Matthew A. Lillegard
Dengxing Lin
Diana M. S. Linehan
Daniel A. Lowen
Joshua Nathan Mandell
Jason Aaron Martin
Stephen Joseph McAnena
Kevin Paul McClanahan
Kirk Francis Menanson
Mark F. Mercier
Richard Ernest Meuret
Jennifer Middough
Scott A. Miller

Ain Milner
Christian Morency
Jarow G. Myers
Brian C. Neitzel
Sean R. Nimm
John E. Noble
Sylvain Nolet
Corine Nutting
Randall William Oja
Christopher Kent Perry
Kathy Popejoy
Ni Qin-Feng
Ricardo A. Ramotar
Leonid Rasin
Sara Reinmann
Sylvain Renaud
Paul J. Rogness
John R. Rohe
Kim R. Rosen
Richard A. Rosengarten
Sandra L. Ross
Brian P. Rucci
James C. Santo
Asif M. Sardar
Jason Thomas Sash
Jeremy N. Scharnick
Parr T. Schoolman
Stuart A. Schweidel
William Harold Scully III
Steven George Searle
Ernest C. Segal
Vladimir Shander
Tina Shaw
Meyer Shields
Theodore S. Spitalnick
Benoit St-Aubin
Scott T. Stelljes
Karrie Lynn Swanson
Varsha A. Tantri
Jonathan Garrett Taylor
Robert M. Thomas II
Laura Little Thorne
Christopher S. Throckmorton
Gary S. Traicoff
Michael C. Tranfaglia
Thomas A. Trocchia
Brian K. Turner
Eric Vaith
Amy R. Waldhauer
Lynne Karyl Wehmueller
Scott Werfel
Dean Allen Westpfahl
Matthew M. White
Apryle Oswald Williams
Dean M. Winters
Jonathan Stanger Woodruff
Yin Zhang
Edward J. Zonenberg

NEW FELLOWS ADMITTED IN MAY 1999

First row, from left: Betsy A. Branagan, Alana C. Farrell, **CAS President Steven G. Lehmann,** Deborah M. King, Michael Shane. **Second row, from left:** Eleni Kourou, Elliot Ross Bum, Dawn M. Lawson, Claudine Helene Kazanecki, Christopher C. Swetonic. **Third row, from left:** Brian Harris Deephouse, Richard Borge Lord, Bruce Daniel Fell. **Not pictured:** Mustafa Bin Ahmad.

NEW ASSOCIATES ADMITTED IN MAY 1999

First row, from left: Larry Kevin Conlee, Jennifer L. Throm, Nathalie Chatbonneau, **CAS President Steven G. Lehmann,** Karen N. Levine, Silvia J. Alvarez, Joseph Paul Greenwood. **Second row, from left:** Vladimir Shander, Yvonne W.Y. Cheng, Nathalie J. Auger, Andrea Elisabeth Trimble, Sally Margaret Levy, Sara Reinmann, Amy Louise Hicks, Joseph John Sacala. **Third row, from left:** Steven A. Cohen, Stephane Brisson, Jason R. Abrams, Paul Jerome Johnson, Terrie Lynn Howard, Anne M. Garside, Emily C. Gilde, Vahan A. Mahdasian. **Fourth row, from left:** Douglas M. Warner, Sean Oswald Curtis Cooper, Paul Edward Budde, Thomas LeRoy Poklen Jr., Jay T. Hieb, Jonathan Stanger Woodruff, Glenn R. Hiltpold, Kirk Francis Menanson.

NEW ASSOCIATES ADMITTED IN MAY 1999

First row, from left: Gary Steven Traicoff, Stephen James Talley, Catherine L. DePolo, **CAS President Steven G. Lehmann,** Conni Jean Brown, Sean Paul Forbes, Annmarie Schuster, Julia Feng-Ming Chu. **Second row, from left:** Burt D. Jones, Thomas S. Botsko, Jo Dee Thiel-Westbrook, Joseph Francis Rosta Jr., Brian Michael Fernandes, Frances Ginette Sarrel, Gwendolyn Anderson. **Third row, from left:** Brian K. Turner, Jeffery Wayne Scholl, Michael A. Pauletti, Daniel George Charbonneau, Jeffrey J. Clinch, Derek A. Jones. **Fourth row, from left:** Paul E. Green Jr., Anthony L. Alfieri, Todd Harrison Hoivik, Todd Douglas Cheema, James M. Gallagher, Jason Thomas Sash.

NEW ASSOCIATES ADMITTED IN MAY 1999

First row, from left: David C. Riek, Dengxing Lin, Sophie Duval, Prabha Pattabiraman, **CAS President Steven G. Lehmann,** Allison F. Carp, Yin Zhang, Seth Shenghit. **Second row, from left:** Derek D. Burkhalter, Michael S. Harrington, Isabelle La Palme, Bryan Hartigan, Sharon Xiaoyin Li, Anthony J. Pipia, Eric John Clymer. **Third row, from left:** Christian Lemay, Mario Richard, Patrick Beaudoin, Jose R. Couret, David W. Warren, Kristen Maria Bessette, Laura Smith McAnena, Christopher Kent Perry. **Fourth row, from left:** Sylvain Perrier, Justin Gordon Gensler, Sylvain Renaud, Robert Allan Rowe, Peter Abraham Scourtis, Jordan J. Pitz, Ronnie Samuel Fowler, Mark R. Frank.

NEW ASSOCIATES ADMITTED IN MAY 1999

First row, from left: Jon S. Walters, Rosemary Catherine Peck, Randall William Oja, **CAS President Steven G. Lehmann,** Janelle Pamela Rotondi, Meredith Martin Woodcock, Borwen Lee. **Second row, from left:** Mark E. Bohrer, Julie Burdick, Amy Lynn Baranek, Karen Ann Brostrom, David Ernest Corsi, Albert Maroun, Mujtaba H. Datoo. **Third row, from left:** Michael Bryan Adams, Jayme P. Stubitz, Leo Martin Orth Jr., David R. Border, John Michael Pergrossi, Jeffery Tim Hay, Fanny C. Paz-Prizant. **Fourth row, from left:** Thomas F. Krause, Christopher David Bohn, John T. Binder, Paul D. Anderson, Robert M. Thomas II, Glenn Steven Hochler, Jeffrey Alan Clements, Steven Bradley Zielke.

NEW ASSOCIATES ADMITTED IN MAY 1999

First row, from left: Kelly A. Lysaght, Sharon R. Corrigan, Carolyn J. Coe, **CAS President Steven G. Lehmann,** Sheri L. Oleshko, Kathleen Frances Robinson, Jason Aaron Martin. **Second row, from left:** Timothy L. McCarthy, Ain Milner, Timothy Michael DiLellio, Genard J. Palisi, Perry Keith Wooley, Sean O. Cooper, David Garrett Shafer. **Third row, from left:** Serge Gagné, Mark Richard Strona, Michael Douglas Nielsen, Anthony Robert Bustillo, David James Belany, John Edward Daniel, Michael W. Morro. **Fourth row, from left:** Ung Min Kim, Travis J. Lappe, Brook A. Hoffman, Kevin Earl Weathers, Bryon Robert Jones, Qing He, Kenneth D. Fikes. **New Associates admitted in May 1999 who are not pictured:** Amy Petea Angell, Anju Arora, Mario Binetti, Jean A. DeSantis, James Robert Elicker, Gregory James Engl, Janine Anne Finao, Theresa Giunta, Todd Bennett Glassman, Brendan Michael Leonard, Kevin M. Madigan, Atul Malhotra, Rasa Varanka McKean, Sarah Kathryn McNair-Grove, John-Giang L. Nguyen, William Dwayne Rader Jr., James C. Santo, Jeremy Nelson Scharnick, Trevar K. Withers.

NEW FELLOWS ADMITTED IN NOVEMBER 1999

First row, from left: James B. Kahn, Robert J. Wallace, David Patrick Moore, **CAS President Steven G. Lehmann,** Patricia Cheryl White, Brian Tohru Suzuki, Michael D. Brannon, Ronald J. Herrig. **Second row, from left:** Christopher Ross Heim, Nathan Schwartz, Hiep T. Nguyen, Matthew C. Mosher, John T. Gleba, Sarah Krutov, David E. Heppen, Richard J. Babel. **Third row, from left:** Gary J. Ganci, Darrel W. Chvoy, Christopher J. Monsour, M. Charles Parsons, Gregory Hansen.

NEW FELLOWS ADMITTED IN NOVEMBER 1999

First row, from left: Thomas P. Gibbons, James L. Nutting, Ann M. Bok, Denise R. Olson, **CAS President Steven G. Lehmann,** Cynthia A. Bentley, Luba O. Pesis, Jeffrey S. Trichon. **Second row, from left:** Mark B. Anderson, Lisa A. Bjorkman, Christopher Edward Olson, Martin S. Arnold, Michael A. Nori, James M. Kelly, Douglas W. McKenzie. **Third row, from left:** Yin Lawn, Karen L. Queen, David E. Marra, Michael Edward Doyle, Michael Boyd Masters, James D. Kunce, Paul W. Mills, Timothy Andrew Davis.

NEW FELLOWS ADMITTED IN NOVEMBER 1999

First row, from left: Rimma Abian, Janet G. Lindstrom, Allison Michelle McManus, **CAS President Steven G. Lehmann,** Siu K. Li, Tracy A. Ryan, Donna M. Nadeau, Anthony E. Cappelletti. **Second row, from left:** Simon Kai-Yip Wong, Kai Lee Tse, Donald M. Gambardella, Kathleen Mary Quinn, Leslie Alan Vernon, Ethan David Allen, William R. Maag. **Third row, from left:** Adam M. Swartz, Jean-Pierre Gagnon, Roosevelt C. Mosley, Gary T. Ciardiello, Jay Matthew South, Scott T. Hallworth, P. Claude Lefebvre.

NEW FELLOWS ADMITTED IN NOVEMBER 1999

First row, from left: James R. Merz, Jodi J. Healy, Dawn E. Elzinga, Randall S. Nordquist, **CAS President Steven G. Lehmann,** Noel M. Hehr, Angela Kaye Sparks, Brian K. Cox. **Second row, from left:** Jonathan Everett Blake, Thomas A. Huberty, Karen Barrett Daley, Bonnie C. Maxie, Christopher S. Downey, Vincent F. Yezzi, Kyle Jay Vrieze. **Third row, from left:** Brian L. Ingle, François L. Morissette, Suzanne E. Black, Jeffrey F. McCarty, Jay C. Gotelaere, Seth Andrew Ruff, Peter F. Drogan.

NEW FELLOWS ADMITTED IN NOVEMBER 1999

First row, from left: Philippe Gosselin, Dom M. Tobey, Wendy L. Witmer, **CAS President Steven G. Lehmann,** Edward H. Wagner, Hany Rifai, Bethany L. Cass, Kevin A. Lee. **Second row, from left:** Teresa K. Paffenback, Jean-François Chalifoux, Rachel Dutil, David Leblanc-Simard, Denis Dubois, Michael C. Schmitz, Chad C. Karls. **Third row, from left:** Martin Carrier, Christopher William Cooney, Nitin Talwalkar, Charles Pare, Rajesh V. Sahasrabuddhe, William P. Ayrcs, John W. Rollins, Kurt S. Dickmann. **New Fellows admitted in November 1999 who are not pictured:** Bryan C. Christman, Claudia Barry Cunniff, Jean A. DeSantis, Mary Ann Duchna-Savrin, Matthew E. Golec, David Thomas Groff, Michael B. Hawley, Mark J. Kaufman, Jean-Sebastien Lagarde, Lee C. Lloyd, Catherine A. Neufeld, David Anthony Ostrowski, Yves Raymond, David L. Ruhm, Bret Charles Shroyer, Matthew Robert Sondag, Benjamin A. Walden, Sheng H. Yu.

NEW ASSOCIATES ADMITTED IN NOVEMBER 1999

First row, from left: Dean Michael Winters, Todd Bennett Glassman, Isabelle Gingras, Maureen Ann Boyle, **CAS President Steven G. Lehmann,** Ian John McCracken, Javanika Patel Weltig, Tony Francis Bloemer. **Second row, from left:** Saeeda Behbahany, Kathleen Marie Farrell, Richard James Engelhuber, Julie-Linda Laforce, Caleb M. Bonds, Michael William Starke, Kevin Paul McClanahan, Douglas Lawrence Dee. **Third row, from left:** Christian Menard, Penelope A. Bierbaum, Robert Andrew Grocock, David Lee Handschke, Stephen James Streff, Jeffrey S. Wood, Weishu Fan.

NEW ASSOCIATES ADMITTED IN NOVEMBER 1999

First row, from left: Josephine L. C. Tan, Brian Roscoe Coleman, Stephanie Ann Gould, Rosemary Gabriel Wickham, **CAS President Steven G. Lehmann,** Apryle Oswald Williams, Linda S. Klenk, Genevieve L. Allen. **Second row, from left:** Tina Shaw, Brendan Michael Leonard, Joshua Nathan Mandell, Karen Lerner Jiron, Rebecca N. Hai, Peter Scott Gordon, Jeremy James Brigham, James P. Lynch. **Third row, from left:** Shangjing Li, Delia E. Roberts, Philip M. Imm, Peter Victor Polanskyj, Aaron Michael Larson, Shawn Allan McKenzie, Kin Lun (Victor) Choi. **New Associates admitted in November 1999 who are not pictured:** Michael D. Adams, Alan R. Clark, Jonathan Mark Deutsch, Richard A. Fuller, Rainer Germann, Robert C. Kane, Ravi Kumar, John B. Landkamer, Josephine Teruel Richardson, Marn Rivelle, Joseph Allen Smalley, David K. Steinhilber.

OBITUARIES

JOHN BEVAN
MARTIN BONDY
JOHN H. BOYAJIAN
J. EDWARD FAUST JR.
ROBERT L. HURLEY
PAUL S. LISCORD JR.
DANIEL J. LYONS
PHILIPP K. STERN

JOHN BEVAN
1917–1999

John Bevan died June 22, 1999, at the age of 82.

Bevan was a resident of Carleton–Willard Village in Bedford, Massachusetts for four years, but had lived for many years in neighboring Lexington. In Lexington, Bevan was active in community affairs, serving as an Appropriations Committee member and Town Meeting member.

After attending Newton High School and Mount Hermon School, Bevan graduated from Wesleyan University in 1938. He attended Harvard Business School for one year and began his career at Liberty Mutual Insurance Company.

In 1942, Bevan volunteered for service in World War II and served as a navigator in the U.S. Army Air Corps. He flew numerous missions in the Pacific theater over New Guinea for which he was awarded the Distinguished Flying Cross. After completing his service, Bevan continued in the Reserves until 1955.

Bevan returned to Liberty Mutual after his military service, and became vice president/lead actuary until his retirement in

1980. Bevan also served on the board of directors of the Lexington Savings Bank and continued his volunteer work at First Parish and FISH.

Bevan became an Associate of the Casualty Actuarial Society in 1951 and a Fellow in 1953. He was a member of numerous CAS committees including the Committee on Social Insurance from 1965 to 1967, the Editorial Committee from 1967 to 1969, and the Committee to Review Election Procedures in 1969.

Ruth (Glynn) Bevan said of her husband, "John always looked forward with enthusiasm to the Society's writings [and to] renewing old friendships from distant parts of the U.S."

In addition to his wife, he is survived by a son, Roger of Ohio; a brother, David of Ohio; and three grandchildren. A son, Geoffrey, predeceased Bevan in 1997.

MARTIN BONDY
1998

Martin Bondy died November 27, 1998.

Bondy, whose contributions to the Society spanned four decades, earned his Associateship in 1953 and attained Fellowship in 1956. Throughout the 1950s and 1960s, Bondy was the author of numerous papers and reviews appearing in the *Proceedings of the Casualty Actuarial Society*. Some of these papers included "The Rate Level Adjustment Factor in Workmen's Compensation Ratemaking" (1956) and the review of "Burglary Insurance Ratemaking" (1967).

Among his many CAS activities, Bondy was a member of the Council (1964), chairperson of the Publicity Committee (1965), member of the CAS Board of Directors (1974–77), and an ex officio member of the Planning Committee (1974). In addition, Bondy was chairperson of the Committee on Loss Reserves (1975–76), consultant for the Education and Examination Committee–Examination (1978–83) and Examination Committee (1984–85), and a member of the Discipline Committee (1993–95).

Bondy was working as an actuary with Royal Liverpool Insurance Group in New York City in 1953, the year he became an Associate. In 1954, Bondy went to work for the New York State Insurance Department in New York City as an associate actuary. He was promoted to principal actuary in 1957. In 1959, Bondy made the move to Mutual Insurance Company (later known as Consolidated Mutual Insurance) in Brooklyn, New York. In the more than six years he was employed with Consolidated Mutual, Bondy served as actuary and was promoted to the posts of assistant treasurer (1961) and assistant vice president and actuary (1962–65). Bondy's next career move to Crum & Forster (first in New York and later Morristown, New Jersey) was the beginning of his longest company affiliation. During his more than 20-year tenure with the company, Bondy served as assistant vice president and actuary (1965–67), vice president and actuary (1968–73), vice president of corporate analysis and planning (1974–78), and senior vice president (1979–87).

In the years following his work at Crum & Forster, Bondy served as senior vice president of corporate planning for Home Insurance Company in New York City (1988–93). His last post before retiring was as senior vice president and chief actuary for Skandia America Group (1994–96), also in New York.

Karl Moller (FCAS 1990), a colleague of Bondy's at Home Insurance Company, remarked that Bondy was a man of many interests. He characterized Bondy as "a gentleman…very good at cards" who had a fondness for the English language. Other former colleagues in the actuarial department of Home Insurance memorialized Bondy, calling him a wise, compassionate, gentle man who inspired others with his humanity, humor, and delight in non-obvious truth.

JOHN H. BOYAJIAN
1917–1999

John H. Boyajian died August 29, 1999, in New Canaan, Connecticut. He was 81.

Born in Melrose, Massachusetts on December 8, 1917, Boyajian attended Northeastern University in Boston. He graduated from Northeastern in 1941 with a degree in mathematics and later did some graduate work in physics.

During World War II, Boyajian joined the Navy and taught midshipmen at Columbia University. He left the Navy in 1946 with the rank of Lieutenant.

In October 1944 he married his wife Jessie. Together they had two children.

Boyajian worked for the National Bureau of Compensation Insurance in New York City from 1946 to 1954. In 1954 he moved his family to San Francisco, where he worked for the California Inspection Rating Bureau until 1961. From 1961 to 1966 he was an actuary for the National Board of Fire Underwriters in New York City. In 1966 he became head actuary for New Jersey Manufacturers Insurance Company (now NJM Insurance Group) in Trenton, where he worked until his retirement in 1982.

In the years following his retirement, Boyajian kept busy with volunteer work at Helene Fuld Hospital in Trenton, New Jersey. Boyajian logged over 5,000 hours of volunteer work there, primarily in patient admittance.

Boyajian became an Associate of the Society in 1950 and a Fellow in 1956. His CAS committee work included service on the Auditing (1965–1966) and Finance (1968–1970) Committees, and as Sites Liaison (1971–1972).

Boyajian and his wife attended many CAS meetings together, often traveling with friends. “[Those were] some of the happiest times of his life," said Candace DeSantes, Boyajian's daughter. "He really loved being an actuary," DeSantes said.

A bowling and golf enthusiast, Boyajian was known for his sense of humor. "Everybody who knew him thought he was really funny," said DeSantes.

Boyajian is survived by his daughters, Lorna Goodrich of Brooklyn, New York, and Candace DeSantes of Westport, Connecticut; four grandchildren; and three sisters, Martha, Flora, and Betty. His wife Jessie predeceased him in 1979.

J. EDWARD FAUST JR.
1925–1996

J. Edward Faust was born on March 1, 1925. He attended the University of Notre Dame and graduated Cum Laude in 1945. After graduating, Faust became a Lieutenant in the Navy, serving from 1945 to 1948. The war ended just before Faust was to be shipped overseas.

In 1948, Faust graduated from the University of Michigan in Actuarial Studies and married his wife, Kathleen.

Faust began his career at the Indiana Department of Insurance in Indianapolis. He then began working at Nelson & Warren Actuaries Consulting Firm in St. Louis, Missouri. He was a member of the Casualty Actuarial Society for 40 years, becoming an Associate in 1956, and a Fellow in 1960.

In 1982, Faust organized his own consulting firm, J. Edward Faust, Jr., in Indianapolis where he continued his work until retirement in 1995. He died December 6, 1996.

When commenting on their busy family life, Mrs. Faust said of her husband that being skilled in math and being father to his children were his greatest accomplishments. Faust is survived by his wife; eight children, Joseph F., Debra, Daniel E., Mary Faith, Carol Anastasia, Eric Anthony, Frederick Martin, and J. Christopher; two brothers; 20 grandchildren; and two great-grandchildren.

ROBERT L. HURLEY
1911–1998

Robert L. Hurley died on October 26, 1998 in Soldiers and Sailors Memorial Hospital in Wellsborro, New York. He was 87.

Hurley was born in Boston, Massachusetts on July 5, 1911, the son of Michael and Anna Lambert Hurley. He was a member of Knights of Columbus and the Lions Club.

Hurley received his Associateship to the Casualty Actuarial Society in 1952 and his Fellowship in 1955. He served as a member of the Education and Examination Committee from 1967 to 1971 and on the Editorial Committee from 1970 to 1972. He also wrote eight papers published in the *Proceedings of the Casualty Actuarial Society*.

Hurley is survived by his two sons, Garrett of Pittsburgh, Pennsylvania and Ian of Riverdale, New York, and four grandchildren. His wife, Gabrielle Hurley, predeceased him in 1981.

PAUL S. LISCORD JR.
1925–2000

Paul S. Liscord Jr. of Peterborough, New Hampshire died February 23, 2000, after a long battle with prostate cancer. He was 74.

Liscord was born on October 29, 1925, in Hartford, Connecticut. He was a graduate of the Loomis School in Windsor, Connecticut and Dartmouth College in 1948. From March 1945 to May 1946, Liscord served in Europe with the Third Infantry Regiment, U.S. Army as a radio technician, achieving the rank of technical sergeant, 4th class.

After World War II, Liscord worked for the Travelers Insurance Company in Hartford, Connecticut from 1948 to 1970, where he became vice president in charge of all casualty actuarial operations within the company. He later joined the Insurance Company of North America in Philadelphia, managing their casualty actuarial department from 1971 to 1975. He also worked for the New Hampshire Department of Insurance in Concord and as chief actuary with the Massachusetts Insurance Rating Bureau in Boston. In 1977, Liscord founded his own consulting firm, Liscord, Ward and Roy Inc. in Concord, New Hampshire.

Liscord was a member of the Casualty Actuarial Society for over 40 years, serving as president in 1973, vice president in 1971, and as a member of the CAS Council from 1968 to 1970. He also served as chairperson on many committees including the Committee on Automobile Insurance Research from 1966 to 1968, the Committee on Sites from 1968 to 1970, and the Long Range Planning Committee from 1974 to 1975.

Liscord enjoyed singing and followed this passion throughout his life. He sang in school glee clubs and at the community Congregational churches to which he belonged. After his retirement in 1990, he sang with the New Hampshire Friendship Chorus, the Concord Vocal Octet, and the Concord Chorale, serving also as its president. His singing talents brought him to Western Europe, Eastern Europe and Russia, New Zealand, and Australia. On two occasions he performed at Carnegie Hall in New York.

Liscord is survived by his wife, and "love of his life," Helen MacDonald Liscord; two daughters, Jean Kelly and Nadine Bothwell; two sons, Paul S. Liscord III and Thomas Liscord; and eight grandchildren.

DANIEL J. LYONS
1905–1997

Daniel J. Lyons died on July 3, 1997. He was 92.

Lyons began his impressive actuarial career after graduating with a mathematics degree from Harvard University in 1926. He worked for three years as an assistant actuary at the Columbian National Life Insurance Company from 1932 to 1934. Lyons then moved to Trenton, New Jersey in 1935 and worked as a chief assistant actuary at the New Jersey Department of Banking and Insurance.

In 1943, Lyons began what was to be a successful 30-year stint with The Guardian Life Insurance Company of American in New York City as an assistant actuary. He was promoted to second vice president in 1949, administrative vice president in 1954, vice president in 1957, and senior vice president in 1960. Lyons served as president from 1964 to 1968, and worked for three more years with Guardian Life as the chairman of the board and chief executive before he retired from the company in 1971.

Lyons remained active in the insurance industry, serving one year as the president of Associated Actuaries Incorporated in Trenton and one year as president at Bankers National Life Insurance Company in Parsippany, New Jersey. Lyons was also a 66-year member of the Casualty Actuarial Society, receiving his Associateship in 1931 and his Fellowship in 1936. He was a member of the CAS Examination Committee in 1989.

Lyons is survived by his wife of 64 years, Irene M. Lyons; two daughters, Jean L. Entwistle of New York City and Irene L. Madden of McLean, Virginia; two sons, Daniel J. Lyons Jr. of Princeton, New Jersey and Paul O. Lyons of Doylestown, Pennsylvania, and his sister, Sister Marion Lyons.

PHILIPP K. STERN
1911–1999

Philipp K. Stern died on April 19, 1999 at his home in Lakehurst, New Jersey.

Born in Paris, France in 1911, Stern moved with his family to Vienna, Austria. Stern attended the University of Vienna, earning a Ph.D. in law before emigrating to the United States in 1939. In 1942 he married his wife Sylvia and together they had three daughters.

An entrepreneur and rating bureau specialist, Philipp Stern became a member of the Casualty Actuarial Society in 1956 when he gained his Associateship. He was a member of the CAS Committee on Automobile Insurance Research (1964) and the author of three papers published in the *Proceedings of the Casualty Actuarial Society*: "Current Rate Making Procedures for Automobile Liability Insurance" (1956); Review of "An Approximation for the Testing of Private Passenger Liability Territorial Rate Levels Using Statewide Distribution of Classification Data" (1964); and "Ratemaking Procedures for Automobile Liability Insurance" (1965).

Throughout the 1950s and 1960s, Stern's professional life was centered in New York City. Stern was an actuary with Mutual Insurance Rating Bureau in New York from 1957 to 1965, and briefly served as actuary-manager for the National Bureau of Casualty Underwriters in 1966. From 1967 to 1969, Stern served as an actuary for the Insurance Rating Board. For most of the 1970s, Stern worked for the New Jersey Department of Insurance in Trenton, where he was an actuary from 1971 to 1976. He later became chief actuary there from 1977 to 1979. With the coming of the new decade, Stern began his own actuarial consulting firm, Philipp K. Stern, Inc., in Newark, Delaware. He retired from his business in 1989.

Stern is survived by his wife Sylvia; daughters Erica Stern of Lakewood, Leda Walker of Morris, New York, and Sheera Stern of Metuchen, New Jersey; and four grandchildren.

INDEX TO VOLUME LXXXVI

INDEX—CONTINUED

Page

INDEX—CONTINUED

INDEX—CONTINUED

INDEX—CONTINUED

Page