CASUALTY ACTUARIAL SOCIETY FORUM

Winter 2000

Including the Ratemaking Discussion Papers and Health and Managed Care Call Papers

CASUALTY ACTUARIAL SOCIETY
ORGANIZED 1914

Printed for the Society by
Colortone Press
Landover, Maryland

NOTICE

The Casualty Actuarial Society *Forum*

Winter 2000 Edition

Including the Ratemaking Papers and Health and Managed Care Call Papers

To CAS Members:

This is the Winter 2000 Edition of the Casualty Actuarial Society *Forum*. It contains eight Ratemaking Discussion Papers, two Health and Managed Care Issues Call Papers, and two additional papers.

The Casualty Actuarial Society *Forum* is a non-refereed journal printed by the Casualty Actuarial Society. The viewpoints published herein do not necessarily reflect those of the Casualty Actuarial Society.

The CAS *Forum* is edited by the CAS Committee for the Casualty Actuarial Society *Forum*. Members of the committee invite all interested persons to submit papers on topics of interest to the actuarial community. Articles need not be written by a member of the CAS, but the paper's content must be relevant to the interests of the CAS membership. Members of the Committee for the Casualty Actuarial Society *Forum* request that the following procedures be followed when submitting an article for publication in the *Forum*:

1. Authors should submit a camera-ready original paper and two copies.
2. Authors should not number their pages.
3. All exhibits, tables, charts, and graphs should be in original format and camera-ready.
4. Authors should avoid using gray-shaded graphs, tables, or exhibits. Text and exhibits should be in solid black and white.
5. Authors should submit an electronic file of their paper using a popular word processing software (e.g., Microsoft Word and WordPerfect) for inclusion on the CAS Web Site.

The CAS *Forum* is printed periodically based on the number of call paper programs and articles submitted. The committee publishes two to four editions during each calendar year.

All comments or questions may be directed to the Committee for the Casualty Actuarial Society *Forum*.

Sincerely,

Dennis L. Lange

Dennis L. Lange, CAS *Forum* Chairperson

The 2000 CAS Ratemaking Discussion Papers and Health and Managed Care Issues Call Papers Presented at the 2000 Ratemaking Seminar March 9-10, 2000 Hotel del Coronado San Diego, California

The Winter 2000 Edition of the CAS *Forum* is a cooperative effort between the CAS *Forum* Committee and two CAS Research and Development Committees: the Committee on Ratemaking and the Committee on Health and Managed Care Issues.

The CAS Committee on Ratemaking presents for discussion eight papers prepared in response to its Call for 2000 Ratemaking Discussion Papers. In addition, the Committee on Health and Managed Care Issues presents eight papers submitted in response to the 2000 Call for Health and Managed Care Issues Papers.

This Forum includes papers that will be discussed by the authors at the 2000 CAS Seminar on Ratemaking, March 9-10, in San Diego, California.

1999 Committee on Ratemaking

Karen F. Terry, *Chairperson*
Charles H. Boucek, *Vice Chairperson*

Shawna S. Ackerman
John P. Alltop
Jean Cloutier
Patrick J. Crowe
Jeffrey W. Davis
Behram M. Dinshaw
Warren H. Johnson
Kenneth R. Kasner
Allan A. Kerin
Joel M. Kleinman
Richard B. Moncher
Kathleen M. Pechan
Joseph Lawrence Petrelli
Barbara H. Thurston
Gregory S. Wilson

Committee on Health and Managed Care Issues

Pamela Sealand Reale, *Chairperson*
Cara M. Blank, *Vice Chairperson*

Theresa W. Bourdon
Brian Z. Brown
James B. Conner
Michael C. Dolan
Bernard R. Horovitz
Alexander Krutov
Dennis L. Lange
John Pierce
Rial R. Simons
Tony King Gwan Tio
Nancy R. Treitel
Robert J. Walling
Thomas A. Weidman
Alfred O. Weller

Table of Contents

Source of Earnings Analysis for Property-Casualty Insurers

Douglas M. Hodes, FSA, MAAA, and
Sholom Feldblum, FCAS, FSA, MAAA

Source of Earnings Analysis for Property-Casualty Insurers

Abstract

Source of earnings analysis has long been a staple of life insurance policy pricing and profitability monitoring. It has grown in importance with the advent of universal life insurance and of similar contracts with non-guaranteed benefits or charges. SFAS 97 now mandates the use of source of earnings analysis for GAAP reporting of universal life-type contracts.

Source of earnings analysis is equally applicable to several lines of property-casualty insurance, such as workers' compensation and personal automobile insurance. An accident of history has restricted it to life insurance. Source of earnings analysis was first developed for allocating policyholder dividends on participating life insurance policies, and it has since been expanded to other policy forms as well. Casualty actuaries have developed their own ratemaking traditions. Casualty actuaries and life actuaries grow up in separate societies with little interaction, and source of earnings analysis has never been extended to the casualty lines of business.

This paper shows the uses of source of earnings analysis for understanding the factors affecting policy profitability. Source of earnings analysis is not a specific ratemaking "method," like the loss ratio method or the pure premium method. Rather, source of earnings analysis is a reporting structure that reveals the sources of gain and loss on a block of business, highlighting errors in the pricing parameters as well as the sensitivity of profit and loss to various pricing factors, and enabling more accurate selection of new parameters and factors.

This paper develops source of earnings exhibits for casualty insurance, using private passenger automobile insurance policies and retrospectively rated workers' compensation policies as examples. The uncertainty in many casualty insurance pricing factors, such as loss development factors and loss trend factors, make source of earnings analysis particularly important for casualty products.

The paper shows how to use the source of earnings exhibits to better analyze the factors driving insurance results. In particular, the paper divides the variance caused by each earnings factor into an estimation error component, which is within the purvey of the pricing actuary, and a random error component, which results from random fluctuations in loss occurrences, inflation rates, or interest rates.

Some sources of gain or loss, such as persistency patterns and investment earnings, are not always included in casualty ratemaking procedures. A complete source of

earnings analysis incorporates (a) an analysis of expected versus actual experience by each pricing factor and (b) an amortization of initial expense and loss costs by policy year. The initial acquisition expense includes the solicitation costs for not taken business, which can be substantial in large account workers' compensation retrospectively rated policies. These costs are not always considered by pricing actuaries, but they have great effect on ultimate profit margins.

Similarly, movements in the achieved interest rate spread is a major factor in life annuity profitability because of the long duration of these policies and the substantial cash accumulation in these policies. Changes in expected versus actual investment income can have a large effect on workers' compensation profitability as well.

Pricing actuaries sometimes say that their indications are best estimates, and they disclaim responsibility for variances of actual from expected. In truth, analysis of the variances from previous years' predictions is one of the best means of improving next year's predictions. Sources of earnings analysis provides the needed *post mortem* to rigorously measure the variances in each source of earnings factor

Source of Earnings Analysis for Property-Casualty Insurers

Section I: Introduction

This paper illustrates source of earnings analysis for property-casualty insurance. Source of earnings analysis is a staple of life insurance policy pricing and reporting, and it is mandated by NAIC regulations or GAAP statements for participating policies issued by mutual life insurance companies, for universal life policies, and for other policies with non-guaranteed benefits or charges.

We discuss source of earnings analysis for private passenger automobile insurance and workers' compensation insurance ratemaking. Because of the complexity of this topic, we focus on issues that are most germane to pricing actuaries for these two lines. Private passenger automobile insurance ratemaking is well suited to source of earnings analysis, since the volume of business is large and the effects of estimation error and random error can be more easily discerned. In addition, private passenger automobile has high retention rates and different acquisition expense costs for new policies versus renewal policies, making profitability highly sensitive to persistency patterns.

Workers' compensation retrospectively rated policies are somewhat analogous to universal life insurance contracts. In both cases, expected profits stem from the margins in the pricing assumptions. The casualty actuary prices the components of the retrospectively rated policy, such as the insurance charge, even as the life actuary prices the components of the universal life policy. In addition, large account retrospectively rated policies have high not taken rates, various premium payment plans, and considerable investment income, which require actuarial expertise for pricing and design. Measuring profitability by comparison of total premiums with total costs may yield little information that can improve the pricing process. Source of earnings analysis is better suited to identifying the causes of superior and inferior performance.

Structure of this Paper

Section II provides a description of source of earnings analysis as applied to life insurance products, with specific reference to two areas: (i) calculation of policyholder dividends by means of the contribution principle for mutual life insurance companies and (ii) the FAS 97 accounting for universal life-type products.

Section III applies source of earnings analysis to private passenger automobile ratemaking. The general framework is outlined, along with a detailed analysis of trend.

The major themes of this section are the differentiation between estimation error and process error; the handling of credibility; the difference between implicit profit margins and explicit profit margins; and source of earnings analysis for investment income.

Section IV applies source of earnings analysis to workers' compensation ratemaking for retrospectively rated contracts. The major themes of this section are static versus dynamic amortization of deferred policy acquisition costs; the source of earnings exhibits showing charged, expected, and actual results; and interpreting the source of earnings exhibits.

Section V summarizes two fundamental implications of the paper regarding pricing paradigms and the effects of random variations.

Section II: Classical Source of Earnings Analysis

Source of earnings analysis was originally used to determine policyholder dividends for permanent life insurance policies sold by mutual insurance companies. With the advent of interest sensitive policies, source of earnings analysis has been mandated for GAAP financial statements: by SFAS 97 for universal life-type contracts and SFAS 120 for participating policies sold by mutual life insurance companies.

Policyholder Dividends: The contribution principle, which is required both by the NAIC model act on policyholder dividends and by the AAA Standards of Practice, mandates that the amount of divisible surplus used to pay policyholder dividends on any block of business reflect the contribution of that block to company earnings.[1] Although simple and elegant, this principle is difficult to apply rigorously, since it requires the actuary to quantify the long-term contributions to profit from calendar year changes in the pricing assumptions.

Persistency Rates: The major elements affecting long-term profitability are persistency rates (or withdrawal rates), interest earnings, and mortality ratios. Let us consider each of these, since they are all applicable to property-casualty business as well. Suppose the expected withdrawal rates are 10% for the second year of a cohort of permanent life insurance policies, but the actual withdrawal rates are 15%. The payment of surrender charges and the takedown of conservative statutory reserves cause an increase in statutory profits in the second year. However, the smaller block of persisting business generally leads to lower profits in succeeding years, which more than offsets the statutory gain in the second year. Policyholder dividends for that block of business must be reduced. Source of earnings analysis helps quantify the equitable change in the dividend rate.

---

[1] See particularly Actuarial Standard of Practice #15, *Dividend Determination and Illustration for Participating Individual Life Insurance Policies and Annuity Contracts*, and Actuarial Standard of Practice #24, *Compliance with the NAIC Life Insurance Illustrations Model Regulation*.

For casualty products priced by traditional ratemaking procedures, we use a simpler adjustment in this paper for persistency changes, though the effects on product profitability are great. Solicitation costs on not taken business, as well as high first year acquisition expenses, are amortized over the expected policy lifetimes.[2] If withdrawal rates increase, the amortization period is reduced and profitability declines.

Interest Earnings: Continuing the previous example, suppose that the expected Treasury bill rate was 6% per annum in the second year of a cohort of permanent life policies but the actual Treasury bill rate was 5% per annum. The change in statutory investment earnings during this year may have been nil, since (i) the coupons on existing bonds have not changed, (ii) bonds are valued at amortized cost in statutory statements, and (iii) invested assets are small in the second year of a cohort of permanent policies. The anticipated change in long-term profitability may range widely, depending on the inflation sensitivity and the duration of the liabilities. For a standard guaranteed cost block of traditional whole life business, the expected long-term profitability would drop, necessitating a decrease in policyholder dividends.

The effects of changing interest rates are more complex for casualty products, since both loss payments and asset returns are sensitive to changes in interest rates and inflation rates. A full source of earnings exhibit, showing the effects of variation in loss cost trends side-by-side with the effects of variation in the investment yield is necessary to judge the net effect on product profitability.

Mortality: Variations in mortality ratios highlight the importance of estimation error versus process error. Suppose that the ratio of actual to expected mortality in the second year of a cohort of business is 150%. If the higher than expected mortality reflects random deaths, policyholder dividends paid to the remaining insureds should not be changed. If the higher than expected mortality reflects a poor quality book of business, the policyholder dividends should be reduced.

For casualty lines of business, loss frequency and loss severity are similar to mortality rates in life insurance or morbidity rates in health insurance. Higher than expected loss frequency or loss severity in any calendar year may reflect either random loss occurrences or estimation error of the expected means. The latter possibility necessitates re-examination of the ratemaking procedure.

---

[2] "Not taken" business is business that is underwritten and where an insurance offer is make but not accepted. The importance of "not taken" business for determining fixed expense provisions by classification is discussed in S. Feldblum, "Personal Automobile Premiums: An Asset Share Pricing Approach for Property/Casualty Insurance," *Proceedings of the Casualty Actuarial Society*, Volume 83 (1996), pages 190-296.). The "Personal Auto Premiums" paper shows the assumptions regarding "not taken" business needed to price the policy. This paper shows the methods to test for variance of actual results from the pricing assumptions.

Amortization of the DPAC

GAAP treatment of deferred policy acquisition costs necessitated a wide application of source of earnings exhibits in GAAP statements for universal life-type policies.[3] In statutory statements, deferred policy acquisition costs are expensed when incurred. In GAAP statements for traditional policies, deferred policy acquisition costs are expensed as the premium is earned. For universal life-type policies, there is no set premium, so one can not amortize the DPAC asset in relation to premiums. Instead, SFAS 97 mandates that the DPAC asset be amortized as a proportion of future expected profits.[4]

To illustrate the use of source of earnings analysis in FAS 97 accounting, consider again the example with an unexpected increase in the withdrawal rate from 10% to 15% in the second year of a cohort of policies. If this cohort consists of universal life-type policies, the DPAC asset would be amortized in relation to future expected profits. Suppose that originally the second year profits were expected to be 10% of all future profits. After the withdrawal rate increase, the actual second year profits increase and the future expected profits decrease, leading to a higher ratio and a larger amount of deferred policy acquisition costs amortized in the second year.[5]

Extension to Casualty Products

Source of earnings analysis is applicable to any insurance product whose returns depend on conditions subsequent to the policy pricing. This is true of all property-casualty products, since their returns depend on random loss occurrences as well as on interest rates and inflation rates.

The profitability of private passenger automobile business also depends (in part) on the persistency of the business, particularly for direct writing insurers.[6] The original pricing of products whose profitability depends on persistency can be done by asset share models. The subsequent monitoring of product performance requires dynamic amortization of the deferred policy acquisition costs and is best accomplished by multi-

---

[3] The unfortunate term "universal life-type" is not an actuarial malapropism; it is the standard GAAP term for policies with benefits or charges that are not fixed.

[4] The amortization of deferred policy acquisition costs in relation to expected profits rather than in relation to premium earnings makes sense for all policies. However, the AICPA did not wish to change accounting practice for existing policies, so the new rules apply only to universal life-type policies.

[5] This is an oversimplified treatment of FAS 97. For a thorough analysis, along with illustrations of the source of earnings exhibits, see Joseph H. Tan, "Source of Earnings Analysis under FAS 97 Universal Life Accounting", *TSA* XLI (1989), pp. 443-506, and Michael Eckman, "Additional Source of Earnings Analysis under FAS 97," *TSA*, Volume LXII (1990), pages 59-81.

[6] See Stephen P. D'Arcy and Neil A. Doherty, "The Aging Phenomenon and Insurance Prices," *PCAS*, Volume 76 (1989), pages 24-44, and S. Feldblum, "Personal Automobile Premiums: Asset-Share Pricing for Property-Casualty Insurers," PCAS (1996), pages 190-276.

year source of earnings exhibits.

Asset share pricing for casualty products is complex, and it is not the intention of this paper to review that topic. Instead, we examine the dynamic amortization of solicitation costs for not taken business in retrospectively rated workers' compensation policies.

Workers' compensation retrospectively rated policies often have premiums that are based on the total exposure but provide insurance coverage only for certain portions of the risk. The cost of the coverage is based on an insurance charge calculation that considers premium bounds, loss limits, the size of the insured, and the class of the insured. The profitability of the book of business depends on implicit margins in the insurance charge and on the investment income from the underwriting cash flows. Source of earnings analysis allows the actuary to monitor the performance of the business in terms of the pricing assumptions.

These two illustrations show the power of source of earnings exhibits to deal with sources of gains and losses that are not adequately reflected in traditional ratemaking and profitability monitoring procedures. However, the primary benefits of source of earnings analysis are applicable to all products. Source of earnings analysis serves as a *post-mortem* of previous reviews, evaluating the accuracy of the various assumptions, and uncovering the causes of poor performance.

Section III: Private Passenger Automobile

The structure of the source of earnings analysis depends on the factors affecting the rates for each line of business. Most life insurance products use a four factor analysis, focusing on withdrawal rates, mortality ratios, interest rates, and expense ratios. For property-casualty products, mortality ratios are replaced by loss assumptions, such as loss development, loss trend, loss frequency, and loss severity assumptions.

There are three levels of the source of earnings analysis.

- The individual factor level shows the application of source of earnings analysis to each earnings factor. For private passenger automobile, we examine loss severity trends in this paper, differentiating between estimation error and process error. For workers' compensation, we examine several earnings factors: non-ratable losses, acquisition costs, and interest earnings.

- The source of earnings exhibits for a single policy or a single policy year combine the earnings factors but without consideration of policy renewals (retention rates). These exhibits are appropriate for blocks of business with low persistency rates, little difference between first year and renewal year loss and expense costs, and low solicitation costs for not taken business. Many independent agency companies ignore persistency rates in their pricing analyses.

- The source of earnings exhibits for a cohort of policies considers both the current policy year and the renewals of existing policies. These are the standard exhibits required for universal life-type policies and for participating policies issued by mutual life insurance companies.

Maintenance expenses are not discussed in this paper. Maintenance expenses are generally stable, and they are more easily analyzed by direct examination than by source of earnings exhibits.

Individual Factor Level: Estimation Error and Process Error

We illustrate the workings of source of earnings analysis with the loss cost trend adjustments used in most casualty rate reviews. For private passenger automobile, which does not use an inflation sensitive exposure base, the trend assumptions are particularly critical for rate adequacy.[7]

Actual insurance results frequently differ from expected results. The source of earnings analysis attributes this variance to the underlying earnings factors (or "sources"). For each factor, there are two potential reasons for the variance: estimation error and process error. Estimation error is the difference between the actuary's forecast and the true expected result. Process error is the difference between the true expected result and the actual realization. These errors emerge over time, starting from the date of the rate review to the final settlement of claims. Estimation error can often be controlled by the pricing actuary, whereas process error is an unavoidable element of actuarial ratemaking.[8]

For the private passenger automobile trend illustration we assume an experience period

[7] For space limitations, we begin *in medias res*, and we continue rapidly through lines of business and ratemaking procedures of various hues: casualty, life, and financial. In practice, source of earnings analysis is overlaid on the ratemaking method. Ideally, we would use a more structured exposition, beginning with the pricing procedure for each line and working through the implicit assumptions and sources of variance in each section of the rate review.

[8] Actuaries often speak of parameter risk versus process risk (see Robert S. Miccolis, "On the Theory of Increased Limits and Excess of Loss Pricing," *Proceedings of the Casualty Actuarial Society*, Volume 64 (1977) pages 27ff and S. Feldblum, "Risk Loads for Insurers," *Proceedings of the Casualty Actuarial Society*, Volume 77 (1990), pages 160-195). These are similar though not identical concepts. For instance, process variance often causes the historical data to imperfectly reflect expected experience. This process variance causes the actuary to misestimate the parameters of the loss distribution, resulting in parameter risk for prospective ratemaking. However, the estimation error for severity trend in this paper *is not affected by the historical experience. The process variance leading to a misestimation of the* parameters of the loss distribution remains process variance in this paper.

Some actuaries categorize risk more finely into process risk, parameter risk, and specification risk. This division is most common in discussions of dynamic financial analysis; see, for instance, Gerald S. Kirschner and William C. Scheel, "The Mechanics of a Stochastic Corporate Financial Model," *PCAS* Vol 85 (1998), pages 404-454, or Hodes, Feldblum, and Blumsohn, "Workers' Compensation Reserve Uncertainty," *PCAS* Vol 86 (1999). Both parameter risk and specification risk would be included in estimation error.

of accident year 1999, and a future period of annual policies written in 2001; the loss trend period is 2.5 years (7/1/99 to 1/1/2002). In the rate review, the trend factors estimated from countrywide fast track data are a +7% per annum severity trend and a +1% per annum frequency trend.[9]

There are three sources of potential error.

1. predicting future fast track trends based on historical experience.
2. applying countrywide fast track figures to a particular state.
3. using loss trend estimates to predict the changes in actual losses incurred.

Several months after the policy year has expired, the source of earnings analysis shows that the actual fast track trends were +8% per annum for severity and +2% per annum for frequency. Our fast track estimates, which we used as a proxy for the actual loss trends, were too low. This is estimation error. The quantification of estimation error is independent of the actual frequency and severity changes in the statewide data.

Discrepancies between countrywide trends and statewide trends are not easily discerned. When there is no change in state compensation systems or other structural characteristics, no difference is normally expected. When there is a change in compensation systems or in other structural characteristics (such as the degree of attorney involvement in insurance claims), trend differences can be significant. To simplify the presentation in this paper, we do not analyze countrywide-statewide differences in expected trend.[10]

We examine the average loss severities and loss frequencies in the experience period and in the new policy period. Our initial numbers are estimates, since (I) the figures for the new policy year are immature and (ii) even for the experience period the loss severities may still be uncertain. We won't have actual loss severity and loss frequency figures for the new policy period until all the policies (not just the policy year) have expired and their data have been collected. For the first source of earnings exhibit, we use some actual data and some revised estimates. For the second and subsequent source of earnings exhibits, we have more complete actual data.

---

[9] Numerous data sources are available for trend estimates. We assume that the pricing actuary uses countrywide fast track data for estimating trend factor, since this allows a clear demarcation between estimation error and process error. In theory, the same two sources of error exist when one extrapolates future trend factors from historical statewide experience, though it is harder to separate the two.

[10] The 1991 compensation system changes in Massachusetts showed the effect of structural changes on expected loss frequency and loss severity; see Sarah S. Marter and Herbert I. Weisberg, "Medical Expenses and the Massachusetts Automobile Tort Reform Law: First Review of 1989 Bodily Injury Liability Claims," *Journal of Insurance Regulation*, Volume 10, No. 4 (Summer 1992), pages 462-514. On the importance of these regional differences as private passenger automobile "cost drivers," see John B. Conners and Sholom Feldblum, "Personal Automobile: Cost Drivers, Pricing, and Public Policy," *PCAS* Vol 85 (1998), pages 370-403.

Suppose that our new loss severity and loss frequency figures indicate a trend of +5% per annum for severity and +4% per annum for frequency, as shown in Table 1.

Table 1: Estimation Error and Process Error

	Estimated Fast Track	*Actual Fast Track*	*Actual S/W Change*
Loss Severity	+7%	+8%	+5%
Loss Frequency	+1%	+2%	+4%

We under-estimated loss severity by 1% (+7% → +8%), and we underestimated loss frequency by 1% (+1% → +2%). For a 2.5 year trend period, this caused the rates to be inadequate by 4.9% [$((1.08*1.02)/(1.07*1.01))^{2.5}$]. This is the estimation error.

The actual loss severity change was +5% per annum, and the actual loss frequency change was +4% per annum. We do not call this the actual trend, since it may be influenced by random losses — either in the new policy period or in the experience period — or it may be a systematic change in loss filing patterns. Lacking other information, we presume that the actual severity trend is +8% per annum, and the actual frequency trend is +2% per annum. There may have been some unusually large claims in the experience period or a lack of large claims in the new policy period, thereby accounting for the low severity trend. Similar random effects may account for the large change in claim frequency.

If compensation system changes and structural changes are not explicitly considered, they are subsumed under the process risk component of the source of earnings exhibits. For instance, there may have been an influx of nuisance claims in the new policy period which are settled for small amounts. The phenomenon has plagued private passenger automobile insurance for the past twenty years, and it must always be considered when the frequency change is large and the severity change is small.

We group all the possible explanations of the difference between the observed patterns in the state and the "hindsight" trend observed in the fast track data as the process error in the trend estimate. This term is not entirely accurate, since not all of the causes of the observed difference are necessarily a result of process error. The intention is simply that this observed difference is not a result of misestimation of the fast track trend.

As the new policy year develops, and as actual data replaces estimates, the observed loss trends may change. The changes can be substantial until the new policy year is fully earned, followed by minor changes as losses are settled. For a single policy, or a single policy year, the primary value of the source of earnings exhibits lies in the first few years. For analyzing a cohort of business whose profitability depends (in part) on persistency of existing business, the year-by-year source of earnings exhibits are critical.

Extending the Exhibits

To analyze the sensitivity of the profits to trend errors, we convert the estimation and process errors into dollar amounts. Assuming $10 million of annual losses and using the figures above, we begin the source of earnings exhibits:[11]

Table 2: Private Passenger Auto Loss Severity (one year)

Date	Projection		Estimation Error		Process Error		Total
12/02	+7%	$0	+8%	-$250K	+5%	+$750K	+$500K

The figures are simplified for ease of presentation. We assume a 2.5 year trend, so a 1% understatement in the trend causes a loss of $250,000 on a $10 million book of losses, as is true for the estimation error in this illustration.[12] Sometimes estimation error is unavoidable; sometimes estimation error results from poor work and can be mitigated by better pricing techniques. The conscientious actuary examines past estimation errors to check for any biases in the ratemaking procedures.

Process error derives from the uncertainties of worldly activity. The presence of process error is the justification of insurance coverage. Nevertheless, the analysis of process error is critical for two purposes.

1. First, it is critical to the analysis of profitability. The management of an insurance company must know whether variance from expected results was predictable or random. Continuing random variances from expected results may indicate that the line of business is highly unstable. Repeated variances in a particular direction indicate possible biases in the ratemaking or underwriting operations.

2. Second, careful analysis of the process error may indicate that certain structural factors are impinging on the insurance environment. Changes in compensation systems and changes in attorney involvement in insurance claims are examples of such structural factors. This analysis is particularly important when state compensation systems change.

Source of earnings exhibits use a multi-year format, particularly when persistency patterns are included. Suppose that by 12/31/2003, the actual severity increase is

[11] We assume that the trend factors in the private passenger automobile rate filing contain no implicit profit margin; see the discussion below in the text.

[12] For clarity's sake, we use rough numbers. "Book of losses" is not a realistic concept, since the size of the losses depends on the trend factors. In practice, the "gain or loss" is the difference in profits under the two trend assumptions. (When examining variances in persistency rates for a cohort of business, the arithmetic is complex, but it is not conceptually difficult.) As discussed further below in the text, we use nominal losses for the trend figures, and we separately quantify the gain or loss from investment earnings. When an increase in trend stems from higher inflation that is associated with higher interest rates, the loss from trend may be offset in part by a gain from interest; see the discussion below in the text.

+6%. A second line would be added to the severity trend source of earnings exhibit:

Table 3: Private Passenger Auto Loss Severity (multiple years)

Date	Projection		Estimation Error		Process Error		Total
12/02	+7%	$0	+8%	-$250K	+5%	+$750K	+$500K
12/03	+7%	$0	+8%	-$250K	+6%	+$500K	+$250K

The dollar variances in each column of the table above relates to the current row, not to the entry in the previous row of the same column. The projection column is the original pricing assumption. Since no implicit profit margin is used in the trend assumption, the original "gain or loss" is $0. The projection columns do not change as additional years are added to the table.

The estimation error columns show the difference between the actual trend rate, based on actual fast track data, from the projected trend rate. The "gain or loss" reflects the variance between the actual trend rate and the projected trend rate translated into dollars of gain or loss in the book of business. In this example, the actual fast track trend is 1% per annum greater than the projected trend rate. For a trend period of 2.5 years and a $10 million book of losses, there is a loss of $250,000 from estimation error.

To keep the exposition simple, the actual fast track trend does not change from December 2002 to December 2003 in the example above. December 2002 and December 2003 are the estimation dates; the fast track trend in each row refers to the same period (July 1, 1999 to January 1, 2003). When the first row of the source of earnings exhibit is completed before final fast track data is available (as is true in this example), the estimation error entries often change between the first and second rows.

The process error columns show the difference from the trend rate as indicated by the fast track data and the actual severity change in the company's ratemaking data for that state. Assuming that there are no structural changes that affect loss severity trends in this state, the difference stems either from random loss occurrences in the historical experience period or from random loss occurrences in the policy period. The average severity in both the historical experience period and the policy period may change as the losses mature, so the variance resulting from process error changes as years are added to the source of earnings exhibit.

Revisions stems from both actual (past) data and revised estimates of the future. Consider the first row in Table 3 above. The "projection" column shows the estimated trend for 7/1/99 through 1/1/2003 at the time of the rate analysis. Some of the fast track trend is actual: if the rate analysis is done in the middle of 2000, then the fast track trend for 7/1/99 through 12/31/99 may be actual and the remaining trend is an estimate.

The "estimation error" column shows the expected trend for this same period at a valuation date of December 31, 2002. Most of the trend is now actual data (7/1/99 through perhaps 6/30/2002) while the trend for 7/1/2002 through 12/31/2002 is a revised estimate.

The source of earnings exhibits trace the replacement of prior assumptions by actual data and by revised assumptions. There is no need to wait until "hard data" come in to form the source of earnings exhibits. For instance, if the actual fast track trend is higher than the assumption for the first half of the trend period, we expect that it will be higher than the assumption for the second half of the trend period as well.

Credibility

Credibility is an important component of casualty actuarial pricing procedures. Life insurance pricing does not use credibility adjustments, thereby facilitating a comparison of expected values with actual values. Source of earnings exhibits are more complex when credibility is used in the ratemaking process.

For other components of the pricing procedure, the actual values are known with hindsight after the policy term expires and the experience is mature. For credibility, there is no actual value. The source of earnings analysis does not compare the initial credibility assumption with a subsequent (revised) value. Rather, the credibility value is used to adjust the initial assumptions.

Credibility is used in a variety of places in ratemaking. For illustration, we focus on statewide credibility factors. The credibility factors adjust the past experience to be a better proxy for the true expected losses in the historical experience period.[13]

A numerical illustration should make this clear. For ease of exposition, we use a pure premium ratemaking framework. Suppose that the underlying pure premium during the experience period of accident year 1999 was $500 per car, based on a rate filing effective July 1, 1998, and intended to be in effect for one year. The new policy period is policy year 2001; that is, the anticipated effective date of the current filing is January

[13] Traditionally, statewide credibility factors are applied to the developed and trended experience loss ratios. This might give the impression that credibility is adjusting the development or trend factors or the future expected values. This is not correct. Credibility factors may indeed be applied to trend factors and development factors, and there are several actuarial papers on this subject. The statewide credibility factors, however, adjust the actual experience to be a better proxy of the expected experience in the past.

The discussion here is based on the "greatest accuracy" justification for credibility, not the "limited fluctuation" justification; see Gary G. Venter, "Credibility," in Matthew Rodermund, et al., *Foundations of Casualty Actuarial Science*, Second Edition (New York: Casualty Actuarial Society, 1992), pages 375-483. Venter correctly notes that the theoretical justification for classical credibility is to limit fluctuations in the rates and that the Bayesian-Buhlmann credibility procedure is designed to optimize rate accuracy. Nevertheless, most actuaries conceive of all credibility procedures as improving the accuracy of the rates. In addition, Mahler convincingly argues that even traditional credibility procedures generally improve expected rate accuracy.

1, 2001. Because of administrative problems, no rate changes were effective between July 1, 1998, and January 1, 2001.

The pure premium trend is 10% per annum. The experience pure premium during accident year 1999 is $600. The credibility for the experience pure premium is 50%; that is, the pure premium used in the ratemaking formula is an equal weighting of the trended experience pure premium and the trended underlying pure premium. How should the source of earnings exhibits reflect the 50% credibility factor?

The trend factor is the same for the experience pure premium and the underlying pure premium. The credibility factor tells us that the true expected loss per exposure during accident year 1999 is a 50:50 average of the information we obtain from the accident year 1999 experience and the rates underlying the accident year 1999 writings.

The rates underlying the accident year 1999 writings are $\$500(1.10)^{0.5} = \$524.40$, since the $500 rates were adequate for the 12 month period from July 1, 1998, through June 30, 1999. The credibility weighted average experience rates are ($600 + $524.40)/2 = $562.20.[14]

On the source of earnings exhibits, this is reflected in the actual loss cost change. The initial trend rate assumption is 10% per annum. The actual trend rate based on hindsight is whatever the trend index reveals after the end of the policy year. The actual loss cost change is the change between $562.20 and the observed pure premium during the policy year.

In sum, the source of earnings analysis accepts the credibility adjustment and tests the loss cost change; it does not test the credibility value itself.[15]

Implicit and Explicit Profit Margins

Actuaries have used both implicit and explicit methods for incorporating profit provisions in the premium rates. For explicit profit margins, best estimate assumptions are used throughout the ratemaking process and a full profit margin is included in the rates. For implicit profit margins, conservative assumptions are used in the ratemaking process and a lower explicit profit margin is included in the rates.

To illustrate the difference, contrast trend factors with discount factors.

[14] For a more complete discussion, see S. Feldblum, Discussion of "The Complement of Credibility" by Joseph Boor, *PCAS*, Volume 85 (1998), pages 991-1033.

[15] This is not to imply that credibility factors are impervious to empirical testing. On the contrary: Mahler, Howard C. Mahler, "An Example of Credibility and Shifting Risk Parameters," *Proceedings of the Casualty Actuarial Society*, Volume 77 (1990), pages 225-308.give three methods for testing the accuracy of credibility factors. However, Mahler tests the accuracy of the credibility estimator. One can not test the accuracy of a particular credibility factor. That is, there is no variance between the actual credibility and the assumed credibility.

- *Trend Factors:* Suppose that fast track data imply a loss severity trend of +5% per annum. This estimate is uncertain, not only because it is a future projection but also because the fast track data are not completely comparable to the ratemaking data (different companies, different states, accident year versus calendar year, closed claims versus incurred claims, and so forth). We presume that the trend rate is probably between 4% and 7% per annum.

  The explicit profit method would use a +5% trend and a full explicit profit margin. The implicit profit method may use a +6% trend and a somewhat lower profit margin. Many pricing actuaries prefer best estimate assumptions for most factors and explicit profit margins in the rates, although rate filing exigencies often compel them to use lower explicit profit margins offset by conservative assumptions.

- *Discount Factors:* Suppose that losses are discounted to present value at the expected risk-free interest rate in a discounted cash flow pricing model. The estimate of future interest rates, based on an analysis of the current yield curve and of any mean-reverting tendencies in the assumed interest rate paths, is 5% per annum. This estimate is uncertain, because we are projecting a future rate and because our interest rate model may itself be flawed. We presume that the future interest rate will probably be between 4% and 6% per annum.

  The explicit method would use a 5% assumed interest rate with a full explicit profit margin. The implicit method might use a 4% assumed interest rate with a somewhat lower profit margin.[16]

  Similarly, suppose that expected investment yields will average 8% over the future policy period. For running an internal rate of return pricing model or a return on capital pricing model, the actuary may choose a more conservative investment yield and a lower cost of capital.[17]

---

[16]The use of an implicit profit margin in the interest rate is not the same as a risk adjustment to the discount rate. For example, Myers and Cohn [1987, op cit] use a CAPM-based risk-adjusted loss discount rate that reflects the covariance of loss returns with market returns, following procedures used by W. Fairley, "Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results," *The Bell Journal of Economics* 10 (Spring 1979) pages 192-210, and R. Hill, "Profit Regulation in Property-Liability Insurance," *The Bell Journal of Economics*, Vol 10, No. 1 (Spring 1979) pages 172-191. The CAPM-based risk adjustment reflects the true present value of the loss payments, not "conservatism" or an implicit profit margin. See also Butsic [1988, op cit], who uses a risk adjustment to the loss discount rate to estimate the true economic value of the loss reserves.

[17] Some actuaries prefer the use of explicit profit margins to better monitor the adequacy of the rates; other actuaries prefer the use of implicit profit margins to prevent overly aggressive pricing. Rate filing requirements in many states influence the type of profit margin. A state that limits the explicit profit margin to an inadequate return on capital may cause insurers to load implicit profit margins into the pricing factors.

Investment Income

The expected investment income earned on policyholder supplied funds and on the surplus funds supporting the book of business is an essential element in pricing insurance products. For interest sensitive products, life actuaries speak explicitly of the spread between the earned interest rate and the credited interest rate. The source of earnings analysis considers the difference between the spread that is actually achieved and the spread that is assumed in the pricing analysis.

Casualty actuaries use a variety of financial pricing models. Often, a financial model is used to determine a target underwriting profit margin, which is then used in the ratemaking procedure. The analyst using the prescribed underwriting profit margin in the rate review may not have participated in determining the adequacy of that margin and may not even be aware of the interest rate assumptions embedded in that margin.

This complicates the source of earnings analysis, but it does not diminish its importance. Indeed, the source of earnings analysis is all the more necessary, since it reveals the additional gain or loss resulting from actual investment earnings being higher or lower than expected.

For the source of earnings exhibits, we need three figures:

⇒ the investment yield assumed by the pricing actuary during the future pricing period (this is the assumed earned interest rate), or IY_0.
⇒ the credited interest rate (CR), or the investment yield used in the pricing model.
⇒ the actual investment yield achieved during the period that reserves are held by the company, or IY_t. The actual investment yield includes dividends, interest, and rents, as well as capital gains and losses. The cleanest way to format the source of earnings exhibits is to use market yields and to include both realized and unrealized capital gains and losses (see below).

The interest spread is most important for the long-tailed commercial liability lines of business, such as workers' compensation and general liability. We must estimate the invested funds for the block of business at each point in time (IF_t).[18] Most casualty pricing models estimate the amount of invested funds by projecting premium collection patterns, loss payment patterns, and expense payment patterns.[19]

[18] Life actuaries use the term "account balance" instead of invested funds. In life insurance and annuities, the funds paid by the policyholder belong to the policyholder and may be withdrawn on demand, sometimes with a surrender charge deducted. In casualty products, the funds paid by the policyholder do not legally belong to the policyholder. We use the term invested funds (instead of account balance) to refer to the financial assets used to fund the unearned premium reserves and the loss reserves.
[19] See Ira Robbin. "The Underwriting Profit Provision," CAS Examination Study Note, 1992, pricing algorithm 7; S. Feldblum, "Pricing Insurance Policies: The Internal Rate of Return Model," Second Edition (Casualty Actuarial Society Part 10A Examination Study Note, May 1992); and Howard C. Mahler, "The Myers-Cohn Profit Model: A Practical Application," *PCAS* Vol 85 (1998), pages 689-774.

The source of earnings analysis quantifies the implicit profit margin in the investment yield assumptions and the subsequent unfolding of the actual profit margin. *Each year's* implicit profit margin in the interest rate assumption equals

> *the invested funds times the difference between the future expected investment yield and the investment yield used in pricing, or* $IF_t(IY_0\text{-}CR)$.

The implicit profit margin in the investment yield assumption is the discounted summation of the annual profit margins.[20]

For example, suppose we are performing a source of earnings analysis on a $10 million cohort of workers' compensation business, with average invested funds of $3 million during the policy year, $4 million the next year, and reducing by $1 million a year until all losses are settled.[21] The company expects an investment yield of 8% per annum, and it prices the business assuming an investment yield of 7% per annum, along with a 12% cost of capital to price the business.[22] The implicit profit margin in the investment yield assumption is shown below. The present values are taken to the middle of the initial policy year (year 0).

Table 4: Source of Earnings Analysis for Interest Spread at Policy Inception

Year	Invested Funds	Expected Invest Yield	Credited Interest Rate	Interest Rate Spread	Interest Rate Margin	PV of Margin
0	3,000,000	8%	7%	0.01%	30,000	30,000.00
1	4,000,000	8%	7%	0.01%	40,000	35,714.29
2	3,000,000	8%	7%	0.01%	30,000	23,915.82
3	2,000,000	8%	7%	0.01%	20,000	14,235.60
4	1,000,000	8%	7%	0.01%	10,000	6,355.18
Total						110,220.89

Between initial policy pricing and final settlement of claims, several items may change.

---

[20] This formula makes the simplifying assumption that IY_0 is the pricing assumption for all future years; that is, the actuary assumes a constant future investment yield.

[21] This progression of the invested funds reflects a policy year exhibit of casualty insurance contracts. With a pre-paid acquisition expense ratio of 20%, a net premium of $8 million paid in up-front on some policies and with premium payment plans on others, and some losses paid out during the first policy year, the average invested funds are about $3 million. The invested funds generally peak about 12 months after inception of the policy year. During the 12 months following the policy year, the remaining premium is paid in and then the invested funds decline to zero as losses are settled.

[22] The cost of capital is the target return on capital. To keep the arithmetic simple, we ignore federal income taxes in this paper. In practice, one can not run source of earnings exhibits for interest earnings without consideration of federal income taxes, since different investments have different tax rates. For prospective policy pricing, one can avoid this problem by using equivalent risk-free portfolios; see Myers and Cohn [1987, op cit.]. The process error in the source of earnings analysis focuses on the defaults and market value changes of risky investments; the assumption of equivalent risk-free portfolios misses much of the analysis.

1. The actual investment yield may differ from the original assumption, since interest rates shift from year to year.
2. The amount of invested funds may differ from the initial assumption.
3. There may be unexpected capital gains or losses.

The new entries in the source of earnings exhibits are a mix of actual figures and revised estimates. For instance, suppose that investment yields rise to 10% per annum during the initial policy year. Year 0 may show 9.5% as the actual average investment yield, and years 1 through 4 may show 10% as the revised estimated investment yield. Based on the actual premium payment plans taken during the policy year, the estimated invested funds may change for all years. Capital gains and losses generally reflect the actual market value changes of securities.

Table 5: Source of Earnings Analysis for Interest Spread after One Year

	Invested	Investment	Credited	Interest	Interest	Capital	PV of
Year	Funds	Yield	Interest	Spread	Margin	Gain/Loss	Margin
0	$2,500,000	9.5%	7%	2.5%	$62,500	-$50,000	$12,500.00
1	$3,500,000	10%	7%	3.0%	$105,000	$0	$93,750.00
2	$3,000,000	10%	7%	3.0%	$90,000	$0	$71,747.45
3	$2,000,000	10%	7%	3.0%	$60,000	$0	$42,706.81
4	$1,000,000	10%	7%	3.0%	$30,000	$0	$19,065.54
Total							$239,769.81

The source of earnings exhibit at the end of the first policy year provides the profitability information that is critical for proper performance measurement and pricing decisions. Investment yields have increased from 8% from the rate review date to 10% by the end of the policy year. Since most of the increase occurred before assets were bought, the capital loss is small but future increases in coupon payments are large.

Inflation Rates and Interest Rates

The full effects of interest rate changes require a combined analysis of assets and liabilities.[23] If inflation rates rise concomitant with the interest rate rise, the loss severity factor will show corresponding increases. The actual loss ratio for the block of business will exceed the target loss ratio, but this loss will be offset by the rise in the investment yield.[24] Traditional profitability measures of loss ratios and combined ratios can be

[23] For traditional source of earnings analysis applied to life insurance policies and annuity contracts, this statement is obvious. It is only on the property-casualty side that pricing actuaries focus on the liability side. This practice may be appropriate for prospective rating methods that use static procedures. It is misleading for source of earnings exhibits for long-tailed lines of business, whose very purpose is to monitor variances from the pricing assumptions.

[24] Cf Robert P. Butsic, "The Effect of Inflation on Losses and Premiums for Property-Liability Insurers," in *Inflation Implications for Property-Casualty Insurance* (Casualty Actuarial Society 1981 Discussion Paper

misleading. Even the statutory measures of total profitability, such as the investment income allocation procedure in the NAIC's Insurance Expense Exhibit, use portfolio investment yields and may be distorted.

Generally, inflation rates and interest rates do not move in lock-step. The source of earnings exhibits provide a year-by-year analysis of the relative gains and losses from inflation and interest, allowing clearer analysis of the product's contribution to the company's performance.

For instance, if there is a general rise in interest rates and inflation rates (not necessarily equal), the actual loss severity change will be larger than expected, leading to a negative profit variance. The actual investment income will be larger than expected as well, leading to a positive profit variance. The net profit variance shows the combined effects of the changes in the interest and inflation rates. This is particularly important for retrospectively rated workers' compensation policies, since inflation has a leveraged effect on non-ratable losses.

Persistency

For the source of earnings analyses required by SFAS 97 for universal life-type policies, persistency is often the most important earnings factor. This makes sense. The source of earnings factors are mortality, maintenance expenses, interest, and persistency. Mortality rates for standard lives are based on fully credible tables, and the rates change slowly from year to year. Maintenance expense costs are relatively low and stable. Interest earnings come from the spread between earned rates and credited rates. Although the earned rates vary significantly from year to year, many companies try to keep the spreads relatively stable.

Persistency rates can only be roughly estimated by the pricing actuary. Variances of estimated from expected are large, and these variances have strong effects on lifetime profitability of the book of business; see the Tan and Eckman papers referenced above.

Similarly, persistency patterns have a large effect on property-casualty insurance profitability, particularly for direct writers. Admittedly, the effects are not as strong as in permanent life insurance, where first year commission rates may exceed the annual premium. Moreover, in private passenger automobile, the loss factors—such as loss development and loss trend—have larger effects than the corresponding life insurance mortality factors. Nevertheless, persistency is significant, and it should be included in the source of earnings analysis.

Ideally, persistency patterns should be incorporated in the ratemaking process for lines of business with high retention rates and differences between first year and subsequent year loss or expense costs by means of asset share pricing models. The source of

---

Program), pages 51-102.

earnings analysis based on an asset share pricing model evaluates the present value of the lifetime profits from a cohort of policies.

For example, if the historical experience used to set the rates has a 90% average persistency rate, and the persistency rate drops to 80% for future business written by the company, the company will show a decline in profitability. The apparent reasons will be higher than expected loss costs and higher than expected expense costs. In truth, expense costs may not have changed and loss costs may be consistent with loss severity indices. Moreover, the full cost of the decline in the persistency rate will not show up until all the old business drops off the company books.

In practice, traditional ratemaking procedures often fail to consider persistency effects. This makes the source of earnings analysis all the more necessary to tease apart the underlying sources of profit or loss.

We illustrate one method in this paper for dealing with the amortization of acquisition expense costs. Traditional property-casualty ratemaking methods combine acquisition expenses with on-going maintenance expenses and treat the sum as either an additive or a multiplicative factor applied to loss costs. This obscures more than it illuminates, and it is not clear why casualty actuaries use these methods.[25] In the illustration here for workers' compensation retrospectively rated business, acquisition costs and solicitation costs on not-taken business are treated separately and amortized over the expected lifetimes of the insurance contracts.

Section IV: Retrospectively Rated Policies

Source of earnings analysis is particularly relevant for retrospectively rated workers' compensation policies. For these policies, the traditional property-casualty accounting framework has severe limitations, which hamper both profitability monitoring and actuarial ratemaking. In this section we apply GAAP (SFAS 97) treatment of universal life-type policies to the corresponding property-casualty policies.

Ideally, actuarial pricing should reflect the underlying economics of the insurance product. Consider first private passenger automobile contracts. In both statutory and GAAP financial statements, earned premium is a revenue, and incurred losses are an expenditure. This is meaningful for profitability monitoring and for ratemaking, since additional earned premium for a block of business signals additional profits and additional incurred losses for a block of business signals decreased profits. Thus, the pricing actuary sets the premium rate (the revenues) based on estimates of the ultimate losses and expenses (the expenditures).

---

[25] This treatment is peculiar to casualty actuarial practice. Life actuaries treat acquisition costs and maintenance expenses separately. The reason for the difference is a combination of accounting practice and inertia. Note that casualty companies do not amortize deferred acquisition costs over more than one year.

For retrospectively rated policies, additional incurred losses generally lead to additional retrospective premiums, with the net effect depending on the premium sensitivity.[26] A change in losses or in premiums does not by itself signal higher or lower profitability.

Accounting for universal life-type policies reflects the economics of these contracts. When the policyholder pays premiums, the monies belong to the policyholders, not to the life insurance company. The insurance company acts as a mutual fund, investing the policyholder's money and deducting a management fee as well as specified charges for insurance services. In GAAP statements for universal life-type contracts, premiums are a deposit, not a revenue.

One may conceive of the workers' compensation retrospectively rated policy in the same fashion. When the insured pays premiums, the insurance company holds the money to pay losses and to cover the various charges, such as the insurance charge and the basic premium charge. If the losses do not materialize, the insurer returns part of the premium to the insured. If additional losses occur, the insurer collects additional premium from the insured.[27]

In sum, generally accepted accounting principles follow the nature of the contracts. For traditional policies (SFAS 60), premiums are revenues and benefits are expenditures. For universal life-type policies (SFAS 97), revenues are the policy charges plus the investment income earned on the account value. Expenditures are benefit payments in excess of the account value, interest credited to the account value, and expenses paid.

The source of earnings analysis in SFAS 97 highlights the causes of gain or loss during the lifetime of a block of business, allowing actuaries to revise the factors used in policy pricing to accord with emerging experience. We demonstrate below how to apply this source of earnings analysis to retrospectively rated workers' compensation policies.

Retro Policies vs Universal Life

The source of earnings analysis for workers' compensation retrospectively rated policies has two major differences from the corresponding analysis for universal life-type policies.

---

[26] See Michael T. S. Teng and Miriam Perkins, "Estimating the Premium Asset on Retrospectively Rated Policies" *Proceedings of the Casualty Actuarial Society*, Volume 83 (1996), pages 611-647, and discussion by S. Feldblum, *Proceedings of the Casualty Actuarial Society*, Volume 84 (1997).

[27] The various charges in a universal life-type policy, such as the mortality charge, the asset management charge, the surrender charge, and the expense charge, are often clearly noted in the policy, particularly if the asset accumulation rate is tied to external investment indices. For the retrospectively rated workers' compensation policy, the pricing actuary sees the individual charges, but the insured may not be aware of the specific components. In addition, because casualty actuaries often avoid an explicit treatment of investment earnings, the investment income earnings factor, as well as the implicit asset management charge, may be nebulous even to the pricing actuary.

1. The *insurance charge* takes the place of the *mortality charge*, and *non-ratable losses* takes the place of *policyholder benefits in excess of the account value*.

2. The SFAS 97 amortization of deferred acquisition costs in relation to expected gross profits, with the year-by-year unlocking of assumptions as actual experience emerges, is complex. It may be justified for universal life policies, where the deferred acquisition costs are 50-60% of the total gross profits for many companies.

   For retrospectively rated policies, a simpler amortization procedure is sufficient. However, the amortization schedule should be dynamic, so that the effects of retention rates (i.e., lapse rates) on profits can be monitored.

Evaluation of Results

Pricing for retrospectively rated policies depends on the evaluation of four earnings factors: (a) investment income, (b) non-ratable losses, (c) expense levels, and (d) retention rates.[28]

Standard reports of premiums and losses do *not* show the expected profits on retrospectively rated policies stemming from these earnings factors. Even more important: they do not show the variations in profit caused by changes in each of these four elements. The pricing actuary has trouble seeing if the ratemaking assumptions reflect the future experience on the book of business. This is the crux of the problem that source of earnings analysis rectifies.

If profits are unexpectedly low, we do not know if the cause is (i) higher than anticipated non-ratable losses relative to the insurance charge, (ii) lower than expected investment income relative to the assumptions used in pricing, (iii) excessive expenses, or (iv) higher than anticipated lapse rates or higher than anticipated not-taken rates. This is particularly important for large account retrospectively rated policies, since the profit margins are narrow and variances in any of these factors have significant effects.

Amortization of Deferred Acquisition Costs

The amortization of deferred policy acquisition costs is essential for monitoring universal life profitability for two reasons:

☛ Deferred acquisition costs are as much as 50% - 60% of gross profits for many

---

[28] Classical actuarial ratemaking procedures focus on the relationship between the insurance charge and the non-ratable losses. This reflects the history of the Casualty Actuarial Society, whose founding fathers were immersed in the theory of retrospective rating. Even today, casualty actuarial candidates spend months learning the intricacies of Table M construction, one of the relics of early 20th century research, and relatively little time on the marketing, financial, and competitive forces that drive the pricing of insurance policies.

universal life contracts.[29] Generally, these products show large losses in the first one or two policy years. Agents' commissions are high in the initial policy year and sometimes also in the first renewal year. Invested assets from policyholder funds are often zero in the initial policy year and low in the first renewal year.

- Retention rates have great effect on long-term profitability. Statutory accounting, however, distorts the effects, since only the surrender charge (a gain) is shown for the current calendar year. Dynamic amortization of deferred policy acquisition costs reveals the actual effects of retention rates on long-term profitability.

The capitalization and amortization of initial acquisition and issue costs is equally important for the analysis of retrospectively rated policies: First year agents' compensation, initial underwriting costs, loss inspection expenses, and policy issue costs form the bulk of many companies' expenditures for retrospectively rated policies.[30]

For large account retrospectively rated business, "not taken" rates can be high. There are a limited number of large workers' compensation accounts in the country. (The "national accounts" workers' compensation business used as the illustration here are risks with annual premiums in excess of one or two million dollars; the exact cut-off varies.) Each account has a risk manager, who puts the account out to bid to the major workers' compensation carriers every five years or so. Each bid from a competing carrier may have only a 10%-20% chance of being accepted, leading to an 80%-90% not taken rate on this business. The costs of developing the bids often are high.

The costs of "not taken" policies must be included with the acquisition costs of a block of business. Some companies spread these costs over related books of business, thereby artificially lowering the profitability of the related books and raising the perceived profitability of the book being priced. For instance, some companies spread the costs of not taken business over the entire workers' compensation line of business.[31]

To properly price this business and to monitor its profitability, the high acquisition expense costs—including the cost of "not taken" policies—must be amortized over the policy lifetimes. It is easy (and tempting) to overestimate persistency rates and to

---

[29] "Gross profits" are the present value of lifetime profits from the block of business before consideration (including amortization) of prepaid acquisition costs; see SFAS 97.

[30] This is especially true for direct writing companies, which pay large first year commissions and low renewal commissions. The illustrations in this paper assume direct writing of workers' compensation and of personal automobile insurance.

[31] This leads to incorrect pricing and marketing decisions. Sometimes there are valid reasons for this practice, as when a new carrier seeks to break into the large account market. More often this practice stems from the pricing actuary's inability to allocate acquisition costs. Among non-actuaries, one sometimes hears the view that solicitation costs for not taken business should be spread evenly over the company's entire business. Actuarial ratemaking requires that these solicitation costs be charged to the block of business under review; see S. Feldblum, "Personal Auto Premiums: An Asset Share Pricing Approach," *PCAS* 1996.

underestimate not taken rates. Source of earnings analysis with dynamic amortization of policy acquisition costs is an effective tool for more accurately pricing this business.

Static vs Dynamic Amortization

Static amortization schedules, like static depreciation schedules, do not change with the passage of time. The rate of amortization or depreciation may vary from year to year, as with double declining balance depreciation schedules, but the amortization rate is not re-estimated as more information is learned about the business.

For instance, if the average policy in a given cohort is expected to persist five years, then one fifth of a policy's deferred policy acquisition costs are amortized each year (assuming a zero interest rate for amortization). If after two years of experience with this cohort of business, the average policy's lifetime is expected to be different from five years, the amortization schedule is not changed.

Static amortization schedules give distorted measures of profitability if the actual persistency rate or the actual investment yield differs from that assumed during pricing. Dynamic amortization allow for revision of the schedules as actual experience becomes known and as future expectations change.[32]

For example, suppose the actuary prices the business assuming that the excess of first year over renewal acquisition costs is 20% of premium, the average policy lifetime is 8 years, that the not taken rate is 20% for this block of business, and that the solicitation costs for not taken business equals 50% of the excess first year acquisition costs for insured business.[33] For simplicity, we assume an amortization interest rate of 0%.[34]

Amortizing the excess first year acquisition costs plus the solicitation costs for not taken business gives a charge of 2.8% of premium each year. These assumptions are uncertain, though they become known with the passage of time. The not taken rates and the solicitation costs for not taken business are known after the new policies are written, and the average policy lifetime can be estimated more accurately two or three

[32] For the universal life-type policies covered by SFAS 97, the deferred policy acquisition costs are amortized in proportion to future expected gross profits. The amortization schedule must be revised whenever actual experience or future expectations differ from initial assumptions for any of three items: persistency rates, investment yield, and expected or actual gross profits. The amortization of deferred policy acquisition costs in relation to expected gross profits is a complex subject in its own right, and it is not dealt with in this paper.

[33] The not taken rate is the percentage of new business contract discussions that are not consummated by a policy. The figures here are overly optimistic. An average policy lifetime of 9 years means a retention rate of about 89%. Of every 100 policies, 89 are renewals and 11 are new business. With a not taken rate of 20%, there are slightly over 2 contract discussions per 100 insureds that do not result in a policy. In practice, the competition for large account business is intense. For each 10 new policies acquired, the insurer might solicit 30 or 40 accounts.

[34] Using the actual investment yield in the amortization schedule obscures the arithmetic and does not change the conclusion.

years after the expiration of the initial policy year (by projecting from early retention rates).

If estimates are not validated, there is often a temptation to estimate optimistically. Suppose that these figures are revised after the new policies are written, for an average policy lifetime of 5 years and a not taken rate of 60%. The annual acquisition cost charge is revised to

Table 6: Solicitation Costs for Not Taken Business

	Assumptions	
	Initial	*Revised*
A. Premium	$100 million	$100 million
B. Excess acquisition costs	$20 million	$20 million
C. Not taken rate	20%	60%
D. Not taken premium [= A * C/(1-C)]	$25 million	$150 million
E. Not taken acquisition costs	$2.50 million	$15 million
F. Total acquisition costs	$22.5 million	$35 million
G. Average policy lifetime	8 years	5 years
H. Annual amortization	$2.81 million	$7.00 million

Retention rates are essential for monitoring the profitability of workers' compensation retrospectively rated policies. Not only are renewal expense costs much lower than first year expense costs, but renewal loss costs are also lower than first year loss costs. Dynamic amortization of deferred acquisition costs enables the pricing actuary to see the effects of retention rates on long-term profitability.

Supporting Surplus

For universal life policies, source of earnings exhibits do not normally consider invested capital. Before the advent of risk-based capital requirements, this approach was reasonable, at least for GAAP statements. Policy reserves often did not significantly exceed the account balance, deferred policy acquisition costs were amortized, and little surplus was needed to satisfy regulatory requirements. In sum, the capital invested in the block of business was small relative to the funds supplied by policyholders.

For workers' compensation, the opposite is true. Substantial amounts of investors' capital is embedded in the undiscounted loss reserves and in the gross unearned premium reserves. Additional capital is needed to meet the NAIC's risk-based capital requirements or rating agency capital formulas.

Both the explicit and the implicit capital contributions should be considered when initially pricing the block of business. At times, some capital contributions are overlooked, either because the pricing model is faulty or because the actuary is not aware of the implied equity flows. The source of earnings analysis includes the actual investment income as one source of gain or loss. This investment income applies to both the policyholder supplied funds and to the capital funds.[35]

In addition to the standard life insurance source of earnings exhibits, the allocated surplus amounts and the target returns should be shown on an aggregate basis, so that the source of earnings analysis can be converted to a return on capital basis. This format highlights the variance between expected and actual return on capital, along with the factors that contributed to this variance.[36]

Charged, Expected, and Actual

For the private passenger automobile source of earnings analysis, we showed three stages for the loss severity trend factors:

1. initial (estimated) trend
2. revised (actual) trend
3. actual loss cost change

The change from estimated trend to actual trend is estimation error; the change from actual trend to actual loss cost change is process error. The same three level analysis applies to loss development factors and loss frequency trends, as well as to other components of insurance ratemaking.

Workers' compensation retrospectively rated business loss component differs in two respects.

A. It is difficult to judge the accuracy of our estimates even after the completion of the policy year. For instance, the insurance charge is based on Table M entry ratios and size of loss distributions, as well as standard trend estimates. The policy year experience tells us the actual non-ratable losses; it does not tell us much about the proper insurance charge.
B. There is an explicit charge for the insurance protection in the rate. For private

[35] At a minimum, the cost of the capital funds is the double taxation on these funds; see Myers and Cohn [1987, op cit.]. Some actuaries assume that there are additional costs of holding capital funds, such as the differential earnings rate resulting from the conservatism in many financial portfolios; see R. W. Sturgis "Actuarial Valuation of Property/Casualty Insurance Companies," *Proceedings of the Casualty Actuarial Society*, Volume 68 (1981), pages 146-159 as well as Miccolis's comments in "An Investigation of Methods, Assumptions, and Risk Modeling for the Valuation of Property/Casualty Insurance Companies," *Financial Analysis of Insurance Companies* (CAS 1987 Discussion Paper Program), pages 281-321.
[36] The determination of the needed capital by line of business and of the returns on this capital from the insurer's various operations are discussed in Douglas M. Hodes, Sholom Feldblum, and Gary Blumsohn, "Workers' Compensation Reserve Uncertainty," *Proceedings of the CAS*, Volume 86 (1999), and Douglas M. Hodes, Sholom Feldblum, and Antoine Neghaiwi, "The Financial Modeling of Property-Casualty Insurance Companies," *North American Actuarial Journal*, Volume 3, Number 3 (July 1999), pages 41-69. respectively.

passenger automobile, we set premiums. For retrospectively rated policies, we set charges.

Accordingly, the source of earnings exhibits reflect the ratemaking procedure. For each "earnings source" in the reporting and evaluation structure, there are three values:

1. the amount charged in the pricing analysis,
2. the expected amount at policy inception, and
3. the actual (realized) amount.[37]

Suppose that a policy is issued on January 1, 2001, with an insurance charge (including the excess loss charge) of $500,000, and with expected non-ratable losses of $450,000.[38] At policy inception, the initial report would show

Table 7A: Workers' Compensation Charged, Expected, Actual

Date	Insurance Charge	Expected Non-ratable losses	Expected Gain	Actual Non-ratable losses	Variance	Actual Gain
1/2001	$500,000	$450,000	+$50,000	---	---	---

Suppose that on December 31, 2001, at the expiration of the policy, the estimated non-ratable losses (including all actuarial bulk reserves) is $470,000. The "variance" is –$20,000, and the "actual gain" is +$30,000. The report would show the following entries for 12/2001:

[37] This structure is analogous to the source of earnings structure for universal life-type policies. For example, non-ratable losses in retrospectively rated policies (that is, the losses which are paid by the insurer and are not reimbursed by the employer), are the analogue of "benefits in excess of account value released" in universal life policies. The insurance charge (including the excess loss charge) in retrospectively rated policies is the analogue of the mortality charge in universal life policies.

[38] The insurance charge, along with the excess loss charge, is amount actually used in pricing the policy. Some actuaries use an insurance charge equal to the expected non-ratable losses, and they include a separate (explicit) profit provision. Other actuaries use conservative factors for the insurance charge and the excess loss charge. The charges in the plan minus the expected non-ratable losses is an implicit profit margin in this earnings factor.

Both methods are common in the property-casualty insurance industry. Life insurance pricing generally uses the latter method, with implicit profit provisions in the mortality and interest factors, and sometimes also in the expense and withdrawal factors. That is, life insurance pricing uses conservative mortality tables as well as a spread between the earned interest rate and the credited interest rate.

We adopt the latter pricing model for the exhibits in this paper. A company that uses the former pricing model, with no "spreads" in the pricing components but with an explicit profit margin, would show zeroes in the initial profit for each source. This does not affect the analysis of gain and loss by source.

In the exhibits, we show all profit margins, gains, and losses as dollar amounts. In pricing the policies, many of these items – such as the insurance charge – are shown as percentages of standard earned premium. This does not affect the analysis in the paper.

Table 7B: Workers' Compensation Charged, Expected, Actual

Date	Insurance Charge	Expected Non-ratable losses	Expected Gain	Actual Non-ratable losses	Variance	Actual Gain
12/2001	$500,000	$450,000	+$50,000	$470,000	–$20,000	+$30,000

The source of earnings exhibits allow the actuary to evaluate the accuracy of the pricing components, providing a better assessment of the pricing procedure. The perception of profitability at expiration of the policy year influences future rates. Multi-year source of earnings exhibits illustrate the effects of claim development, investment yield changes, and persistency changes on the profitability of the book of business. For instance, if actual non-ratable losses increase to $515,000 by December 31, 2002, the report would show

Table 7C: Workers' Compensation Charged, Expected, Actual

Date	Insurance Charge	Expected Non-ratable losses	Expected Gain	Actual Non-ratable losses	Variance	Actual Gain
1/2001	$500,000	$450,000	+$50,000	---	---	---
12/2001	$500,000	$450,000	+$50,000	$470,000	–$20,000	+$30,000
:	:	:	:	:	:	:
12/2002	$500,000	$500,000	+$50,000	$515,000	–$65,000	–$15,000

The illustration above assumes that the non—ratable losses expected when initially pricing the policy are less than the insurance charge. In practice, various relationships may be used. The discussions above of the loss severity trend factor and the investment earnings factor imply that

1. when explicit profit margins are used, the initial variance between the expected amount and the factor used in pricing is zero, and
2. when implicit profit margins are used, the expected amount is less than the factor used in pricing.

Insurance Charge

The situation is more complex with the insurance charge. The insurance charge is stated in nominal dollar terms, not in present value terms. A zero dollar initial variance is actually an implied profit margin, since the insurance charge is collected before the excess losses are paid. Retrospectively rated policies can be priced in one of several ways:

1. Present values of losses may be used for determining the insurance charge. Although this method may seem natural, it is not commonly used, since the maximum and minimum premiums are in nominal dollar terms.
2. Ultimate values of losses are used, but the insurance charge is reduced for the expected investment income on the excess losses. This is the method implicitly used in some bureau plans. The insurance charge is stated as a percentage of standard premium, which may have a profit factor that takes into account expected investment income. The resultant insurance charge may be less than the expected (nominal) excess losses. The implicit assumptions are that (a) the profit margin in the standard premium truly takes into account all investment income and (b) the loss payment pattern for excess losses has an average payment date similar to that for all losses. Both assumptions are dubious; in particular, excess losses have slower payment patterns, leading to implicit profit margins in the insurance charge.
3. The insurance charge is based on ultimate losses, and a separate investment income factor is calculated based on all insurance cash flows (both ratable and non—ratable losses), which reduces the basic premium charge.[39]

For simplicity, this example assumes a single policy written on January 1, 2001. Actual reports would be for blocks of policies, such as all large account business written by a particular sales office in policy year 2001. Since non-ratable losses have great random fluctuation, a report showing variances is meaningful only on a block of business basis. The subsequent examples are for a block of policy year 2001 business.

Expenses

Expenses are divided into two components:

1. Underwriting and acquisition expenses, including solicitation costs for not-taken business
2. Policy maintenance expenses, including loss adjustment expenses

The effects of acquisition and underwriting expenses on profitability depends on expected versus actual not taken rates and renewal rates. To reflect traditional source of earnings analysis, we group these under the "persistency" factor. The effects of other expenses on profitability depends on the efficiency of company operations and of loss department procedures as well as on the litigiousness of the claim filing population.

Combining the Earnings Factors

[39] There are endless variations in pricing retrospectively rated policies. For instance, when pricing wide swing plans for large accounts that wish to avoid paying expense fees to the insurance company, the insurer may use ultimate losses to calculate the insurance charge, add a large loss conversion factor, and have no other expense charges. Pricing practices that are intended to conceal the actuary's expectations hamper source of earnings analyses as well.

The complete source of earnings exhibits show the variances by earnings factor. The first row in the table shows the contribution to profitability from each factor in the pricing assumptions. Subsequent rows show the variance resulting from actual data and revised estimates. (The term "variance" is used here in the accounting sense, meaning the difference between expected and actual.)

Table 8: SOE Analysis for Retro Policies ($000)

Valuation Date	Non-ratable Losses	Interest Earned	Persistency	Maintenance Expenses	Explicit Profit	Total Profit
1/1/2001	$2,000	$2,500	–$1,500	$750	$1,250	$5,000
12/31/2001	$1,400	$3,400	–$2,500	$750	$1,100	$4,150
12/31/2002	$2,100	$3,600	–$2,900	$750	$1,100	$4,650

Pricing Assumptions

At January 1, 2001, the inception of the policy year, the figures show the (implicit and explicit) profits margin embedded in the pricing assumptions. Most of the expected profit is built into the pricing components. The pricing actuary has set insurance charges that exceed the expected non-ratable losses by $2 million. In addition, the company expects about a one year float on insurance funds (between premium collection and loss payment) for which the policyholders are not given full credit. Specifically, the actual investment income is expected to exceed the investment income assumed in pricing by $2,500,000.

The company expects actual maintenance expenses (including loss adjustment expenses) to be $750,000 below the amount assumed in pricing. In addition, the company builds in an explicit profit component of $1,250,000 for this book of business.[40]

The company expects to lose money from high solicitation costs on not taken business. Much of this money is recouped from acquisition expense charges in the basic premium. The amount that is not recouped is a negative implicit profit margin of $1,500,000. It is hard to persuade policyholders that they should reimburse the costs of soliciting other business, so it is difficult to explicitly charge for this cost in the premium.

Underwriting

The first row shows the pricing assumptions at the beginning of the policy year. Rarely are all pricing assumptions realized in the policies actually sold. The second row shows

[40] The company wishes to show the low profit margin to policyholders, not the full implicit plus explicit profit margin. For many books of large account business, the explicit profit margin is zero. The company shows the policyholder a pricing analysis with no apparent profit. The company expects to earn profits on the book of business from the conservative pricing assumptions. Some companies even show negative profit margins to their consumers, at times even professing that they are losing money simply to retain a valued account.

the revised earnings factors at the end of the policy year. The variances from initially expected profits stem from two causes: (a) the charges actually embedded in the policy components may differ from those originally anticipated by the actuary and (b) fluctuations in losses or shifts in the financial environment may affect the costs actually incurred by the company.

The changes from underwriting are as follows. Interest rates have risen and the marketplace has softened, but the company underwriters have adhered closely to the actuarial pricing recommendations. The rising interest rates led to a revised estimate of excess losses, since an anticipated rise in inflation has a leveraged effect on higher layers of loss. This reduces the implicit profit from non-ratable losses by $600,000. A few insureds were given premium credits, reducing the explicit profit margin by $150,000. Because of the soft market, not-taken rates increased, leading to a additional $1 million loss from unfulfilled solicitation costs. Interest rates rose before the company received the policy premiums or invested them, leading to an additional $900,000 implicit profit from the interest spread.[41]

Much of the first year revisions stem from underwriting changes in the policy proposal. Thus, revisions in the profit from non-ratable losses stem from changes in the insurance charge in the policy, and revisions in the profit from acquisition expenses stem from unexpected not taken rates.

Actual Experience

Subsequent revisions stem primarily from unanticipated changes in the financial environment, the insurance marketplace, or from random loss occurrences. For instance, the 12/31/2002 row shows an increase in the expected profits from non-ratable losses. By December 31, 2002, all policies have run their course, and there have been fewer large losses than expected. This may be a result of stringent underwriting or of random loss fluctuations.[42]

The December 31, 2002, figures are actual figures in part and estimates in part. For instance, the investment yield in 2001 and 2002 is known; the effect of acquisition costs on policy profitability still depends on future persistency rates.

The source of earnings exhibits are updated until most of the losses have been settled or until subsequent changes in estimated earnings are not material. For the first few years, the changes can be significant, particularly for earnings from non-ratable losses, interest, and persistency.

---

[41] The pricing actuary must take care to reflect the higher interest rate, and the potentially higher inflation rates, in the insurance charge. If this is not done, the implicit profit margin from non-ratable losses may be overstated.

[42] Most years show somewhat fewer large losses than average; a few years show significantly more large losses than average. The skewed distribution of claim sizes leads to source of earnings gains or losses on different books of business.

Non-ratable losses: Projection of ultimate losses has always been the actuary's task. When pricing retrospectively rated contracts, however, some actuaries rely on published industry figures, as contained in Table M data from rating bureaus. Individual company data are sometimes considered insufficiently credible for revising Table M figures, and the needed adjustments for inflation and for changes in the size of loss distribution are considered complex. In fact, Table M charges should be reviewed periodically to ensure their adequacy. The source of earnings analysis provides a hindsight view of insurance charge adequacy that can be invaluable for the pricing actuary. The challenge for the pricing actuary is to determine from the emerging experience how much of the variance stems from estimation error and how much stems from process error.

Interest: The earnings from interest on retrospectively rated contracts depend on several factors: the investment yield actually received, the investment yield used to price the policy, the payment dates for losses, and the collection dates for premiums. Large accounts often seek cash flow plans to retain more of the investment income for themselves, and their plans may be individually tailored for the insured. For these large accounts, the pricing actuary may have to determine the expected earnings from interest on a plan by plan basis.

As noted above, some of the expected investment earnings may be incorporated in the insurance charge. For instance, the insurance charge may be expressed as a percentage of standard premium, and the standard premium may be adjusted for expected investment income. The source of earnings analysis follows the pricing analysis; it does not dictate it. However, the pricing actuary should be aware of the implicit profit margins in each earnings factor in order to properly monitor profitability.

Persistency: For large account retrospectively rated business, the solicitation costs for not taken business and the persistency of newly acquired business have large effects on overall profitability.[43] The source of earnings analysis ensures that pricing actuaries incorporate their effects in the ratemaking formulas.

[43] The full effects of interest rate changes and persistency changes take several years to play out. Some pricing actuaries disclaim responsibility for interest rate changes, not taken rates, and persistency rates, since traditional casualty actuarial ratemaking procedures do not deal with these items. The common disclaimer is that "the actual investment yield is the responsibility of the Investment Department; we simply use the projections that they provide us." Similarly one hears that "the actual persistency rate, or the actual not taken rate, is the responsibility of the Marketing Department or of the sales force; we simply use the projections that they provide us." This retort is disingenuous. The source of earnings analysis does not bring investment policy or marketing philosophy under the purview of the actuary. Nevertheless, just as the reserving actuary does not rely solely on the claims department's loss estimates, the pricing actuary can not rely solely on others' estimates for the basic input parameters.

Section V: Conclusions

Two themes run through this paper. One theme underlies the workers' compensation illustration and the other theme underlies the private passenger automobile example. Neither theme is new; both have been expressed in other forms by life actuaries, accountants, and statisticians. When seen from the perspective of source of earnings analysis, however, they imply a major revision of casualty actuarial pricing. We summarize the two themes below and their implications for practicing actuaries.

Pricing Paradigms

A premium-loss pricing paradigm currently dominates casualty actuarial ratemaking. The actuary determines policy premiums to cover expected losses and expenses.

With the policy revolution of the 1980's, life actuaries moved to a credit-charge paradigm. The new interest-sensitive policies were unbundled into various components. The actuary determines explicit charges and credits for the policy components, which may be rearranged into full policies to meet customer needs.

The credit-charge pricing paradigm is extremely flexible, and it is increasingly being used for large account commercial lines ratemaking. The account purchases a customized policy with a variety of specialized components: sublines, deductibles, premium payment plans, retrospective rating, loss engineering services, claims handling services, excess coverage, and so forth.

The actuary prices the components, which are assembled by the underwriter into the policy. For instance, the actuary determines the appropriate insurance charge for a set of plan parameters, or the appropriate interest credit for a given plan type and premium payment pattern. Source of earnings analysis enables the actuary to monitor the adequacy of the charges and credits.

The shift from a premium-loss pricing paradigm to a credit-charge pricing paradigm brought "universal" contracts to the life insurance industry. We may conceive of universal policies as (in effect) retrospectively rated contracts where the premium adjustment depends on the investment yield achieved, not on the loss experience.[44]

By unbundling the policy into its components, the casualty actuary can offer universal policies for lines with long term claim payments, such as workers' compensation. The actuary sets the investment spread; the actual premium for the coverage varies with the investment income realized. Such policies would be particularly attractive to large

[44] There are differences, of course. Universal policies allow more management discretion in setting the credited interest rate; workers' compensation retrospectively rated policies have contractually determined premium adjustments. Universal contracts depends on the insurer's investment yield or on an external interest index; retrospectively rated policies depend on the individual insured's loss experience.

accounts seeking aggressive investment returns. Pricing for unbundled policy components is directly tied to source of earnings analysis.

Random Variations

Actuaries often attribute differences between expected and actual results to random loss fluctuations, to unforeseeable changes in inflation, or to unanticipated market pressures on underwriters and agents. The work pressures on actuaries are so great, and the potential causes of adverse results are so diverse, that many pricing actuaries never examine the variances in past results. In short , some actuaries believe that their time is too valuable to be spent re-examining their past analyses.

In truth, efficient examination of past results is a requisite for accurate prospective pricing. The source of earnings exhibits enable the actuary to quantify the contribution of each earnings factor to overall changes in profitability and to differentiate between estimation errors and process errors within the earnings factors. This "policy post-mortem" may reveal biases in earnings factors or unstable pricing procedures.

Ratemaking is prospective; we price next year's business, not last year's business. The pricing actuary succeeds by peering into the future, not by looking back.

Yet our ratemaking procedures are not infallible. Often our methods are defective and our predictions are erroneous. Ever afraid of looking back, we try to outrun the errors.

We can not outrun our errors. If we never look back, we never know the causes of our errors. We never learn if a variance of actual from expected results from random loss fluctuations or from poor ratemaking assumptions.

Appendix: Implementation Issues

Source of earnings analysis is not a theoretical exercise intended for pure actuaries and academic journals. In life insurance pricing, source of earnings analysis is part of the practicing actuary's repertoire. Similarly, this paper is written for the practicing casualty actuary.

The source of earnings procedures described here are foreign to most casualty actuaries, and they require data that is not always kept by casualty insurers. This latter characteristic is true for many new actuarial procedures.

Imagine a large, multi-line foreign insurer writing long-tailed commercial lines of business. The insurer keeps loss data only by calendar year, and it sets loss reserves by claim adjusters' estimates. The insurer notices that its reserves seem to be perpetually inadequate, and it hires a North American casualty actuary to analyze the problem and to recommend a solution.

The actuary informs the insurer that the required analysis is straightforward and asks to see the company's accident year loss triangles. The company's management is confused; they say that they don't keep accident year data. The actuary requests policy year or report year data, but only calendar year aggregate data are available.

> *Well*, says the actuary, *we must form accident year loss triangles by line of business, as well as by subline, by state, and by type of policy.* The management of the company agrees.

Any practicing actuary can supply the denouement of this tale. The actuary spends months trying to create the necessary triangles, with (ostensibly) full support of the insurer's management, but the efforts come to naught. Revising company data systems is an staggering undertaking. Certain data are simply not available; in some lines of business, the accident date may not even be coded in the electronic claim files. Other bottlenecks are human. No-one has the time or the persistence for this task.

Most reserving actuaries can not conceive of an insurer writing long-tailed lines of business without keeping accurate accident year experience for reserve estimation. Yet the effort to first create an accident year reporting system is enormous. Unless the insurer already appreciates the importance of accident year loss triangles, the insurer is unlikely to expend the effort to create the system.

In the real world, the situation is worse. Practicing actuaries are busy, busier than Alice's White Rabbit. These busy actuaries are forever computing things, crunching numbers, forming endless exhibits. There is never time to review previous work, since current tasks are always pressing.

All too often, the busy actuaries are busily computing numbers which never get used, numbers which do not accurately reflect the values that they purport to measure. The busy actuaries do not realize this, because they are always too busy to evaluate the accuracy of their work.

Time and again, we have looked at the work of some of our colleagues—pricing actuaries, reserving actuaries, and valuation actuaries—and pointed out fundamental errors that negated the value of their efforts. At first there is disbelief, then denial: could it be that months of work were wasted? Eventually comes grudging acceptance, perhaps hastened by the authors' reputations in the actuarial community. Finally the actuaries run off to correct the procedures; they works evenings and weekends to get the project completed on time.

This is the actuary's destiny: the incessant computation of complex exhibits that bewilder the audience and sometimes entrap even the actuary, so that when errors creep in and lead the results astray, no one can distinguish right from wrong.

Source of earnings analysis is crucial to good actuarial work. Source of earnings analysis asks whether the assumptions are borne out by actual results. Some assumptions, like trend factors, development factors, credibility factors, seem trivial. The practicing actuary says: "How can one get these factors wrong?" The practicing actuary shakes his head in disbelief and walks away. But the authors have seen months of highly sophisticated work on trend factors, development factors, and credibility factors that led to erroneous results, unbeknownst to the busy actuaries. Source of earnings analysis enables the practicing actuary to examine the accuracy of the efforts.

Other assumptions are more elusive. The pricing actuary's rate indications rely on investment income assumptions, persistency patterns, acquisition cost assumptions, and loss discount rates. Sometimes the assumptions are explicitly worked into the underwriting profit margin or the underwriting expense ratio; sometimes the assumptions are implicit in the actuary's target loss ratio or target combined ratio. Year after year these assumptions are repeated in the rate reviews. Perhaps the assumptions are supported by extensive "actuarial research," which is all too often a combination of intensive number crunching and sloppy statistics. Rarely—if ever—does the actuary examine the validity of the assumptions.[45]

[45] Sometimes the results are humorous. (i) Casualty actuaries have produced a plethora of financial pricing models, many of which are at odds with financial theory. With no way of checking their validity, rate makers use these models over and over again. (ii) Auto pricing actuaries are among the busiest actuaries there are, churning out rate indications in state after state, repeating the cycle year after year. Yet the incessant churning often misses the true cost drivers of auto insurance losses; see John B. Conners and S. Feldblum, "Personal Automobile Insurance: Cost Drivers, Pricing, and Public Policy," *Proceedings of the CAS*, Volume 85 (1998), pages 370-404.for a re-examination of why costs are higher or lower in different areas. Similarly, source of earnings analysis forces the actuary to rethink the

The practicing actuary may complain that it is difficult to implement the source of earnings analysis for a particular factor, such as the interest earnings factor or the persistency factor. What the actuary is saying is that it is hard to determine whether the factors being used are correct. Let us rephrase this: if it is hard to determine whether the factors are correct, then it is quite possible that the factors are not correct. If the factors are not correct, then not only has the actuary wasted much time computing these factors, but the actuary has wasted even more time performing the analyses that rely on these factors. Source of earnings analysis is not an impediment to productivity; it is crucial to making the busy hours become productive hours.

Data Availability

One of the most common complaints about source of earnings analysis is that the data are not available. This complaint is made about many new techniques, with one difference for source of earnings analysis: the data that are needed for source of earnings analysis are the data that are crucial for policy pricing.

Consider the discussion of workers' compensation retrospectively rated policies in this paper. Almost invariably the pricing actuary says:

> *We don't have the data needed for the analysis of expenses. We don't keep track of our not taken rates, we don't quantify the solicitation costs for the not taken business, we don't separately evaluate the first year acquisition costs, and we don't keep records of policy persistency.*

We wonder: *If you don't know your expenses, how do you price the business?*

The pricing actuary adds:

> *We don't have the data needed for the analysis of the interest factor. We know when the losses are paid, on average, but we don't have a good handle on the premium collection pattern. We have incurred loss retros and paid loss retros, and we have all sorts of premium payment patterns; we don't know when the average premium comes in either for the aggregate book of business or for particular groups of policies. We don't know when the expenses are paid; all we have are IEE aggregate figures by calendar year. We have estimates of new money rates, but we don't know how much we actually earn on a given book of investments. We simply don't have the data to quantify the amount of interest we actually earn.*

We wonder: *If you don't know your interest earnings, how do you price the business?*

---

assumptions used in the rate reviews.

The answer to our questions is straightforward: *We price the business as well as we can, using estimates and guesses when we don't have data.*

If an assumption is not material, then it can be ignored in the source of earnings exhibits just as it is treated glibly in the pricing analysis. A good example is maintenance expenses, which we ignored in this paper.

If an assumption is central to the pricing analysis, such as the acquisition expense assumption or the interest earnings assumption, then it can not be ignored in the source of earnings analysis. But it can not be ignored in the original pricing analysis either. The source of earnings analysis tells the actuary the work that must be done. It is amusing to watch pricing actuaries credibility weight loss development link ratios that are computed to three decimal places even while they are oblivious of the acquisition expenses or the interest earnings on their book of business.

Estimation Error and Process Error

For forty years, actuaries have debated the issues of process risk, parameter risk, and specification risk. Some of our readers complain that the risk categorization in the *paper is not refined enough. Others complain that one can not easily separate the* errors into the categories in the paper.

We do not wish to intrude on this debate. We have discussed these issues in other papers, and there is no gain from repetition here.

But the central idea of the paper bears repeating. Mere identification of the variances of actual from expected is not sufficient. We must determine (as best we can) the cause of the variance. If the cause is process error, such as random loss fluctuations or random stock market movements, then there is little that the actuary can do to avoid the error. But if the error stems from other causes, whether estimation error or "parameter risk" or "specification risk," then the pricing actuary should attempt to correct the errors, minimize the errors, or ensure that they do not repeat themselves.

In sum, we do not try to specify which items are estimation error and which are process error. The pricing actuary performing the rate analysis is better equipped to classify the errors than we are, since the classification depends on the type of rate analysis and the line of business. The objective is as stated above: to separate the errors which stem from random fluctuations from the errors which are attributable (at least in part) to the estimation procedures.

Investment Income

The earnings factor causes problems for many practicing actuaries. The criticism generally takes the following form:

The source of earnings analysis presupposes some sort of investment income assumption in the rate analysis. But that is not how we develop rate indications. We price to a target combined ratio, or a target underwriting profit provision. This target is not chosen by the pricing actuary doing the rate review. It is set by the chief actuary (or by company management) after reviewing the recommendations of the research actuary (or the research department). The research actuary uses an internal rate of return pricing model, or a Myers-Cohn discounted cash flow model, or a Butsic risk-adjusted loss discount model, to determine the target combined ratio. Even in these models, there is no simple interest assumption: we have the internal rate of return, or the Myers-Cohn CAPM adjusted discount rate, or Butsic's risk adjusted rate. Our pricing procedure does not fit into the source of earnings mold.

This criticism is dismaying. It has been twenty years since actuaries first began using financial pricing models for casualty insurance products. The parameters of these models—such as the assumed investment yield, the risk adjustment, the surplus assumptions, the assumed equity flows—greatly affect the final premium rate. Yet many actuaries who are expert in other pricing issues still can't figure out what their pricing model says. They can tell you the effect of a one point increase in the assumed trend factor, but they can't tell you the effect of a one point increase in the assumed discount factor.

Once again, source of earnings analysis is part of the solution. The source of earnings analysis asks two questions:

1. How much investment income does the pricing model assume the company will receive?

2. How much investment income does the company actually receive?

Some pricing models explicitly consider the investment income of the company stemming from the insurance operations; examples are the internal rate of return pricing model. Other pricing models focus on loss discount rates instead of on investment income rates. Examples are the Myers-Cohn discounted cash flow model and Butsic's risk-adjusted loss reserves discount model.[46]

[46] For summaries of these pricing models, see Myers, Stewart and Richard Cohn, "A Discounted Cash Flow Approach to Property-Liability Insurance Rate Regulation," in J. David Cummins and Scott E. Harrington (eds.), *Fair Rate of Return in Property-Liability Insurance* (Boston: Kluwer*Nijhoff Publishing, 1987), pages 55-78; Butsic, Robert P., "Determining the Proper Interest Rate for Loss Reserve Discounting: An Economic Approach," *Evaluating Insurance Company Liabilities* (CAS 1988 Discussion Paper Program), pages 147-188; Robert P. Butsic and Stuart Lerwick, "An Illustrated Guide to the Use of the Risk-Compensated Discounted Cash Flow Method," *Casualty Actuarial Society Forum* (Spring 1990), pages 303-347; S. Feldblum, "Pricing Insurance Policies: The Internal Rate of Return Model," Second Edition (CAS Part 10A Examination Study Note, May 1992); Ira Robbin, "The Underwriting Profit Provision," *Casualty Actuarial Society Part VI Study Note* (1992).

In fact, both of the latter two models assume investment earnings at a risk-free rate. The risk adjustment to the loss reserves discount rate serves to compensate the insurer for its underwriting risk. In the source of earnings exhibits, the assumed interest earnings are the interest earnings at the current risk-free rate.

Insurance Charges

Retrospectively rated workers' compensation policies seem the ideal candidates for source of earnings analysis, because the actuary sets charges for each component separately. There is a basic premium charge for underwriting expenses, a loss conversion charge for loss adjustment expenses, a tax charge for premium taxes and state assessments, and an insurance charge for the cost of non-ratable losses. The source of earnings exhibits would compare the charges in the policy with the actual costs incurred by the insurer.

Once again, investment income is the problem. Casualty actuaries have generated numerous models for pricing retrospectively rated contracts. Yet the models are built on a nominal loss foundation. The current "Table M" formulas use nominal loss values, not the present values of the losses.

The nominal loss models are used because the loss limits and the premium limitations are expressed in nominal dollars. The rationale for ignoring investment income in the pricing formula is that the insurance charge is a factor applied to the standard premium. The underwriting profit provision in the standard premium takes into account the expected investment income.

This rationale is applicable only when the insurance charge is indeed based on a standard premium, as is true for the NCCI retrospective rating plan. It is not applicable when companies separately combine an insurance charge with other expenses for their large account business.

Even when an NCCI type plan is used, the rationale assumes that the cash flow patterns are the same for prospectively priced business as for retrospectively priced business. This is rarely the case, and the difference can be substantial.

Many large companies judgmentally reduce the final premiums for expected investment income. Alternatively, these companies judgmentally reduce the component charges for expected investment income.

Actuarial Rates and Market Prices

Some readers have commented that the actuarial indications are not the problem. The problem is that the sales force or the underwriters cut the prices below the indications, either to meet peer company competition or to retain valued customers.

The source of earnings analysis explicitly incorporates such price adjustments. Most commonly, a market decision to revise the charged price is shown as an adjustment to the explicit profit provision in the rates. For instance, if the actuary's indications assume a profit, after incorporation of investment income, equal to 8% of premium, and the underwriter grants a 10% premium reduction, then the revised explicit profit provision is a negative 2%.

The standard critique of this source of earnings analysis is that price cutting is not done arbitrarily. The 10% rate reduction may have been offered to retain market share, to keep a customer, or to keep down fixed expense costs. The source of earnings analysis does not tell us if the 10% rate reduction is justified.

It has been emphasized throughout this paper that both pricing and profitability measurement must be done using "lifetime" methods. Ideally, policy pricing is done by asset share analysis that considers deferred policy acquisition costs, changes in loss costs over time, and policy persistency rates. Similarly, source of earnings analysis should incorporate a persistency factor, and it should examine the cohort of policies from original inception.

This does not mean that we can examine policy profitability only after the policies have been in existence for several years. On the contrary: source of earnings analysis enables us to examine long-term profitability reasonably quickly, since we can examine the extent to which original pricing assumptions have been validated by experience.

This is an introductory paper, and we have not attempted to show source of earnings exhibits for a cohort of policies, using assumed and actual persistency rates. These source of earnings exhibits are meaningful only if the pricing analysis explicitly incorporates persistency factors. If the pricing analysis is deficient, then the source of earnings analysis will be deficient as well.

Classification Rates

One reviewer of this paper has commented (paraphrasing):

> *The paper deals with statewide rate indications. But we don't actually price based on the statewide rate indications. We use classification relativities and territorial relativities in private passenger automobile; we use partial pure premiums by classification to develop rates in workers' compensation.*

This is correct. The first draft of our paper included a section on classification ratemaking. We excluded that section because several other actuaries have already dealt with this topic in well-thought out analyses. These actuaries include Glenn Meyers, Roger Hayne, and Howard Mahler.

We do not dismiss the work of these actuaries; their analysis is good, and it complements the source of earnings exhibits. However, source of earnings analysis has value to company management, in addition to its value for the pricing actuary. Company management is concerned with variances from planned results on an aggregate basis, such as line of business or state within line of business; source of earnings analysis deals directly with this issue.

Combined Effects

One reviewer (Ruy Cardoso) commented upon the potential non-linear effects. This is a much debated issue in traditional life insurance source of earnings analysis. We skipped over this issue because of its complexity. We discuss the problem here for those practicing actuaries who seek to implement source of earnings analysis at their companies.

We illustrate the problem with an example. Suppose that the developed and trended losses are $100 million. The source of earnings analysis shows that the loss development factor should have been 10% higher and the loss trend factor should have been 10% higher. A rote application of the procedure discussed in this paper would show a (negative) gain of –$10 million from development and a similar –$10 million from trend. In truth, the total variance is –$21 million, not –$20 million. When there are multiple non-linear factors in the ratemaking formula, the problems becomes more complex.

This problem is a technical one, not a conceptual one. Actuaries use three types of solutions:

1. *Assign the linear component of the variance to the individual factors, and assign the non-linear components to a general "combined" bucket.* This solution is easy, but it is unsatisfying to many actuaries.

2. *Determine the order of application of the ratemaking factors, and determine the variances by the order of application.* At first glance, this solution seems ideal. In truth, this solution is arbitrary, since there is no inherent "order" to the calculations. For example, do we trend the developed losses or do we develop the trended losses? Most actuaries choose the former because that is the order in most elementary ratemaking texts. But the latter is mathematically identical to the former, and it has as much intuitive rationale as the former does. (Cf. C. F. Cook, "Trend and Loss Development Factors," *PCAS*, Volume 57 (1970), pages 1-14.)

3. *Spread the non-linear components over the individual factors on a formula basis.* This method is the most sophisticated, but it is the most complex.

In sum, the mathematics is not as simple as one might infer from the text of this paper.

When the total variance is small, the non-linear components (or the "second order" components) are small enough that they do not affect the study. When the total variance is large, one of the above procedures should be used for the non-linear components.

Loss Drivers

One reviewer, an experienced and astute pricing actuary for private passenger automobile insurance (John Conners) has pointed out several areas where further analysis would be useful. We paraphrase one of this comments below, though Mr Conners wrote this not as a critique but as an additional subject to be treated:

> *You discuss trend for private passenger automobile. Trend factors we can deal with; that's not the problem for the pricing actuary. Our problems lie with loss development and with weather induced losses.*
>
> *Most pricing actuaries use accident year data with incurred loss chain ladder development factors. They rarely supplement their analyses with paid loss development or with examination of frequency and severity. Moreover, they often use countrywide development factors for individual states.*
>
> *Numerous factors affect these results. Estimates of ultimate losses may be distorted by internal company changes, such as changes in case reserving philosophy, as well as by external changes, such as changes in attorney involvement in auto liability claims.*
>
> *Weather related losses are a significant concern for auto pricing actuaries. Changes in weather conditions—a cold winter versus a mild winter—can affect auto liability losses. It is difficult to examine historical data and project future expected losses when weather has a large and sometimes unpredictable effect.*

Mr Conners is correct, though our paraphrase is a bit misleading. Mr Conners did not intend this as a critique of source of earnings analysis, as we pointed out above. He is saying that the traditional private passenger automobile ratemaking techniques are not optimal, since they ignore some important factors that are crucial to estimating rate needs. These problems are not picked up by source of earnings analysis.

We fully agree. Perhaps this is a happenstance of actuarial education. Philipp Stern's seminal paper on private passenger automobile ratemaking has been read and studied by actuarial students for 30 years, and subsequent papers begin with his framework. Stern uses aggregate incurred loss chain ladder development; Stern does not discuss the effects of weather. Had Stern used paid loss development, or had he examined frequency and severity separately, or had he analyzed the effects of weather, our standard private passenger automobile ratemaking techniques might be different.

This is true for all aspects of actuarial practice. We tend to think of our procedures are the "natural" method of determining our results, when in fact they are the arbitrary results of a choice made 50 years ago and never changed. Actuarial students often seek meaning in the different credibility procedures and standards used in different lines of business and areas of practice. Sometimes the only "meaning" is a rating bureau review done half a century ago, whose procedure has been repeated year after year by a continent of actuaries.

Even more surprising are the differences between casualty and life actuarial sciences. Two separate actuarial societies developed separate techniques for analyzing the same problems. Credibility theory was nurtured among casualty actuaries and has only rarely been applied in life actuarial sciences, despite its obvious applications. Similarly, source of earnings analysis developed among life actuaries, and this paper is its first application to casualty lines of business.

Premium Trend Revisited

Timothy L. McCarthy, ACAS, MAAA

Abstract

Premium trend has been an integral part of the ratemaking process. The *Statement of Principles Regarding Property and Casualty Insurance Ratemaking* lists it in its enumeration of considerations for trends. However, current models for estimating the premium trend have been limited to an exploration of changes in the base exposure. Limiting the premium trend to simply reflect changes in the base exposure can produce a biased indication, as internal loss trends implicitly reflect distributional shifts underlying the rating plan, while the exposure based premium trend fails to incorporate such changes. A methodology for determining premium trend that expands beyond the traditional methods is discussed and the theory underlying the proposed methodology is developed.

The author would like to thank Chuck Boucek, Barbara Thurston, and Greg Wilson for their assistance in reviewing this paper. Their many comments and suggestions improved it immeasurably.

Premium Trend Revisited

David Brockmeir has written a paper entitled *Homeowners Premium Trend* [1]. Brockmeir's paper discusses different methodologies for adjusting Homeowners' premiums to bring them to current coverage levels. While Brockmeir's Static II Method and Dynamic Method are improvements over the static method traditionally used in property ratemaking procedures, they still are an incomplete model for adjusting Homeowners premiums. This paper will discuss some of the weaknesses of current premium trend procedures employed in the property lines, provide an alternative method, and demonstrate how this alternative premium trend method can (and why it should) be applied to other lines of business.

A Very Simple Example

Consider the following very simple example of the normal process followed in a typical rate review. For simplicity, let us assume that these are private passenger automobile, bodily injury, basic limits, single class (i.e. adult operators) data. The basic data outlined below would first be obtained.

Table 1
Premium and Untrended Losses For Indication

Year	Premium @ Present Rates	Developed Losses	Loss Ratio
A	428,571	300,000	70.0%
B	442,857	310,000	70.0%
C	457,143	320,000	70.0%

In the table above, the losses are not trended. The following data, internal to the book of business, are used to trend the losses.

Table 2
Trend Data[1]

Year	Exposures	Paid Losses	Pure Premium
A	2,000	300,000	150
B	2,000	310,000	155
C	2,000	320,000	160

The above data generate a trend of approximately +3.3%. The losses would then have the trend applied to them and the overall (generally, statewide) indication would be developed. Table 3, below, summarizes the development of the trended loss ratio at current rates which would be used to develop the indication.

Table 3
Premium and Trended Losses For Indication

Year	Premium @ Present Rates	Trended and Developed Losses	Loss Ratio
A	428,571	341,604	79.7%
B	442,857	341,714	77.2%
C	457,143	341,468	74.7%
Total	1,328,571	1,024,786	77.1%

If the company has a permissible loss ratio of 70%, and all expenses are variable, then a +10.1% rate increase is indicated.

[1] The general standard for automobile ratemaking is to use at least twelve quarters of fiscal calendar year paid data, and regress. However, only three years calendar years are used as the example is intended to remain simple

The next step in this very simple rate review would be to perform a territorial analysis. Table 4, below, summarizes data needed to perform the territorial analysis.

Table 4
Territorial Premium and Loss Data

	--- Territory A ---			--- Territory B ---		
		Pd	Loss		Pd	Loss
Year	Premium	Losses	Ratio	Premium	Losses	Ratio
A	285,714	200,000	70.0%	142,857	100,000	70.0%
B	314,286	220,000	70.0%	128,571	90,000	70.0%
C	342,857	240,000	70.0%	114,286	80,000	70.0%
Total	942,857	660,000	70.0%	385,714	270,000	70.0%

From the data above we can see, upon application of standard loss ratio analysis, that no change is indicated for the territorial relativities.

The Problem

In the above example, each of the three years had 70% untrended, developed loss ratios at present rates. Similarly, in each year each territory had a 70% loss ratio. However, a positive +3.3% annual loss trend, developed using the data internal to the book of business has generated an overall indication of +10.1% for the line.

An examination of the territorial data in Table 4 shows that the premiums at present rates are increasing in Territory A, while in Territory B the premiums at present rates are decreasing. Table 5, below, summarizes the exposures which underly the premiums and losses in Table 4.

Table 5
Exposures Underlying Territorial Premium and Loss Data

Year	Territorial Exposures		
	A	B	Total
A	1,000	1,000	2,000
B	1,100	900	2,000
C	1,200	800	2,000

Table 5 shows that, although the total exposures have remained unchanged over the experience period, there has been a shift away from Territory B to Territory A. (One can verify readily from the data in the above tables, that the base rate used to calculate the Territory A premiums at present rates is $285.71, and that the base rate for Territory B is half that of Territory A.) The loss trend developed using the data in Table 2 is misleading. Although the average losses for the book are increasing annually, there is no economic or social trend which is driving the increase in the average loss costs. Instead, a demographic shift, which may or may not be peculiar to the company, from one territory to another, is driving the change in the average loss costs. Since the base rates anticipate the cost differentials which exist between these territories the formula for developing the rate level indications must be revised, so that an overall loss trend which is reflecting only a demographic shift does not drive the indications.

A Discussion of the Alternatives

Most actuaries may feel somewhat insulted by the example, and state that they would recognize such a shift, and make allowances for it. However, the preceding section where the indication is developed is entitled "A Very Simple Example" for good reason. In the real world, pricing is not as simple as that which is shown in the example. Random variation in the losses frequently obscures relationships like those which are so readily evident in this simple example. Rarely will one encounter a situation wherein there is absolutely no social or economic inflation impacting a line of insurance like that shown in the example. Additionally, business pressure to complete one rate review, and move on to the next one, can create an environment wherein time constraints inhibit both the discovery and exploration of shifts akin to that created above.

Can an actuarial model be developed that accounts for the bias in our ratemaking model caused by the shift in the distribution of exposures by territory?

Three alternatives are readily available:

1. **Eliminate Loss Trends from the Indications**

   This is not a realistic option. Although "A Very Simple Example" has shown that bias can be introduced into the indication with a simple distributional shift, the complete elimination of the loss trend to address such shifts replaces a ratemaking process with one bias with a ratemaking system with a new bias. That is, if inflationary and / or demand shifts are occurring in the book if business, then an indication that ignores these economic forces is a biased indication, albeit with a different bias. Additionally, the proposal to eliminate loss trends from the ratemaking process is a non-starter as it is contrary to the "Statement of Principles Regarding Property and Casualty Ratemaking".

2. **Temper the Loss Trends by Reflecting Known Shifts which Impact Losses:**

   This would appear to be a desirable solution. That is, as in the case of the Very Simple example, one could weight the denominator of the pure premiums with the territorial relativities. However, the complexity of this solution makes it undesirable.

   First, these indications were developed with pure premium trends only. Suppose one wishes a more complete analysis of the loss trends through the exploration of both the changes in the inflationary impact (severity trend) and the changes in the demand impact (frequency trend) on the book of business.[2] This would require recognition of both the severity and frequency impacts in the development of the territorial relativities. Thus, to completely analyze severity one would need a severity based relativity to weight the claims, and a frequency based

[2] Diamantoukos [2] recognizes the need for separate analysis of exposure related to both frequency and severity is his discussion of Bouska. "The best solution to approximating the true exposure in some cases might be to utilize more than one exposure base. Two exposure bases might be used, one for frequency and the other for severity."

relativity to weight the exposures, a very complex requirement.

Second, the Very Simple Example considered the book of business to be written *only* for a single class of insureds. That is, our simplifying assumption was this book was comprised of basic limits, single class data. If a book of business has a mix of classifications / dimensions, then relativities that are both frequency and severity based need to be available by all the classifications / dimensions.

The need to weight the denominator of the loss trends with frequency and severity based relativities make the appropriate adjustment to the loss trends quite problematic, making this alternative undesirable.

3. **Don't Adjust the Loss Trend, but Develop a More Sophisticated Premium Trend, which Reflects All Premium Related Changes:** The territorial relativity is part of the premium charge. The intent of introducing a premium related trend into the indication, which reflects the territorial differences in our example, is to provide a financial statistic to state our premium and loss projections on a more common level. Such an adjustment should eliminate, or significantly reduce, the bias seen above. As premium trend is a consideration explicitly enumerated in the "Statement of Principles Regarding Property and Casualty Ratemaking", the inclusion of such a trend in indications

previously lacking premium trend, or an improvement upon current premium trend procedures, result in indications that better comply with ratemaking principles.

A More Sophisticated Premium Trend

It was proposed above that incorporating a more sophisticated premium trend procedure into the ratemaking process will accomplish the desired effect of eliminating the bias in our indication.

The proposed premium trend is simply the average premiums at present rates. The use of the average premiums at present rates eliminates much of the bias in the indication. A theoretical justification for its use is provided later in the paper.

The use of this procedure is a departure from traditional methodologies. Traditionally exposure trends have been used as a surrogate for premium trend. Bouska [3] requires that the exposure base have a continuous, linear, multiplicative relationship to the losses. Homan [4] develops a premium trend which considers only amount of insurance which is comparable to the Static Method I discussed by Brockmeir. Chernick [5] reflects in Private Passenger Auto Physical Damage the "linear" nature of model year. Feldblum [6] discusses the impact of changes to payroll in developing a premium trend for

Workers Comp. While for General Liability, Graves and Castillo [7] discuss the use of payroll and receipts in premium trend.

The actuarial goal of developing a premium trend with each of the above is to account for the changes to income emanating from a change in the exposure base. However, the use of the term "Premium Trend" is a misnomer, insofar as it fails to account for the impact of variables that act upon both historic and projected collection of revenue. That is, to the extent that the premium trend fails to recognize the impact that changes in the distribution of insureds across classifications has upon the collection of income (i.e. premium), it fails to adequately represent the change to premium.

In discussing the exposure base, Bouska commented upon what the exposure base was not. She stated that the exposure base was not a rating variable, and noted that, unlike the exposure base, a rating variable has a discrete, non-linear relationship. If these discrete, non-linear relationships are changing over time, then failing to account for them in lines where a traditional exposure based premium trend is included in the indication, generates a biased premium trend, and, hence, a biased indication. In "A Very Simple Example" we also see that failing to account for the change in the distribution of discrete, nonlinear relationships in a line of insurance that has not incorporated premium trend in its development also generates a biased indication.

Let us close this section with a restatement of the Very Simple Example using a premium trend based upon the average premium at present rates. Table 6, below, summarizes the average premiums at present rates.

Table 6
Average Premiums at Present Rates

Year	Prem @ Prst Rate	Exposures	Avg Prem @ Prst Rt
A	428,571	2,000	214.29
B	442,857	2,000	221.43
C	457,143	2,000	228.57

The average premium at present rates increases annually at +3.3%, which should not be surprising, given that the loss trend was identical and driven solely by the territorial distribution shift. The revised loss ratios for the indication, using the premium trend, are developed below in Table 7.

Table 7
Revised Loss Ratios Using Premium Trend

Year	Trended Prem @ Prst Rates	Trended and Developed Losses	Loss Ratio
A	488,005	341,604	70.0%
B	488,163	341,714	70.0%
C	487,812	341,468	70.0%
Total	1,463,980	1,024,786	70.0%

Given the previously stated permissible loss ratio of 70%, no change is indicated.

A Theoretical Justification

The above resolution of the indications developed from "A Very Simple Example" produces a result that is intuitively appealing, but does not provide a rigorous justification for the use of average premiums at present rates for premium trend.

To provide a justification, we return to our initial example. Let us break down the components of the total premium the year *i* in territory *j* (P_{ij}) into its basic components.

Let

x_{ij} = *The exposure for year i in territory j;*
t_j = The current territorial relativity for territory *j*;
n = The number of territories; and
r = The current base rate

Then the premium in year *i* for territory *j* is

$$P_{ij} = r * t_j * x_{ij}$$

The total premium in year *i* is

$$P_i = \sum_{j=1}^{n} r * t_j * x_{ij} \qquad (1)$$

In "A Very Simple Example" there is no inflation over time, and the territorial relativities are perfectly priced. We define the loss cost drivers as follows:

Let

f_{ij} = The frequency in year i for territory j; and
s_{ij} = The severity in year i for territory j.[3]

Thus, the losses in year i for territory j (L_{ij}) are

$$L_{ij} = f_{ij} * s_{ij} * x_{ij}$$

Then the total losses for year i (L_i) are defined by the equation

$$L_i = \sum_{j=1}^{n} f_{ij} * s_{ij} * x_{ij} \qquad (2)$$

In our simple world with no inflation, let us assume that our territorial relativities are defined using pure premiums (i.e. there are no fixed expenses). Let z represent the base territory. Then the territorial relativity for territory j is defined by

$$t_j = \frac{f_j * s_j}{f_z * s_z} \qquad (3)$$

Let us now incorporate (3) with (1) and (2) to develop a loss ratio for a year i.

$$L_i \;/\; P_i = \sum_{j=1}^{n} (f_{ij} * s_{ij} * x_{ij}) \;\Big/\; \sum_{j=1}^{n} r * [(f_{ij} * s_{ij}) / (f_{iz} * s_{iz})] * x_{ij} \qquad (4)$$

[3] In the example, the loss costs do not vary by year, so that at this juncture differentiating the frequency and severity by year is superfluous; however, as we expand upon the example, the ability to differentiate frequency and severity by year will become important.

Because we are considering only a single year's experience, we can drop the *i* subscript, and (4) can be simplified to

$$L_i \ / \ P_i = (f_z * s_z) \ / \ r \qquad (5)$$

Equation (5) is simply the base territory's pure premium divided by the base rate. This is the permissible loss ratio, which comports well with the data seen in "A Very Simple Example". That is, in our non-inflationary environment where the product is properly priced by territory, we would expect that the loss ratio would equal the permissible loss ratio.

Now, consider the impact of our earlier application of loss trend in "A Very Simple Example". Let us, for simplicity, consider the loss trend as a one year ratio of the pure premiums. Then we define our loss trend factor, q, where x_i represents the total annual exposures, as

$$q = \left[\left(\sum_{j=1}^{n} (f_{(i+1)j} * s_{(i+1)j} * x_{(i+1)j})\right) / x_{(i+1)}\right] \Big/ \left[\left(\sum_{j=1}^{n} (f_{ij} * s_{ij} * x_{ij})\right) / x_i\right]$$

But frequency and severity in our non-inflationary world are the same for the years i and (i + 1), so that our loss trend simplifies to

$$q = \left[\left(\sum_{j=1}^{n} (f_j * s_j * x_{(i+1)j})\right) / x_{(i+1)}\right] \Big/ \left[\left(\sum_{j=1}^{n} (f_j * s_j * x_{ij})\right) / x_i\right] \qquad (6)$$

If we multiply (6) by unity in the form of [$(r/\,f_z s_z)$] / [$(r/f_z s_z)$] both the numerator and the denominator are converted to premiums at present rates. Using the relationship in (3) we define our premium trend factor, *p*, for simplicity, as a one year ratio of premiums at present rates.

$$p = \left[\left(\sum_{j=1}^{n} (r * t_j * x_{(i+1)j})\right) / x_{(i+1)}\right] \Big/ \left[\left(\sum_{j=1}^{n} (r_j * t_j * x_{ij})\right) / x_i\right] \quad (7)$$

Thus, after multiplying the loss trend by unity, we obtain the premium trend and can see that in this non-inflationary environment the loss trend equals the premium trend. Given our simplifying assumptions about the book of business, our premium trend simply reflects the change in average territorial relativity. In this non-inflationary environment when we apply the loss trend (6) to the loss ratio (5), we can now see that we are simply adjusting the permissible loss ratio to reflect the change in the average territorial relativity. A biased indication results when we fail to adjust the premiums for this change in average territorial relativity.

A Return to the Real World

"A Very Simple Example" had four assumptions that are not encountered in the real world.

1. Products are priced with more than one rating variable.

2. Products are not perfectly priced to begin with.[4]
3. Loss trends do occur in the real world.
4. Random fluctuation occurs.

With regard to multiple rating variables, the equations above can be adjusted to account for them, and the resultant multiple summations yield the same result.

With regard to products not being perfectly priced, we can introduce an error component, e_j, into our loss equation (2). This error component varies by territory, under the assumption that each territory has the potential to be inaccurately priced. Note, however, that it does not vary by year, since the current relativities are being applied to bring the premiums to present rates.

To account for loss trends we introduce the loss trend component, q_{ij}, into loss equation (2). This trend component accounts for both the demand and inflationary changes. If one wanted to account for frequency and severity changes independently, then two such components could be introduced. Note that we have allowed the demand and inflation changes to vary both by year and territory. From experience we know that inflation is not constant over time and that it can vary regionally.

[4] Some might argue that this assumption is not true, working for a company that always takes its fully credible indicated rates. But even a company that does such in the competitive insurance market place still is not guaranteed it charges the "correct" rate. For a fully credible rate using the classic 1,082 full credibility standard is still not "perfect". The company adopting such a rate is still only 90% confident that is within 5% of the correct rate.

Random fluctuation in the losses can be accounted for with the component R_{ij}.

Thus, we can now restate the losses for year *i* in terms of components that occur in the real world.

$$L_i = \sum_{j=1}^{n} R_{ij} * q_{ij} * e_j * f_j * s_j * x_{ij} \quad (8)$$

The frequency and severity components in (8) represent the *a-priori* average loss costs that would be assumed when the initial rate review began, and thus do not vary by year. Dividing two consecutive years of average losses produces the change in average annual losses. One would face an enormously complex task if required to develop the random error, trend, and pricing components independently. However, if one were to divide the ratio of the average annual losses by the change in average premium, then one would have the change in losses not attributable to premium trend, or, a pure loss trend. Thus we can define the pure loss trend, for a year[5], *Q*, as the ratio of the loss trend divided by the premium trend. Thus, using our knowledge that f_j and s_j do not vary by year and the relationships implicit from (6) we obtain

---

[5] This concept of a "pure loss trend, for a year" is an extremely poor term. A trend cannot be developed using two years of data. This term has been created for illustrative purposes only for the ensuing equation. Realistically, the pure loss trend would be developed by dividing a selected loss trend, developed using more than two points, by a premium trend developed using multiple years of points.

$Q = [(L_{(i+1)} / x_{(i+1)}) / (L_i / x_i)] \;/\; [(P_{(i+1)} / x_{(i+1)}) / (P_i / x_i)]$

$$= \frac{\left(\sum_{j=1}^{n} R_{(i+1)j} * q_{(i+1)j} * e_j * f_j * s_j * x_{(i+1)j}\right) / \left(\sum_{j=1}^{n} R_{ij} * q_{ij} * e_j * f_j * s_j * x_{ij}\right)}{\sum_{j=1}^{n} \left(f_j * s_j * x_{(i+1)j}\right) / \sum_{j=1}^{n} \left(f_j * s_j * x_{ij}\right)} \qquad (9)$$

Before the premium trend procedure was introduced, a potential alternative for addressing the bias in "A Very Simple Example" was to *"temper the losses by reflecting known shifts which impact losses".* Dividing the loss trend by the premium trend provides such a tempered loss trend (9), our pure loss trend.

Thus, the use of premium trend developed using average premiums at present rates provides two important tools to the actuary and company management. First, analysis of the premium trend provides information on how the income stream is being impacted by distributional shifts. Second, dividing a traditionally developed loss trend by the premium trend, provides information on how average loss costs are changing independent of distributional shifts accounted for in the premium.

Before closing out this section, let us consider an additional advantage to using the average premiums at present rates for premium trend. Homan develops Homeowners indication using the $100 deductible. For other lines of insurance, such as private passenger physical damage, this procedure is also employed. An advantage to the recommended procedure is that it eliminates the need to either adjust premiums and losses to a common deductible or examine only a particular deductible's experience. The inclusion of all deductible experience eliminates a bias in the overall rate level indication. Implicit in the use of a single deductible is the inference that the deductible relativities are correct. If this inference is incorrect, then the overall indication is biased by the error in the deductible relativities[6].

Assume, for example, that the higher deductibles are inadvertently underpriced. Then a statewide indication developed at the $100 deductible would understate the overall rate need. While the appropriate action would be to analyze the deductibles (and, for that matter, every rating variable) at every review to ensure proper rating, time constraints frequently prohibit such analysis. Thus, the recommended premium trend procedure, while creating subsidies within rating variables, still enables the overall correct premium to be developed.

The Trend Period

In Homan, the premium trend is extended to the average day of writing for the period in which the rates will be in effect. Under the proposed methodology, the premium trend has been calculated using the average earned premiums at

[6] If pricing analysis is performed with low deductibles, this is even more problematic, since large deductibles are subject to significant leveraging of their losses in an inflationary environment (Hogg and Klugman [8]).

present rates. For consistency, the trend is extended from the average date of earning in the experience period to the average day of earning in the period for which the rates will be in effect.

Written premiums could be used for the determination of the premium trend. In this example, the earned premiums at present rates have been determined by multiplying the earned exposures by the current base rates and relativities. The primary disadvantage to using the average written premiums at present rates is the need for additional calculation, if the earned premiums at present rates are calculated using the earned exposures.[7] If written premiums at present rates are calculated, and used to determine the premium trend, then the trend period extends from the average day of writing in the experience period to the anticipated average day of writing for the period for which the rates will be in effect.

Measuring Premiums at Present Rate Levels

We have seen how the use of premiums at present rates can produce a more accurate premium trend. The methodology to develop the premiums at present

[7] This need for additional calculation stems from the inherent mismatch that occurs in any comparison of earned premiums to written exposures. That is, the rate of change is a function of the exposure base, and in a dynamic book of business the rate of change measured using written exposures can be different from the rate measured using earned exposures. If the rate of change is different between the written and earned exposure bases, additional calculations are required to both approximate this difference in the rate of change, and adjust written rate to an earned rate. Of course, if the earned premiums at present rates are calculated using written exposures that are re-rated and then the written premiums are earned no additional calculation is needed.

rates should be the extension of exposure technique.

It is still common for premiums to be brought to present levels using the geometric or parallelogram technique described by McClennahan [9]. Using average premiums at present rates developed with the geometric method to determine the premium trend produces distorted results.

The parallelogram method presupposes an even distribution of writings. Historically, when this even distribution of writings has been discussed it has been more in reference to the timing of when risks are written (i.e. whether they are written evenly throughout the year or whether seasonality impacts the level of writings). Implicit in the even distribution of writings presupposed in the parallelogram technique is the even distribution of the types of policies being written.

When the rate changes that underlie the parallelogram technique are measured, they are measured against the distribution of risks in effect at the time of the proposed rate change, frequently the most recent year's written premiums at present rates. The parallelogram technique applies that change to past writings, with the application of an overall change to the premiums implicitly assuming that the distribution of rating variables in effect historically is the same as the distribution of rating variables when the rate change was measured. If any rating variable distributional shifts are occurring (e.g. amount of insurance, class, deductible) then the average premiums at present rates developed with the parallelogram technique fail to capture the complete nature of the change to the income stream.

If the resources are unavailable to completely rerate a book of business using the current rate manual, then the change in current average premium relatitivities can be applied to the average premiums at present rates developed with the parallelogram technique to obtain a more complete approximation.

Loss Trend Issues

Let us consider for a moment the traditional property ratemaking procedure. In a traditional property ratemaking process the premium trend has been limited to an exploration of changes in the dwelling coverage amounts (Brockmeir and Homan). That is, the ratemaking methodology has considered the change in the exposure base to be the sole contributor to changes in premium income. The loss trend has relied almost exclusively on external trends, specifically, the Boeckh and modified CPI external indices. Since inflationary pressures are driving the changes in coverage amounts, the traditional ratemaking procedure has contemplated the need to balance the changes in premium income with the changes in losses as measured with external indices[8]. However, to the extent

[8] Note, however, that even if the external indices measure cost inflation perfectly, they still may not be a reliable surrogate for loss trends.

that non-coverage amount[9] distributional shifts are impacting the premium income, the traditional property ratemaking process may be biased.

Consider a book of business that, in addition to coverage level changes, is experiencing a shift to higher deductibles, lower cost territories, improved protection classes, and a construction-type shift from Frame to Brick. Each of these shifts has a downward impact on premium income. The traditional premium trend calculated using the changes in coverage amounts (i.e. tempered or non-tempered amount of insurance relativity changes) should be modified by the changes in each of the aforementioned relativities. However, if the loss trend is not also modified with these non-coverage related relativity changes, then the indication would be overstated.

Assume that the non-coverage related relativities each produce –2% annual

---

First, they fail to account for changes in demand / frequency. That is, they do not reflect how the claims process is impacted by increasing or decreasing claims consciousness on the part of the insurance consumer.

Second, the indices do not necessarily reflect claims inflation, but overall inflation. For example, the Boeckh index is used to measure how <u>total</u> building costs are changing over time. In property insurance few claims are total losses. Thus, there is a potential mismatch between the claims process, which is driven by partial losses, and the external index, which addresses total building costs. Consider a state where the predominant type of claim is roof losses from hail If the components of roof construction are inflating more rapidly than the other component costs of building, then the loss trend would be understated, as equal weight is given to the roof and non-roof components in the external index, but the claims process is more heavily weighted with roof material purchases.

Finally, they fail to account for rate related distributional shifts.

[9] Coverage Amounts are considered to be the dwelling face amount throughout this discussion, because the traditional premium trend procedure relates solely to the dwelling face amount. Note that although the deductible impacts the amount of coverage an insured possesses, for purposes of this discussion deductible is not included under the "coverage amount".

change in premium. Then their combined effect reduces premium income –7.8% annually. As mentioned above, applying this change to premium trend, but providing no similar adjustment to the loss trend will overstate the indicated rate change. If one assumes that the current rate relativities are correct, then one can apply the change in the non-coverage related relativities to the loss trend to bring the projected premiums and losses to a distributional balance.

Under the assumption that offsetting the selected external loss trend indices with the change in the non-coverage amount relativities, one readily sees that these changes to the relativities offset one another in both the premium and the losses. One might argue that the beauty of the traditional property ratemaking procedure lies in the simplicity of not needing to concern the actuary with the changes in the non-coverage amount relativities. This would be an incorrect assessment on two counts.

First, failing to recognize the non-coverage amount related premium and loss trends can produce a biased indication, even if these non-coverage relativities are perfectly priced. In our example, the non-coverage amount related premium trend is negative. The fixed expenses remain unchanged despite this perfect pricing. Thus, an indication using a fixed expense loading would be understated if the premiums and losses were not adjusted as described. Contrariwise, if the non-coverage related premium trend were positive an excessive rate would be developed without the described adjustment.

Second, the external indices may well provide inaccurate estimates of changes to insurance losses, as they fail to consider changes in demand, the mismatch between the general housing inflation rate and the insurance claims indemnification process, and changes in the distributional mix. Thus, the use of the external indices, while simplifying the ratemaking process, will produce biased indications to the extent that they are an inaccurate surrogate for claims inflation.

Let us consider further the impact of applying the change in the non-coverage amount premium relativities to the external loss trend. The assumption made in applying these changes was that the product was perfectly priced. However, we introduced an error term, e_j, to our loss function, equation (8), because we know that products may not be perfectly priced. Let us assume that our protection class relativity in this property example is inaccurately priced, and that if correctly priced, the change in the protection class relativity would be –3%, rather than the –2% reflected in our current premium trend. Ignoring the random error term in equation (8), if internal loss trends were used, then they would reflect the –3% annual change in loss costs from the protection class distributional shift. That is, the impact of the protection class shift would cause the losses to deflate more rapidly than the premium. Applying the protection class portion of the premium trend to losses assumes that the losses and the premiums are deflating at the same rate, and will produce an excessive indication.

The preceding paragraph raises a critical issue associated with both the use of internal loss trends and our adjustment to external loss trends, and merits additional amplification. In "A Very Simple Example" there was no loss trend outside the distributional shift by territory. Additionally, the territories were perfectly priced. Consider the component, (e_j * f_j * s_j * x_{ij}), from our loss equation (8), where the perfect pricing assumption was eliminated. In successive periods, this component measures the change in losses due to both the change in the distributional shift from the assumed underlying frequency and severity of j (i.e. the assumption that the product is perfectly priced) and the error in the estimated frequency and severity (i.e. the recognition that the perfect pricing assumption is violated). That is, the internal loss trends measure much more than the inflation impacting the claims process. In addition to measuring the impact of inflation upon the book, the internal loss trends measure the change in average loss costs due to distributional shifts, without regard to how properly these rate-related distributions are priced. The use of external trends fails to consider how changes in the rate-related distributions are impacting the losses, and even with the proposed adjustment, the external indices will not capture changes in losses due to incorrect pricing.

We have outlined problems that exist with the use of external data to use as loss trends in property ratemaking. When the proposed premium trend procedure is applied to the premium, the need to reflect the changes in losses due to distributional shifts introduces additional difficulties. Thus, it would appear that internal data would provide a more accurate indication. Traditionally, external loss trends have been used in the property ratemaking process, in part, because catastrophic claims make internal data difficult to use. McCarthy [10] provides a method through which catastrophes can be removed from property data, enabling the development of loss trends using data internal to the book of business.

The adjustment to the premium trend procedure has applied the non-coverage amount distributional changes to both the premiums and losses. This was done under the assumption that the coverage related premium changes and the external loss trends were measuring similar issues related to inflation. This requires a two-step premium and loss trend procedure. An additional advantage to using internal loss trends is that the premium trend and loss trend procedures are simplified

Observations

Brockmeir's Static Method II and Dynamic Method provide more accurate methodologies for ascertaining the impact to premium of coverage level (i.e.

amount of insurance) changes in the Homeowners line of business. However, they fail to measure the impact to premium of changes to non-coverage related premium-affecting distributions. The proposed methodology provides for a more complete analysis of all the components related to premium.

When one considers the results of personal lines products in the second half of the 1990's, a question arises as to if some of the profitability issues associated with these lines throughout the latter half of the decade are related to a biased ratemaking process. One cannot ignore the impact that increased catastrophes have had upon Homeowners, but even when catastrophes are removed from the experience Homeowners profitability has still lagged Automobile's.

Have population demographic shifts exacerbated Homeowners profitability? If we assume that the growth in Homeowners has come in more urban, highly protected areas, and if the premium projections within the indications failed to account for the reduced revenue associated with this shift, then the indication's projected loss ratios would be understated, and the indicated premiums inadequate. The same population shift would generate results in Auto somewhat akin to those seen in "A Very Simple Example". That is, in Auto the indication's projected loss ratios would be overstated, resulting in excessive indicated rates.

The population demographic shift is but one example of how a ratemaking procedure that fails to account for shifts in average premiums at present levels can result in inaccurate indications. It, by no means, provides a complete explanation of why we see the divergence in these personal lines results as the twentieth century comes to a close.

In Workers Comp and General Liability, adjustments to the premium trend to reflect the changes in classification changes over time at the current rates may result in more consistent internal loss trends. Graves and Castillo note that the ISO is developing external indices to be used in General Liability indications due to a dissatisfaction with the internal data. A more sophisticated premium trend procedure, which accounted for the classification, territorial, and exposure changes may produce more stable and reliable pure loss trends, reducing the need for the external indices.

Although the premiums used in the loss ratio trends for Workers Comp are brought to present rates using the parallelogram method, some of the value of this trend may stem from a partial reflection of changes to classification distributions, that are not part of the general loss trends generally developed for the line. Indeed, the Loss Ratio Trend in Workers Comp resembles the pure loss trend introduced in this paper.

Conclusion

While the model presented should result in the development of more accurate rate level indications, it still is incomplete. The model assumes that all expenses are variable. To the extent that rate relativities reflect fixed expense loadings, there is still a mismatch between the prospective premiums and the prospective losses upon which a rate level indication is based. Additionally, the analysis has used the same exposures for the loss trend as that for the premium trend. The use of the same exposures essentially assumes that the losses are paid immediately upon occurrence, ignoring timing issues associated with the loss payments.

Premium trend has been an integral part of the ratemaking process. However, current models for estimating the premium trend have been limited to an exploration of changes in the base exposure. Limiting the premium trend to simply reflect changes in the base exposure can produce a biased indication, as the loss trends implicitly reflect distributional shifts underlying the rating plan, while the premium trend fails to incorporate such changes. The proposed methodology for developing premium trend is a theoretically sound approach to redress such mismatches.

References

[1] Brockmeier, D., "Homeowners Premium Trend", Casualty Actuarial Society *Forum*, Including the Ratemaking Call Papers, 1994, pp. 439-484.

[2] Diamantokous, C., "Discussion of Paper Published in Volume LXXVI – Exposure Base Revisited," *PCAS* LXXVIII, 1991, pp. 201-214.

[3] Bouska, A., "Exposure Base Revisited," *PCAS* LXXVI, 1989, pp. 1-23.

[4] Homan, M.J., "Homeonwers Insurance Pricing", Pricing Issues, Casualty Actuarial Society *Discussion Paper Program*, 1990, Vol. II, pp. 719-780.

[5] Chernick, D., *Private Passenger Automobile Physical Damage Ratemaking,* CAS Study Note Reading

[6] Feldblum, S., "Workers' Compensation Ratemaking," September 1993, CAS Study Note Reading

[7] Graves, N.; and Castillo, R., "Commercial General Liability Insurance Ratemaking for Premises and Operations", *Pricing Issues*, Casualty Actuarial Society *Discussion Paper Program*, 1990, Vol. II, pp. 631-696.

[8] Hogg, R.V.; and Klugman, S.A., *Loss Distributions*, 1984.

[9] McClenahan, C.L., "Ratemaking," *Foundations of Casualty Actuarial Science* (Third Edition), Casualty Actuarial Society, 1996, Chapter 2, pp. 25-90.

[10] McCarthy, T., "A Frequency Based Model for Excess Wind in Property Ratemaking", *Casualty Actuarial Society Forum*, Winter 1998, pp. 209-240

The Impact of Personal Credit History on Loss Performance in Personal Lines

James E. Monaghan, ACAS, MAAA

Introduction

At the time of this writing, a process of both education and debate is occurring with regard to the use of personal credit history in the underwriting or rating of personal lines insurance policies. The insurance industry, the NAIC, and other interested third parties are all involved in educating both themselves and each other on such issues as correlation, multivariate correlation, causality and the social or actuarial appropriateness of using this tool in either underwriting or rating. Although the scope of regulators is more finely focused on rating, the recent trend towards tier rating and the utilization of multiple rating companies by members of the insurance industry has blurred the distinction considerably between the two. The use of personal credit history in personal lines insurance has therefore, through its manifestation in underwriting, gone largely unnoticed until recent years. The rapid increase in its use has brought credit history to the forefront of debate in many jurisdictions, in addition to its use in quasi-rating schema.

The development and use of third-party scoring algorithms for credit evaluation, and the proprietary nature of such models, has made it difficult for regulators, companies, agents and customers to get a firm grasp of the underpinnings of automated risk evaluation based on credit history. Apparently, it is not only actuaries who occasionally take the position that "if I can't touch it, is it actually real?" The key issues under debate are the existence (or non-existence) of a correlation between past credit history and expected loss levels (and which variables are responsible for that correlation) and the establishment of causal links for such correlation. Both will be addressed here, although only the former can be statistically analyzed. Causality will be addressed on an informational (and necessarily subjective) basis. The key questions that will be addressed in this paper are:

1) Is there a correlation between credit history and expected personal lines loss performance?
2) If so, which specific criteria within a credit file are indicative of abnormal loss performance (favorable or unfavorable)?
3) If this correlation exists, is it merely a proxy, i.e., is the correlation actually due to other characteristics (which may already be underwritten for or against, or rated for)?
4) As a corollary to 3), are there dependencies between the impact of credit history on loss performance and other policyholder characteristics or rating variables?
5) What are the ramifications of utilizing such data for underwriting and/or ratemaking?

Research Database Construction

The data utilized in researching the relationships between credit history and private passenger automobile loss experience was assembled from several sources. All policies originally written during calendar year 1993 were first identified. Earned premiums for the calendar/accident years 1993 through 1995 were then appended for all coverages. The longest exposure period for any given policy is therefore 36 months, in the case where the policy was written on January 1st, 1993 and remained inforce through December 31st, 1995. All policies were included in the database, regardless of whether or not they remained inforce through the end of the experience period, making the shortest possible exposure period for any given policy one day. Hence policies are not homogenous in either length of exposure or in coverages afforded. Also of note is the fact that the company did not utilize credit information in underwriting or rating of policies during this time period.

Incurred losses were then added, where incurred loss was defined as the sum of paid losses, case reserves, supplemental reserves on case (which are established to cover adverse development on known losses), loss expenses and salvage and subrogation recoveries. These losses were evaluated as of June 30th, 1996 for the exposure period January 1st, 1993 through December 31st, 1995. Incurred losses during accident year 1993 therefore had 42 months of development, those during accident year 1994 were developed 30 months, and those during accident year 1995 were developed 18 months. All earned premium and incurred loss were determined at the policy level, i.e., accumulated for all vehicles insured on the policy at any time during the experience period and for all coverages afforded on those vehicles.

Data was then appended to each policy record that defined the underwriting and rating characteristics of the policy at the time of initial writing. This dataset contained such information as number of drivers, number of vehicles, prior accident and violation activity, state of residence, residence type and stability and prior insurance carrier information. Some of these variables certainly would have changed value during the experience period for many risks. In order to provide predictive value, information was compiled which related to the conditions in effect at the time of writing.

The dataset was sent to a national credit vendor to append archived credit histories for each match that could be found. These credit histories were retrieved from credit files archived at the time each policy was written (or at the nearest three-month interval). Each record was then stripped of any identifying information (i.e. policy number, name, address) in order to ensure compliance with the Fair Credit Reporting Act. This action permitted analysis of the data without knowledge of the identity of any individual risk. Again, in order to provide predictive value, information gathered was pertinent to the conditions in effect at the time each policy was originally written. The credit information added to the dataset contained all of the information in the insured's credit file. The original listing of policies contained approximately 270,000 records. Matches were obtained on approximately 170,000 of those. This "hit rate" is rather low; recall, however, that many of the policies were no longer actively insured by the company and address and other information could have been outdated.

Queries were then constructed and run against this database, accumulating earned premium and incurred loss during the experience period for various combinations of policy characteristics. In fact, thousands of such queries were run, evaluating the loss ratio and loss ratio relativity of given subsets of data relative to others and to the whole. These subsets each contained one or more variables from the two groups underwriting/rating characteristics and credit characteristics. The database had a grand total of $394 million in earned premiums for all records combined. The results of these queries, and the conclusions that could be drawn from them, shed light on the startling foundations of the credit scoring models: the individual credit characteristics. A data dictionary containing the description of all fields utilized in the results contained herein can be found in the Appendix.

Limitations and Difficulties

The construction of the database caused some inherent difficulties in interpretation and also rendered most traditional ratemaking methodologies unusable. The dataset was not compiled with the intent of applying ratemaking methods and principles. Since the process of risk selection occurs on a policy basis, the data was compiled to be utilized in that setting; loss ratio relativity is the only meaningful measure of performance expected to arise from these data.

The credit file utilized was associated with one individual, although many policies have more than one covered driver. This individual was the named insured. The named insured may or may not have been the individual involved in prior accident or violation events, and may or may not have been involved in subsequent losses during the experience period. This difficulty arises from the use of policy level data. The question remains unanswered as to what kind of loss experience one can expect from, for example, a married couple with significantly different credit histories (as can be expected with policies written on recently married persons).

Another difficulty encountered was determining the appropriate method of binning the data, particularly where the independent variable was of the continuous type (dollars, for example). Any data grouping of a continuous variable will have greater stability when larger bins are employed. Many different bin groupings were used in such cases, although only one will be shown here for each example.

Results of Data Queries

The database contained a large number of variables relating to underwriting characteristics, rating characteristics and credit information. Space limitations preclude presenting information about most of the queries that were run and results obtained. A sampling of this data will be reviewed and discussed. The first section will contain information about individual credit characteristics. All earned premium and incurred loss dollars will be shown in millions unless otherwise specified. The aggregate loss ratio for the entire database is 76.3%; this number is higher than average for the private passenger auto industry but recall this is premium and loss experience during the first (at most) 36 months of experience from a block of newly written policies. New business in general produces higher loss ratios than longer-tenured business.

1. Amounts Past Due (APD)

APD is defined as delinquent amounts that are uncollected as of the report date. This amount is the sum of all delinquent amounts on the credit file, regardless of how many accounts are delinquent. A scheduled payment must be at least 30 days late before it appears on the credit file as delinquent. Note that there is a significant amount of premium volume in the categories below $10. This is due to a logistical difficulty with the data: some records contained the value $0, others were blank. In order to run queries, the data must be uniformly formatted, yet there could have been statistically significant differences in results for "blank" versus $0. Therefore, all records with blanks were assigned a value of $1. The premium and loss dollars in the categories below $6-20 should be considered included with $0.

APD	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio	Fitted Relative LR
$0	$ 257.7	$ 180.9	70.2%	0.92	
1-2	45.8	31.8	69.3%	0.91	1.03
3-5	6.5	4.9	75.9%	1.00	1.07
6-20	4.7	4.4	94.0%	1.23	1.11
21-50	5.5	4.8	87.5%	1.15	1.16
51-99	5.8	5.8	99.7%	1.31	1.19
100-199	7.7	7.3	95.9%	1.26	1.22
200-499	12.0	11.1	92.7%	1.22	1.25
500-999	10.2	10.9	107.2%	1.41	1.28
1K-2K	10.1	9.9	97.2%	1.27	1.31
2K-5K	12.5	12.6	100.5%	1.32	1.35
5K-10K	7.8	8.3	106.1%	1.39	1.38
10K +	7.7	7.6	99.8%	1.31	1.41
Total	$ 394.0	$ 300.4	76.3%	1.00	

A linear regression performed on loss ratio relativity vs. logarithm of APD generated a coefficient of 0.83. The t-statistic for 99.5% significance level with 10 degrees of freedom is 3.17; the t-stat for this dataset is 5.65. Thus the null hypothesis that slope of the regression is 0 is rejected with 99.5% certainty. A less statistical observation would be that loss ratio increases as the APD increases, but the change is very small compared to the large jump in loss ratio from around 70% for $0 to the mid-nineties at almost any value greater than $0. This is somewhat counter-intuitive, as one might speculate that small delinquencies should not have the same impact as large ones. Recall, however, that what is being measured is impact on loss ratio, not credit worthiness or any other characteristic. Since the causal links are not established, preconceived notions should be considered with skepticism.

2. Derogatory Public Records (DPR)

DPRs include such items as bankruptcies, federal, state or municipal tax liens, civil judgments and foreclosures. The presence of a DPR on a credit file also has significant impact on future loss performance. This should come as no surprise, as this variable is the one that has been utilized in the personal lines industry for the longest time and is the most widely accepted.

DPR	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio	Fitted Relative LR
None	$ 358.6	$ 264.7	73.8%	0.97	1.04
1	22.4	21.6	96.5%	1.27	1.18
2	7.1	7.4	104.2%	1.37	1.33
3 or more	5.9	6.7	114.1%	1.50	1.54

Linear regression on number of DPR vs. relative loss ratio generated an R^2 value of 0.95. The loss ratio for all DPR that had an outstanding liability on the file of greater than $0 is 102.2%, (relativity = 1.34) and premium volume of $31.1. Although many will not be surprised that there is a correlation with this variable, the size of the difference in loss ratio may confirm the underlying reason for its historic use.

3. *Collection Records*

A collection record is generated when responsibility for collecting a delinquent account (or trade line as they are generally referred) is transferred to a collection agency In general, this occurs when a delinquency is more than 120 days past due. Collection records can, however, occur for delinquencies that are not associated with a trade line, i.e., in the case of a utility bill.

Collections	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio	Fitted Relative LR
0	$ 364.6	$ 270.1	74.1%	0.97	1.05
1	19.0	18.5	97.5%	1.28	1.21
2	5.5	6.0	108.4%	1.42	1 37
3 or more	5.0	5.9	118.6%	1.56	1.61

R^2 value for the regression of number of collections vs. relative loss ratio is 0.96. The loss ratio for any collections with outstanding liability greater than $0 is 107.6% with a premium volume of $22.3. The results for this variable are very similar to those for DPR. Although there is increasing loss ratio for increasing number of collections, the largest jump in loss ratio occurs between 0 and 1.

4. *Status of Trade Lines*

Each trade line is given a rating based on its current status. A rating of 0 indicates no information is available, while a rating of 1 indicates that the most recent payment made was as agreed, or no more than 30 days past the payment due date. Status codes 2-5 are used to indicated trade lines where the most recent payment made was 30-59, 60-89, 90-119, or over 120 days past due, respectively. Codes 7-9 are used to denote such situations as accounts which are being paid under a wage earner plan, are in repossession, have been written off as bad debt, and others.

Condition	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio
All trade lines not rated 2-5	$ 314.8	$ 227.3	72.2%	0.95
At least 1 trade line rated 2-5	79.2	73.1	92.3%	1.21

Condition	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio
All trade lines not rated 7, 8 or 9	$ 334.1	$ 240.8	72.1%	0.95
1 or more trade line rated 7, 8, or 9	59.8	59.6	99.6%	1.31

If these two types of ratings are viewed exclusively, the following results are obtained:

All trade lines rated 1	$ 286.7	$ 198.8	69.3%	0.91
1 or more rated 2-5, none 7-9	47.5	42.1	88.6%	1.16
1 or more rated 7-9, none 2-5	28.1	28.5	101.5%	1.33
1 or more of each type	31.7	31.0	97.8%	1.28

When combining both types of trade line status, Note the difference between this variable and APD: APD refers to amounts that are currently delinquent, whereas status refers to the account evaluation based on the most recent payment made.

5. *Age of Oldest Trade Line*

This variable measures the time between the report date and the oldest date that any trade line was opened. Trade lines include more than just revolving-type accounts; home improvement loans, installment loans, car loans and mortgages are also considered trade lines. The years listed in the following table reflect the fact that the database involved policies written in 1993.

Year of Opening/ Age of Oldest Line	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio	Fitted Loss Ratio Relativity
1963 & Prior (30+ yrs)	$ 9.6	$ 6.4	66.4%	0.87	0.79
1964-1968 (25-29 yrs)	24.4	14.7	60.2%	0.79	0.85
1969-1973 (20-24 yrs)	41.0	29.4	71.8%	0.94	0.91
1974-1978 (15-19 yrs)	68.3	48.9	71.5%	0.94	0.97
1979-1983 (10-14 yrs)	82.9	60.5	73.0%	0.96	1.03
1984 (9 years)	26.5	20.2	76.2%	1.00	1.07
1985 (8 years)	26.4	20.6	78.2%	1.03	1.08
1986 (7 years)	23.2	19.3	82.9%	1.09	1.09
1987 (6 years)	21.2	19.8	93.3%	1.22	1.10
1988 (5 years)	18.9	15.9	84.2%	1.10	1.11
1989 (4 years)	16.5	12.8	77.6%	1.02	1.13
1990 (3 years)	14.0	12.2	87.2%	1.14	1.14
1991 (2 years)	10.4	9.6	92.5%	1.21	1.15
1992 (1 year)	10.7	10.2	95.0%	1.25	1.16

The t-statistic for the dataset is (5.86); the t-stat for the 99.5% significance level for 12 degrees of freedom is (3.06), thus the null hypothesis that the slope of the regression is zero is rejected at the 99.5% confidence level. The linear regression on years since opening and relative loss ratio generated an R^2 value of 0.86. Here is a correlation that has drawn skepticism: are these results arising merely from the age of the insured, rather than the age of the oldest trade line? This question will be answered in the multivariate section using driver age data, but one can nevertheless deduce that if younger drivers are responsible for the poorer loss results in the lower section of this table, then the same results should be found in the class experience for those ages. This is not true for policies in this dataset, nor is it true for the insurance industry as a whole.

6. *Non-Promotional Inquiry Count*

A strong relationship was also found between loss ratio and non-promotional inquiry count. An inquiry is posted to an individual's credit history file any time that file is reviewed. Many such inquiries are made for direct mail marketing campaigns, which are not requested by the insured. These inquiries are excluded from consideration, and only those that arise from the activities and requests of the insured are included. Federal law prohibits the maintenance of inquiry records for longer than 24 months, at which point they are purged by the credit bureaus.

Number of Inquiries	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio	Fitted Loss Ratio Relativity
0	$ 130.9	$ 92.9	71.0%	0.93	0.92
1	82.7	58.4	70.6%	0.93	0.96
2	55.1	40.9	74.2%	0.97	0.99
3	37.4	28.8	77.0%	1.01	1.03
4	24.9	20.8	83.4%	1.09	1.07
5	17.5	15.2	87.0%	1.14	1.11
6	12.0	9.7	80.6%	1.06	1.15
7	8.7	7.9	90.8%	1.19	1.18
8	6.0	5.3	87.7%	1.15	1.22
9	4.4	4.8	110.0%	1.44	1.26
10	3.2	3.2	100.1%	1.31	1.30
11-15	7.6	8.2	108.6%	1.42	1.41
16 or more	3.7	4.4	117.5%	1.54	1.60

The t-statistic is 9.51; the t-statistic for 11 degrees of freedom for the 99.5% significance level is 3.11. The correlation coefficient for the regression is 0.94. Once again, a single characteristic from an individual's financial management history has a surprisingly large and consistent impact on loss ratio, even in the smaller premium volume cells.

7. *Leverage Ratio on Revolving-Type Accounts*

This variable is calculated as the ratio of the sum of all revolving debt to the sum of all revolving account limits. Trade lines such as mortgages and installment loans are excluded due to the difference in the nature of such accounts. Since leverage ratio is a continuous-type variable, it was difficult to determine how to define data bins.

When the data was initially reviewed, it was found that the loss ratio relativity for leverage ratio = 0% was 1.04, while the relativities for leverage ratios below 10% were in the 0.75-0.90 range, and subsequently rose as leverage ratio increased. This anomaly occurred due to the fact that records with limits of $0 caused a zero divide, and were given a default leverage value of 0%. Therefore, the table displays a more detailed breakdown of records with 0% leverage, due to the marked difference that was evident in loss ratio impact where limits were low or zero.

Leverage Ratio	Revolving Limits	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio	Fitted Loss Ratio Relativity
0%	$0	$ 20.3	$ 20.0	98.4%	1.29	
0%	$ 1 – 499	8.6	8.0	93.0%	1.22	
0%	$500 or more	35.8	23.2	64.9%	0.85	0.84
1-10%		91.6	58.9	64.3%	0.84	0.85
11-39%		91.6	65.0	70.9%	0.93	0.92
40-60%		41.8	31.5	75.2%	0.99	1.01
61-80%		30.5	24.8	81.2%	1.07	1.08
81-100%		24.6	21.7	88.1%	1.16	1.14
101% or more		49.0	47.3	96.6%	1.27	1.26

T-statistic for this dataset (excluding the low-limit, 0% leverage group) is 26.3, using weighted means of the leverage ratio ranges. The 99.5% confidence t-stat is 4.03. The R^2 value is 0.996. The practice of some insurance companies of utilizing the characteristic 'possession of a major credit card' as an underwriting criteria for company placement seems justifiable when the top segment of this table is considered. This depends of course on the average rate level of the writing company.

8. Revolving Account Limits

This variable is the denominator in the calculation of leverage ratio discussed previously. It is the sum of credit limits for all revolving-type trade lines on the report for a given individual.

Revolving Limits	Earned Premium	Incurred Loss	Loss Ratio	Relative Loss Ratio
$0	$ 41.5	$ 39.4	95%	1.25
$1 - $500	9.8	8.6	88	1.15
501-1000	13.0	12.5	96	1.26
1001-1500	12.0	10.3	86	1.13
1501-2000	11.2	10.8	96	1.26
2001-2500	10.0	8.1	81	1.06
2501-3500	18.8	15.3	81	1.07
3501-5000	26.0	20.6	79	1.04
5001-7500	36.2	28.2	78	1.02
7501-10 K	31.4	24.5	78	1.02
10 – 15 K	50.8	34.8	69	0.90
15 - 20 K	37.7	24.0	64	0.83
20 – 25 K	27.6	19.0	69	0.91
25 – 30 K	18.7	12.9	69	0.91
30 – 40 K	22.0	13.5	61	0.80
40 – 50 K	10.9	7.3	67	0.88
50 K +	16.4	10.7	65	0.85

Correlation coefficient for this regression is (0.78), using midpoints of the limit ranges. The first conclusion that could be drawn is that this correlation only duplicates the one already discussed in the leverage ratio section. This will be addressed in the multivariate section. Another conclusion that has been drawn is that this variable is directly correlated to personal income, and use of revolving limits in any underwriting or rating program is discriminatory towards lower income individuals (disparate impact). This may or may not be true; the data does not contain income information. It would be erroneous however, to assume that all people with low revolving limits are also low-income. Many people choose not to use credit; others may have substantial income but low revolving limits due to the fact that they cannot obtain such credit lines based on their past bill payment performance.

Many other individual variables were reviewed from the credit file. Some exhibited correlation to loss ratio at various significance levels, others had no such correlation. Those displayed thus far, however, show a systematic predictive power that requires explanation and understanding.

Causality

Explanation of these correlations, for the most part, cannot be found in the data assembled for this research. I would be remiss, however, if I did not at least attempt to set down those arguments which could be made suggesting reasonable causal links between an individual's bill paying history and expected loss experience for insured losses under a private passenger auto insurance policy.

Before listing such arguments, it is first appropriate to review the Actuarial Standards of Practice #12, entitled "Concerning Risk Classification". The relevant section is 5.2, which states the following:

> **5.2 Causality – Risk classification systems provide a framework of information which can be used to understand and project future costs. If a cause-and-effect relationship can be**

established, this tends to boost confidence that such information is useful in projecting future costs, and may produce some stability of results.

However, in financial security systems, it is often impossible or impractical to prove statistically any postulated cause-and-effect relationship. Causality cannot, therefore, be made a requirement for risk classification systems.

Often, the term "causality" is not used in a rigorous sense of cause and effect, but in a general sense, implying the existence of a plausible relationship between the characteristics of a class and the hazard for which financial security is provided. For example, living in a river valley would not by itself cause a flood insurance claim, but it does bear a reasonable relationship to the hazard insured against, and thus would be a reasonable basis for classification.

Risk classification characteristics should be neither obscure nor irrelevant to the protection provided, but they need not exhibit a cause-and-effect relationship.

Clearly, the operative word in this Standard of Practice is irrelevant, as the historical data in question is not obscure. Therefore, arguments must be put forth which, despite being speculative, are reasonable statements that a reasonable person would find relevant.

Why would an individual who has current or past difficulties with meeting financial obligations be expected to have above-average costs to an auto insurer? Since there is an administrative expense associated with the processing of insurance premiums and related transactions, it can be argued that subsequent lapses in the individual's payment history is a direct cost to the insurer. This cost would fall under the category of expenses, however. The focus here is loss costs.

Maintenance

The argument has already been made, and often, that auto insurers' underwriting practices are created for risk selection, and one characteristic that is viewed as favorable for selection is described in various quarters as "stability" or "responsibility". Few, however, could give an objective definition of how one could measure such a characteristic, but historically many customer characteristics have been utilized as an assumed proxy for this nebulous attribute, such as home ownership, marital status, number of vehicles, coverage and limits selected, etc. It is entirely possible that a person's current and historical management of debt is another indicator that could be utilized to identify this quality. If a person manages their financial affairs responsibly such that debts are paid on time, they may also take the same approach to the maintenance of other aspects of their lives, including their automobile. A vehicle kept in good working order and condition is less likely to be involved in an accident than one that is not, all other things being equal. Such an individual may also take greater care in operating that vehicle.

Morale Hazard

The CPCU textbook "Personal Insurance" defines morale hazard in the following way:

Morale hazard is a condition that exists when a person is less careful because of the existence of insurance. Morale hazard does not involve an intent to cause or exaggerate a loss. Instead, the insured becomes careless about potential losses because insurance is available. Leaving the keys in an unlocked car or allowing fire hazards to remain uncorrected are examples of morale hazard. Morale hazard results in additional losses that drive up the cost of insurance because of injuries and damage that could have been prevented."

The previous discussion of responsibility could lead to the argument that individuals who are careless in the management of finances also present a morale hazard in the area of automobile insurance.

Claims Consciousness

An insurer's loss experience measures dollars of loss which are paid on claims that are filed. The number of claims filed is less than the number of accidents that actually occur. Consider two risks that are identical in all ways (from an insurer's perspective) except for the fact that one manages their financial affairs much better than the other does. The risk who has a troubled financial history and condition is much more likely to be in debt and to a larger degree; the need for capital to satisfy financial obligations has a bearing on decisions made in many areas of his/her life. Suppose for example, that these two risks are both involved in an auto accident, involving no injuries, but causing property damage to their own vehicles which is some nominal amount (say, $100) more than the deductible. The risk whose financial condition is more sound has a disincentive to file the claim. It may impact his/her rates at the next renewal; the time and effort involved may not be even worth the compensation obtained. The risk with the poorer record of financial management has a greater incentive to file the claim and obtain the compensation, as it has greater value to that individual.

Fraud: Increased Severities

Continuing with these same two risks, consider now the situation in which the damage to property was much greater than the deductible; the vehicles each sustained damage measuring in the thousands of dollars. If an auto repair technician suggested a relatively easy way of recouping the deductible for the insured, or the benefits of padding the repair costs, the individual under the greater financial pressure would be more susceptible to acquiesce. This does *not*, however, imply that risks with poor bill-paying histories have any less integrity than other risks. Some people would never commit fraud on any level; others would do so with no need for provocation or encouragement; still others could be convinced to do so only under the proper conditions. This argument only implies that any individual who *could* be induced to participate in this level of fraud would be more likely to do so if they were under financial pressure from other sources.

Fraud. Increased Frequencies

The presence of severe financial pressure could also produce claims that would not have existed otherwise. There is some segment of the population that either does or could view the insurance mechanism as a financial opportunity. Fraudulent claims in the form of staged accidents, phantom claimants, phantom vehicles or arson are a way that an individual can extract funds from the insurance mechanism. Once again, this argument does not imply anything about the integrity of a risk with poor bill-paying history. What it does assert is that an individual with severe financial pressure could look to all possible sources of funds to alleviate that pressure. Therefore, any individual who was capable of committing this type of fraud is more likely to do so given the existence of that financial pressure compared to the absence of it.

Stress

The assumption is made here that individuals who are under financial pressure from debt exist under a greater level of stress than average. This stress could exist from the associated worries over future impact of financial condition. Individuals under such stress may be less focused on proper operation of a motor vehicle and make them more susceptible to accidents resulting from chance occurrences or distraction. It would be useful if there were some other condition which could produce this same level of stress, for which loss data was available, to strengthen the argument. A few currently coded customer characteristics could be considered candidates. One such variable is number of children under the age of 16. One must first make the assumption that risks with three or more children under the age of 16 have a higher level of stress than average. Whether or not one agrees with that probably depends on whether or not they are a parent! In any case, the loss ratio for such risks reviewed in a 1993 research study was over 20 points higher than average. Another possible variable candidate could be self-employed risks. The added responsibilities and worries of a small business owner could imply that their level of stress is higher than average. From that same 1993 study, self-employed risks had a loss ratio which was roughly 15% higher than average.

It is important to make note that this list is not suggested as a menu from which to select the one correct answer. It is likely that the impact on losses of financial management history is a cumulative impact of some or all of these situations, as well as others not listed here.

Multivariate Analysis: Underwriting Characteristics

There have been many assertions made, in the absence of data, about this relationship between loss experience and credit history. The following comes from the NAIC's "Credit Reports and Insurance Underwriting", dated December 14, 1996:

> **"There still is insufficient data to prove to all regulators' satisfaction whether credit history ... are or are not valid indicators ... independent multivariate analysis, a statistical method some regulators view as necessary, has not been performed." (p. 15) "Some regulators suggest that an unbiased and reasonably precise multivariate analysis is necessary to determine the actual rating factor.... They ask whether a person's credit history is truly correlated with future loss experience or whether it is a spurious correlation?" (p. 17)**

It is beyond the scope of this paper to determine whether or not the loss ratio method is appropriate to analyze this particular database. This method is questioned in the aforementioned NAIC report; the assertion is made that small errors in pricing for a number of rating factors could add up to a fairly significant overall pricing error, making loss ratios a biased measure. For purposes here, it is assumed that differences in relative loss ratio are due to differences in expected average loss costs after adjustments for individual premiums, and that this method is a reasonable way of measuring such differences when reviewing more than one variable simultaneously.

The utilization of the factors discussed earlier when performing multivariate queries tended to produce premium volumes in the individual cells which were smaller than desired for credible results. Strict credibility adjustments could not be performed, due to the fact that a) claim counts were not contained in the data and b) the premium and loss on each record arose from all coverages combined. In order to generate larger premium volumes, the credit variables were combined into four mutually exclusive profiles. These profiles were designed to achieve significant loss ratio differences and significant premium volumes described by each. Group A is defined by those characteristics producing the highest loss ratio, i.e., derogatory public records, collection records and large amounts past due. Group D is defined by those characteristics producing the lowest loss ratio, i.e. low leverage ratio, high age of oldest trade line, good account ratings, etc. The precise definitions of the four groups are contained in the appendix. These profiles will be used in this multivariate section for the sake of simplicity and brevity. Each individual credit characteristic was reviewed in conjunction with the underwriting and rating variables described herein. The variables discussed here are a sampling of all those reviewed; they were selected based on assumed relevance. The overall performance of these four profiles is as follows:

Group	Earned Premium	Incurred Loss	Loss Ratio	Loss Ratio Relativity
A	$ 74,279	75,333	101.4%	1.33
B	158,922	124,723	78.5%	1.03
C	69,043	47,681	69.1%	0.91
D	91,746	52,688	57.4%	0.75

Prior Driving Record

The loss performance of various prior driving record combinations is influenced by two significant factors: the underwriting practices of a given company and the experience modification system utilized in rating. Earned premium and incurred loss were aggregated for risks based on their prior accident and violation activity (in the three year period before they were originally written) and based on credit category (A-D):

Prior Driving	Group A		Group B		Group C		Group D		All Groups	
Record	Prem	LR	Prem	LR	Prem	LR	Prem	LR	Prem	LR
No incidents	28.4	93%	66.0	71%	30.7	64%	45.8	53%	170.9	68.6%
1minor*	8.0	94%	17.3	68%	7.5	68%	8.4	50%	41.2	69.4%
1 at-fault accident	3.7	101%	7.7	74%	4.1	68%	5.9	65%	21.4	75.2%
1 non-fault acc.	6.6	109%	14.8	81%	7.3	70%	9.9	70%	38.7	80.7%
2 minors*	2.5	86%	6.0	59%	1.9	41%	2.4	43%	12.8	58.7%
2 incidents (any)	6.5	108%	13.5	96%	6.6	82%	7.9	64%	34.4	88.2%
All other (more Than 2 incidents)	18.6	114%	33.7	95%	10.8	83%	11.5	66%	74.6	93.1%

* *minor refers to a minor moving violation*

The favorable overall performance of the category '2 minor moving violations' can be attributed to both underwriting practice and experience modification surcharge system of the company from which this data was obtained. Of note here is the marked consistency of the loss ratio relationships across credit groups, regardless of prior driving record. Loss ratio relativities, calculated relative to each driving record sub-group, display this consistency:

Group	A	B	C	D	All Groups
No incidents	1.36	1.04	0.93	0.77	1.00
1minor moving violation	1.36	0.98	0.98	0.72	1.00
1 at-fault accident	1.35	0.99	0.90	0.87	1.00
1 non-fault accident	1.35	1.00	0.87	0.86	1.00
2 minor moving violations	1.47	1.01	0.69	0.74	1.00
2 incidents of any kind	1.23	1.08	0.93	0.73	1.00
All other (> 2 incidents)	1.22	1.01	0.89	0.70	1.00
Total	1.33	1.03	0.91	0.75	

Of particular note in this table is the wide difference in performance between clean driving record/poor credit history risks (93%) vs. poor driving record/good credit history risks (66%).

Age of Driver

It could be argued that the loss experience for poorer credit history risks is influenced by driver age distribution. If a disproportionate percentage of young drivers are contained in Group A, then credit history is merely substituting for age. However, as stated earlier, this would only be true if loss experience for younger drivers was adverse, which is not the case. There is a distributional difference in the four groups by age, but the loss experience relationships across credit groups is again robust:

Age of	A		B		C		D		Total	
Driver 1	Prem	LR	Prem	LR	Prem	LR	Prem	LR	Prem	LR
< 25	$ 3.8	121%	$23.6	75%	$ 1.4	51%	$ 1.9	53%	$ 30.8	78%
25-34	21.1	103%	55.8	79%	22.6	66%	8 9	63%	108.4	80%
35-39	13.0	100%	21 8	81%	12.9	65%	13.0	54%	60.7	76%
40-44	12.4	109%	18.5	82%	10.4	76%	15.6	52%	57.0	79%
45-49	9.8	93%	14.6	83%	8.2	·76%	14.8	58%	47.4	76%
50-59	9.2	97%	14.4	78%	7.9	68%	16.5	53%	48.0	71%
60+	3.8	110%	8.3	75%	4.9	81%	20.0	67%	37.1	75%

Some of the individual cells in this table have significantly lower premium volumes than prior tables; they are shown nonetheless for completeness. Clearly, age of driver is not the cause of the poor loss experience in Group A.

Age of driver was also reviewed in conjunction with many of the individual credit variables. For example, the following is the cross-hatching of relative loss ratios for age of driver and non-promotional inquiry count:

Inquiry Count	Age of Driver 1 ⇒ Under 30	30-39	40-49	50-59	60+	Total
0-3	1.01	0.95	0.95	0.87	0.92	0.95
4-7	1.09	1.07	1.18	1.06	1.38	1.12
8-15	1.22	1.34	1.32	1.43	1.69	1.33
16+	1.48	1.88	1.25			1.56

(values are not shown for cells with premium volume less than $ 0.5 M)

The variable age of oldest trade line, reviewed earlier, could have a relationship to losses that is dependent upon age of operator. When these two variables were combined, the impact exhibited independence:

Age of Oldest Trade Line	Age of Driver 1 ⇒ 26-30	31-35	36-40	41-45	46-50	51-55	56-60	60+	Total
< 7 years	1.15	1.23	1.19	1.43	1.25	1.19	1.44	1.15	1.15
7-9 years	1.02	1.03	1.01	1.20	0.96	1.07	0.87	0.92	1.05
10+ years	0.90	0.93	0.94	0.95	0.94	0.87	0.89	0.98	0.93

Classical Underwriting Profile

Historically, the underwriting function has identified and selected for various combinations of characteristics. The risk groups exhibited lower than average frequency of loss, which in the absence of premium adjustments, produced more profitable results. One such profile is the married, multicar, homeowner risk with clean driving record. In an effort to produce a favorable loss ratio within Group A, this characteristic was evaluated:

	Married multicar homeowner				All risks NOT married multicar homeowner			
Group	Clean Driving Record		All other		Clean Driving Record		All other	
A	$ 10.2	97%	10.6	102%	$ 27.8	92%	$ 25.6	113%
B	22.3	77%	20.2	85%	62.9	69%	53.4	88%
C	14.5	76%	13.5	76%	24.4	58%	16.7	74%
D	20.2	57%	16.0	58%	34.4	50%	21.2	70%
Total	67.3	74%	60.3	79%	149.5	67%	116.9	88%

Again, it is important to keep in mind that these results are heavily influenced by underwriting practice at the time of writing by a given company; this can influence column totals. The underwriting function, however, had no knowledge of the information that defines credit groups A-D, and the relationships across these groups are again consistent.

Rating Territory

A key concern voiced by regulators in at least a handful of states is the potentially disparate impact that the utilization of credit history in underwriting or rating could have on lower income urban risks. This paper will not address whether or not income levels in urban areas are in fact lower than suburban or rural areas. The issue of rating territory, however, was analyzed. Although rating territory was not a variable in the original database, subsequent state profiles were developed for inforce policies in order to determine distribution of risks by credit characteristics (again using the Groups A through D) in a sampling of states. The exposure distribution shown below exhibited no clear-cut disparate impact on urban territories when compared to non-urban territories:

State	Exposure Distribution Type	A	Group B	C	D
Connecticut	Urban	14%	32%	12%	42%
	All Other	13	29	13	46
	Total CT	13	30	12	45
New York	New York City	10	26	8	55
	Other urban	14	23	11	52
	All other	13	25	13	49
	Total NY	13	25	12	50
Ohio	Urban	14	20	12	54
	All other	10	19	16	54
	Total OH	11	20	15	54

Data is also available for many other underwriting characteristics, including number of vehicles, number of drivers, residence type, residence stability, job stability, prior insurance type, gender, marital status and many others. These characteristics were also queried against the individual credit variables, in addition to queries run against the four groups utilized above. The results were very similar. There were no variables that produced even roughly uniform results across the credit characteristics.

Multivariate Analysis: Credit Characteristics

Another group of variables that was analyzed is credit characteristics in combination with other credit characteristics. This is necessary to ensure that no dependencies or cross-correlations exist within these characteristics. As with the other analyses, this group contains many cross combinations that were reviewed; only a sampling will be discussed here.

Leverage and Revolving Limits

It was noted in single variable section that leverage ratio could be duplicating the impact of revolving account limits. When reviewing the numerator of leverage, revolving balances, it was found that there was virtually no relationship between that variable and loss ratio (R^2 value of 0.04). The array of loss ratio relativities (for all cells with premium greater than $ 0.5 M) for leverage ratio versus revolving limits shows the independence of their impacts:

		Leverage Ratios ⇒						
Revolv. Limits	Selected Midpoint	0% 0.00	0-50% 0.25	50-75% 0.625	75-100 0.875	100% + 1.20	All	Correl. Coefficient
$ 0	0	1.27	1.25			1.18	1.25	
1-999	500	1.02	1.01	1.35	1.38	1.34	1.21	0.87
1K-3K	2000	0.96	1.11	1.15	1.23	1.33	1.16	0.97
3K-5K	4000	0.78	0.99	1.04	1.19	1.34	1.05	0.98
5K-10K	7500	0.77	0.95	1.11	1.13	1.25	1.01	0.97
10K-25K	17500	0.78	0.83	1.09	1.07	1.07	0.88	0.88
25K +	35000	0.65	0.85	0.87	0.95	0.98	0.86	0.92
Total All		1.08	0.89	1.07	1.16	1.24	1.00	0.74
Correl Coefficient		-0.72	-0.74	-0.80	-0.90	-0.86	-0.87	

Note the consistency of the coefficients in both directions. This would not exist if one variable simply proxied for the other. In more general terms, risks with high leverage ratios have poorer loss performance than those with lower leverage ratios, regardless of limits; risks with low revolving limits have poorer loss performance than those with higher limits, regardless of leverage ratio.

Derogatory Public Records and Collections

Given the similarity of the distribution and loss results of these two characteristics, it might be expected that there is overlap between the two, i.e., individuals that exhibit one type of record commonly exhibit the other. This did not turn out to be the case:

DPR	Collections	Earned Premim	Loss Ratio	Loss Ratio Relativity
0	0	$ 339.2	72%	0.95
0	1	17.2	94%	1.23
1	0	13.7	96%	1.25
	Total 1any	30.9	95%	1.24
0	2	4.8	93%	1.22
1	1	3.1	88%	1.15
2	0	3.4	107%	1.41
	Total 2 any	11.2	96%	1.26
	Total 3 or more	12.6	117%	1.53

Each variable produced poor loss results regardless of whether or not the other variable was present. Both variables also had significant distributional volume.

Leverage Ratio and Inquiry Count

If the basis for the relationship between credit history and loss performance can be attributed to a more general characteristic, one might refer to that characteristic as financial stress, distress or duress. Since leverage ratio and high inquiry count can be expected to occur under such situations, it is reasonable to assume that there may be some overlap between these two variables also. As with the other multivariate combinations that are reviewed, it is important to keep in mind the distinction between distributional imbalance and loss ratio imbalance. In the driver age vs. credit group (A-D) table, there is a clear distributional imbalance, with older drivers being disproportionately represented in the best performing credit group. The loss ratio impact, however, remains consistent across credit groups and is not offset by the inclusion of age. This is also true to a lesser degree in the table of loss ratio relativities below: risks with higher leverage ratios are disproportionately represented in the higher inquiry count groupings, but the two-way impact on loss ratio remains:

Limits:	<500	>500			Leverage Ratio		
Inquiries	0%	0%	1-50%	50-75%	75-100%	100%+	Total
0	1.25	0.74	0.86	0.94	1.01	1.04	0.93
1-3	1.27	0.87	0.86	1.03	1.05	1.26	0.96
4-6	1.23	1.12	0.95	1.21	1.57	1.30	1.10
7-10	1.24		1.20	1.36	1.22	1.35	1.25
11+			1.18	1.28	1.54	1.99	1.46
Total	1.26	0.85	0.89	1.07	1.16	1.24	1.00

Trade Line Counts and Status

In addition to searching for variables that duplicated loss ratio impact within the credit characteristics, bivariate tables were reviewed to determine if some variables partially mitigated those impacts. For example, trade line status showed a strong impact earlier. One could argue that the impact of any trade line not rated 1 would diminish as the total number of trade lines increases. That is, if just one trade line is not in good standing, should that not have less significance for a risk with many trade lines, compared to one with only a few? The following table reveals that this appears not to be true generally:

Total Trade Lines	Total Rated 2 through 9	Earned Premium	Loss Ratio	Loss Ratio Relativity
1	0	$ 12.9	78%	1.02
	>0	3.3	116%	1.52
2	0	9.7	88%	1.15
	>0	6.2	103%	1.34
3	0	9.0	72%	0.94
	>0	8.1	93%	1.22
4	0	13.2	68%	0.89
	>0	5.1	90%	1.18
5	0	13.8	69%	0.91
	1	2.3	101%	1.32
	2 or more	3.0	104%	1.37
6	0	14.3	72%	0.94
	1	2.3	94%	1.23
	2 or more	3.1	117%	1.54
7-8	0	31.4	67%	0.88
	1	4.7	96%	1.26
	2	2.4	103%	1.36
	3 or more	4.3	105%	1.38
9-10	0	31.4	66%	0.87
	1	4.6	101%	1.33
	2-3	3.7	95%	1.25
	4-6	2.5	88%	1.16
	7 or more	0.5	134%	1.76
11-15	0	67.7	66%	0.86
	1	9.9	76%	0.99
	2-3	7.2	91%	1.19
	4-6	5.2	88%	1.15
	7 or more	2.3	106%	1.39
16 or more	0	75.6	69%	0.91
	1	13.5	89%	1.16
	2-3	8.5	82%	1.07
	4-6	5.5	99%	1.29
	7 or more	6.3	97%	1.27

Derogatory Public Records and Collections: Age and Amount

Another area of concern for both regulators and the insurance industry is the severity of a given event and its age. It is common practice for other variables, such as prior claims, to be evaluated differently based on their severity or amount paid. Thresholds are established to determine whether or not experience modification surcharges should apply in such cases. The age of a claim is also an important consideration in making underwriting decisions for private passenger auto applications. This concept is being applied to credit characteristics as well, as insurance companies apply different criteria to both age and amount when it comes to such items as DPRs and collections. The most commonly used vendor scoring algorithm also applies lesser weights to older events. This research database unfortunately was not large enough to have sufficient premium volumes in all the sub-groups, but those that have substantial weight indicate that severity and age may not be nearly as relevant factors as the existance of the record itself:

Age of Event	*Event=Collection* Premium	Loss Ratio	*Event=Derog. Public Record* Premium	Loss Ratio
Within 12 months	$ 5.8	110%	$ 7.6	103%
12-24 months	7.3	108%	7.5	93%
24-36 months	5.7	102%	6.1	107%
36-48 months	3.7	100%	4.7	106%
48-60 months	2.9	90%	3.6	111%
60-84 months	3.8	99%	5.9	92%
No collection records	364.7	74%	No DPR 358.9	74%

Amounts	*Event=Collection* Premium	Loss Ratio	*Event=Derog. Public Record* Premium	Loss Ratio
$0	$371.7	74%	$362.9	74%
$1 - $49	3.6	98%	6.9	95%
$50 - $99	3.7	102%	0.2	-
$100 - $499	9.6	106%	4.4	99%
$500 or more	5.4	120%	19.6	106%

Again, there were hundreds of other combinations of variables reviewed and analyzed; these have been provided as a sample. What has arisen is a significant number of variables within the credit history of an individual each of which has independent influence on private passenger auto loss experience. Such an environment lends itself most readily to a scoring-type mechanism, as the variables can be assigned independent weights that can be accumulated for an overall impact estimate for a given potential applicant. But the social and regulatory acceptability (or lack thereof) of these relationships has made it such that univariate scoring models are not viewed as the most favorable way of treating this particular set of data.

Other Impacts: Retention

One of the variables that was included in the research database was an indicator which designated whether or not a policy was still inforce at the end of the experience period, December 31st, 1995 (anywhere from 24 to 36 months since policy inception). The length of time that an auto policy remains inforce has a direct relationship to overall profitability, both from a loss and an expense standpoint. Characteristics that indicate better policy retention therefore indicate better expected experience over the lifetime of the policy.

The credit characteristics reviewed showed that in general, risks with better bill payment histories were retained at a higher rate than those with poorer bill paying histories. The reason for non-renewal was not available, therefore policies could have been no longer active due to a variety of reasons such as price shopping, underwriting cancellation, non-payment of premium, or any other reason for which a policy can normally cease to be inforce The following table shows percentages of policies still inforce at the end of the experience period for various categories:

All policies	48%	Number of Inquiries = 0	51%
		1-3 inquiries	48%
Policies with no collection records	49%	4-6 inquiries	44%
One collection record	36%	7-10 inquiries	41%
2 or more collections	30%	11 or more inquiries	33%
No derogatory public records	49%		
One DPR	38%	Leverage = 0 ($0 limits)	33%
Two or more DPR	33%	=0 ($1-$500 limits)	39%
		= 0 (limits > $500)	51%
Amounts Past Due = $0	52%		
$1 - $20	52%	0% - 50%	53%
$21 - $100	40%	50% - 75%	47%
$101 - $499	36%	75% - 100%	44%
$500 or more	33%	100% or more	38%

It could appear as though the increase in losses and the deterioration of retention are two effects of the same cause. This is not the case, however, as the loss ratio variation by, for example, number of collections still exists within both subsets of policies: those that remained inforce at the end of the experience period and those that did not. The loss ratios for policies still inforce are 72%, 101% and 114% for risks with none, one, or two or more collections, respectively. The same values for policies that did not remain inforce throughout the experience period are 80%, 93% and 113% for risks with none, one, or two or more collections. This pattern is true for other variables as well. This is a second way in which credit history can impact loss experience.

Homeowners Line of Business

A database was constructed to analyze the impact of credit history on loss experience for the homeowners line of business. The procedure was nearly identical to that described above for the auto line of business, with the exception that the policies included were those originally written in policy years 1993 and 1994. In addition to obtaining the credit data at the time the policy was written, similar data was obtained on those same policies at later dates. This was done in an effort to determine what percentage of risks experience significant changes in their bill-paying profiles over time. Policies were not included in the study from other miscellaneous property lines such as renter, condominium, dwelling fire and landlord policies.

There are some differences in the two datasets. This homeowners database contains $120 million in earned premium and has an overall loss ratio of 64.1%, excluding catastrophe losses. The loss ratio is 79.2% with those catastrophe losses included. The experience period was extended to December 31, 1996 for the policies originally written in 1994, making the experience period 36 months for both policy years. For the majority of the writing period, 1/1/93 through 12/31/94, the company that wrote the policies did not use credit as an underwriting or rating tool. Approximately 10% of the policies were written after such a program was implemented in the underwriting area. During the experience period, all policies inforce were re-underwritten using credit score. While no action was taken directly due to the score, some policies received condition and maintenance reviews and had inspection reports ordered, if such reports were not ordered upon first issuance of the policy. Also, rating territory was included in this database from the outset.

There were striking similarities between the auto and home databases with regard to credit impact on loss experience. The most significant difference seemed to be that derogatory information on a credit report for a homeowners policy had a more severe impact on loss performance (Group A below). If premium and loss are aggregated according to the same Groups A through D as was done with the auto line of business, the results are as follows, with the auto experience displayed again for comparison (premiums are in millions and loss ratios exclude catastrophes for homeowners):

	Homeowners			***Auto***		
Group	Earned Premium	Loss Ratio	Loss Ratio Relativity	Earned Premium	Loss Ratio	Loss Ratio Relativity
A	$ 17.6	111.7%	1.74	$ 74.3	101.4%	1.33
B	41.4	66.5%	1.04	158.9	78.5%	1.03
C	11.9	54.5%	0.85	69.0	69.1%	0.91
D	49.1	47.4%	0.74	91.7	57.4%	0.75
Total	120.0	64.1%		394.0	76.3%	

The similarities between the loss ratio relativities for these profiles lends credence to the assertion that the impact of bill paying history on insured losses transcends line of business, and is not a characteristic attributable only to property policies and claims associated with them. Note that there is a much larger premium distribution in group D for homeowners, the best performing group. This could arise due to a variety of reasons. The same derogatory characteristics that make up Group A are considered in a loan or mortgage application, so a homeowners policy applicant has already (at some point) undergone a screening process based on credit history. The company's underwriting program during the experience period likely decreased the volume of group A policies in the cohort, increasing the proportional amount of Group D.

Individual Credit Variables

The review of individual variables will not be discussed in depth here, as many of the results were parallel with those obtained from the auto study. A handful of examples will be displayed. Compare these with the tables for auto on pages 3 through 5.

Amounts Past Due

APD	Earned Premium	Loss Ratio	Relative Loss Ratio
$0	$ 106.7	58.9%	0.92
$1 - $20	0.9	67.8%	1.06
$21 - $100	2.1	69.2%	1.08
$101-$500	3.5	100.0%	1.56
$501 +	6.8	124.9%	1.95

Collection Records

Number of Collections	Earned Premium	Loss Ratio	Relative Loss Ratio
0	$ 112.0	59.7%	0.93
1	5.2	125.3%	1.95
2+	2.9	124.9%	1.97

Derogatory Public Records

Number of DPRs	Earned Premium	Loss Ratio	Relative Loss Ratio
0	$ 105.4	57.7%	0.90
1	8.0	99.3%	1.55
2	3.0	122.5%	1.91
3+	3.6	125.1%	1.95

Age of Oldest Trade Line

Age in Years	Earned Premium	Loss Ratio	Relative Loss Ratio
< 1	$ 2.3	115.8%	1.81
2 – 3	3.0	68.7%	1.07
4 – 5	5.1	70.9%	1.11
6 – 7	8.3	77.6%	1.21
8-10	19.6	73.8%	1.15
11-15	26.6	60.5%	0.94
16-20	23.6	65.3%	1.02
21+	30.2	48.9%	0.76

Non-Promotional Inquiry Count

Number of Inquiries	Earned Premium	Loss Ratio	Relative Loss Ratio
0	$ 82.2	60.4%	0.94
1	19.5	59.5%	0.93
2	8.1	65.9%	1.03
3	4.1	84.2%	1.31
4-6	4.3	96.8%	1.51
7-10	1.3	106.7%	1.66
11+	0.5	261.2%	4.07

In nearly all characteristics reviewed, it was found that the range of the variable that was correlated with poorer loss experience produced more severe values for the homeowners line than for auto. The linear correlation coefficients for the above tables for loss ratio relativity were 0.95 for APD (0.78 for logarithm of APD versus loss ratio relativity), 0.81 for collection records, -0.74 for age of oldest trade line and 0.93 for non-promotional inquiry count.

Multivariate: Underwriting and Credit Combinations

As with the auto line of business, queries were run to produce premium and loss data for various combinations of risk characteristic and credit characteristic. For purposes of credibility, the credit characteristics were grouped into the same profiles shown above, Groups A through D. A sampling of those results are shown here.

Prior Loss History

At the time of application, an effort is made to determine if there were prior losses filed on the residence. This information arose either from a property CLUE (Comprehensive Loss Underwriting Exchange) report or from the interview with the applicant. Note that the loss ratio across credit levels is not that much different for risks with prior losses compared to those risks with no such prior losses. This is due to a) underwriting practice of the company writing the business and b) relatively less complete information in property CLUE than is present in the auto CLUE system and the state motor vehicle record histories combined.

	Risks with no prior losses			**Risks with at least 1 prior loss**		
Credit Group	Earned Premium	Loss Ratio	Relative Loss Ratio	Earned Premium	Loss Ratio	Relative Loss Ratio
A	$ 15.6	111.2%	1.73	$ 1.9	115.5%	1.80
B	37.7	66.7%	1.04	3.8	64.4%	1.00
C	11.0	56.2%	0.88	1.0	35.3%	0.55
D	43.6	45.7%	0.71	5.5	61.0%	0.95
Total	$ 107.9	63.6%	0.99	$ 12.2	68.6%	1.07

Town Class or Protection Class

Loss experience in the form of loss ratio relativities for credit groups A through D are evaluated within the various protection class designations and is shown below. Values are not shown for cells that possess a premium volume below $500,000.

Protection Class	Credit Profile Group A	B	C	D	Total
1	1.30	0.68		0.65	0.77
2	1.63	1.06	0.84	0.66	1.00
3	2.15	1.20	0.92	0.77	1.14
4	1.61	1.03	0.93	0.71	0.97
5	1.95	0.92	0.72	0.83	1.00
6	1.48	0.88	0.55	0.79	0.90
7		0.63	0.42	0.42	0.79
8		0.67		1.26	1.31
9		1.72		0.48	0.97
10					
Total	1.74	1.04	0.85	0.74	1.00

There is much more fluctuation for individual cells for this dataset compared to the auto line due to both the overall smaller premium volume and the greater volatility of homeowners losses. The consistency across the profile groups is still quite evident for various protection classes, and the relativities decrease monotonically wherever there is significant premium volume in the cells.

Liability Limits

During the two-year period of policy writing, the company wrote an approximately equal proportion of $100,000 and $300,000 liability limits on homeowners policies. A much smaller volume of premium was written with other limits of liability. The base premium was set based on the former limit, and the latter was offered as additional optional coverage.

	Liability Limit = $100,000			**Liability Limit = $300,000**		
Credit Profile Group	Earned Premium	Loss Ratio	Relative Loss Ratio	Earned Premium	Loss Ratio	Relative Loss Ratio
A	$ 9.7	115.5%	1.80	$ 6.4	100.3%	1.56
B	20.4	63.3%	0.99	17.5	70.4%	1.10
C	5.7	59.4%	0.93	5.2	48.4%	0.75
D	21.1	50.9%	0.79	23.2	43.7%	0.68
Total	$ 56.9	67.2%	1.05	$ 52.2	60.1%	0.94

Note the steady shift in distribution of premium between the two limits by group. The premium distribution of the $100,000 limit for the four groups (A through D) is 60%, 54%, 52% and 48%, respectively. Risks with poorer bill paying histories are more likely to choose the lower liability limit, even though the cost of this additional coverage was less than $10 in most cases.

Bill Mode

The two most common forms of payment of homeowners insurance premiums are direct bill, in which the policyholder pays the premium directly, or mortgagee bill, where the financial institution which holds the note on the property pays the premium.

Credit Profile Group	Direct Bill Earned Premium	Loss Ratio	Relative Loss Ratio	Mortgagee Bill Earned Premium	Loss Ratio	Relative Loss Ratio
A	$ 7.7	117.2%	1.83	$ 7.8	103.9%	1.62
B	19.9	69.5%	1.08	17.4	62.6%	0.98
C	5.3	56.7%	0.88	5.5	54.3%	0.85
D	27.7	46.8%	0.73	15.2	48.2%	0.75
Total	$ 60.6	64.0%	1.00	$ 46.0	63.9%	1.00

Rating Territory

As with the auto line, premiums and losses were aggregated by rating territory by assigning characteristic definitions to each rating territory, designating each territory as urban, suburban or rural. This designation was done by eye, without any objective definition of urban (such as population density); major urban areas were designated as such, satellite territories around urban areas and smaller population centers were referred to as suburban, and the remaining regions were called rural. Although there was little credibility when this data was reviewed at the state level, there was sufficient volume when premiums were accumulated by territory type across states. The credit-defined groups showed consistent impact on losses within each group, and there were only slight distributional differences. Only the largest 12 states were included in this query; these states made up roughly two-thirds of the premium volume of the entire sample.

Credit Profile Group	Urban Earned Premium	Relative Loss Ratio	Suburban Earned Premium	Relative Loss Ratio	Rural Earned Premium	Relative Loss Ratio
A	$ 2.8	1.23	$ 7.3	1.99	$ 2.3	1.31
B	7.0	1.07	18.6	1.02	5.1	1.14
C	1.6	0.96	5.7	0.91	1.3	0.57
D	5.5	0.64	23.4	0.80	6.8	0.66
Total	$ 16.9	0.95	$ 54.9	1.04	$ 15.5	0.91

Motility

In order to understand the migration of risks from one credit profile to another over time, additional data was added to the homeowners database. Credit files from future dates were included, which were taken from archived records approximately 12 months after original writing date, and again at 48 months after the original writing date. For this discussion, the same four credit profiles will be used as in the above exhibits.

Group A, the poorest performing profile, was populated with 10,737 policies written in 1993. Of these, 84% still had Group A characteristics 12 months later, and 66% of those risks were still categorized as Group A 48 months later. 20% had migrated to Group B,and the remaining 14% to C and D. This is not surprising, given that 2 of the 3 criteria for Group A are maintained for many years on the credit file (derogatory public records and collections).

Group B was not as stable over time, significant portions of the population migrated in both directions. Of the original Group B in 1993, 67% were still in the group 12 months later, and 36% 48 months later. At that time, 31% had moved to D, 12% to group C, and 21% to A.

Group C was the least stable. Since this group is defined by better than average characteristics, it is not surprising that as those characteristics continue to improve, much of the distribution migrates to Group D. Only 50% of the group still had the Group C characteristics 12 months later, and only 11% at 48 months. 65% of the entire group migrated to Group D in four years. This is not surprising due to the fact that one of the differences between C and D is age of oldest trade line; for those risks that did not qualify as D, time can be the only factor necessary to cause a migration over the subsequent 3 year period. (Again, refer to the Appendix for exact Group definitions.)

Group D, the best performing group, showed the most stability. Risks with the best credit profiles are more likely to maintain those profiles over time. Of the 23,248 policies in this group, 87% still met the criteria for D 12 months later, and 78% met those criteria 48 months later.

This data was not collected on the original auto cohort, so the above data is for homeowners only. It does provide some indication about the necessity of updating the review of credit profile for the purpose of rating and/or underwriting.

Implications and Other Related Issues

The impact of credit history on expected loss performance is a major factor influencing whether or not this variable should be utilized in the rating of personal lines insurance premiums. There are, however, many other relevant issues that must be considered.

The credit history contains a large amount of data. The impact on loss performance has been measured in this study as if arising from a single variable, which is one particular accumulation of the credit data. There is of course an enormously large number of ways in which the data can be combined for this purpose of measurement. When the variables are inspected, individually, one finds that there are some that are historic, and cannot change until they are purged from the record (i.e., derogatory public records, collection records, inquiries and delinquent payments). Others contain information about current conditions, such as account status, current balances and limits, and overdue amounts. The method of combination of these variables will determine where the model falls in the responsiveness versus stability spectrum. This study has shown that both types have strong influences on loss performance. How they are combined is currently an open field for individual insurers' discretion. This study utilized a mutually exclusive profiling technique; scoring models can and do utilize a large number of variables, giving numeric weights to each individual characteristic which are then added to obtain a total. Either method can be accomplished using a wide range of variable counts.

An important gap in this study is the impact of credit history on loss performance for customers who have been insured with the same company for a number of years. Recall that the data was assembled from new policies written in a give policy year, and the subsequent three-year loss experience. This data cannot show if long-term customers who have similar credit characteristics are expected to have the same differences in loss performance. The creation of a rating factor based on credit history can affect renewal customers as well as new customers, yet there is currently no data publicly available to my knowledge that shows such relationships. Without such data, it would be speculative at best to assume that the relationships hold true regardless of tenure. Studies have shown that long-tenured customers produce far better loss experience than new customers. Opinions vary as to whether this is due more or less to two (or more) dominant factors which can cause such improvement: 1) the fact that longer term customers have more experience in operating a motor vehicle or maintaining a home, and 2) that the underwriting function of a given company will selectively non-renew poor performing risks, which could not be identified accurately in the underwriting process when the policies were originally written. The research done with this data has shown that longer-tenured customers tend to have better credit profiles than newer customers. This is one variable, policy tenure, that could be both distributionally and loss performance-linked to credit history.

The question as to how often the credit history needs to be reevaluated is also of concern. Although the motility information above indicates that there is a fair amount of stability over time for credit conditions, there is still significant change that occurs within such distributions. Each reevaluation will cause the creation of an additional inquiry record on the file. Although such inquiries should not be utilized for evaluation, there is no guarantee that all financial institutions and other users of credit data will ignore their existence. When such a reevaluation occurs, there is also the question as to which risks should experience premium adjustment. Is there reasonable justification for an individual risk to experience an increase in premium solely due to a change in a variable within the credit file? A different type of database construction technique would be required to answer such a question.

From an actuarial standpoint, questions arise concerning the nature of the variable. The literature is replete with admonitions concerning the use of variables that are, or can be, under the control of the insured. Although the historic variables are not under the control of the insured, certainly those that measure current conditions are. Worth considering, however, is the argument that such control is not nearly as relevant as other rating factors that are not utilized for this reason. An individual who has a poor history of timely bill payment, and is under a considerable debt load is already experiencing detrimental effects from these conditions. Such conditions are causing economic penalties in the form of monthly interest payment, or debt service, and can also result in higher interest rates charged for credit lines, installment loans and mortgage loans. There already exists a financial disincentive to maintain financial management habits that produce these conditions. Will a difference in auto or homeowners insurance premiums cause a change in such habits, where these other economic disincentives have not? It is likely, in my opinion, that the magnitude of the premium difference would not be as large as the sum of all other financial consequences of such a credit profile in most cases. This may mitigate the concern over the control the risk appears to have over the data contained in the credit file.

Another area of concern that is related to variable control is data accuracy. Reports as to the accuracy of credit history data vary widely depending upon the source. Credit bureau sources quote data accuracy values in the 99% to 100% range. Some consumer groups have quoted this number to be as low as 30% to 40%. This discrepancy is due to the way in which errors are measured. One could obtain the first result if errors were considered to exist only in cases where a) an adverse decision was made for a financial transaction, b) the customer inquired as to the credit data, c) discovered an error, d) contacted the creditor to correct the error, and e) the financial institution reversed the decision based on that correction. Dividing the number of such events by the entire credit warehouse would produce a very high level of accuracy. To produce the second, much lower values, one could simply count every possible error within the file, including seemingly irrelevant errors such as street name misspellings, and divide this count by the total number of records. Neither is a very good measure of data accuracy. For all parties concerned to get a true understanding of accuracy, a good method of measurement must be established. In any case, the utilization of credit history for rating requires the insurance industry to assist its customers by informing them of the method for resolving true inaccuracies on record, and taking those corrections into account through reevaluation.

An outstanding issue that will likely remain outstanding is causality. Although arguments were put forward earlier in this paper which attempted to link financial management responsibility and future expected loss levels, such arguments are unsupported, even if reasonable, speculation. The arguments of causality are generalized; in fact the difference between one rate level and another charged to a given individual could be different due to only one particular variable within the credit file. That individual may ask for an argument of causality pertaining only to the one characteristic that separates him or her from the next lower rate. Such questions may never be answered with statistical causality, even if the entire credit file (however that is aggregated) can be demonstrated to be causal in a way that goes beyond the mathematical correlations.

The issue of acceptance of credit history data in personal lines insurance has more obstacles than mere causality. The social and regulatory acceptance of such data in the rating of personal lines insurance may be restricted for other reasons. Arguments have already been made that indicate that some groups consider its use invasive, and that credit-based rating is a breach of privacy, regardless of its strength as a tool to reduce rate subsidies between risks. The auto line of business has considered past driving record to be a key factor in underwriting and rating. One key characteristic of prior accidents is negligence, i.e., whether the accident was the fault of the insured or not. It is natural for some people to immediately apply this concept to credit history as well. Credit files contain information about derogatory events that an individual may feel are perfectly explainable. Such explanations are commonplace in the area of mortgage financing, where an event is not considered if there is a suitable explanation for its existence in some cases. The key difference, however, is that the use of this data for rating or underwriting is not done for the purpose of credit worthiness. It is not done for the purposes of judging character, lifestyle, integrity or financial soundness. The purpose is to segregate risks by different levels of expected losses only, a point which may be difficult to communicate.

It may be easier to obtain regulatory acceptance comparted to social acceptance with regard to the use of credit history as a rating tool. The NAIC White Paper on the use of credit in underwriting, referred to earlier, makes several specific statements which indicate their deference to rating, rather than underwriting. The use of credit in rating requires the filing of a rating plan with supporting documentation. It permits inspection of content by both regulators and consumers. Such filing gives a regulatory body the evidence required to give valid statistical response to constituents who may call to inquire or register a complaint.

The data reviewed in this study produced clear evidence of a strong correlation between credit history and future loss performance. The understanding of this relationship, and its acceptance, have grown rapidly over the last few years. This understanding has come primarily in the form of scoring model results. Hopefully, this paper will serve as a starting point in an effort to place more detailed information from credit history, other than scoring models, and the relationship such data has to personal lines losses, in a public forum. This effort is necessary in order to promote greater understanding of the driving forces behind this relationship, and can only serve to improve the quality of discussion during future debates on the ways in which it will be utilized.

APPENDIX

1. Data Fields

Policy Variables included and reviewed:
State transfer indicator
Policy Tier
Original policy written month, day and year
Acive status indicator
Months of coverage
Writing company
Original producer code
Risk state
Vehicle type
Non-standard indicator
Number of vehicles
Number of operators
Number of potential operators
Payment plan
Residence stability
Residence code
Residence type
Number of years employed
Prior insurance code
Number of vehicles financed
For each driver:
- Age
- Gender
- Marital status
- Occupation code
- Number of years licensed
- Driving record: fault losses, non-fault losses, moving violations
- Comprehensive losses

Earned premium
Incurred losses

Variables included from National Credit File:
Trade Record: Subscriber code, date opened, high credit, date verified, date reported, date closed, date paid out, associated code, payment pattern, current balance, amount past due, account type, current manner of payment (status), credit limit, terms, maximum delinquency date, maximum delinquency amount, number of months 30-59 days past due, 60-89 days past due, 90+ days past due, loan type, dispute code, collateral field, duplicate indicator, account number, short subscriber name.
Inquiry Record: Subscriber code, inquiry date, type, loan type, loan amount.
Public Record: Date reported, amount, public record type, date paid, assets, liabilities, attorney, plaintiff, docket number.
Collection Record: Date reported, subscriber code, amount owed, status, date paid, creditor name.
Summary Record: Number of inquiries, trades, collections, public records, manner of payment totals for each status code.

2. Definitions of Credit Profiles Used in Exhibits

Group A: Existance of any of the following: Derogatory public record with liability amount >$0, collection record, or amount past due of $500 or more.

Group B: Does not meet any other group criteria.

Group C: No DPR or collection records, no APD; no trade lines with status codes other than 0 or 1, leverage ratio on revolving accounts less than 60%, age of oldest trade line at least 7 years.

Group D: Same as group C, plus nonpromotional inquiry count less than 4 and age of oldest trade line at least 10 years.

Using Generalized Linear Models to Build Dynamic Pricing Systems

Karl P. Murphy, Michael J. Brockman, and Peter K. W. Lee

Using Generalized Linear Models to Build Dynamic Pricing Systems for Personal Lines Insurance

by

Karl P Murphy, Michael J Brockman, Peter K W Lee

1. Introduction

This paper explains how a dynamic pricing system can be built for personal lines business, whereby profit loads and risk premiums can be tailored to the individual behavioural characteristics of the customer.

The approach has been developed for a free and competitive rating regulatory environment, although the techniques can be reverse-engineered to provide customer value models in markets where rates are controlled.

We use the statistical technique, generalized linear models (GLMs), for estimating the risk premium and price elasticity components of the model. These techniques are well established in the British and European markets and are recently becoming more widely used in the United States.

The objective is to use as much information as input to these models in order to establish which risk factors are the most predictive. In our experience, every company is different; in their size and organisation, in their underwriting and marketing philosophy, in their claims handling, and in their means of distribution. It is essential that the pricing models built reflect these characteristics.

The paper is intended to be practical, hence we have kept the theory to a minimum, quoting other papers or literature where a more theoretical exposition is required. In particular, much of the contents of this paper follow on from the work of Brockman and Wright which sets out the theory of generalized linear models and its application for personal lines pricing. Since Brockman and Wright, the use of GLMs has become much more common. Whilst GLMs are being widely utilized in the UK and Europe, we do not believe that the results are being fully exploited. In this paper we will be explaining how the results of GLMs can be more effectively employed.

The structure of the paper will be as follows: firstly we are going to discuss the main statistical technique used for predicting behaviour (both claims and demand), namely generalized linear models. We will then look at how GLMs can be used in assessing the risk premium for a policy, followed by a discussion on using GLMs to assess price elasticity curves. Next we will take an overall look at combining the supply and demand models. Finally, we will describe a possible pricing algorithm which brings both the cost and demand sides of the equation together.

2. Generalized Linear Models

2.1. Traditional rating versus a multiple regression approach

The insurance industry is quite unique in that, unlike most manufacturing companies, both the cost of writing a policy and the demand for the product are highly dependent on the characteristics of the individual to whom the policy is sold. In addition, the range of factors which affect both claims experience and demand is very large, and this creates problems when deciding on what pricing differentials to apply between different groups.

The traditional way to determine relativities by rating factor is to look at series of one-way tables, either focusing on the relative risk premiums or the relative loss ratios. Holler, Sommer and Trahair (section 1.2) provide a nice example of why one-way reports give very misleading results because of the different mix of business underlying the levels of each factor. This problem should be fairly familiar to most readers, so we do not propose to dwell on it here.

One solution is to use some form of multiple regression approach which removes any distortions caused by different mixes of business. A flexible approach, and which we have found to be very useful in practice, is a regression method known as generalized linear models (GLMs). Many different types of models which suit insurance data fall under this framework, and the more familiar "classical" regression (which assumes that the response variable is a linear combination of the explanatory variables and is normally distributed) is a subset of GLMs.

The additional benefit of using GLMs over one-way tables is that the models are formulated within a statistical framework. This allows standard statistical tests (such as χ^2 tests and F tests) to be used for comparing models, as well as providing residual plots for the purpose of model diagnostic checking.

2.2. Specification of GLMs

We outline below the fundamental algebra behind GLMs. The notation that we use is slightly informal, in that we have omitted estimate indicators ^ as well as omitting vector indicators, but this should not detract from the key results.

GLMs take the following form:

$$y = h(X\beta) + \text{error}$$

where h is a monotonic function, and y is distributed according to a member of the exponential family of distributions. This family includes the normal, Poisson, gamma, inverse Gaussian and binomial distributions. As we will see later, the results can be extended to a more general family.

Notice that this gives a much more flexible model form than the classical case. y is now a function of a linear combination of the explanatory variables rather than just a direct linear combination, and y is not constrained to be normally distributed.

On a point of terminology, X is known as the design matrix. The dimensions of X are n by p, where n is the number of observations available, and p is the number of parameters being estimated. The parameters can either be continuous variables or categorical variables, a distinction which we will detail later. The algebra is the same for both types of variables.

$X\beta$ is know as the linear predictor, and is a straightforward linear combination of the estimated parameters. The linear predictor is usually denoted by η, and is of dimensions n by 1.

$h(X\beta)$ is known as the fitted values, and simply transforms the linear predictor. It is usually denoted by μ.

We discuss the main features of GLMs below.

2.3. Link function

Although the model is usually specified as $y = h(X\beta) + \text{error}$, it is usual to refer to the inverse of h as being the link function, and this is usually denoted by g. If we want to fit a multiplicative model, then $h(x) = \exp(x)$, so $g(x) = \log(x)$, so we have a log link function.

Other common link functions include the identity link function ($h(x) = x$, $g(x) = x$) and the logit link function ($h(x) = 1 / (1 + \exp(-x))$, $g(x) = \log(x / (1 - x))$).

Notice that by fitting a model with a log link function, we ensure that the fitted values are non-negative, and by fitting a model with a logit link function the fitted values fall between 0 and 1. These properties are often very appropriate for insurance data – risk premiums are always non-negative and renewal rates always fall between 0 and 1. Under classical regression, the fitted values can take on any value.

2.4. Variance function

As we will see when it comes to estimating the parameters β, the key feature of the distribution of y is given by the variance function.

The variance function links the variability of y to the fitted values μ. This is the second key difference with classical regression. Under classical regression, the variance of y is constant, no matter what the fitted value. With GLMs, the variance can vary with the fitted values.

Common variance functions are as follows:

Normal:	1 (i.e. constant variance)
Poisson:	μ
Gamma:	μ^2 (i.e. constant coefficient of variance)
Inverse Gaussian:	μ^3
Binomial:	$\mu(1 - \mu)$

These variance functions should be quite intuitive to anyone familiar with these distributions.

The property of non-constant variance is often quite appropriate for insurance data. Suppose, for example, a model is being fitted to the average cost of claims. Under a normal error structure, the standard error for an observation with a fitted value of $100 might be (say) $10. Equally, the standard error for an observation with a fitted value of $1,000 would also be $10, as it would be for an observation with fitted value $10,000. It is much more intuitively appealing to have a proportionate standard error, so the standard error for a fitted value of $100 might be 10% (i.e an absolute standard error of $10), for $1,000 would also be 10% (absolute standard error of $100), as it would be for $10,000 (absolute standard error of $1,000). This proportional standard error model has a constant coefficient of variance, and so is a gamma model. In this way, therefore, the variance function can be initially selected based on knowledge of the likely variance relationship of the process being modelled. This can then be tested statistically using standard diagnostic checks. We have found from experience that the correct selection of variance function significantly improves the robustness of the parameter estimates.

2.5. Estimation of the β parameters

β is usually estimated via an iterative process, where the r^{th} estimate of β is given by the following equation:
$\beta^r = (X^T W^{r-1} X)^{-1} X^T W^{r-1} (z_i^{r-1} + \eta_i^{r-1})$

where

W^{r-1} (n by n) = $\text{diag}\{ 1 / [g'(\mu_i^{r-1})^2 V(\mu^{r-1}_i) / w_i]\}$
$V(\mu^{r-1}_i)$ = variance function
$\eta^{r-1} = X \beta^{r-1}$
$\mu^{r-1} = h(\eta^{r-1})$
z_i^{r-1} (n by 1) = $g'(\mu_i^{r-1})(y_i - \mu_i^{r-1})$
w_i = weight vector

This equation looks quite complicated, but simplifies quite considerably for many common models.

For example, for a normal model with an identity link function, the W matrix simplifies to a diagonal matrix of the observation weights, and $z_i^{r-1} + \eta_i^{r-1}$ reduces to y, and so we get the familiar general linear model (as opposed to *generalized* linear model) equation:

$\beta = (X^T W X)^{-1} X^T W y$

For a Poisson model with a log link function, the diagonal entries of W reduce to $\mu_i^{r-1} w_i$, and z_i^{r-1} simplifies to $(y_i - \mu_i^{r-1}) / \mu_i^{r-1}$.

For a gamma model with a log link function, the diagonal entries of W reduce to w_i, and z_i^{r-1} simplifies to $(y_i - \mu_i^{r-1}) / \mu_i^{r-1}$ (as for the Poisson model).

The iterative process continues until a statistic known as the *deviance* converges. The deviance is a function of the log likelihood statistics, and is a function of the y values and the fitted values μ. The analogous statistic under classical regression is the sum of squared errors. The concept of "deviance" is described in McCullagh and Nelder (p 33).

2.6. Extending the range of variance functions

We listed in section 2.4 above the variance functions for most of the common distributions. The theory works just as well for a slightly broader range of distributions, whereby the variance function takes the form μ^α, where α can be any specified value. This family can be useful if any diagnostic checks on the goodness of the model indicate that the fitted distribution may not be appropriate. A value of α between 1 and 2 is often used to model risk premium data; this corresponds to the Tweedie distribution. Mildenhall (sections 8.3 and 8.4) discusses this broader range of distribution further.

2.7. Categorical and continuous variables

The use of categorical variables allows separate parameters to be fitted for each level of a rating factor. For example, we can fit a separate parameter for each territory in the rating structure; if we used a continuous variable we would be imposing some sort of relationship (such as a linear or polynomial relationship) across territories, and this may not be appropriate. Indeed, some factors may not have any natural continuous relationship to use.

Factors that have a natural continuous scale can also be treated as categorical variables. For example, a policyholder age scale can be grouped by bands of ages, and each band considered to be a categorical level. If the trends in the categorical level parameters indicate that a continuous scale may be appropriate, then the categorical variable can be converted to a continuous variable. An example of this is given below.

In terms of the design matrix, when fitting a categorical variable we need a column for each level of each rating factor being fitted (with a slight adjustment as described below). The columns then take on the value 1 or 0, depending on whether the particular observation contains that level or not. If we included a column for each level of the rating factor, the design matrix would contain a linear dependency (because for any one categorical factor the sum of all levels would sum to 1), and so one level is omitted from each categorical variable. This level is known as the base level;

we usually choose this level to be the one with the greatest volume of business. In a sense, the base level is subsumed into the first column of the design matrix (the intercept term).

For example, if we are fitting a model with two factors, the first being territory (with levels A, B and C), and the second being policyholder age (grouped into <25, 25-39, 40-59, 60+), then the design matrix would look something like:

Intercept	Territory B	Territory C	Policyholder Age < 25	Policyholder Age 25-39	Policyholder Age 60+
1	0	0	1	0	0
1	1	0	1	0	0
1	1	0	0	1	0
1	0	1	0	1	0
1	0	1	0	0	0
1	0	0	0	0	0
1	1	0	1	0	0
1	1	0	0	0	0
1	1	0	0	1	0
1	1	0	0	0	1
1	0	1	1	0	0
1	0	1	0	1	0
1	0	0	0	0	1
1	1	0	1	0	0
1	1	0	0	1	0
1	1	0	0	0	1

The base level for Territory has been chosen as level A, and the base level for policyholder age has been chosen at level 40-59. The choice of base is somewhat arbitrary, but is often taken to be the level with the most numbers of observations. The parameter estimates are only affected in that they are now relative to a different base; the fitted values are identical.

If having estimated the parameters for policyholder age it was felt that the parameters were broadly linear, then the best straight line relationship can be derived by changing the design matrix. If the average policyholder ages within each level are 22, 32, 50 and 69 (say) then the design matrix changes to:

Intercept	Territory B	Territory C	Policyholder Age
1	0	0	22
1	1	0	22
1	1	0	32
1	0	1	32
1	0	1	50
1	0	0	50
1	1	0	22
1	1	0	50
1	1	0	32
1	1	0	69
1	0	1	22
1	0	1	32
1	0	0	69
1	1	0	22
1	1	0	32
1	1	0	69

In other words, we are replacing the categorical (three parameter) policyholder age scale with a (one parameter) linear relationship.

2.8. Practical example

Below is an extract from the output of a generalized linear model fitted to an average cost model.

EMBLEM X \ ..\Theft average cost fac Theft Average Cost (Fitted Parameters)

File Specify Fit View Tools Window Help

Parameter Number	Name	Value	Standard Error	Standard Error (%)	Alias Indicator (%)	Exp(Value)
1	Mean	6.3092	0.03007	0.5		549.6307
-	Sex (Male)					
2	Sex (Female)	0.0091	0.01227	135.8		1.0091
3	Sex (Unknown)	0.1591	0.17526	110.1		1.1725
4	Policyholder Age (≤ 17)	-0.2724	0.40211	147.6		0.7616
5	Policyholder Age (17)	0.1668	0.06514	33.1		1.1815
6	Policyholder Age (18)	0.1911	0.05665	29.6		1.2106
7	Policyholder Age (19)	0.0962	0.07252	[illegible]		1.1002
8	Policyholder Age (20)	0.1604	0.07038	43.9		1.1740
9	Policyholder Age (21)	0.1424	0.05500	36.6		1.1531
10	Policyholder Age (22)	0.1432	0.05940	41.5		1.1539
11	Policyholder Age (23)	0.1135	0.03491	30.7		1.1202
12	Policyholder Age (24)	0.1745	0.03571	20.5		1.1907
13	Policyholder Age (25)	0.1000	0.03302	33.0		1.1052
14	Policyholder Age (26)	0.1241	0.03111	25.1		1.1321
15	Policyholder Age (27)	0.0013	0.02959	2,196.2		1.0013
16	Policyholder Age (28)	0.0264	0.02646	100.2		1.0268
17	Policyholder Age (29)	0.0360	0.02622	[illegible]		1.0367
18	Policyholder Age (30)	0.0013	0.02937	2,347.4		1.0013
19	Policyholder Age (31-32)	0.0324	0.02067	[illegible]		1.0330
20	Policyholder Age (33-35)	0.0100	0.01949	194.2		1.0101
-	Policyholder Age (36-40)					
21	Policyholder Age (41-45)	-0.0153	0.02116	138.6		0.9848
22	Policyholder Age (46-50)	0.0251	0.02442	97.4		1.0254
23	Policyholder Age (51-55)	0.0014	0.02962	2,082.3		1.0014
24	Policyholder Age (56-60)	0.0817	0.03691	45.2		1.0851
25	Policyholder Age (61-65)	0.0208	0.05029	242.2		1.0210
26	Policyholder Age (66-70)	-0.2100	0.06280	29.9		0.8106
27	Policyholder Age (71-75)	0.0507	0.09125	179.8		1.0521
28	Policyholder Age (75-80)	-0.1782	0.06150	34.5		0.8368
29	Policyholder Age (81+)	-0.2947	0.06606	22.4		0.7447
30	Rating Area (Y)	0.0908	0.16470	181.4		1.0950

The model fitted has a log link function and a gamma error structure. Because of the log link function, the parameters are on a log scale (the final column calculates the exponential of the parameter values).

The "mean" parameter value (6.3092) represents the intercept parameter, and the exponential of this number (549.6) is the fitted value for a base risk. For this particular model, a base risk is a male aged 36 to 40 (plus appropriate bases for the factors not shown in the extract).

To derive the fitted value for a non base risk, simply add up the parameters for the appropriate levels and then take the exponential this number. For example, the fitted value for a female age 23 is exp(6.3092 + 0.0091 + 0.1135) = 621.3. We can derive the relative claims experience for any one factor by simply looking at the exponential of the parameter. For example, the female relativity is 1.0091 (=exp(0.0091)) compared to males.

The "Standard Error" column gives us an idea of the variability of the parameter estimates.

2.9. Other GLM concepts

We have described above the key concepts underlying generalized linear models. There is naturally considerably more theory underlying the process, as well as other concepts, which we have not discussed.

Concepts which we have not mentioned include residuals graphs, offsetting models to fix parameter values, joint modelling techniques, model diagnostics, dealing with sparse data, and statistical tests. The interested reader is referred to McCullagh and Nelder for a much more thorough explanation.

3. Risk Premium Modelling

Brockman and Wright explore in some detail the use of generalized linear modelling for risk premium analyses. We summarize below some of their main conclusions, and discuss in more detail issues concerning the modelling process and allowing for large claims.

3.1. Splitting the risk premium

Brockman and Wright (sections 2, 3 and 5) recommend that the risk premium is split down into frequency and severity (average cost) by each type of claim covered under the policy. The reasons for doing this include, for example:

- response variables for frequency and severity follow different statistical distributions. Usually a Poisson error structure is used for a frequency model and a gamma error structure for a severity model. A log link function is usually used for both frequency and severity
- a greater insight into the underlying cause of claims experience variability is provided
- certain models are inherently more volatile that others. For example, the average cost of liability claims is likely to be much more volatile than the frequency of own damage claims. By modelling total risk premium rather than splitting it into its constituent parts, we would not be able to identify whether an apparently anomalous trend is the result of a random fluctuation in liability average cost or a genuine trend in the own damage frequency

In a sense, by "best" fitting models with the least parameters to each component process, we are maximizing the information being extracted from (often sparse) datasets.

3.2. Risk premiums and GLMs

Risk premium modelling fits very naturally within the generalized linear model framework, especially when split into its constituent parts (i.e. frequency or average cost by claim type).

The response variable is taken to be the frequency or average cost, with the rating factors available used as the explanatory variables. Usually the rating factors are initially classified as categorical variables; it is a simple task to convert them into continuous variables if it is appropriate, as is described in section 2.7. In addition, we recommend that "time" is included as a rating factor (usually accident year or accident quarter). The parameters indicated by this factor will provide estimates of inflation rates and changes in frequency levels from period to period. Furthermore, we can use the time parameters to test the consistency of parameter trends over time. We discuss this further in section 3.3.

It is possible to build factors into the analysis which are not directly asked of the policyholder, or that are collected but not used in the rating tariff. These might include, for example,

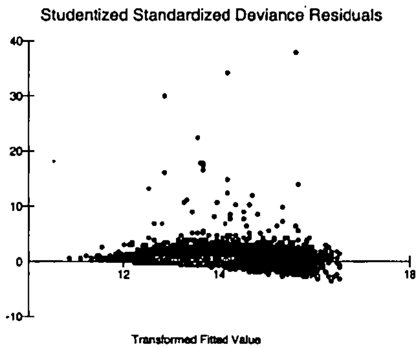
- the length that they have been with the company
- whether or not they have other policies with the company
- the distribution method that attracted the policyholder
- information from external data sources (such as socio-demographic data or credit rating information)

Frequency models are usually fitted assuming a log link function with a Poisson error structure. Claim frequencies fall quite naturally into a Poisson process, and we have found empirically that this is an appropriate model. Using a log link function ensures that the fitted frequencies are positive. Brockman and Wright (section 2.1) go into further detail about why a multiplicative model is appropriate for frequency models.

Average cost models are usually fitted assuming a log link function with a gamma error structure. Once again using a log link function ensures positive fitted values. By having a gamma error structure we have a constant coefficient of variance, and so the standard errors around the fitted values are proportional to the fitted values. For example, if for a fitted value of $100 the standard error is $10, then for a fitted value of $1,000 the standard error would be $100.

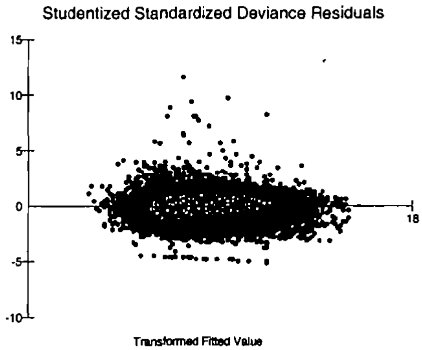
The appropriateness of the error structure can be tested by looking for any evidence of heteroskedasticity within the model residuals. If there is any evidence of the residuals fanning inwards or outwards when plotted against the fitted values, then a different error structure may be more appropriate.

For example, below is a graph showing the deviance residuals for an average cost model (fitted using a Normal error structure). The residuals appear to be fanning outwards, so the variability of observations with higher fitted values is greater than those with lower fitted values. The variance function for a Normal model is 1 (i.e. μ^0), and the shape of the residuals suggests using a variance function with a higher power (such as μ^2).



Deviance residuals are explained in McCullagh and Nelder (p 39). The residuals have been "studentized" and "standardized" (see McCullagh and Nelder section 12.5). The x-axis is a transformation of the fitted values for each observation (in this case, twice the log of the fitted value).

The residuals for a gamma model fitted to the same data are shown below. The obvious fanning outwards has been removed, and this graph would suggest that the gamma model is an appropriate error structure for the data.



3.3. *The modelling process*

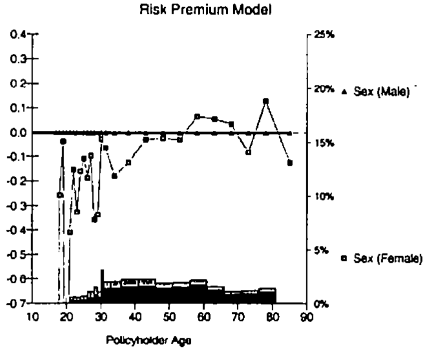
As with any statistical model, we want to find the model which replicates the data quite closely, yet contains as few parameters as possible, i.e. we wish to find the most parsimonious model.

The steps involved in doing this include the following:

- removing rating factors from the model which do not seem to affect variations in claims experience
- grouping levels within rating factors to reduce the number of parameters fitted (e.g. grouping policyholder age into under 25s and over 25s)
- fitting curves to variables which have a natural continuous scale
- making the model more complicated by fitting interaction or combination terms

It is interesting to dwell a bit more on this last point. If within the data there is an interaction effect, then we want to reflect this in the model, even though it involves increasing the number of parameters in the model. An interaction effect occurs when the relative claims experience of a particular factor depends on the level of another factor. For example, a common interaction to find is between policyholder age and gender, whereby the claims experience for young females is considerably better than that for young males, but the experience of older females is either the same or worse than for older males. If an interaction term were not included then there would be a constant difference between males and females across all ages.

The graph below shows the relative claims experience for females compared to males by policyholder age for an overall risk premium model. Although there is an element of randomness, there is a strong upward trend by policyholder age (i.e. the experience of females relative to males worsens with increasing age). A simplification to the interaction would probably be made, to smooth out the random effect.



Interactions are usually tested before simplifying the model; otherwise there is a danger of removing a factor which should be included in an interaction.

There are several ways of assessing whether a simplification to the model (or indeed a complication via fitting interactions) is appropriate. These include:

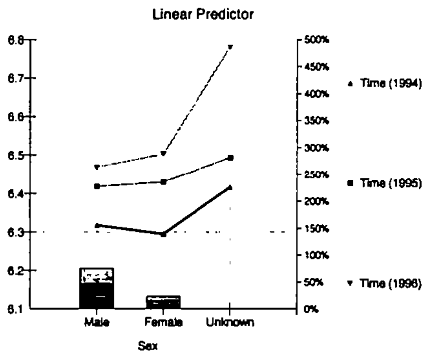
(i) looking at the relative size of the parameters. If the model is only indicating a 0.5% difference in experience between two levels of a rating factor, this would probably be judged to be not significant

(ii) looking at the standard errors of parameters. As a rough rule, if the parameter estimate is within two standard errors of the base rate, then there is no statistical difference between the two

(iii) doing a formal statistical test between two nested models (i.e. where one model is a subset of the other). This is achieved by comparing the change in deviance with the change in degrees of freedom between the two models. If model 2 is a subset of model 1, then model 2 will have a larger deviance, but more degrees of freedom. Our recommended statistical test is to compare the change in deviance with a χ^2 distribution with degrees of freedom being the difference in degrees of freedom between the two models

(iv) fitting an interaction between the factor being considered and the time factor to see the consistency of any trends between different accident years. If the trends are inconsistent over the period being investigated then it is questionable as to whether the trend will exist in the period to which the model is being projected. Equally a consistent trend over time will give more confidence that the same trend will occur in the future.

The graph below shows an interaction fitted between time and gender for an average cost model. Each line represents a separate accident year.



For two out of the three years, the male experience is better than the female experience, but for the other year females have better experience. It is unclear as to what the relative experience is going to be in the following year.

Sometimes a particular trend will be shown to be significant under one test, but not for another. Which test to believe is entirely a matter of judgement. Perhaps the one over-riding criteria is that any trend must make sense and there has to be a rational explanation for it. A good question to ask when making these judgement calls is whether the trend could be explained to an underwriter and whether he/she would believe it.

Another practical point is that eventually the model will be populating a rating tariff, and so the modeller should bear in mind the structure of the tariff when doing the modelling.

3.4. Combining models

Once each frequency and average cost model by claim type has been modelled, it is usual to combine the separate component models to generate one final model.

This can be achieved by calculating the fitted values for each component model for each observation, adding each element together to get a total risk premium, and then fitting a final smoothing model to this total risk premium.

It is usually necessary to make adjustments to the individual components before combining. These adjustments include:

(i) adjusting frequencies and average costs to make allowance for any under- or over-reserving within the data. In addition, we would usually allow for a lag between the period of investigation and the data extraction date so that most claims will at least have been reported

(ii) adjusting frequencies and average costs for projected future trends (i.e. frequency changes or inflation).

(iii) a large claim adjustment (especially if the period of investigation does not contain typical large claims experience). We discuss large claims in more detail in section 3.5.

For the purposes of combining the models, we usually update the claims experience to the latest accident year and project on from there. Effectively at this stage we are not treating time as a rating factor.

The trends from the risk premium model quite naturally are very smooth because each individual component model has been smoothed, and any residual plot should show little variation between the actual and fitted values. Any interactions present in the component models will also be present in the risk premium model. In addition it is possible that additional interactions will be generated from the combining process. This is because the proportions by each claim type across a factor may vary according to the level of another factor.

At this final smoothing stage, it may be necessary to fix the relativities of some factors as certain parts of the rating tariff cannot be changed (perhaps for regulatory or marketing reasons). This process is known as "offsetting". By offsetting we are fixing the parameters for one or more factors, and allowing the other factors to find their correct relativities given this offset. A common example of this in the United Kingdom is offsetting the no claims discount (or bonus malus) scale. Many companies have a fixed scale, and underwriters are reluctant to adjust it. No claims discount (NCD) is very highly correlated with policyholder age, and so by offsetting the NCD scale the policyholder age scale will adjust to compensate. There is a danger that if NCD is not offset, but the fixed scale is used in the rating tariff, then certain segments will receive a double loading or double discount. The effect of the offsetting depends on the degree of correlation with the other factors and the difference between the "theoretical" and offset scale.

3.5. Large claims

The average cost of liability claims is highly variable (especially in the United Kingdom where there is no limit on the size of liability). This makes it difficult to pick out trends by rating factor when modelling the liability average cost.

One way of dealing with large claims is to separate out large and small claims. Fitting a model to the small claim average cost will give more stable results than modelling total average cost.

An additional large claim loading is then required so that the total risk premium is not understated. One method might be to simply increase the small claim average cost by the ratio of total liability average cost to small liability average cost. However, this ignores the possibility of certain segments having a higher propensity to have large claims (for example, young people might have a higher propensity to have large claims than older people). To test this we can formulate a GLM with the following structure. The rating factors are once again the explanatory factors. This time, however, the process is a binomial process, with the number of trials being the number of liability claims (both small and large), and the response variable being the number of liability claims that are large claims. If any difference is shown by any of the rating factors, then the adjustment needs to be made factor-dependent.

Note that it is only necessary to make the adjustment factor-dependent if for certain segments a higher proportion of liability claims become large claims. For example, young people are likely to have a higher liability claim frequency than old people. The factor-dependent adjustment is only required if a higher proportion of those claims are large claims.

3.6. Building in external data

It is always possible to supplement the knowledge known about the policyholder by attaching on information from external sources (such as socio-demographic information or credit rating). This is often achieved from the policyholder's zip-code. The external factor is then treated just like any other rating factor.

Many of the external databases available are quite highly correlated with other rating factors already included in the existing rating structure. When evaluating the significance of this information it is essential to do it within a multi-variate framework. This ensures that correlation issues are properly addressed and understood. Simply looking at a one-way table will give very misleading results if the external information is correlated with other factors.

3.7. Data volumes and sparsity issues

Our philosophy when fitting GLMs is to include as many factors as is practical into the model and not to group levels within factors to any great extent. The effect of this is to generate quite large data-sets when modelling claim frequencies, with most of the cells having no claims in them. Those that do will usually only have one claim associated with them. We therefore have a sparse data-set.

Sparsity is not necessarily a problem provided that the model is correctly specified and we have found generally that the parameters from a GLM are robust under sparsity conditions. Standard errors produced, however, are less robust, and care needs to be placed when interpreting them. Furthermore, model diagnostic checking is more difficult since the residual plots will not follow the classical shape of being symmetrically spread around zero. Rather most residuals will be small and negative (i.e. those cells which have no claims), and the rest will be large and positive. However, this is not to say that the model is a poor fit – this is the expected shape of the residuals.

Some people believe that the data should be reduced in order to "improve" the look of the residual plots and produce good fitting models, either by limiting the number of factors in the model or by restricting the number of levels within a factor. The extreme example of this is to reduce the data down to one point, the mean and fit a model to this. By definition it will produce a perfect fitting model, but this tells us nothing about the true within-cell variation in claims experience.

By collapsing the data (by reducing the number of factors in the data, or grouping levels within factors to broader bands) we are deceiving ourselves that we have a better statistically fitting model. The within-cell variation arising from individual exposures will be hidden. More importantly, detail about the variations by rating factor is lost.

We believe, therefore, that the analysis should be carried out with as many relevant risk factors in the model as possible. We will typically have around thirty factors in our models for risk premium and renewal/conversion analyses. The model parameters can then be reduced intelligently as part of the analysis rather than by making prejudgements of the data before the analysis is carried out.

Residual plots can still be used for model validation by simply collapsing the residuals after the model has been fitted using the collapsed fitted values and actual values in broader cell definitions. These residuals will be the same as if the data had been collapsed in the first place, and we do not lose the detail of the uncollapsed model.

4. Estimating Price Elasticities

Section 3 above concentrated solely on estimating relative claims experience by rating factor (essentially the supply side in economic terms). This is the traditional area with which actuaries have been involved.

Just as important to the whole business process, however, is the ability to understand how customers respond to price and price change (i.e. the demand side of the equation). Similar techniques to those used for estimating risk premiums can also be used to estimate price elasticity functions for individual customers, and to identify areas of the account with good or bad conversion/renewal experience.

There are two aspects to demand-side modelling, namely new business (or conversion) demand and renewal demand. There are some differences in the types of analyses performed on each, but the fundamental framework is the same for each.

4.1. Demand modelling as a generalized linear model

Renewal rates and conversion rates once again fall naturally into the generalized linear modelling framework. The explanatory factors are the rating factors, plus additional factors which are discussed in more detail below. The response variable is whether or not the policy converts/renews.

The link function is usually taken to be a logit link function, so the model is of the form

$$y = 1 / (1 + \exp(-X\beta)) + \text{error}$$

By having a logit link function we are forcing the fitted values to fall between 0 and 1 (as they should do). Alternative link functions include the probit link function and the complementary log-log function, although in practice the chosen link function makes little difference to the fitted values. The response variable is quite naturally a binomial (0/1) response.

The process of simplifying the model is the same as that for modelling the risk premium element, in that we are trying to find the most parsimonious model for the data. Similar statistical tests and judgmental issues need to be considered.

Because of the 0/1 nature of the response variable, the results from demand models tend to produce less variability than that in risk premium models, and the results derived are often quite smooth. Less data is necessary to generate very good results; indeed demand information is much more dynamic than risk premium information, so it is important that the experience used is not too out of date to be useful.

4.2. *Conversion modelling*

The modelling of conversion (or new business) rates can only be achieved effectively if full information of both quotations that converted to policies and quotations that were declined is known. Currently this can only be achieved effectively by direct response operators; usually companies operating through intermediaries or agents only find out about the quotations that are converted to policies.

Information available includes all questions asked during the quotation (i.e. the traditional rating factors), plus the premium offered. Other information that the insurer might know about the proposer includes:

- information provided by external data providers (such as socio-demographic or credit information)
- the source of the business
- whether or not the proposer has existing policies with the insurer
- information about the market premium available to the proposer (we will discuss this in further detail when discussing renewals)
- methods of payment
- marketing/campaign information

The explanatory factors in the model are the rating factors that are asked of the customer plus any additional information, for example, those listed above. The response is 1 or 0 depending on whether the quotation was accepted or not.

One question about the modelling process is whether or not the premium quoted should be included in the model. The answer to this question depends on the purpose of the analysis. If the purpose of the analysis is to derive price elasticity functions, then the quoted premium should be included in the analysis, and the derived parameters represent the elasticity curve. However, if the purpose of the analysis is to identify the relative conversion rates by segment, then usually only the rating factors would be included, and the parameter estimates derived will indicate the relative conversion rate by segment.

There are some practical considerations to bear in mind when extracting data for a conversion analysis. These are briefly outlined below.

(i) Often one proposal will result in several quotations, for example if the proposer requests alternative quotes under different deductibles. For the accepted quotes we know what deductible is taken, but for the declined quotes we do not. For these quotes a consistent rule is required so as not to double-count quotations (e.g. take the first quotation generated for the declined quotes).

(ii) If quotations are guaranteed for a period of time, then a time delay is required before doing an analysis to allow for development lags.

(iii) Many insurers will refuse to provide a quotation on certain risks that do not fall within their underwriting guidelines. Usually these should be identified and removed from the analysis.

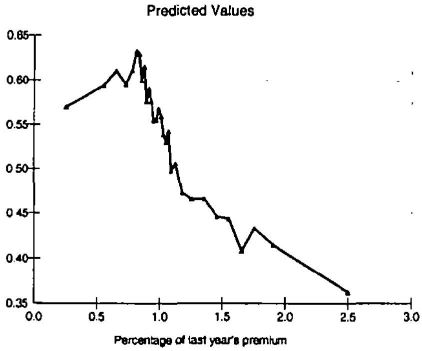
4.3. Renewal modelling

Renewal modelling is in many ways similar to conversion modelling, but usually more information is known about existing policies. In addition, all insurers will know the details of renewals invited and accepted, and so a renewal analysis is not confined to those insurers (mainly directs) who have information about accepted and declined quotes.

Critical to the renewal behaviour of the policyholder is the price movement over the previous year, and so one of the explanatory factors used will be the ratio of this year's premium to last year's. This is a continuous variable, but as discussed earlier this would usually be initially banded into a categorical variable.

It may be necessary to adjust the ratio to allow for the policyholder's expectations of price movements. For example, the policyholder might expect a significant price decrease if they have hit what they perceive to be a critical age. If this decrease is not given, they may shop around elsewhere for alternative quotes. It is impossible to determine exactly what the policyholders' expectations are, but one possible adjustment is to calculate the ratio of this year's premium using this year's rating factors to this year's premium using last year's rating factors (assuming that the policyholders' expectations are reflected in the rating structure). In the UK, the most important adjustment would be for movements up or down the no claims discount (or bonus/malus) scale because policyholders are very aware of like premium movements as the result of having or not having a claim during the policy year.

The graph below shows the shape of a typical elasticity curve. The x-axis is the percentage change in premium on renewal, the y-axis the renewal rate. The graph is plotted for a particular customer segment and this determines the level of the graph. The shape of the graph can also depend on customer segment if the model includes segment related elasticity interaction effects.



Notice that where there has been a large price decrease, the curve tends to flatten off (below 0.8). This is because no matter how cheap this year's premium is, there will always be some policyholders who do not renew (for example if they no longer require auto insurance). Indeed, if they have a very large price decrease they may feel aggrieved by the premium they paid the previous year and seek insurance elsewhere anyhow. Equally, some policyholders are very price inelastic, and for them they will accept virtually any price increase. The steepest part of the graph is around 1 where the premium change has been very small. This makes intuitive sense; the likely impact of a 55% premium increase compared to a 50% premium increase is going to be less than a 10% increase compared to a 5% increase.

Another important input into the model looks at the ratio of the premium quoted to an alternative "market" premium. This ratio is primarily used as an index of competitiveness and is not an absolute measure of renewal rates. There are many reasons why a policy may renew despite this ratio exceeding 1, and equally there may be reasons why a policy may not renew despite this ratio being less than 1. These include:

- depending on how the index is constructed, the policyholder may not have all quotes available, or alternatively may have additional quotes not included when making up the index
- the policyholder may have been with the insurer for many years and is reluctant to switch
- the policyholder may be happy with the brand and good customer service received and so may be willing to pay extra
- the competitive index may not be perfect

The way that the index is constructed depends very much on the distribution channel used and the availability of other insurers' rates. In the United Kingdom, the intermediary market accounts for approximately 60% of the market, with direct response insurers accounting for most of the rest. Quotation systems are available that allow premiums to be calculated across the whole intermediary market for each renewal, and an appropriate competitive level can be chosen. It may not be necessary to choose the cheapest – often the third cheapest, or an average of the top five quotes is more appropriate. In this situation, 40% of the market has been ignored. However, the competitiveness level is only being used as an index, so if the relative competitiveness of the direct market remains constant the model will still be predictive into the future. Even if the direct writers target particular segments of the market (such as older people), the parameter estimates for older people will adjust accordingly.

If other insurers' rates are publicly available, then it is a matter of calculating all alternative rates for each renewal, and choosing an appropriate premium level. Again this appropriate level may not necessarily be the cheapest premium.

In the situation where market premiums are not available, a market index can be constructed from the results of a conversion analysis. Conversion rates are, very often, the best indicator of competitiveness since this is a direct measure of how the company's enquiry profile is responding to their price; an index can be constructed by estimating the expected conversion rate as if the renewal quote was actually a conversion quote.

The importance of market competitiveness depends very much on the distribution method and the easy availability of alternative quotations for the renewals. In the UK, a large proportion of business is written through intermediaries, and so on renewal the intermediary may re-broke if they feel that a better price is available with another insurer. In this case, competitive positioning is very important and is very predictive of renewal rates. For a direct operator, however, the story may be very different. Although the policyholder may be able to get better quotations elsewhere, these quotations are not as readily available; in other words the policyholder is less aware of a better price, and (although still important) competitive positioning is less predictive.

Because of the importance of this factor, and because of the dynamic nature of market premiums, it is necessary to update renewal analyses quite frequently (perhaps monthly or quarterly), and certainly more frequently than a risk premium analysis (risk premiums will not change so rapidly over time).

If a market premium index is not available, then the other parameters will adjust in the areas where the quoted premium is more or less competitive. However, as the market competitive levels change in different segments, so too will the parameter estimates. The model is likely to change quite significantly over time, leading to unstable parameter estimates.

There are many similarities between conversion and renewal analyses, both in terms of the theoretical underpinning and the results that come out of each. Renewal rates tend to be much higher than conversion rates (i.e. renewals are much less price elastic than new business quotations), but we have found in practice that many of the trends seen in a conversion analysis are similar to that in a renewal analysis (albeit at different levels).

5. Developing Optimal Pricing Structures

Section 3 and 4 discussed how to use GLMs to estimate the risk premium and customer demand. In this section we shall start looking at how we can use GLMs to develop optimal pricing structures.

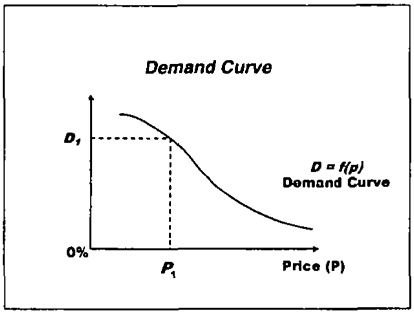
In simplistic terms there are four basic components of estimating the premium to be charged; firstly the risk premium, secondly the direct policy related expenses, thirdly a contribution to fixed overhead expenses and finally a profit margin.

Most premium rating structures do not, however, allow for each of the above elements explicitly. It is normally the case that the profit load and risk premium relativities for different customers are combined in some way. Consequently, in a competitive market, marketing discounts and/or rating action in response to competition become indistinguishable from rating action taken in response to changing claims experience.

However, if it were possible to anticipate an individual customer's response to the new business or renewal terms offered, then a probabilistic approach could be adopted in setting both the contribution to overheads and the profit margin in order to maximise expected profit or some other form of corporate objective. The model needs to be flexible enough to take into account the individual characteristics of each policyholder and be able to respond to the dynamics of the market place. Under certain market conditions or market segments it is quite plausible that the profit loads can be negative.

5.1. Demand curves

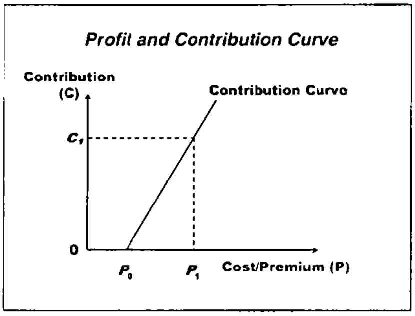
Just considering a one year time horizon initially, the demand for insurance (be it from new business or for renewing policies) for any individual customer is given by the following graph.



Demand curves are generally naturally downward sloping, and the shape and level of the curve depends of the individual characteristics of each customer, on the pricing action taken over the last year (for renewal business), and on the competitive positioning of the market segment involved. The level and shape of the curve will also be significantly different for new business quotations and renewal quotations. As discussed in section 4.1, these functions can be estimated for each customer using GLMs.

5.2. Contribution curves

Ignoring fixed expenses, the contribution curve for a policy looks as follows:



The contribution curve is a 45° line anchored at a point P_0. P_0 represents the expected claims cost arising from the policy (i.e. the risk premium) plus an allowance for direct expenses associated with the policy. P_0 is often referred to as the "breakeven" premium.

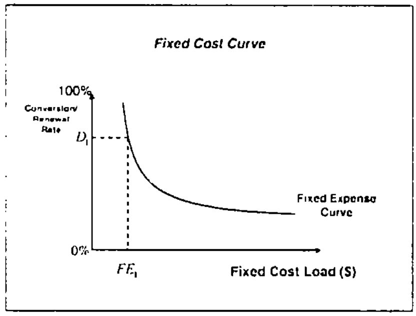
Again, P_0 is very heavily dependent on the risk characteristics associated with each policyholder, because the expected claims costs and associated expenses are heavily dependent on the policyholder. Given that a policy has been written at premium level P_1, the contribution to fixed costs and profit will simply be $P_1 - P_0$.

Again, we have discussed in section 3 how the risk premium and therefore P_0 can be estimated for each customer using GLMs.

5.3. Fixed costs

There is more than one way of building into the rating structure an allowance for fixed expenses. One possible way is to carry out some form of projection of business volume and convert this to a per policy cost. This method is slightly circular in that the volumes written depend on the prices, but equally the loading for expenses in the prices depends on volume. Clearly if more business than expected is written the per policy contribution to fixed costs reduces, the converse being the case if less than expected business volumes are written.

If the probability of conversion/renewal for any individual risk could be predicted, then this could be taken into account in the load for fixed costs prior to issuing the conversion/renewal quote. If a particular policy had a high probability of conversion/renewal compared to average, then the policy contributes a greater expected amount towards fixed costs than if the probability of conversion/renewal was expected to be lower than average. The fixed expense load for each policy can take this into account which can therefore be expected to be inversely related to the probability of conversion/renewal. The graph below gives a typical shape of fixed expense curve.

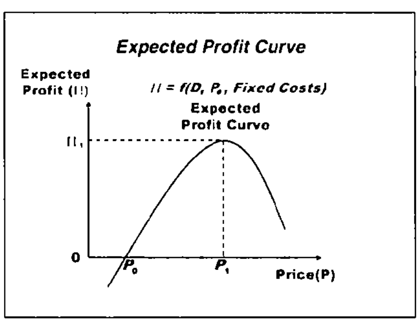


Building in fixed expenses in this manner has the intuitive appeal of attracting policyholders with higher persistency rates.

An alternative way to deal with fixed costs is to ignore them; instead of having an expected profit curve as described below, this changes to an expected contribution to fixed costs and profits curve. It then becomes necessary to check that fixed costs are covered across all policies.

5.4. Expected profit curve

Given the basic components of the demand function and the corresponding costs, it is possible to combine these in order to generate an expected profits curve. The shape of the resulting curve is given in the graph below.



The price that equates to the maximum profit (or minimum loss) can easily be derived, in this case P_1. Other "profit" functions can be estimated. For example, the corporate strategy may not be maximization of profit but rather to achieve satisfactory profit subject to minimum levels of business. These alternative scenarios can easily be calculated by optimising other suitable functions which may place greater value to certain components of the process. Time horizons greater than a year can also be allowed for, and this is discussed further later.

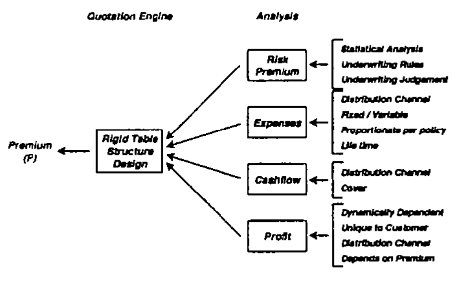
6. Implementing an Optimal Pricing Structure

6.1. The need for flexible pricing engines

We have detailed above the derivation of two of the most important elements of the price, namely the expected claims cost and the demand for the product. We have also briefly looked at fixed expenses. There are of course other considerations, such as a more in-depth treatment of variable expenses, investment income, cash flows and claims development, return on capital etc, but we do not propose to explore these within this paper in any more detail. Instead we will now look at how we can build these different rating components into a tariff system. The statistical analyses undertaken are normally tailored to provide information to help populate the existing tariff structure.

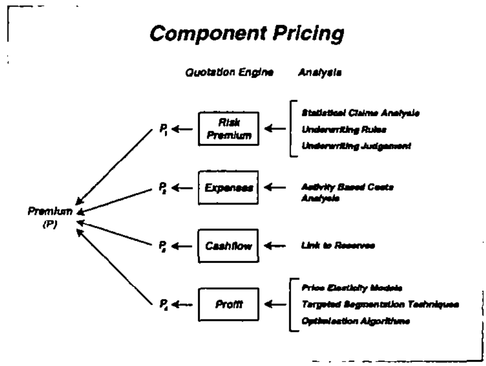
The chart below illustrates how typically the various statistical analyses are currently incorporated into rating tariffs.

Typical Pricing Structures



As you can see, insurers do carry out very sophisticated analyses on each individual element of the premium, but much of this sophistication is compromised in order to fit into a rigid table design from which the premiums are calculated. Much of the analytical benefit is therefore lost at this stage because the lookup tables that are used to derive the actual premium rates are a compromise of the results of the different analyses undertaken. In particular, profit margins cannot be explicitly monitored, leading to a merging of the loads and discounts given for a variety of reasons, for example as a reaction to competitive pressures, for marketing strategy, and for adverse claims experience. This makes it very difficult to manage the profit versus demand equation.

In this paper we have chosen to tackle the problem of implementing pricing structures from a different angle, i.e. how do we best exploit the information obtained from the statistical analyses? We therefore see the need for a much more flexible pricing engine that calculates the premium in several stages. This is shown schematically below.

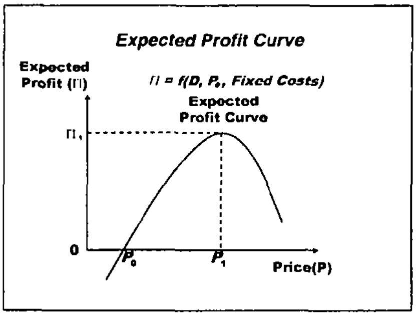


The quotation engine separately calculates each individual element of the premium, and the total premium only calculated at the end. We refer to this approach as component pricing. This approach facilitates much more focused management information. Reports can be designed to specifically manage each premium component. For example, the risk premium component can be directly monitored against the emerging claims cost. Employing such rigorous control on the "cost" components of the premium provides a foundation to exploit the demand for the product given the competitive pressures that exists at the time. At certain parts of the underwriting cycle we may strategically decide to write some segments of the account at a loss. The most important principle here, though, is that informed strategic decisions must be regularly made.

Another way to look at this problem is as a profit optimization process. The optimization will be subject to corporate goal constraints. We shall look at the profit optimization process in more detail in the next section.

6.2. Optimizing Profit

The profit optimization process is the most interesting part of a component pricing system. Whilst the other parts of component pricing can be driven from a series of look-up tables, profit loadings are best implemented by an optimization algorithm. The algorithm will be based on the expected profit curve introduced in section 5, which is reproduced below.



Let us first consider the simple case where the insurer's objective is to maximize the expected profit over a one-year time horizon. The profit optimization process would then look at each customer individually. For any given customer the expected profit curve is estimated by calculating the profit and expected demand (be this conversion or renewal probability) for a series of different premium values. The expected profit at each premium value is then calculated by multiplying the profit with the expected demand. The premium value at which the expected profit is maximized can then be read off. The process can then be repeated for each customer.

The above algorithm simplifies the problem somewhat. It is unlikely that the insurer's objective will be to maximize profits over a one-year time horizon. There may also be some overall corporate goals to consider. For instance, a target of maximizing business volumes whilst achieving satisfactory profitability may be a preferred objective. The algorithm can be easily adjusted to allow for the desired objectives. One way of making the adjustment is to define a utility function that explicitly maps a trade-off between volumes and profitability. The utility function effectively adjusts the expected profit curve so that a new maximum can be derived.

Another factor which the insurer might want to build into the equation is profits made by cross-selling other products. This can be done by either reducing the breakeven premium P_0 by a fixed amount for each policy, or by building a cross-sell model to reflect different cross-sell propensities for different customers.

6.3. Controlling the structure

With such a complex rating structure, it may be viewed as being a difficult task to keep track of and to control the parameters input. In many ways, however, the structure is easier to control than the more traditional lookup table structure because each element of the price can be separately monitored.

The key to successful control involves storing each element of the premium charged in the management information system. This facilitates two things.

Firstly, the actual experience of each component can be monitored against the expected experience. For example, the expected risk premium can be compared against the actual claims experience; the expected conversion and renewal rates can be compared against actual conversion and renewal rates (the latter can be monitored much quicker because of the slow emergence of claims experience).

Secondly, a profit forecast can easily be produced for each month's written business. This will give an early indication to management of reduced levels of profitability; furthermore, the cause of the reduction in profitability can easily be identified. For instance, previously if there was an increase in written policies in a month, the marketing manager would be quite happy because he would be seen to have carried out his job; however, the effect on the bottom-line profitability would be unknown for some time.

By monitoring each element of the price separately, different people or departments can be made accountable for achieving certain goals. The actuaries/statisticians would be held responsible for achieving a 100% loss ratio based on the risk premium component of the total premium: the marketing department would be held responsible for maximizing profits (if they were in control of setting profit levels). Previously if premium rates were too low it would be unclear whether those rates were too low because the actuary underestimated claims experience or whether competitive pressures forced rates down.

7. Adapting the Process in Other Environments

It may not always be possible to implement fully the ideas discussed in section 6 in terms of profit optimization because of regulatory or other restrictions. However, the techniques developed in this paper can be used in a variety of other ways.

One possibility is to reverse engineer the whole process, and use the expected claims costs and demand behaviour to identify the most valuable customers under a set of restricted premium rates. These could be ranked in order of value and scored in terms of customer value. It may be then possible to use marketing means to attract those more desirable customers. For example, it may be possible to market to particular segments in an attempt to improve conversion/renewal rates for those customers, or to be used as the customer value measure in customer relationship management (CRM) programs.

Even if non-pricing means are unavailable, the techniques can be used to effectively cost the consequences of particular pricing action taken, both in terms of expected claims cost and the volumes of business written. The models developed will be based on each individual customer's characteristics and will move dynamically with changing market activity, providing early feedback on pricing decisions taken.

8. Conclusions

We have discussed in the paper the use of generalised linear models (GLM's) and their application in personal lines insurance. We have shown that GLMs have a model and variance structure that closely reflects many of the processes that we often find in insurance. In our experience this leads (if GLM's are used correctly) to reliable and robust parameter estimation. We have also found that these techniques work well on large datasets with many millions of cells and large numbers of risk factors.

Furthermore, we have shown in this paper, how GLMs can be used to estimate both the risk premium and demand functions for individual customers. We then showed how these could be combined and used to set optimal profit loads tailored to the behaviour of each individual customer. Alternatively, the profit load can be estimated for any given premium. The concepts can be extended to cover longer time horizons to create life-time pricing models. Alternatively corporate goals can also be built into the optimization algorithms. In addition, we have demonstrated how, by re-designing the rating tariff, the results of GLMs can be more effectively utilised. The re-designed rating tariff leads to a more rigorous control and management information environment.

Clearly there is more work that can be done. The modelling of the demand function is a new area and the predictiveness of models suggested needs to be more thoroughly road-tested. However, our experience to date is encouraging. Also there are several approaches that could be adopted to defining the optimization algorithm and again these need to be more fully explored. Furthermore, our approach can also be defined in terms of a stochastic environment by incorporating the standard errors of estimates into the algorithms. We hope, however, that this paper will sow the seeds of thought for more ideas and research in these areas.

References

MJ Brockman and TS Wright, *Statistical Motor Rating: Making Effective Use of Your Data*, (JIA 119III)

KD Holler, D Sommer and G Trahair, *Something Old, Something New in Classification Ratemaking With a Novel Use of GLMs for Credit Insurance*, (Casualty Actuarial Forum, Winter 1999)

P McCullagh and JA Nelder, *Generalized Linear Models*, (Chapman and Hall, 1989)

S Millenhall, *A Systematic Relationship Between Minimum Bias and Generalized Linear Models*, 1999 Proceedings of the Casualty Actuarial Society

European Auto Insurance Pricing Considerations

Karen E. Schmitt, FCAS, MAAA

European Auto Insurance Pricing Considerations

This paper provides a broad overview of auto pricing in Europe. The change of markets from a cartel pricing situation to complete pricing freedom is explored. Items that should be considered in developing correct rates are listed and illustrated in turn with examples from the author's own experience abroad. The list and examples are meant to help the reader avoid mistakes stemming from commonly made assumptions. Finally, an approach to pricing is suggested for success as a multinational personal lines insurer.

This paper is a discussion of my personal observations when examining European markets. My feeling is that there are two somewhat contradictory views held by most US actuaries about foreign markets. The first is that these markets are somehow exotic; the second is that they are exactly the same as the US. If either of these conflicting thoughts captures your initial impression, then I feel you are completely right. As I have become acquainted with several local markets, sometimes flirtatiously and other times with more serious intentions, similarities and surprises abound. The constant allure lies in that the unique twist will arise in a completely different arena compared to the last country or market.

One thing to take as a given, given you are an outsider, is that you need to learn as much as possible about the whole of the operation, insurance market and culture in order to understand how to price the business. There are implicit assumptions rattling around in each of our heads that say 'this type of loss is always covered,' 'direct sales works this way, ' etc. We are probably unaware of how many assumptions we hold. But if you seek a complete description by interviewing insurance company professionals in all of the functions, you will soon learn which assumptions you have been holding and push them aside.

The image of the 'international' CAS actuary seems centered on the actuary who works on the large deals or excess business. He or she may work in London market reinsurance, on international insurance programs for the likes of GM and more recently, finite and alternative risk transfer programs. The home base for this actuary could be the US or not. The actuary doing this type of business would have little opportunity to see things in a local way. Nothing allows you to get down to the grass roots of a market as effectively as examining how John Doe or Hans Muster buys insurance.

I do not claim to be an expert on any of the local markets used here as examples. That would take much more experience in each. Nor is there a claim to make any description complete. That would take a large book that would need to be updated daily. Just as the US market has constant changes through the likes of new catastrophe mechanisms, new regulations, and changes in distribution channels, so do European markets. In fact, the rate of change in continental Europe is often more intense than the US. Or at least, has that potential in the very near future. The main cause is the relatively recent change that is simply called 'deregulation' throughout Europe and most of the rest of the world.

The UK motor insurance market really cannot be lumped together with the rest of Europe. It has been deregulated much longer, is much more competitive with the domination of direct sales, and has unique coverage design. It is also difficult to say how the UK is very much like US, except that many companies have very refined rating schemes and underwriting rules driven by a tough, competitive market. This foreign market is quite fluid, partly since rates can be changed daily. Many of the pricing issues that face a US actuary everyday, due to the regulatory processes, are absent.

Deregulation

The word deregulation in Europe means that prices are no longer fixed by an industry cartel or a government agency, but by each insurance company individually. A fixed price structure was the rule in Europe (excluding the UK) until the late 80's or the early 90's depending upon what particular country you are examining. These prices were often simply called the 'tariff.' Each country had its own unique tariff and policy structure with perhaps more similarity between countries with the same language, cultural and legal roots. Coming out of deregulation, these fixed prices embraced all possibilities, being excessive in some countries and inadequate in others. Some lands had very refined rating plans based on statistical methods and data, whereas others had much rougher rating schemes and little science to back them up.

The dates of deregulation for various European countries range from the late 1980's to the early 1990's. The exception is the UK, which has been deregulated for decades. The impetus for deregulation was the formation of the European Union (EU) as an entity with free markets, which of course would preclude pricing cartels. The latest European country to get rid of cartel rates was Switzerland in 1996. While not an actual member, Switzerland has significant treaty relationships with the EU, which often means that Swiss laws must be harmonized.

Each country has its own story to tell as to how deregulation has changed the actual pricing behavior. The cartel mentality has been slow to dissipate. This is to be expected since, for insurance professionals, the whole way of thinking about their job had to change.

It is not necessarily true that deregulation of prices will make for lower consumer prices. An example is Italy, where it is well known that the rates have been too low and that auto insurers are losing money. With deregulation, Italian insurers can raise their rates to the correct level and have indeed made some upward adjustments. But can they raise them enough in the short term to stop the bleeding? Since market share is the mantra of most companies, it seems unlikely that rates will rise immediately to the correct level. On the other hand, in markets where the rates have been historically too high, there is room to slash prices and lure away customers. The question in either situation is still, what is the correct economic price level and when will it reached?

One side effect left over from the days of the cartel tariff is that there are very few local pricing experts in any of these markets today. Reinsurers seemed to be the only ones with any pricing freedom and this is the one area where the pricing knowledge is well developed. Otherwise pricing has been more or less an academic subject. This is a huge problem for companies that are suddenly openly competing with one another.

Deregulation means likewise no regulation of the prices charged to consumers. There are no filing laws.

Rate Regulation

When US actuaries learn a little bit about Europe, they think that pricing is easy since a company can charge whatever it wants. While this freedom is granted, it has yet to be taken up by writers in many markets. It must also be recognized that along side of this freedom is the uncomfortable feeling of flying in the dark. You may complain in your own company about inadequate information, but compare this to (practically) none. In the case of the US, there are the services of the ISO to provide data, which has great value despite whatever weaknesses may exist. In Europe, there is freedom but little support. There is at least one market where a reinsurer's consulting arm has developed detailed pooled data for auto pricing, but this sort of statistical service is unique in Europe.

One problem generated by the lack of regulation is the proliferation of rates used by each company. It is not uncommon to use many different sets of rates for old customers (to 'preserve' the business) and apply an entirely different set of rates for new business. There may also be new territorial designations or car classification systems arising in the meantime, meaning that coding is not even consistent. Some companies will grant discounts if a customer requests a lower rate, which makes managing the price more difficult. While the concept of greater flexibility is appealing, in an undisciplined environment you may find an undecipherable tangle of rates. Taming this beast will become a part of any pricing exercise.

Rate regulation has been a boon to the US and Canadian actuarial profession. Even though the North American filing processes are tedious, it is difficult to get around needing a qualified professional to justify the rates charged. This is perhaps another reason why the concept of pricing actuary in Europe is slow to arrive. There simply is no requirement for rates to be set by an actuary. The reality is, if a market is moving rather slowly, an experienced underwriter can make a reasonable guess at rates that make a profit. This is how much of continental Europe operates at present.

It would be a mistake to think that rate justification or regulation will never arrive in Europe. Portugal requires companies to have an actuary and is already beginning to ask insurers to prove that rates are adequate, which is only a first step. In Ireland, where court awards and premiums are very high, consumerist branches of the government are starting to get interested in rate justification. To avoid any risk of being told what to price, it would be wise for insurers to be proactive and learn to price their own business.

Distribution

In many countries you will hear about brokers for personal lines, which seems unusual for personal lines but is akin to our independent agent. There some countries in which the tied agent (exclusive agent) concept is completely unknown, such as the UK, but others in which they dominate such as in Switzerland. Similarly the level of direct sales, mostly via telephone but also the Internet, varies immensely by country. The UK has an ever-increasing portion of its sales being made by the direct channel after the entry of Direct Line and others in the early 90's.

In some countries, the insurance agent may be your next door neighbor, literally. Insurance is not only sold through the traditional professional agent or company employee. In Luxembourg, for instance, selling insurance is a way to make a little extra money. Think of Tupperware. It is hard not to buy Tupperware from your relatives. There is a lot of natural loyalty built into such a distribution system. In Switzerland, regular employees of insurance companies are actively encouraged to sell insurance on the side as well.

Other than the cost structure, there are control issues at stake in the distribution methods. With direct sales, any rating system that the systems and operators can manage can be implemented. Brokers move the business around for the best price, to keep commissions high, to keep the client loyal to him and not to an insurance company. The broker may be given the discretion to offer five to ten percent discounts in order to attract sales, which would need to be monitored and managed somehow. The brokers may also be using pricing software that forces a different sub-optimal rate structure. For example, UK brokers have created a quotation system which does not allow for prices to be quoted on the full postcode (a full postcode is somewhere in between the five and nine digit US zip codes). The direct writers, on the other hand, have full post code capability built into their quoting engines.

Distribution leads to an interesting terminology issue. Commissions are often add-ons. You might hear someone talk about a target loss ratio of 80% or 90%, or some other equally improbable sounding figure. You need to question what is included in the losses and exactly what the premium covers.

Cover/ Policies

The UK and its neighboring Republic of Ireland offer policies, which are confusingly called comprehensive and third party. There is a saying that America and England are two nations divided by a common language. Anyone who visits France would naturally expect everything to have a different name. This is also true for 'motor' insurance in Great Britain.

When a comprehensive motor policy is discussed, this simply means that all the types of loss are covered (the car is covered on a comprehensive basis). Liability losses are covered as well as theft, damage to your own car, and glass breakage. A third party only policy is just that, coverage for bodily injury and property damage caused by the insured. A third party fire and theft auto policy is in between in terms of coverage.

There are very few options on excesses (deductibles) in the UK and Ireland, but there may be mandatory excesses for certain types of drivers, applying even to third party losses. These excesses may not actually be a fixed amount agreed at purchase, as we are accustomed to seeing. Instead, it may depend upon who is driving the car at the time of the accident. A policy may state, for instance, that if an accident occurs in which the driver of the car is 18 years old, there is a mandatory excess of £300. So, if Mom buys a policy with no excess and includes her teenage son on the policy, there is still potential an excess to be applied.

The customer who chooses the comprehensive auto policy is considered the preferred customer, and additional benefits and rating features are designed to attract more of these customers. This is not dissimilar to the pattern in the US markets: customers with better loss results are often those who wish to protect their assets more fully whereas the high-risk drivers (especially the less careful driver) will select bare bones coverage needed to satisfy the legal requirements.

This design of the policy has natural implications for the price. Comprehensive and third party policies are priced separately in the best companies, modeling the individual loss types and aggregating to an indivisible premium for each type of policy. This in some ways addresses the customer profile more fully than merely adding an amount for each of the coverages purchased regardless of the combination selected. Or perhaps the best way to express it is that the UK insurers address the preferred / non-preferred customer issue in a different way.

The other European countries I have examined thus far have coverage designs more like the US. Liability coverage is purchased, although usually presented as a single coverage unlike the typical bodily injury / property damage split. A common thread is the prevalence of unlimited liability coverage or such high limits that it is practically unlimited. Additionally, one of two variations of physical damage (full or partial) may be bought. In these countries, there are several deductible options to choose from, similar to the US. Innovative insurers are adding more types of physical damage coverage to select from.

A natural implication of the unlimited coverage and fewer deductibles is that there is less need for ratemaking techniques specifically to model censored or truncated losses.

One thing that you cannot assume about other countries is that policy length is either six months or one year. There are markets where contents insurance contracts are 10 years long and auto contracts can be long term as well, up to five years. The premium is fixed at a certain level throughout the 5 or 10 years, which is only altered by movements on the bonus/malus scale (this scale is also fixed over the life of the policy). A customer must give notice well in advance of the anniversary date to break a contract and often must pay a large fee to cancel. Competing insurers sometimes pay this on behalf of the policyholder to entice them to change companies. In Germany, a premium change at the anniversary date means that the policy is now 'new', and that the customer is free to change insurer without penalty. Whereas yearly rate changes on renewals are 'business as usual' in the US, this is a very uncomfortable thought for a German insurer. A customer's reaction to a rate increase may be quite different when he is used to a level premium.

Taxes

As mundane an issue as taxes is, it has a great effect upon insurance. In some countries, there is a very high sales tax on automobiles, making your auto a very valuable possession indeed. What type of behavior or mentality would this generate? Would there be an attitude that the insurance should be cheap because the buyer has already spent so much? Would there be a greater propensity to buy very complete physical damage coverage, a willingness to pay more for that coverage than the liability?

An interesting tax issue is found in the UK market. In the not too distant past, the UK had horrendously high income tax rates. As a tax efficient way to reward employees, company cars were given to many levels of employee and are now considered a standard benefit for a working professional. The net effect of this is that the personal lines market is smaller than would otherwise be the case, and that the average personal lines driver profile may be different than what we expect.

Another tax story comes from Germany. Oscar Lafontaine, the new finance minister in late 1998 had been designing tax schemes to impose a large tax burden on German domiciled insurance companies, recognizing the hidden profits that these companies have. Some of the major players made threats to move to more tax-friendly European Union countries. In the end, the finance minister resigned (not just over this issue), but the issue of the fat reserves of German companies has now been made quite public and the end result is unknown. Certainly, the EU formation is forcing more things to be harmonized, and tax is one of them.

An interesting thing I have observed is that premium tax is often an extra. Prices are quoted and shown as premium plus the tax amount.

Car Ownership and Use

What about car ownership? It is certainly almost a truth of American society that every adult owns a car except in some eastern metropolises. It is so much a given, that it is perceived as a right, which leads to much of the politicization of auto insurance. While auto insurance is an important product line in other countries, it receives nowhere near the political or regulatory attention as it does in the US. This could be because of basic differences in assumptions, meaning basic cultural differences. Numbers are hard to come by and are not often directly comparable; so let's look at some assumptions we make.

A family living in suburbia or small town America has two cars, at least. There is one for each parent and perhaps a car for teenagers to use or a sports or other special occasion car. The parents both need cars because they both work or, if the wife stays home, she needs to drive to the market to buy large amounts of groceries once a week or to drive the children to play groups, etc. (Please forgive me if these assumptions offend, but we need to start with an example.) The work place is likely in some suburban office center, which is conveniently near some highway exit but is not (or not well) serviced by public transport. There is almost certainly no direct route of the public transport from home to work.

A couple living in the suburbs of Zurich may not have two cars. Public transit is very good, so it is not necessary to drive to work. Public transit is also seen as a way for everyone to travel, not just those who cannot afford a car (often the US attitude) or city commuting. It is safe and clean and runs throughout the day. Most work places are near public transit stops. Switzerland is much more densely populated than the US, and so requires mass transit in order not to have constant traffic jams (which it has anyway). There are many natural barriers in almost all major cities in Switzerland such as lakes, rivers, and mountains which mean fewer ways for everyone to get from A to B. Although the trends are changing, most women do not work after having children, and by tradition, marketing is a daily event with stores within easy walking distance. Factor in the much higher cost of car ownership starting from the purchase price, heavily taxed gasoline (roughly US prices time 4) and high labor costs for repair and maintenance, which makes car ownership much more of a burden. An interesting fact is that Switzerland has a relatively high level of car ownership compared to the rest of the European continent despite all of the obstacles and reasons to not have one.

I use this one example to illustrate how it may be false to assume the same level of car ownership in other countries. It may also be false to assume to what degree cars are being used to commute versus other activities, which would preclude adopting US relativities for type of use.

Reserving

With a US background it was difficult at first to see how the completely different rules governing statutory reporting and local accounting standards would have such a large effect on pricing automobile insurance. Reserving to the best estimate is a given in the US market, with some rather limited latitude for conservatism or optimism. This practice had not been the rule in other countries, which may have reserving practices which consistently under or over estimate the ultimate pay out. These could either be the statutory rules or the local convention. Obviously, the reliance upon the reserves as the correct view of reality when it is not leads to incorrect pricing. This happens in companies with little actuarial, IAS (international accounting standards) or US GAAP experience.

The Germanic countries (Germany, Austria and Switzerland) are required to have equalization reserves, the sole purpose of which is to smooth out the results over time but which are bounded at the low end by zero. Therefore, the reserves of many companies are quite fat. Equalization reserves could easily become a means of financing an all out pricing war to gain market share or to keep prices low to maintain position. Having such a comfort margin in the reserves however, means that the management's incentive to move towards better more actuarial pricing will be slow to build.

Another interesting twist is the definition of IBNR. In much of Europe, the traditional task has been to estimate the pure IBNR. Estimation of the redundancy or deficiency of the case reserving was not done. A less than full understanding of the reserve need just adds barriers to actuarial pricing.

Bonus / Malus

This is a topic to which many European actuaries have contributed in the academic literature. To the uninitiated, bonus is a discount for good claims experience and malus is a surcharge for bad. Each country has its own scale. Most insurers in a given country still tend to use the same one (some are creating their own schemes), but there are many differences country to country. Even in the UK, where price freedom has existed for a long time, you will find remarkably similar bonus systems (the idea of malus has gone out the window). The level of highest bonus in the UK (around 50 to 60%) is reached after 4 to 6 years, making it more of a marketing scheme rather than a way to truly distinguish risk. In contrast, it can take up to 25 years to earn the best rate with a German insurer. The bonus / malus systems have developed as a market tradition and the customer values highly the bonus earned. The customer is so concerned with maintaining bonus status that it had become possible to sell bonus protection profitably. An individual insurer is unlikely to abandon bonus / malus any time soon no matter what the statistics may say.

Claims Handling

We once had a data set from a European company with losses, claims handling expenses and premiums and came to the conclusion that the claims expenses were either wrong or mislabeled. This was an incorrect conclusion stemming from not knowing the facts. There was an option for the customer to hire an independent adjuster at the expense of the insurer and these expenses were often higher than the claim itself, which of course causes problems for the insurers in cost containment.

In the US, there has been a general trend toward more controlled claims handling through preferred repairer programs, glass repairers, and claims specialization, to name a few. These programs are being taken up in other countries as well. Companies with UK ties or US will likely take successful claims innovations and modify them to their markets to save money and avoid putting in rate increases.

Just like the reserving actuary, the pricing actuary needs to understand the claims handling in order to do the job right.

Legal system

Needless to say, each country has its own legal system and speed of change in implementing new laws. These new laws will impact trend picks and coverage design just as in the US. Despite the unlimited liability coverage afforded in many European countries, there is still much less of a deep pockets attitude than in the US. However, other countries are busy catching up with American ideas and ideals and unfortunately, the trend toward using lawyers more often is one of them.

Other Considerations

In order to round out the checklist of what to examine when investigating another country, the following are listed with no guarantee of completeness. It is always recommended not only to interview employees of the company and examine the rate book but also to read the policy itself. Interesting facts can be found in these policies, such as a listing of the items taken into consideration in the pricing (Germany), a description of the entire bonus-malus system including the level, percent discounts/surcharges and the rules for the moving up and down the scale (many countries). All this has quite a restrictive effect on the pricing actuary.

Other things to investigate are:

Payment methods and interest
Assigned risk or other requirements to insure
Insurance regulation on price, coverage, claims handling, etc.
Rating plan
Data availability

Pricing in a Multinational Company

A company offering personal lines in many countries is much more multinational than international. International somehow implies that the boundaries have been transcended and a unified product, service or image is offered. Multinational emphasizes that we must operate within each country in a unique way, even if there are some elements that go across country borders. The multinational approach allows us to honor the local cultures, laws, distribution systems, and economy. This approach was necessary before deregulation and the creation of the

European Union and will continue to be important in order to succeed. The greater strength, however, lies in the multinational company that can also use information and intelligence from one market to the benefit of another.

We considered if it were possible or even desirable to have an ISO-type setup as a model for rate making in Europe. Would it be possible to have a center that creates rating plans that are applicable to all countries, and send them out for implementation? There is an appeal to this. After all, if the physical damage relativity from a Mazda to a Ford is X in both New York and Louisiana, why not France and Italy? A car is a car, we think, but is it? Consider that imports in the US are generally more expensive to repair than domestics. Not only is Ford an import in (most of) Europe, an Audi is a domestic car in Germany and an import in France.

While learning about the Association of British Insurers (ABI) rating used in England and observing how well they work there in ranking the cars, we considered this as an alternate tool to the ISO approach. In Ireland, the opinion was that the ABI ratings did not work there for much the same reason as ISO: English cars are cheaper to repair in England than imports. But English car parts are just as expensive as those from anywhere else, since all cars and their parts are import items in Ireland. Perhaps a tool such as the ABI rating would be more transferable if it varied the code by cause of loss, such as the car groupings in Germany do. This may be another reason the ABI does not suit Ireland. Ireland has a larger proportion of it losses in bodily injury and the ABI assignment is the same for each type of loss covered. An interesting observation is that it is more common to vary liability premium by the type of car driven in Europe than it is in the US. When modeling the losses for liability in the UK, differences for ABI grouping in liability losses are fully recognized and incorporated. In Germany, most insurers use a table which assigns a car group to each make and model not only for physical damage but for liability as well. Although I have been out of the US for a few years, I have heard of reluctance to move forward on this issue because of conflicts with safety issues and marketing.

Unfortunately it is not an easy step to unify vehicle groupings across countries. Further to the above complications is that each country has a different idea of how to create vehicle groups: England as above, Germany using many more groupings but the assignment of a car to a group varies by coverage. Many other countries use horsepower or value in combination with the number of seats.

The rating relativities for driver characteristics would seem intuitively different from land to land since there are legal, cultural, demographic and economic differences that affect this. There is further the complication that each country and even market within a country is at a different point in pricing and data collection. Whereas the Germans may use and code insured' and 'insured and spouse' a British insurer may use 'insured' and 'insured plus one other driver' or 'insured plus spouse.' So the existing rating plans diverge from each other as well as from the 'married' and 'single' categorization that we North Americans would like to impose. In Ireland, it is presently possible to hold a provisionary license for several years due to the backlog in driver testing facilities. How do you compare the experience of the Irish temporary license holders to drivers in other countries who typically hold a temporary license for a much shorter period? And the list of differences continues to grow. A couple of these little quirks can be handled if you are trying to bring things together. However, there are just so many differences that it soon becomes impossible, in addition to the result being, most probably, incorrect.

Within a multinational company, another tactic is needed. Rather than imposing a new standard rating structure on each local market, a better approach would be to impose a standard for analysis. This is an ideal starting point, remembering that while most of these local companies have their own data systems and structures, acceptable pricing criteria, policy forms, etc., they are unlikely to have actuarial pricing tools. This would be a logical first step in drawing the most value out of the existing resources. It helps to imagine the set up of a multinational company. Each local company having in the not too distant past a cartel rate with given rating criteria and coverage which were developed independently of the other countries, each company having its own local management team and identity. Given this as the usual set up, along with the need to abide the local regulations and the cacophony of languages, what multinational is likely to have a centralized actuarial function for personal lines already in place? And rather than imposing a center with little knowledge of each market, wouldn't a better model be to build a bridge to enable experts to work together?

This common analysis approach would facilitate a sharing of experiences and potentially data, which has always been an incredible obstacle for the multinational insurers. It may be that there is potential for data to be shared across country borders. That would be of great benefit for any insurer since historically so little has been done with such data. In addition to sharing knowledge about the business, the development of pricing capabilities would be accelerated with a common analysis process.

For this common approach to work, it has to focus on building the rates from the bottom up. As I see ratemaking in the US, much work is put into making sure each state's coverages are correctly priced along with the territorial relativities for each coverage. Other rating plans are updated only every few years. More than once I have seen large changes in relativities causing price disruptions under such a procedure. This seems to occur whether the company is a well-staffed company using its own rating plans or the ISO itself. This does not mean that another system would necessarily avoid large changes, but that a method that allows more frequent review on a detailed basis could help.

Another reason to work more from the detailed level is that speed of change is expected to be dramatic. Pressures to turn a profit on the current year's results will come. Consolidations and takeovers will continue. Companies with sophisticated parents or branches will begin to apply more refined rating plans, driving others to either do the same or leave the market. Profits in the divisions handling larger clients will continue to thin as too much capital courts them, focusing companies on the money making potential of the less sexy areas, such as personal lines.

Similar to the Americans, the Europeans are entering the newer distribution systems of group, direct phone and Internet sales. These distribution channels demand rigor in getting the details correct and the greater data accumulated, without the broker filter, invites statistical analysis. On the other hand, if an agent is not acting as a screen for the company, if suddenly the customer profile is different to that in your prior books, money could be lost quickly. And, with the data you are able to gather for your own proprietary use, it would be almost criminal not to exploit it to its full potential.

Working at the detailed level of analysis brings value to the smaller units in the multinational insurer as well. This is due to the ability to draw on the resources of other countries' analyses. However, a detailed analysis of similarities and differences would be required in order to select the proper base country for the various loss types and rating plans.

Lack of rate oversight allows for more creativity and indeed more science in the methods applied. A reason cited for not taking certain approaches in the U.S. is that they may be too difficult to explain to the regulators. Or, that methodology should be kept consistent to assure regulatory approval of price changes. Methods can be constantly refined in an unregulated market without such concerns. Because of this freedom, Europe may very well be the first place that more long-term focused pricing models take off. Pricing models could take into account the pure premium expectations of a policyholder for a time horizon longer than just one year, factoring in the renewal propensity. It would be a great leap forward if we could bring these longer-term pricing considerations into property and casualty insurance. In addition to the rating freedom, wrapping long-term thinking into the price may be easier to sell in continental Europe where, with the history of multi-year policies, the focus has not been quite as short-term as in the US.

Given the differences in the markets, data structures, data availability, and the need to work from the detailed level, those companies who will profit the most are those who tend carefully to their data. Those who survive will not only have appealing theories but hard facts to back them up. In the short-term, assuring access to the masses of information locked within the current customer profiles will be a success criteria. In the medium to long-term, the key will be to build bridges either in terms of the actual data or the information drawn out of that data.

Summary

The ones who succeed in 'going international' in personal lines markets are those who are able to transcend the usual tendency to take truths about one market and (blindly) apply them in others. Instead, they will think 'multinationally.' Each country has its own culture, economy and laws. Each locally grown operation has its own rating structure, coverage, distribution system, etc. The key will be a flexible approach in how the integrated pricing and marketing is designed. Structure is better placed in required basic analysis tools, which are robust enough to cover a major portion of the facts needed to make good decisions. Data collection will be essential. The trend towards continual efficiency and consolidation will almost necessitate these points for any company who has aspirations of success as a multinational personal lines insurer.

Ratemaking Considerations for Multiple Peril Crop Insurance

Frank F. Schnapp, ACAS, MAAA,
James L. Driscoll, Thomas P. Zacharias, and
Gary R. Josephson, FCAS, MAAA

Ratemaking Considerations for Multiple Peril Crop Insurance

Submitted by:

Frank Schnapp A.C.A.S., M.A.A.A.
Director, Actuarial Analysis and Research
National Crop Insurance Services

James L. Driscoll
Senior Actuary
U.S. Department of Agriculture
Risk Managment Agency

Thomas P. Zacharias
Executive Vice President
Actuarial, Statistical and Information Services
National Crop Insurance Services

Gary R. Josephson F.C.A.S, M.A.A.A.
Consulting Actuary
Milliman & Robertson, Inc.

Abstract

Multiple Peril Crop Insurance (MPCI) is a unique public/private market insurance product. This paper is intended to provide an introduction to the MPCI ratemaking process, as well as a discussion of some of the political and economic forces affecting the program.

The paper will provide a description of the coverage offered under the program and an overview of the ratemaking methodology. Specific challenges relating to the catastrophic nature of the coverage and the geographical influences on loss exposure will be discussed. In addition to current ratemaking techniques, which involve a credibility weighting of county experience with the experience of adjacent counties, the paper will discuss alternatives, including:

- fixed territorial groupings of counties,
- spatial smoothing, and
- spatial credibility techniques.

The paper will discuss unique aspects of the product and the ratemaking process, including:

- the role of the Federal Government in supporting the program,
- the high correlation of experience among exposures, and.
- the use of econometric models and non-insurance data in validating experience.

The paper will also discuss some of the recent changes in the federal crop insurance program and how these are reflected in the rate process.

I. Introduction

The Federal Crop Insurance Program

The Federal crop insurance program is a joint effort of the Federal government and private industry. The insurance product, which is known as Multiple Peril Crop Insurance (MPCI), was created to serve the needs of farm producers in the era of the Oklahoma dust bowl. For many years, the participation among farm producers was very modest despite subsidies provided by the Federal Crop Insurance Corporation (FCIC). In order to increase participation, Congress authorized private insurers to sell, service, and underwrite MPCI coverage beginning in 1980. This enabled Crop Hail insurers to market a product which previously had competed against their own.

Since 1998, the sale and underwriting of MPCI coverage has been completely privatized. The current Federal role in the program consists of three essential activities: establishing the MPCI rates and rules, subsidizing the premium and the administrative costs of the program, and administering the reinsurance mechanism for the participating insurers. These activities are managed by the Risk Management Agency (RMA) of the United States Department of Agriculture (USDA). Subsidization of the program is necessary in order to keep the cost of coverage affordable to the individual farm producer. The reinsurance mechanism, implemented in the Standard Reinsurance Agreement (SRA) between insurers and the FCIC, is necessary to protect insurers from severe or catastrophic losses. By combining the marketing efforts of the private sector with the financial strength and support of the Federal government, the Federal crop insurance program has become much more successful in achieving its aim to provide financial protection to farm producers. An indication of this success is that in cropyear 1998 the MPCI program provided coverage on 181 million acres (almost 70%) of U.S. cropland, insured $27.9 billion in crops, and generated a total premium of almost $1.9 billion.

Public Policy and Federal Crop Insurance

Federal crop insurance was authorized by the U. S. Congress in the 1930s as a pilot program. It was one of several public policies to assist agriculture's recovery from the Great Depression and the Dust Bowl years.

This legislation followed several failed attempts to offer such insurance commercially. Costs for salaries and other operating costs of the program were paid from the U.S. Treasury, and persons taking the insurance paid the full risk premium. Insurance was restricted primarily to major crops in principal producing areas, with annual premium volumes well under $100 million. Operations were managed completely by the Government.

In the 1970s, free disaster assistance protection for certain crops was authorized as part of price support legislation affecting agriculture. By the late 1970s, the dichotomy of this coexistent public assistance -- one free and the other partially subsidized -- resulted in passage of the Federal Crop Insurance Act of 1980. This Act made the free assistance unavailable if crop insurance was available. To make insurance more attractive, the risk premium was partially subsidized. This Act also authorized the Government to reinsure commercial insurance companies that sold and serviced the Federally-developed insurance policies at the Federally-approved premium rates. Additional subsidies were authorized to pay the operating expenses of those companies.

Following this Act, more crops and growing areas became eligible for insurance. Premiums increased from $156 million in 1980 to $436 million in 1988. However, the 1988 premiums represented only about 18 percent of acres planted to principal crops, a level that proved inadequate to withstand demands for disaster assistance.

Beginning in 1988, several years of adverse weather conditions affected different parts of the country, culminating in the floods of 1993 that impacted urban areas as well as agriculture. Several ad hoc assistance bills (i.e., temporary rather than permanent measures) were enacted in 1988 and subsequent years. These ad hoc measures typically paid more benefits to producers who had insurance and also required beneficiaries to purchase insurance the following year. By 1994, premiums had increased to nearly $950 million and insured acres approached 40 percent of planted acres.

In 1994, Congress again amended the enabling legislation for crop insurance. Insurance was required as a condition of eligibility to receive benefits available under other Federal programs for agriculture. A level of coverage intended to provide benefits only in the event of catastrophic losses was introduced and offered to producers for a minimal fee. This legislation also increased the subsidy for those persons who carried higher coverages than this minimum. In 1995, premiums increased to $1,550 million and over 80 percent of planted acres were insured.

Although the mandatory purchase of crop insurance was rescinded for 1996, the level of crop insurance sales remains high. In 1998, premiums reached a record $1,875 million and insured acres approached 70 percent of planted.

During the 1990s, the Congress also authorized subsidies and reinsurance for commercially developed insurance products. The first of these was offered in 1996. It modified the traditional insurance plan that indemnified only losses in yield so that changes in market prices for the insured commodity also could result in an indemnity. A commercial product from a second company, again one that includes risk of changes in market prices, was introduced in 1997. A third commercial product has been introduced for 1999. More are anticipated in future years.

The thrusts of public policy during the past two decades have been twofold: encourage farmers to actively manage the risks they have in farming and to encourage commercial insurance companies to be more active in this market. Both thrusts have been successful. More acres are insured, premiums are at record levels, and private companies are much more involved with providing this coverage.

A Description of Multiple Peril Crop Insurance

MPCI coverage is designed to insure the yields of farm producers over an entire growing season on an all risks basis. The primary cause of loss is weather, either for a single identifiable event or over an extended period. More specifically, perils include wind, rain, drought, hail, fire, prevented planting due to too much rain, flood, disease, insects, cold, frost, or any other reason for low yields. Due to the high damageability of crops, coverage is only provided in excess of a large deductible. The MPCI program currently provides coverage for almost 100 crops in all 50 states, but not including the District of Columbia. The program is gradually being extended to cover additional crops not currently insured. Crops currently being evaluated in pilot programs include cabbage, sweet cherry, winter squash, wild rice, and watermelon.

Since the Federal crop insurance program is a public/private partnership, public policy considerations have a significant influence on the operation of the program. For example, insurers are required to accept all applicants. In addition, the farm producer selects the amount of coverage to be purchased. Since Congress regularly evaluates the operation of the program, pricing and policy design decisions may differ from those that would be made if MPCI were solely a private insurance program. Public policy considerations may also result in unanticipated changes to the coverage after the policies are sold.

The MPCI premium is computed as product of the published rate and the exposure. Generally, the premium is paid at the end of the cropyear. The MPCI exposure is the liability measured in hundreds of dollars. The liability represents the total insured value of the crop, and is the product of:

the APH yield
the acres planted
the selected Coverage Level
the Base Price for the crop, and
the Price Election percentage.

The first element in this calculation, the APH yield, is based on the producer's Actual Production History. The APH represents the producer's normal yield, and is based on 4 to 10 prior years of yield information.

The number of acres planted may be estimated by the producer at the time the policy is issued, in order to have an estimate of the premium. Subsequent to planting, the producer must file a report of the actual acreage planted. Since the acreage is verified at the time of loss, overreporting will not increase the producer's indemnity payment. Underreporting of acreage will result in a penalty because the total value of production on all of the producer's acreage, including any unreported acreage, is compared to the insured liability to determine the indemnity at the time of a loss.

The Coverage Level represents the producer's deductible, with a Coverage Level of 75% meaning that the insurance pays nothing if the loss is less than 25% of the value of the crop. Coverage Levels ranging from 50% to 75% are currently available, with 85% coverage being offered in some counties on a pilot basis.

The Base Price represents the price of the crop at the start of the growing season, and is established by RMA based on the latest market prices. The Price Election percentage allows the producer to further modify coverage by insuring the crop at a lower value than the Base Price. The producer may choose to insure production at any level from 60% to 100% of the Base Price for the crop. For example, a lower Price Election percentage may be selected if the producer wishes to insure only the cost of planting rather than the full value of the crop. The Base Price and the Price Election percentage are also used to determine the value of the crop for loss indemnification.

The following example reviews the steps in determining the liability, premium, and loss. The Producer Premium Percentage Factor in Step 2 is taken from a countrywide Table and represents the premium subsidy factor for the selected Coverage Level and Price Election percentage. The premium subsidy is discussed further in the next section.

Step 1: Determine the Liability		
Acres Planted	500	acres
Actual Production History (APH)	120	Based on producer's past bushels/acre
Coverage Level Selection	75%	Selected by producer
Base Price	$3.00	$ per bushel, established by RMA
Price Election	100%	Selected by producer
Liability	$135,000	= 500 x 120 x 75% x $3.00 x 100%

Step 2: Determine the Premium		
Liability	$135,000	from above
Rate	$2.00	per $100 of liability
Risk Premium	$2,700	= $135,000 x $2.00 / 100
Producer Premium % Factor	0.765	Subsidy factor for selected Price Election
Producer Paid Premium	$2,065	= $2,700 x 0.765
Step 3: Determine the Loss		
Actual Amount Harvested	40,000	bushels
Value of Production	$120,000	= 40,000 bu x $3.00/bu x 100% price
Liability	$135,000	from above
Indemnity Payment	$15,000	= $135,000 - $120,000
Summary: The producer expects to harvest 60,000 bushels and insures 45,000 bushels. Since 40,000 bushels are actually harvested, the indemnity represents the value of 5,000 bushels at the Base Price of $3.00 per bushel. The market price for the crop at the time of harvest is not considered in this calculation.		

Unlike standard Property and Casualty contracts, MPCI coverage is not triggered by an event. Instead, the indemnity payment is determined after the crop is harvested. The producer's actual production is multiplied by the Base Price, adjusted by the Price Election percentage, to determine the value of the crop. This value is compared to the Liability under the contract. If the value of the crop is less than the Liability, the producer is paid the difference.

The standard MPCI policy insures the producer for a loss of yield, not a loss of revenue. The policy includes no protection against the risk that the market price at harvest will be different from the Base Price established at the start of the growing season. If the market price is lower than the Base Price, the producer's total revenue will be less than was anticipated at the start of the growing season. The producer can obtain protection from crop price changes during the growing season through a variety of mechanisms, including forward contracts, futures, and options. A recent innovation is the development of revenue contracts which extend the standard MPCI coverage to include market price protection. For the producer, the simplicity of purchasing protection against fluctuations in crop prices as part of the MPCI coverage has proven to be very popular, with more than 13% of all crop insurance premium arising from revenue coverages in just the third year since their inception. The design and rating of revenue contracts is an interesting subject which is beyond the scope of this discussion.

Unique Features of the Multiple Peril Program

The involvement of the Federal government in the MPCI program creates a social insurance program which operates on different principles than a privately underwritten insurance market. The important differences include producer premium subsidies, insurer expense reimbursements, and pricing for an underwriting loss. This section discusses these differences and other unusual characteristics of the program.

The MPCI program offers two levels of coverage, known as Catastrophic and Buy-up. The Catastrophic level of coverage protects against only the most severe outcomes, such as a complete crop failure. Specifically, Catastrophic coverage reimburses producers only when the actual production falls short of 50% of the APH yield, with the loss of yield evaluated at a 55% Price Election percentage. The most a producer could collect under this coverage would be 27.5% of the expected value of the crop. The producer premium for Catastrophic coverage is completely subsidized except for a $60 administrative fee per county per crop which is paid to the Federal government. However, an "imputed" premium is established which represents what the producer would pay if no subsidy existed.

Buy-up coverage allows the producer to purchase additional coverage at a partially subsidized price. However, the Buy-up and Catastrophic coverages are priced and sold as different deductibles rather than as distinct products. The premium at all deductibles is subsidized by a dollar amount determined from the cost of the catastrophic coverage.

For the 1999 cropyear, additional Federal financial assistance of $430 million will be provided to encourage the purchase of more adequate amounts of coverage. The additional financial assistance originated with an emergency farm bill recently passed by Congress. Preliminary estimates are that this will result in a 30% reduction to the Buy-up coverage premium. Since the Catastrophic premium is completely subsidized, the producer paid premium for Buy-up could decrease the full 30%. However, a significant percentage of this savings is being used to purchase higher levels of Buy-up coverage, as was intended when this additional financial assistance was offered.

Congress has expressed a desire to eliminate emergency Federal disaster assistance and to use the Federal crop insurance program as the primary mechanism for directing aid to producers. This objective is consistent with recent international trade agreements which restrict the types of subsidies which nations can provide to producers. Insurance is considered to be a form of farm income protection which does not distort producers' market incentives to grow particular crops, and for this reason is excluded from the treaty restrictions. In comparison, disaster assistance provides an incentive for the producer to grow as large a crop as possible. The expectation that Congress would protect producers from unanticipated losses would encourage excessive planting, resulting in reduced crop prices.

The rates established by RMA do not include a loading for insurance company expenses. Instead, insurers are compensated by the Federal government for their expenses in a separate arrangement. Currently, Congress has authorized an expense reimbursement of 24.5% of the premium for the Buy-up coverages. The expense reimbursement is intended to compensate an insurer for its commissions, administrative expenses, and all loss adjustment expenses. In comparison, a loss adjustment reimbursement of 11% applies to the imputed premium for the Catastrophic coverage. State premium taxes do not apply to MPCI premiums. The reimbursement percentage has been reduced significantly in recent years. This reduction was partially justified based on the high crop prices in the mid 1990's, since high prices result in an increase in the expense reimbursement payments without a corresponding increase in insurers' actual expenses. The reimbursement is not intended to generate a profit for private insurers. However, the reduction in crop prices in 1998 and 1999 has not resulted in an offsetting increase to the expense reimbursement percentage. As a result, the actual expenses of the crop insurance industry now exceed the amount of the expense reimbursement according to one study.

Another unusual aspect of the program is that the MPCI rates are currently established to produce a long term loss ratio of 107.5%. Since the premiums are collected and the losses are paid at the end of the cropyear, little or no investment income can be earned. As a result, the program is not designed to produce an operating profit on a direct basis for participating insurers. To encourage private participation in the MPCI program, the reinsurance arrangements in the SRA have been designed to enable insurers to earn a reasonable profit on a net basis. The financial and operational details of the SRA are complex and are beyond the scope of this discussion.

The Standard Reinsurance Agreement between insurers and the FCIC is designed to transfer much of the crop insurance risk to the Federal government. Previously, the SRA required insurers to reinsure their exposures by county and crop. Beginning in 1998, it permitted individual policies to be reinsured. Unlike the expense reimbursement percentage, the SRA is negotiated between RMA and private insurers.

In situations in which RMA broadens the MPCI coverage subsequent to the final date for policy revisions, insurers may experience greater losses than they would have otherwise anticipated. These revisions may also arise too late for insurers to cede the affected exposures to the SRA. These situations are negotiated between RMA and the insurance industry, with an occasional recourse to litigation.

Rates for Individual Producers

MPCI rates are established for combinations of county, crop, and farming practice. Certain crops such as wheat may be rated by the variety, such as winter wheat vs. spring wheat vs. durum wheat. Farming practices differ by crop and location. An example of a farming practice would be the distinction between irrigated vs. non-irrigated crops.

Since a producer may plant several crops or use more than one farming practice, the producer is rated for each distinct crop and practice. Also, coverage for certain practices in selected counties may be unavailable in order to prevent adverse selection against the MPCI program. For example, coverage for non-irrigated extra long staple cotton is unavailable in certain counties in Texas.

The rate structure for an individual county is fairly simple. For a given crop and practice, the two key rating characteristics are Coverage Level and Rate Class. Coverage Level generally ranges from 0.50 to 0.75 in increments of 0.05. The rates for lower Coverage Levels are less than those for higher Coverage Levels since a low yield is less likely than a more normal yield. Rate Class represents a subdivision of the APH yield range. Studies have shown that producers with lower than average APH yields also have significantly higher variability of yield. Since MPCI coverage protects against lower than expected yields, these producers would have relatively greater losses than producers with average yields. For example, both a low yield and a high yield producer may purchase 75% Coverage Level, but the low yield producer is more likely to have a poor crop, resulting in more claims than the producer with the higher and more stable yield. For this reason, different Rate Classes are established for producers with different APH yields. The number of Rate Classes depends on the crop. For many crops, the APH range is generally subdivided into Rate Classes R01 through R09, with R05 representing the typical yield. The APH ranges corresponding to each Rate Class are determined by defining R01 as any yield below 50% of the average yield, defining R09 as any yield above 150% of the average yield, and defining the remaining Rate Classes using bands of equal width.

Other rate adjustments include a credit for insuring the producer's entire operation as compared to insuring individual fields and a credit for a Hail and Fire exclusion, which may be useful if the producer elects to purchase Crop Hail coverage in combination with MPCI. Disregarding these exceptions, the final MPCI premium is developed as the product of the appropriate rate, the Liability, and the Producer Premium Percentage Factor corresponding to the selected Coverage Level and Price Election Percentage.

Overview of Current Ratemaking Methodology

MPCI ratemaking follows a pure premium approach, with each crop analyzed separately. At present, only the experience arising from the standard MPCI APH yield coverage is included in the analysis. The experience arising from the MPCI revenue contracts is being considered for inclusion in future analyses. The first step in the analysis is to convert the losses for each county to a base level. The second is to stabilize the results for each county by capping the largest pure premiums. Third, the pure premium is smoothed over a local neighborhood. Next, the pure premium is adjusted to include for a risk factor and to spread back the losses eliminated by capping. The resulting pure premium is compared to the current rate in order to select the base rate change and the final rates for each Coverage Level and Rate Class.

The base level to which the loss experience is converted is the 65% Coverage Level. Paid claims are converted to the 65% level simply by restating the value of the loss for the difference in the deductibles. A further adjustment is needed for claims eliminated by the deductible. For policies insured at less than 65% coverage, the losses eliminated by the deductible are estimated from the severity distribution for policies with higher Coverage Levels.

Stabilization of the pure premiums is accomplished by the use of an 80/20 rule. Since 20 years of experience are currently used in the analysis, the 16 smallest pure premiums (80% of 20 years) are considered to be normal. The remaining four years of experience (20% of 20 years) are capped at the largest value among the 16 normal pure premiums. This rule has been selected judgmentally, based on a study of 1948 through 1979 experience for corn and wheat. This study found that the 80/20 rule resulted in a larger reduction of variance relative to the reduction in expected losses than the two alternatives of 75/25 and 70/30 which were considered. The indicated pure premium for each county is selected as the straight average of the capped pure premiums for all the years in the experience period. The pure premium is not adjusted for trend since trend is expected to have an equal impact on losses and the liability exposure.

The previous steps produce a preliminary value for the base pure premium for each county. However, even with the use of 20 years, the experience is not sufficiently credible to establish rates due to the large uncertainty in the expected value. This uncertainty can be observed from the countrywide loss ratios for all crops combined in Chart 1. The magnitude of variation in the loss ratios is much larger than that normally experienced in Property/Casualty coverages. The variability in the losses is significant even on a countrywide basis for all crops combined. The variability at a county level for a single crop is much greater.

In order to produce a more stable and more reliable pure premium, the smoothed pure premium for each county is determined as a weighted average of the indicated pure premiums over all nearby counties. This technique is known as the concentric circle method. Since counties do not possess a uniform, orderly arrangement, RMA has predetermined which counties are included in each concentric circle. The weights for each county are based on the liability of each county, and are computed separately by year. The rationale for the concentric circle method is that the causal or statistically correlated factors that determine the experience for each county operate on a broad geographic basis. A drought, for example, will typically affect an area much larger than an individual county. In comparison, standard actuarial ratemaking procedures tend to disregard the spatial relationships among rating territories, assuming instead that territorial experience is independent of the experience of other nearby territories.

After applying concentric circle smoothing, the smoothed pure premium is increased by a factor of 1.14, which is intended to satisfy the Congressional requirement that the rates be adequate to pay expected losses and build a "reasonable reserve." RMA has defined a reasonable reserve as an amount sufficient to achieve financial adequacy over a 10 year period at an 85% confidence level, evaluated on a countrywide basis. This loading may be understood to be an adjustment for risk to ensure the long term financial viability of the program. It should not be interpreted as an adjustment to the historical experience to more accurately estimate the expected losses, which may be potentially underestimated due to the absence of a catastrophic year in the 20 year experience period. The development of the factor will not be discussed here.

The final pure premium is determined by adding a statewide loading for the losses that were removed by the 80/20 rule to the smoothed pure premium loaded for the safety factor. A charge for losses that were excluded from the analysis, arising from prevented planting and other causes which are not directly related to yield loss, is also included. The final pure premium is divided by the current rate to produce a normalized loss ratio. The normalized loss ratio is compared to a judgmentally predetermined schedule of rate changes centered at the mandated target loss ratio of 107.5%. For example, if the normalized loss ratio falls between 90% and 115%, the current rate may not be revised. If it falls between 80% and 90%, a 5% rate reduction may be indicated, and so on.

Once the new base rate for a county is determined, rates for each farming practice must be developed. All practices within a county had been combined for ratemaking purposes. However, practices such as irrigation have a significant influence on yield variability, and consequently on the expected losses. The indicated rates for each practice are determined by multiplying the new county base rate by factors which reflect the relative riskiness of each practice relative to the county average. The factors are based on insurance data drawn from larger geographic areas as well as on the relative importance of the various practices within the county.

The rates by Coverage Level and Risk Class are determined by applying factors which are uniform for all states and crops, with minor exceptions. All rate increases are limited to no more than 20% in accordance with Federal law.

The rates established by these procedures are for coverage provided on an optional unit basis, meaning that each field is insured independently of any other field farmed by the producer. The producer also is permitted to insure the production of all these fields in total. Since this option diversifies the risk, a reduced rate is provided.

II. Challenges in Ratemaking

Catastrophic Nature of the Coverage

MPCI can be considered to be a catastrophic form of coverage. For an individual producer, MPCI compensates the producer for a portion of his loss when his yield is abnormally low. In this sense, MPCI is a high deductible product. However, when one producer has a poor year because of climatological factors, it is likely that many other producers will also have a poor year. This strong correlation of the experience between exposures limits the insurer's ability to reduce its risk through diversification. As a result, even the statewide MPCI experience can vary dramatically between years. For example, Chart 2 of Iowa experience shows two years since 1980 with loss ratios in excess of 350% and another year with a loss ratio in excess of 200%. If the experience were examined over the past 5 years only, Iowa would seem to be a very profitable market.

Since weather is the primary determinant of MPCI experience, exposures which are located in geographical proximity to one another will be highly correlated. This can be observed in the similar historical pattern of loss ratios for Iowa corn as compared to Iowa soybeans as shown in Chart 3. In years with severe weather, exposures separated by even greater distances can have similar experience. This can be observed in Chart 4, which compares the historical experience of Iowa, Minnesota, and Missouri.

Another perspective on the catastrophic potential of MPCI coverage can be obtained by examining a simulated distribution of producer outcomes rather than the loss ratios in aggregate. Chart 5 provides an illustration of how a weather induced shift in yields of -10% can result in much greater frequency of claims. Consider a producer whose APH yield is 120 bushels. The probability of this producer experiencing a loss in excess of a 25% deductible, i.e., an actual yield of less than 90 bushels, is under 16% in a normal year. This probability increases to 25% if weather results in a 10% reduction in yields. As a result, the expected number of claims would rise by 60%, six times as great as the change in the expected yield.

The limited ability to eliminate the risk through diversification affects the ratemaking analysis by increasing the uncertainty of the expected pure premium. One method used to address this uncertainty is to include many years of experience in the analysis. Currently, 20 years are used, and this will be increased in future reviews. A second is to limit the extreme losses from the analysis of individual counties, and a third is to smooth the pure premiums over a broader geographic region, as is currently done with the concentric circle method.

Geographical Influences on Farming and Risk

This section provides summary information on MPCI and farming in general as an introduction to current ratemaking issues.

Chart 6 shows the distribution of MPCI premium by major crop groupings over the period from 1980 through 1998. For example, cotton premium has increased in share in recent years. This may be due to high crop prices resulting in higher production, both of which result in higher insured values and greater premiums. In contrast, the premium market share for soybeans has decreased in recent years. Chart 7 shows the Herfindahl index, defined as the sum of the squares of the market shares for each crop grouping, which demonstrates that the shifts between crops being grown has not led to an increase in concentration. On an individual state basis, the corresponding market share and Herfindahl index exhibits (not included) would show a persistence of the preferred crops over time. This is primarily due to the limitations that climate places on certain crops. Another reason is governmental disincentives, recently eliminated, which discouraged producers from planting different crops.

The insured liabilities for 1998 in millions of dollars for the four major crops are shown in Charts 8 to 11. Average yields per acre for these crops are shown in Charts 12 to 15. The maps indicate where each crop is grown and how productive it is. For example, the availability of water for irrigation in California and Arizona has resulted in very high cotton yields. Texas cotton producers do not have access to inexpensive water, resulting in much lower yields. Despite this, the bulk of the nation's cotton crop is grown in Texas. The primary reason that cotton is grown in Texas is that this is the most productive use for the land.

The concentration of exposures in limited geographical areas, as for Texas cotton, is one reason for the non-independence of the loss experience. Not only weather conditions, but soil types, elevation, and rainfall can be expected to be similar for exposures situated in close proximity to one another. The question of how much influence geography plays on the expected losses is an important issue. This will be discussed with regard to the recent increase in participation in the MPCI program.

In the past decade, the participation of producers in the MPCI program has roughly doubled. Most of this increase occurred in 1995 due to a federal requirement, now rescinded, that producers purchase insurance coverage in order to qualify for other government programs. This increase in exposure is thought to result in a wider spread of risk, which should lead to more stable loss costs and less risk for insurers. There is also a question whether the spread of risk should result in lower loss costs and lower rates. This would be the case if the doubling of the insured exposures has reduced any adverse selection operating against the program. In some sense, these expectations have been proved true by the experience. Chart 1 shows that the countrywide loss ratios in the period from 1994 through 1998 to be much lower than in any year from 1980 through 1993. However, this argument disregards two key factors. The first is general weather conditions, which have been very good in recent years. Except for a drought in Texas during 1998, weather has not resulted in major disruptions to farming. The Deputy Chief Economist of the USDA has reported to the Senate "with the exception of regional loss events like the drought in Texas and parts of the South in 1998, most of the country has enjoyed relatively benign weather since 1995." The El Nino and La Nina events of the past two years have had little impact. For this reason, the experience from 1994 through 1998 should be expected to be excellent.

The second factor is the effect of geography on experience. If the producers who have purchased MPCI coverage only in the last five years are in close proximity to the previously insured producers, their experience should be expected to be similar, solely due to the common influences of weather, soil types, elevation, and other factors. Insuring additional exposures which are similar to and highly correlated with other insured exposures may not result in a significant reduction to the loss costs or to the risk. At an extreme, if all exposures within a county were 100% correlated, then the experience of a county would model the experience for each producer within the county. Consequently, the variability of yields for the county over time would be a reasonable proxy for the variability of yields for the individual producer. Since the variability of yields (or more specifically, the shortfall in yields) for each individual producer determines the loss payments under MPCI, the variability of county yields should be highly correlated to the county's historical loss costs. This is demonstrated for Iowa corn experience by county on Chart 16, where the measure of variability of yields is defined as the 100 times the coefficient of variation of yields over time. This evidence supports the idea that the yields, and hence the losses, of individual producers are strongly influenced by external factors. Consequently, producers results are highly correlated, which would suggest that the recent increase in exposures may have limited influence on MPCI loss costs or risk.

The previous discussion also raises the possibility of predicting expected loss costs based on yield information. Chart 17 shows the relationship between the aggregate loss costs and the yield for a given year. This relationship could be used to provide an estimate of the expected loss costs based an estimated distribution of the yields. Past yields by county could be trended to reflect productivity improvements in order to obtain an estimate of the distribution for the current year. However, the coefficient of variation of yields is a reasonable alternative to using the distribution of the trended yields. The coefficient of variation distills the distribution of yields into a single number for each county, and appears to be effective in predicting the loss costs. This makes it possible to consider either approach as a technique for estimating the expected loss costs whenever past loss experience is not available. This could also be used to test the indicated loss costs for reasonableness. Furthermore, counties with high coefficients of variation of yields are those in which farming is more uncertain. Not only are these counties expected to cluster together, but it is likely that the uncertain growing conditions will apply to all crops grown within the county. As a result, the variability of yields for one crop may be a means for predicting the variability of yields for a crop being newly introduced or with minimal loss history.

The Effect of Increased Producer Participation

The most recent five year period has produced very good results for the MPCI program, as seen in Chart 1. The same five year period has seen much greater producer participation in the program than in earlier years. One interpretation of these results is that the increased participation has brought lower risk producers into the program, reducing the adverse selection and improving the experience. The issue to be addressed here is whether this conclusion is justified. The approach to be taken will be to examine whether farming risk and insurance risk have diminished in recent years.

Chart 18 shows countrywide yields for Corn from 1980 through 1997. The years 1983, 1988, and 1993 all show abnormally low yields relative to the preceding and subsequent years. The long term trend in corn yields is +1.7% per year. The chart also includes the fitted yield curves. The second exhibit, Chart 19, shows the absolute value of the residuals from the first regression. If farming were becoming less risky, yields might be expected to follow the long term trend line on Chart 18 more closely than in the past. As a result, the absolute value of the residuals on Chart 19 should decline over time. A fitted trend line is included on this chart to show that the data does not have a strong downward trend. The absolute residuals in the final three years are below the trend line, but this is also true for the first three years as well as for the period from 1989 through 1991. Also, the t statistic for the slope of the fitted line is not statistically significant at the 95% confidence level. That is, the fitted line is essentially flat. This analysis does not support the conclusion that farming risk is less than in past years.

A similar analysis can be performed for the countrywide loss costs for corn from 1980 through 1997, as shown in Chart 20. The test of the residuals in Chart 21 leads to a conclusion similar to that for yields, that there is no significant decrease in insurance risk over this period. Again, the slope of the fitted line is not statistically significant from zero at the 95% confidence level.

This issue can be considered from another perspective. The discussion of spatial and intertemporal correlation, presented below, provides a means for evaluating how strongly the experience in one county is correlated to the experience in adjacent counties. This would imply that external factors which operate over large geographic areas are the source of much of the risk in farming. Consequently, an increase in producer participation in a given county may not have a significant effect on the overall riskiness of the MPCI program. The location of the farm rather than the skill of the individual producer may be the primary determinant of the risk. If this conclusion is true, it would imply that the improved MPCI experience in recent years may due to good weather conditions rather than increased producer participation.

III. Future Considerations for Ratemaking Analysis

Using Non-Insurance Information in Ratemaking

An important aspect of MPCI coverage is the linkage between the loss experience and non-insurance information. A potential use of this external information is to test the reliability of the indicated rates. Another is to provide a means to develop rates for counties in which past experience for the crop is not available.

The first link to be considered is that between MPCI pure premiums and yields. For individual producers, the yield determines the indemnity payment. Since the results for individual producers are strongly influenced by the weather, the loss experience and the yields among producers are highly correlated, even when aggregated to the county or state level. An analysis of this relationship at the statewide level for Iowa Corn is provided below.

Consider Chart 17. The illustrated relationship between the natural logarithm of the loss ratio and the natural logarithm of yield has been fit to the straight line:

$$\ln y = a + b * \ln x$$

This is equivalent to:

$$y = e^a * x^b$$

where x is the yield and y is the loss ratio. The best fitting curve has parameter values a=34.0 and b=-6.3. In economic terms, this formula describes the elasticity of the loss costs relative to the yields. The interpretation is that a 1% increase in the average yield for Iowa Corn results in a -6.3% change in the loss ratio. As a result, even a small change in yields has a highly leveraged impact on losses.

The volatility of the loss ratios implied by the elasticity coefficient highlights one of the difficulties in MPCI ratemaking. When losses vary widely between years, as for MPCI, the uncertainty in the estimate of the mean pure premium will be large. However, the additional information provided by the relationship between losses (pure premiums) and yields can improve the analysis in the following manner.

Suppose that the distribution of the yields over time is known or can be estimated. For example, the actual yields for the past 20 years could be considered. The distribution of yields for the coming year can be estimated by applying trend factors to yields from past years. Using the known relationship between yields and pure premiums, estimates of the potential pure premium outcomes for the coming year can be determined. The average of these outcomes is an estimate of the expected pure premium.

While it is unlikely that this technique would be used as part of the standard ratemaking process, several aspects may prove useful. One simple use is to identify data processing errors by identifying years in which the losses and yields are not consistent. A second use is to estimate pure premiums when insufficient loss information is available. For example, when a new crop is introduced into a county, it may be possible to estimate the variability of the yields based on the variability of the yields for other crops or other counties. A third use of this technique would be in smoothing past experience. Since a large portion of the losses are produced in a few abnormally poor years, the number of abnormal years in the experience period has a strong influence on the average pure premium. Because the number of abnormal years is always an integer, the average pure premiums can increase or decrease sharply when an abnormal year enters or leaves the experience period. These fluctuations can be reduced by taking the distribution of yields or pure premiums into consideration. For example, the expected number of abnormal years can be used in place of the actual number, with the corresponding severity based on a larger body of experience.

A second link to consider is that between MPCI loss costs and geographic and climatological information. This would involve the use of econometric modeling techniques, but without the need to consider time dependency as was the case with yields. The concept is that the yields and the variation in yields, and hence the pure premiums, are related to the suitability of the land for the crops being grown. For example, average county pure premiums could be modeled as a function of independent explanatory variables, such as average annual rainfall, growing days, soil type, and elevation. The advantage of this form of analysis is that is can be used to estimate the pure premiums even if yield experience is not available.

Potential Enhancements to Ratemaking Techniques

The previous section considered the use of non-insurance information in testing or modeling insurance experience. However, non-insurance information is generally considered to be supplementary to rather than as a replacement for insurance experience. The following discussion considers several approaches to improving the accuracy and increasing the stability of the rates based solely on insurance experience.

One recent proposal for improving the accuracy of the rates is to create fixed rating territories consisting of adjacent counties with similar agronomic characteristics. The rating territories could vary depending upon the crop being rated. The rationale for this proposal is that it would eliminate a perceived problem with the current concentric circle technique, due to the potential inclusion of experience from neighboring counties having dissimilar agronomic characteristics.

A more technically demanding approach is known as spatial smoothing. For example, the current concentric circle method is a simple form of spatial smoothing. A more sophisticated approach, known as locally weighted regression smoothing, has previously been introduced into Crop Hail insurance ratemaking by Dr. Michael Lewis. The advantages of this technique are its ability to produce smoother results than the concentric circle method and its ability to take spatial correlation into account. This technique may also eliminate the need to spread excess losses for each county across the state. A detailed explanation of the spatial smoothing process is included in the appendix.

A third approach for improving MPCI ratemaking analysis is to extend the concept of credibility to consider the spatial and intertemporal correlation between territories, that is, to create a spatial credibility model. This model is still in the early development stage, and its development is deferred to the appendix.

The spatial credibility technique is similar to spatial smoothing in that the experience in nearby counties is given more weight than that of more distant counties. However, it may allow for greater local fluctuations than spatial smoothing would produce. Each county's loss cost is used to the extent it is credible, with the remainder of the credibility being assigned to nearby counties, based on the relevance of their information.

The spatial credibility technique permits the loss costs to be estimated even for counties with little or no experience by taking advantage of the redundant information in nearby territories. In comparison, the classical credibility technique is generally applied to the indicated price changes, which may reduce or eliminate the spatial correlation between territories. As a result, classical credibility produces price changes for small territories that are similar to the statewide price change. Consequently, if these territories had been previously misrated, the classical credibility approach may not correct the misrating.

IV. Conclusion

Crop Insurance is a unique public/private market insurance product. In addition to public policy considerations, ratemaking needs to consider risk elements that are not common to other property/casualty coverages. This paper has intended to provide an introduction to the crop insurance product, and an overview of the ratemaking methodology.

Appendix

Spatial Smoothing

Locally weighted regression smoothing ("loess") in two spatial dimensions can be considered to be analogous to linear regression using a single independent variable. For linear regression in a single variable, the straight line which best fits the data is found, where the best fit is determined by minimizing the sum of squares of the residuals. This produces a curve of a known functional form, $y = a + bx$, which is linear in the parameters a and b. In comparison, loess finds fitted values using a local regression technique. The fitted values produced by this process are a surface of the form $z = a + f(\text{latitude, longitude})$, where z is a transformation of y. Unlike linear regression, the shape of the fitted surface is not describable using a known functional form, i.e., it is a non-parametric surface. The transformation used to modify the loss costs is $z = \log(y/(1-y))$. Since the MPCI loss costs are bounded between 0 and 1 and tend to be closer to the low end of the range, this transformation produces a less skewed dependent variable, which helps to improve the quality of the fit. In addition, the form of the transformation guarantees that the fitted loss costs will be non-negative.

The data used in the loess procedure is a single value for each location. In this example, the data is the average yield for Iowa corn by county over the period from 1981 through 1997. For ratemaking, the average loss costs over the experience period would be used instead. By using the average for each county, all intertemporal correlation is eliminated from the analysis. Even though the experience in adjacent counties may be correlated over time, this is not considered to be essential in estimating the expected values. Instead, spatial smoothing is only concerned with the spatial correlation of the data. The underlying concept is that the yields or loss costs change smoothly over space, and that knowledge of the yields in nearby counties provides redundant information which can be used to produce a better estimate of the expected value of the variable in each county.

The loess procedure determines the fitted value z by performing a local regression for each county. The transformed loss costs are fitted to the independent variables of latitude and longitude including an interaction term. The set of fitted values over all counties defines the fitted surface. Each local regression includes only those points that are in the neighborhood of the point being fit, where a neighborhood consists of the nearest k% of points in the sample space. A small value of k results in greater local accuracy, whereas a large value of k results in a smoother surface. For this analysis, 100% of the data points have been included in each neighborhood. However, greater weight is assigned to nearby counties than the more distant counties by the use of the tri-cube formula $(1 - d^3)^3$. The distance d between any two counties x_0 and x_i is defined as $d = |x_0 - x_i| / \max(|x_0 - x_k|)$, where the denominator is computed over all values k within the neighborhood. Here, each point x_i represents the joint latitude and longitude coordinates at the center of a specified county. The coordinates of the county midpoints must be transformed to the Euclidean coordinate system using a distance preserving projection prior to their use in the loess procedure.

The best fit local regression for each county is determined by using a maximum likelihood technique under the assumption that residuals have a normal distribution with constant variance. The residuals are weighted by the actual cumulative liability (i.e., exposure) for each county in order to improve the accuracy of the smoothed results. As in actuarial credibility, the loss costs of counties with larger weight, as measured by cumulative liability, reflect their own experience to a greater degree than counties with smaller weight. The amount of smoothing produced by this process can be evaluated from a comparison of Charts 22 and 23, which show cotton loss costs for Crop Hail insurance in the southeastern states. The perspective on these maps is looking west from the Atlantic Ocean, with Florida shown to the left and North Carolina to the right.

Despite the complexity of the description of spatial smoothing, it can be implemented very efficiently in the S-Plus programming language. In the following programming statement, lo represents the loess function, while latitude and longitude are the projected coordinates of the county midpoints. The span of 1 defines each neighborhood as consisting of the nearest 100% of points in the sample space. Also, the dependent variable uses the untransformed loss costs, with the transformation being performed by the logit function.

```
gam(loss.cost ~ lo(longitude, latitude, span = 1),
family = quasi (link = logit, variance = "constant"),
data = your.data, weights = liability, na.action = na.omit)
```

From a practical standpoint, the primary weakness of spatial smoothing is its complexity, which increases the difficulty of explaining the results to insurance regulators. Since the analysis cannot be reproduced in a spreadsheet, the reliability of the results cannot be easily confirmed.

The loess procedure is actually a simple form of spatial smoothing. More sophisticated forms of spatial smoothing have been developed, but these require a knowledge of the field of spatial statistics.

Development of the Spatial Credibility Model

Classical credibility theory can be used to develop a best estimate for a territory by weighting the territorial average loss cost with the overall statewide loss cost. Generally, credibility is applied to the indicated changes in loss costs instead. However, the classical credibility formulas are developed under two assumptions which are not valid for MPCI. The first is that the true territorial expected loss costs are independent of one another, i.e., that there is no spatial correlation. The second is that the intertemporal random fluctuations in one territory are independent of the random fluctuations in other territories.

For MPCI, the spatial correlation between counties as a function of distance can be described using a variogram. The first step in the preparation of the variogram is to calculate the statistic Y_{ij} as ½ of the squared difference in pure premiums for each pair of counties (i,j). The distance between each pair of counties is also required. Given this information, distances are grouped into ranges and the average of all Y_{ij} within each range is determined. This produces a variogram similar to that shown in Chart 24 for yields. The variogram is an estimator for $E[(X_i-X_j)^2/2 \mid \text{counties } i,j \text{ in distance range } k]$, where X represents pure premium. The variogram is also a proxy for the spatial covariance. Notice that $E[(X_i-X_j)^2/2] = \frac{1}{2} VX_i + \frac{1}{2} VX_j - Cov(X_i,X_j) + \frac{1}{2} (\mu_i - \mu_j)^2 = \sigma^2 - Cov(X_i,X_j)$, where all X_k are assumed to be from the same distribution. That is, a small value for the variogram implies a high value for the spatial covariance. The chart shows that the variogram is low for nearby counties and gradually increases as the distance increases, within a certain range.

Similarly, the intertemporal correlation of a county's experience to that of its nearest neighbors is also greater than its correlation to more distant counties, as shown in Chart 25 for Iowa corn in Adams county. The average intertemporal correlation across all counties as a function of distance is shown in Chart 26. Each type of correlation needs to be taken into account in a spatial credibility formula.

The spatial credibility formula determines the best estimate of the projected loss cost for county t in future year "0" using a linear combination of the known observations. Using Formula 4.1 from Chapter 7 of the "Foundations of Casualty Actuarial Science" text, the objective is to determine the coefficients which minimize:

$$E[X_{t0} - (a_0 + \Sigma_{i,u}\, a_{iu}\, X_{iu})]^2$$

where t is the county being evaluated. Here, X_{iu} represents the loss costs in county i (from 1 to N) in year u (from 1 to n). It will be assumed that the loss costs can be decomposed into three components:

$$X_{iu} = m + R_i + Q_{iu}$$

The first component, m, represents the mean loss cost over all counties. R_i represents the variation of the individual county loss costs from the overall mean. These are selected such that the average deviation over all counties is zero, $E(R_i) = 0$. Individual year random fluctuations are represented by Q_{iu}, with the average of the Q's over all years and counties being zero, $E(Q_{iu}) = 0$. The R's are assumed to be independent of the Q's. These assumptions imply that the overall expected value $E(X_{iu})$ is m and that the expected value for a particular county i is $E(X_{iu}|R_i) = m + R_i$.

In order to determine the coefficients, the partial derivations of the expected value with respect to the coefficients are set to zero, which produces the following equations:

(1) $$E(X_{t0}) = a_0 + \Sigma_{i,u}\, a_{iu}\, E(X_{iu})$$

(2) $$Cov(X_{t0}, X_{jv}) = \Sigma_{i,u}\, a_{iu}\, Cov(X_{iu}, X_{jv})$$

for each j and v. The county t is assumed to be fixed.

Equation 1 can be evaluated as:

$$m = a_0 + \Sigma_{i,u}\, a_{iu}\, m$$

which yields:

$$a_0 = m\,(1 - \Sigma_{i,u}\, a_{iu})$$

Notice that the sum of the coefficients equals 1.00 as in classical credibility if a_0 is replaced by ma_0.

Equation 2 requires that the covariances be evaluated. This can be done by considering the following identity:

(3) $$Cov(X_{iu}, X_{jv}) = E[Cov(X_{iu}, X_{jv}|R_i,R_j)] + Cov[E(X_{iu}|R_i,R_j), E(X_{jv}|R_i,R_j)]$$

But,

$$E(X_{iu}|R_i,R_j) = m + R_i$$

and:

$$E(X_{jv}|R_i,R_j) = m + R_j$$

This permits the second term on the right side of equation 3 to be simplified:

$$Cov[E(X_{iu}|R_i,R_j),\ E(X_{jv}|R_i,R_j)] = Cov(m + R_i,\ m + R_j) = Cov(R_i,\ R_j)$$

It will be assumed that:

$$Cov(R_i,\ R_j) = f(d(i,j))$$

That is, the spatial covariance is a function of the distance d between counties i and j. If another pair of counties are separated by the same distance, the value of f will be assumed to be identical. For simplicity in notation, f(d(i,j)) will be replaced by f(i,j).

The first term on the right side of equation 3 can also be evaluated:

$$E[Cov(X_{iu},\ X_{jv}|R_i,R_j)] = E[Cov(m + R_i + Q_{iu},\ m + R_j + Q_{jv}|R_i,R_j)] = E[Cov(Q_{iu},\ Q_{jv})]$$

When $u \neq v$, the independence of the experience between different years implies that the covariance of the random fluctuation term is 0. When $u = v$, it will be assumed that:

$$E[Cov(Q_{iu},\ Q_{ju})] = g(d(i,j))$$

That is, the expected intertemporal covariance is a function of the distance between the two counties. For simplicity, g(d(i,j)) will be replaced by g(i,j). By defining δ_{uv} as 0 when $u \neq v$ and 1 when $u = v$, the first term on the right side of equation 3 can be expressed as:

$$E[Cov(X_{iu},\ X_{jv}|R_i,R_j)] = \delta_{uv}g(i,j)$$

As a result, equation 3 can be expressed as:

$$Cov(X_{iu},\ X_{jv}) = \delta_{uv}g(i,j) + f(i,j)$$

Inserting this into equation 2 gives:

$$Cov(X_{t0}, X_{jv}) = \Sigma_{i,u}\, a_{iu}\, Cov(X_{iu}, X_{jv}) = \Sigma_{i,u}\, a_{iu}\, [\delta_{uv} g(i,j) + f(i,j)]$$

or:

$$Cov(X_{t0}, X_{jv}) = \Sigma_i\, a_{iv}\, g(i,j) \; + \; \Sigma_{i,u}\, a_{iu}\, f(i,j)$$

Notice that the left hand side of this equation can also be evaluated as:

$$Cov(X_{t0}, X_{jv}) = \delta_{0v} g(t,j) + f(t,j) = f(t,j)$$

This produces the following simplification of equation 2:

(4) $$f(t,j) = \Sigma_i\, a_{iv}\, g(i,j) \; + \; \Sigma_{i,u}\, a_{iu}\, f(i,j)$$

for all values of year v. Defining $b_i = \Sigma_v\, a_{iv}$ and summing both sides of equation 4 over v produces:

$$n\, f(t,j) = \Sigma_i\, g(i,j)\, \Sigma_v\, a_{iv} + n\, \Sigma_i\, f(i,j)\, \Sigma_u\, a_{iu}$$

or:

$$n\, f(t,j) = \Sigma_i\, g(i,j)\, b_i + n\, \Sigma_i\, f(i,j)\, b_i = \Sigma_i\, b_i\, [g(i,j) + n\, f(i,j)]$$

which represents N equations (j = 1 to N) in N unknowns b_i. It should be observed that the values of b_i depend on the covariance functions f and g, but not on the loss costs. We will assume that these equations can be solved for the values of b_i.

Substituting the known values of b_i into equation 4 yields:

(5) $$f(t,j) = \Sigma_i\, a_{iv}\, g(i,j) \; + \; \Sigma_i\, b_i\, f(i,j)$$

or:

$$\Sigma_i a_{iv} g(i,j) = f(t,j) - \Sigma_i b_i f(i,j)$$

Since all of the terms on the right hand side are known and depend only on j, the right hand side can be written more simply as c_j:

$$\Sigma_i a_{iv} g(i,j) = c_j$$

This is a system of Nn equations in v and j, with Nn unknowns a_{iv}. This can be written as the product of the transpose of NxN matrix $[g(i,j)]$ with Nxn matrix $[a_{iu}]$. Since the Nxn product matrix $[c_{ju}]$ has $c_{ju} = c_j$ for all u, its column rank is 1. Assuming that matrix $[g(i,j)]$ is non-singular, the column rank of matrix $[a_{iu}]$ must also be 1. This means that each column is a multiple of the first column, that is, $a_{iv} = k_v a_{i1}$. Substituting this into the previous formula results in the conclusion that $k_v = 1$ for all v. This conclusion can also be reached intuitively by noting that the right hand side of the equation is independent of v. This permits the symbol a_i to be used in place of a_{iv}, so that:

$$\Sigma_i a_i g(i,j) = c_j$$

for each value of j. An immediate solution to this system of equations can be obtained by observing that $b_i = \Sigma_v a_{iv} = \Sigma_v a_i = n a_i$. Since the values of b_i are known, $a_{iv} = a_i = b_i / n$. The value of a_0 can also be determined from $a_0 = m (1 - \Sigma_{i,u} a_{iu}) = m (1 - \Sigma_i b_i)$. This also shows that the coefficients depend on the county but are independent of the year.

The reader can confirm that this result is consistent with the classical credibility formula under the assumption of the covariance structure $f(i,j) = \delta_{ij} \sigma^2$ and $g(i,j) = \delta_{ij} s^2$.

It should also be noted that the functions f and g, or more properly, f_t and g_t, represent the spatial covariance and the intertemporal covariance in a neighborhood A_t around county t. The neighborhood needs to be large enough to reliably estimate f_t and g_t, but small enough to represent the covariance structure near county t. In practice, it may be appropriate to assume that f_w and g_w are essentially identical to f_t and g_t for all counties w in a neighborhood B_t of t. Further refinements in the model could be achieved by permitting the functions f and g to depend on location and direction, rather than on distance only.

A simplification of the spatial credibility result can be obtained by approximating f(i,j) and g(i,j) with discrete valued functions. Recall that these functions depend solely on the distance between counties i and j, not on their specific locations. For example, let f and g take distinct values for a series of distances such as 0 to 50 miles, 51 to 100 miles, 100 to 150 miles, and so on. This would enable the references to specific counties for the functions f(t,j), f(i,j) and g(i,j) in equation 5 to be replaced by a small number of values. If it can be assumed that the functions f and g are independent of county t, and if all counties formed a uniform pattern such as a rectangular grid, then the resulting b_i coefficients would be independent of t. In this situation, the same coefficients would be applied for all counties, eliminating the need to reevaluate the coefficients for each individual county t. It may be advantageous to superimpose a rectangular grid to replace the actual county structure in order to achieve this simplification. The result would be a smoothing process analogous to the current concentric circle technique, with more weight given to nearby counties than to those that are further away.

References

Arlinghaus, S.L.,
Practical Handbook of Spatial Statistics,
CRC Press, 1996

Driscoll, James L.,
"Changes in Rate Making for Federal Crop Insurance" Internal Memorandum
Federal Crop Insurance Corporation, USDA

Eves, Howard,
Elementary Matrix Theory,
Dover Publications, Inc., 1966

Glauber, Joseph W.,
"Statement of Joseph W. Glauber
Deputy Chief Economist, U.S. Department of Agriculture
Before the Committee on Agriculture, Nutrition and Forestry
United States Senate
March 10, 1999"

Kaluzny, S.P.; Vega, S.C.; Cardoso, T.P.; Shelly, A.A.,
S+SPATIALSTATS User's Manual, Version 1.0, February 1996
MathSoft, Inc.

Kitanidis, P.K.,
Introduction to Geostatistics: Applications to Hydrogeology
Cambridge University Press, 1997

Knight, Thomas O. and Coble, Keith H.,
"Survey of U.S. Multiple Peril Crop Insurance Literature Since 1980"
Review of Agricultural Economics, Volume 19, Number 1, Pages 128-156

Lewis, Michael,
"Louisiana Loss Costs"
Internal Memorandum,
National Crop Insurance Services

Milliman & Robertson, Inc.,
"Analysis of Catastrophe Provisions – Task 3"
Contract 12-27-129-372
Prepared for Federal Crop Insurance Corporation, September 9, 1983

PricewaterhouseCoopers LLP,
"Federal Crop Insurance Program Profitability and Effectiveness Analysis, 1999 Update"
Prepared for Morrison & Hecker L.L.P.on behalf of National Crop Insurance Services, Inc., May 1999

S-Plus 4 Guide to Statistics,July 1997
MathSoft, Inc.

"Crop Insurance for Cotton: Premium Rates and Related Matters"
Prepared for U.S. Senate in response to Report accompanying S. 2159 Agriculture, Rural Development, Food and Drug Administration, and Related Agencies Appropriation Bill, 1999
by Federal Crop Insurance Corporation of United States Department of Agriculture, August 11, 1999

Venter, Gary G.,
Chapter 7, "Credibility"
Foundations of Casualty Actuarial Science, Second Edition, 1990

Zacharias, Thomas P.,
Chapter 7, "Impacts on Agricultural Production: Huge Financial Losses Lead to New Policies"
The Great Flood of 1993, Causes, Impacts, and Responses, Stanley A. Changnon, Ed.
Westview Press, 1996

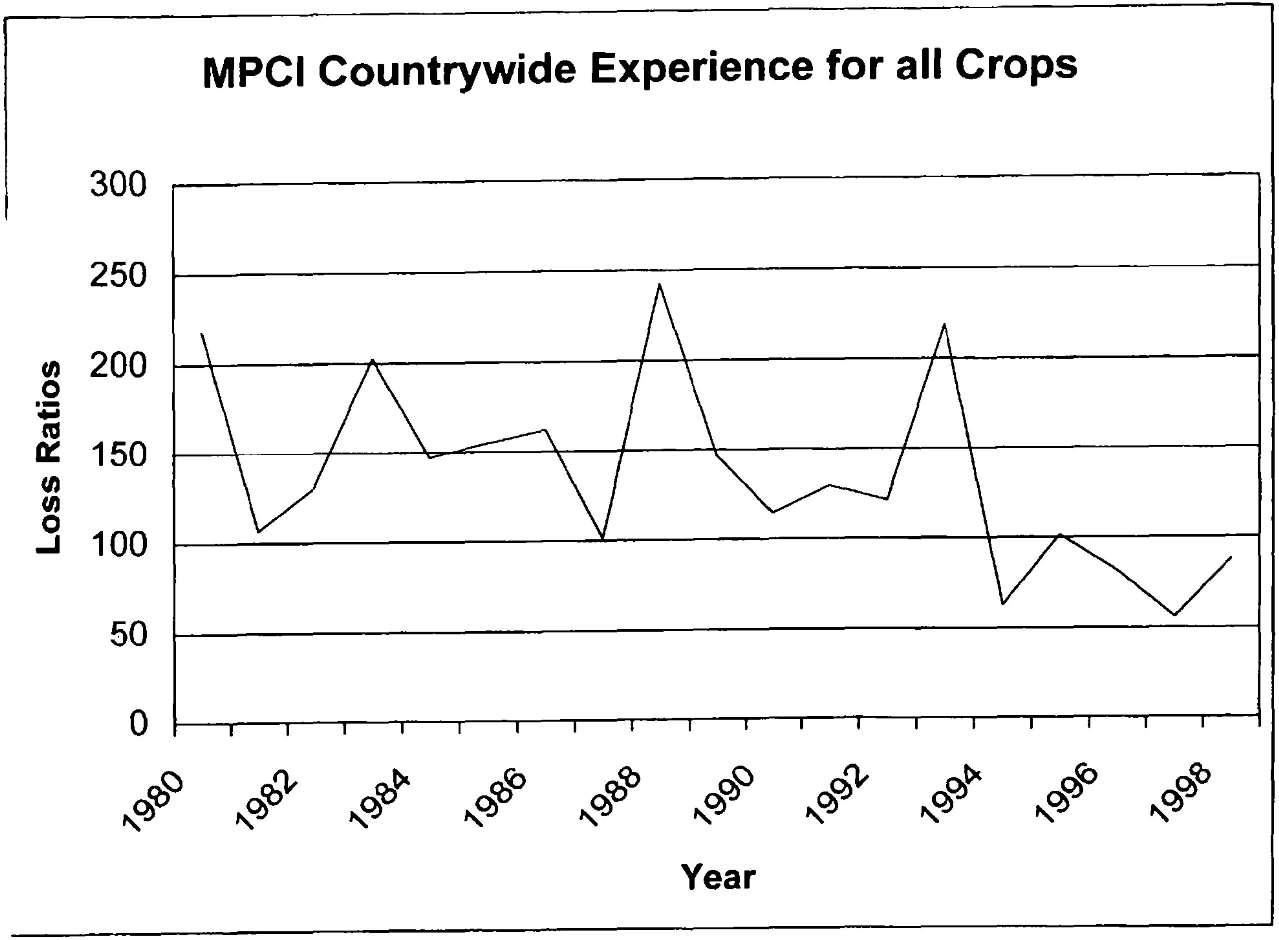

MPCI Countrywide Experience for all Crops
Loss Ratios
300
250
200
150
100
50
0
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
Year

Exhibit 1

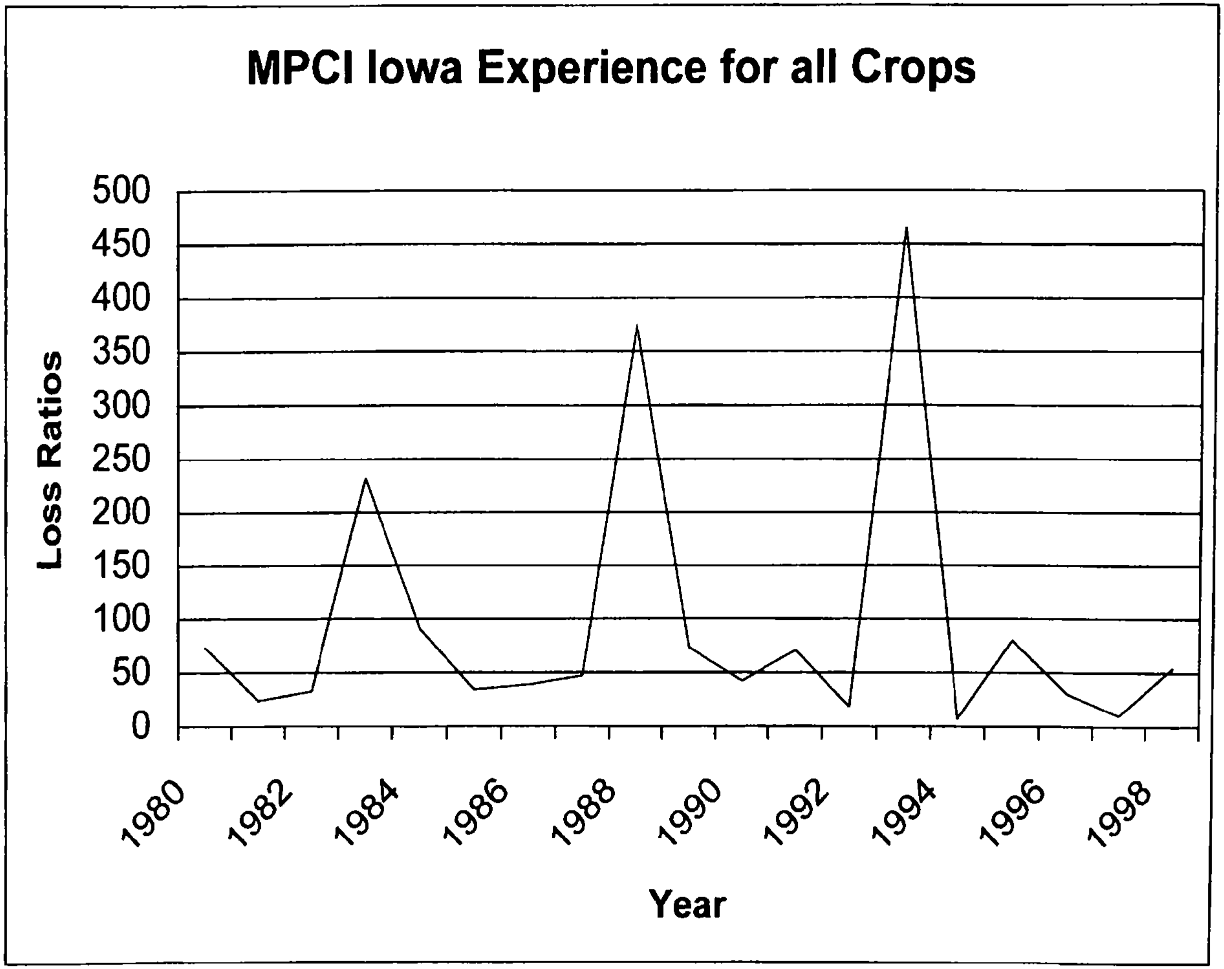
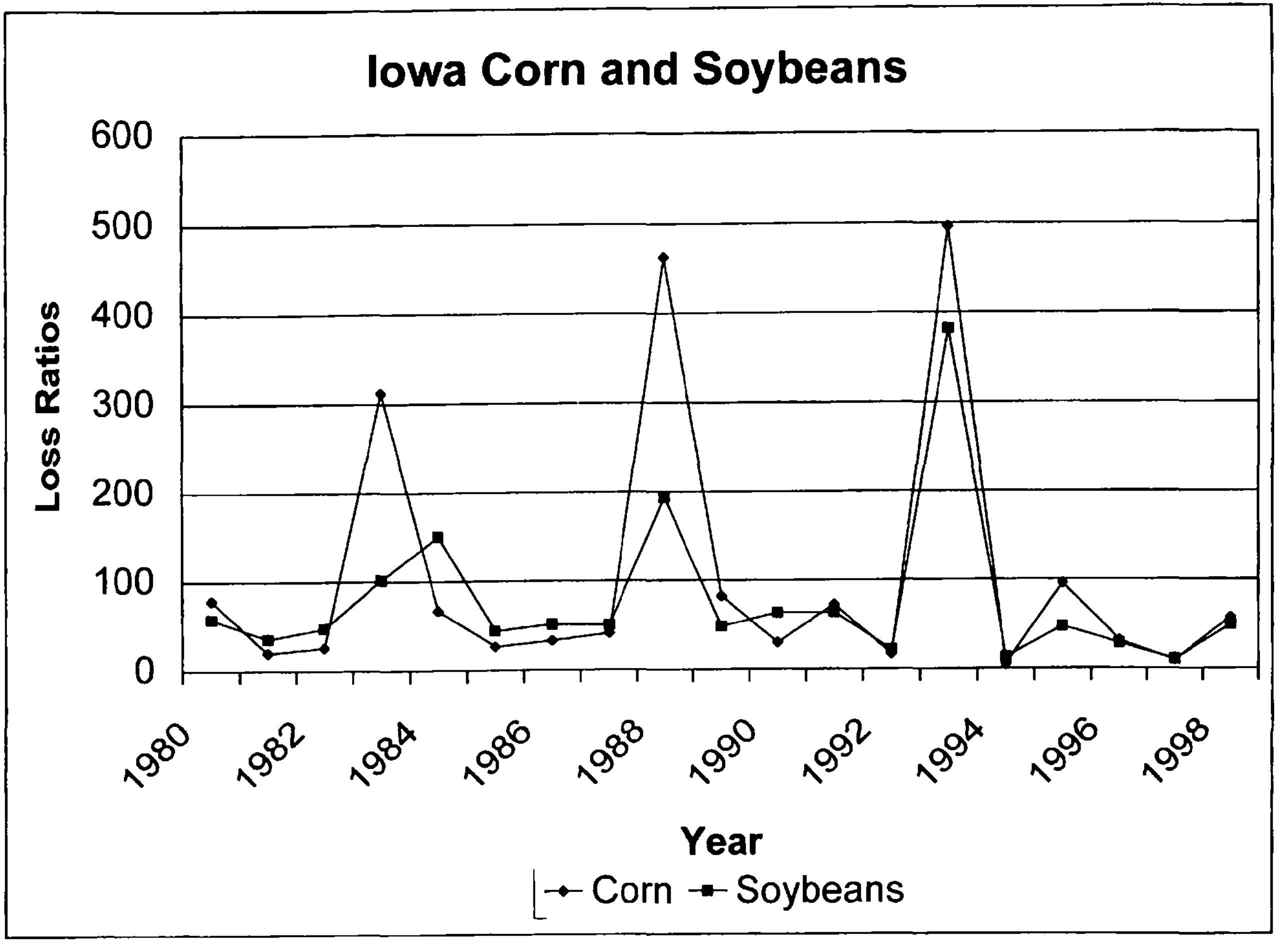

MPCI Iowa Experience for all Crops
Loss Ratios
500
450
400
350
300
250
200
150
100
50
0
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
Year

Exhibit 2



Iowa Corn and Soybeans
600
500
400
300
200
100
0
Loss Ratios
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
Year
Corn
Soybeans

Exhibit 3

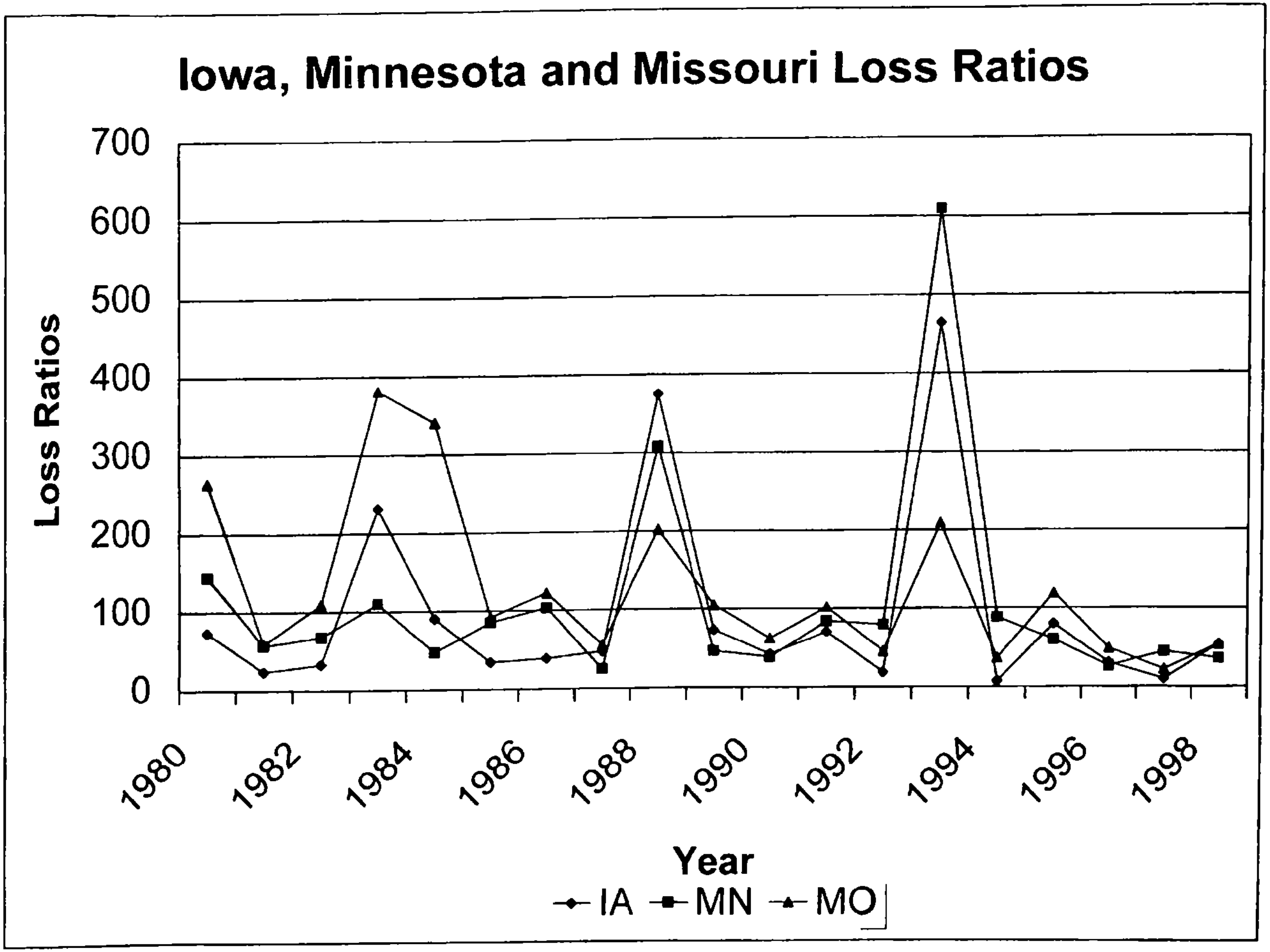

Iowa, Minnesota and Missouri Loss Ratios
Loss Ratios
0
100
200
300
400
500
600
700
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
Year
IA
MN
MO

Exhibit 4

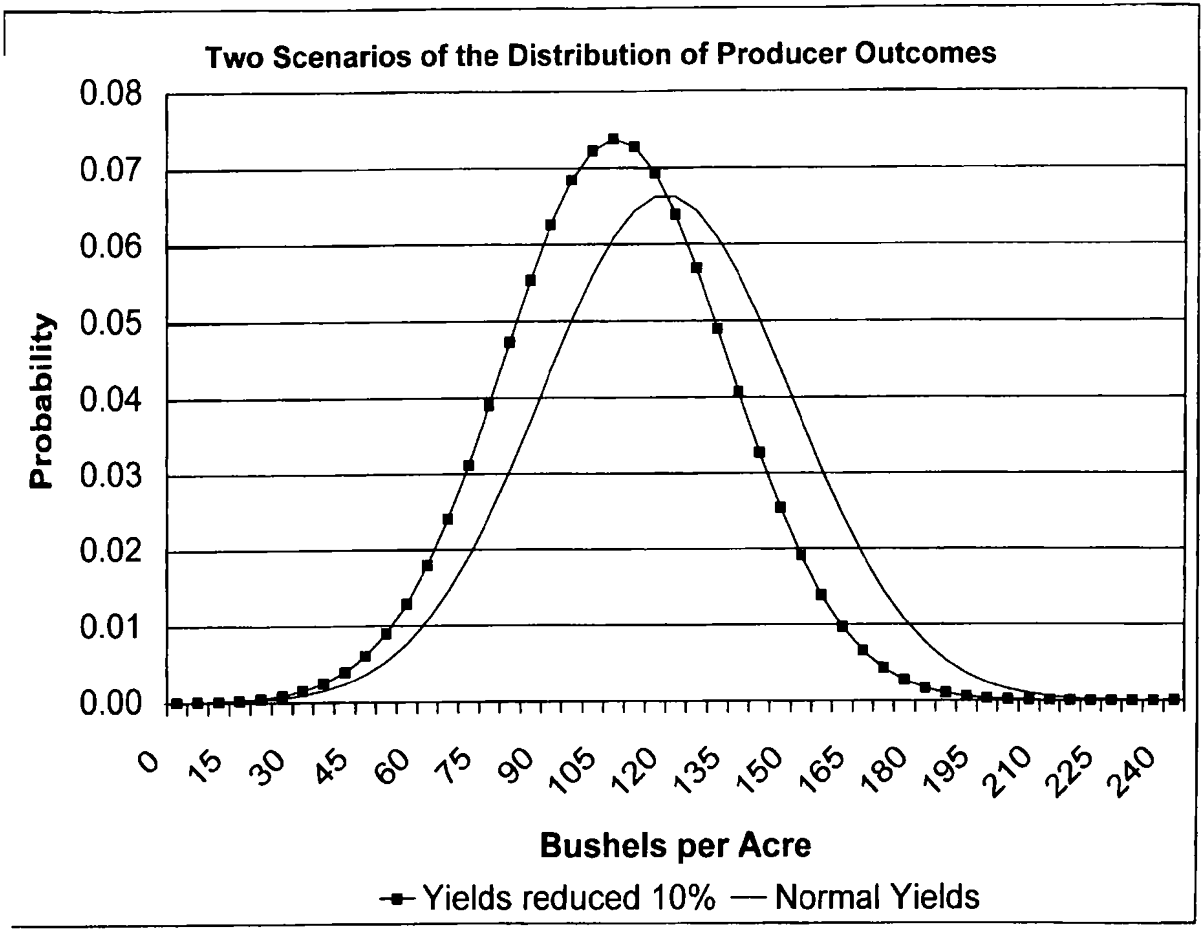

Two Scenarios of the Distribution of Producer Outcomes
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
Probability
0
15
30
45
60
75
90
105
120
135
150
165
180
195
210
225
240
Bushels per Acre
Yields reduced 10%
Normal Yields

Exhibit 5

MPCI Cumulative Market Shares by Crop

Crops from bottom to top are Corn, Cotton, Grain,Soybeans, Tobacco, Other

STATE=Countrywide

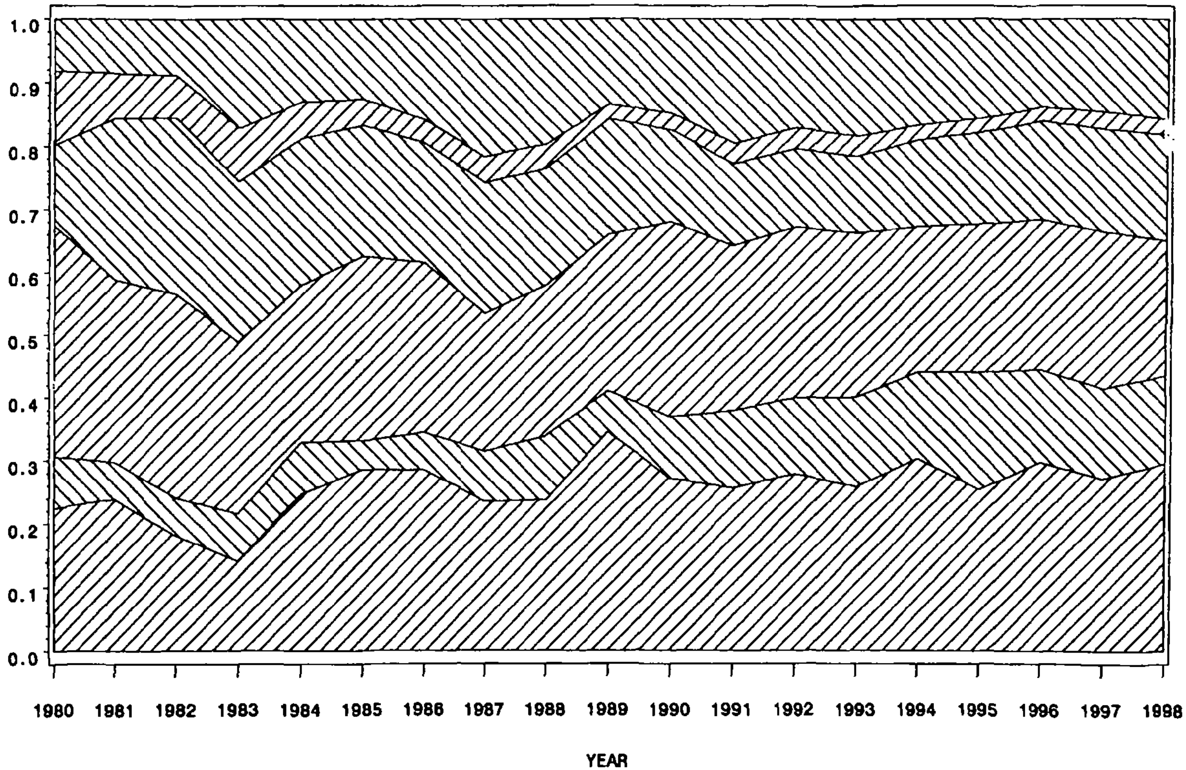


Exhibit 6

Herfindahl Index for Crop Hail Data

STATE=Countrywide

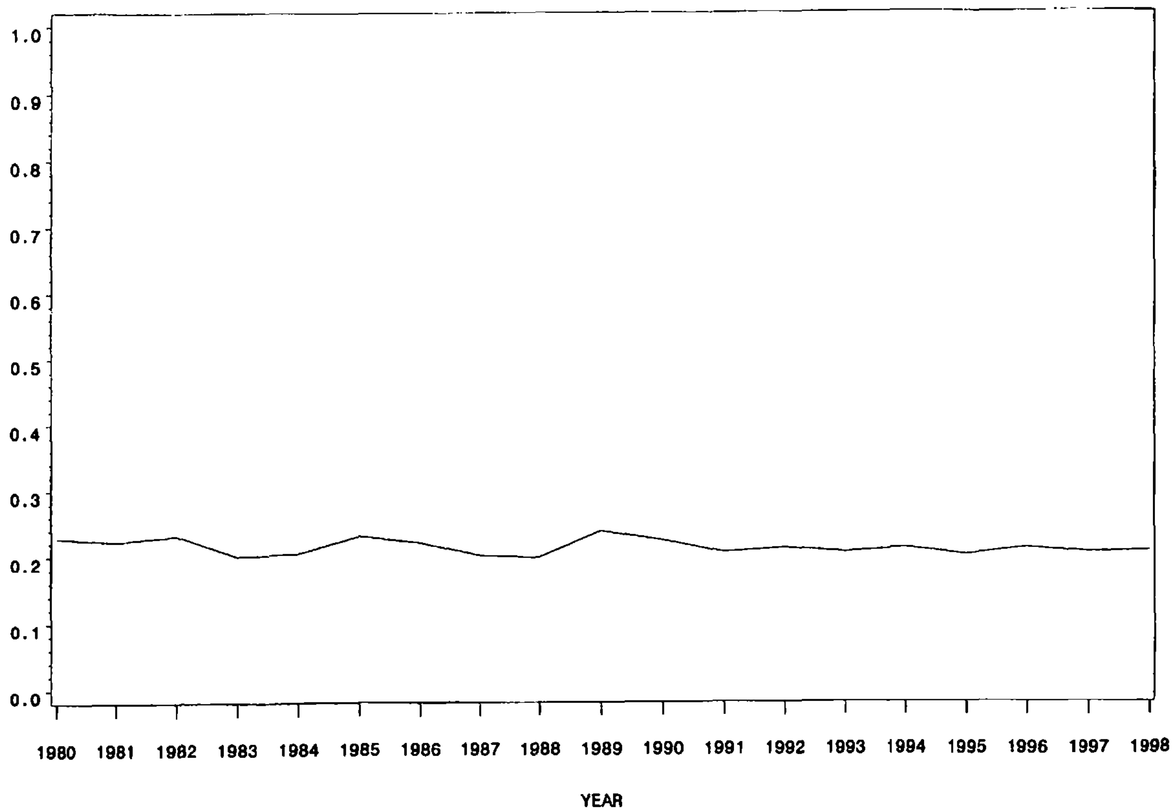


Exhibit 7

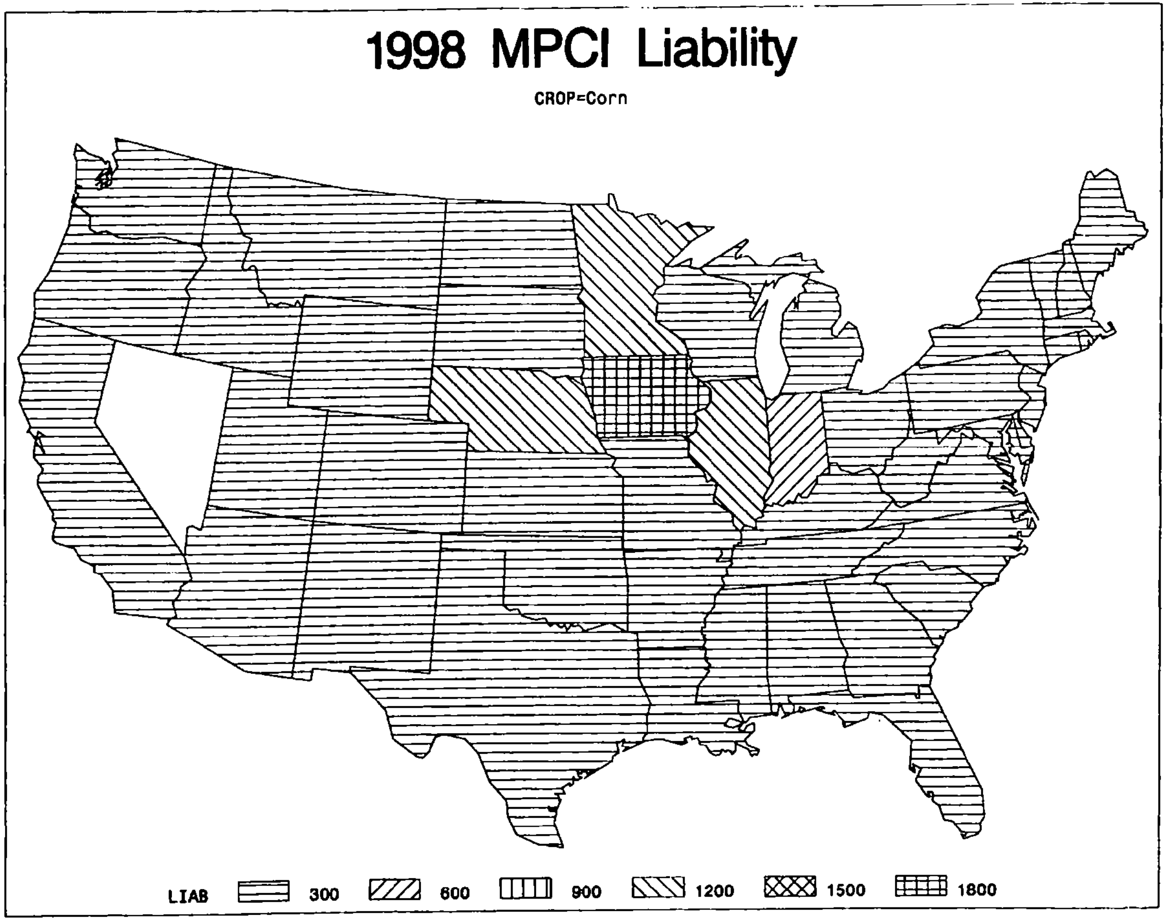

1998 MPCI Liability
CROP=Corn
LIAB
300
600
900
1200
1500
1800

Exhibit 8

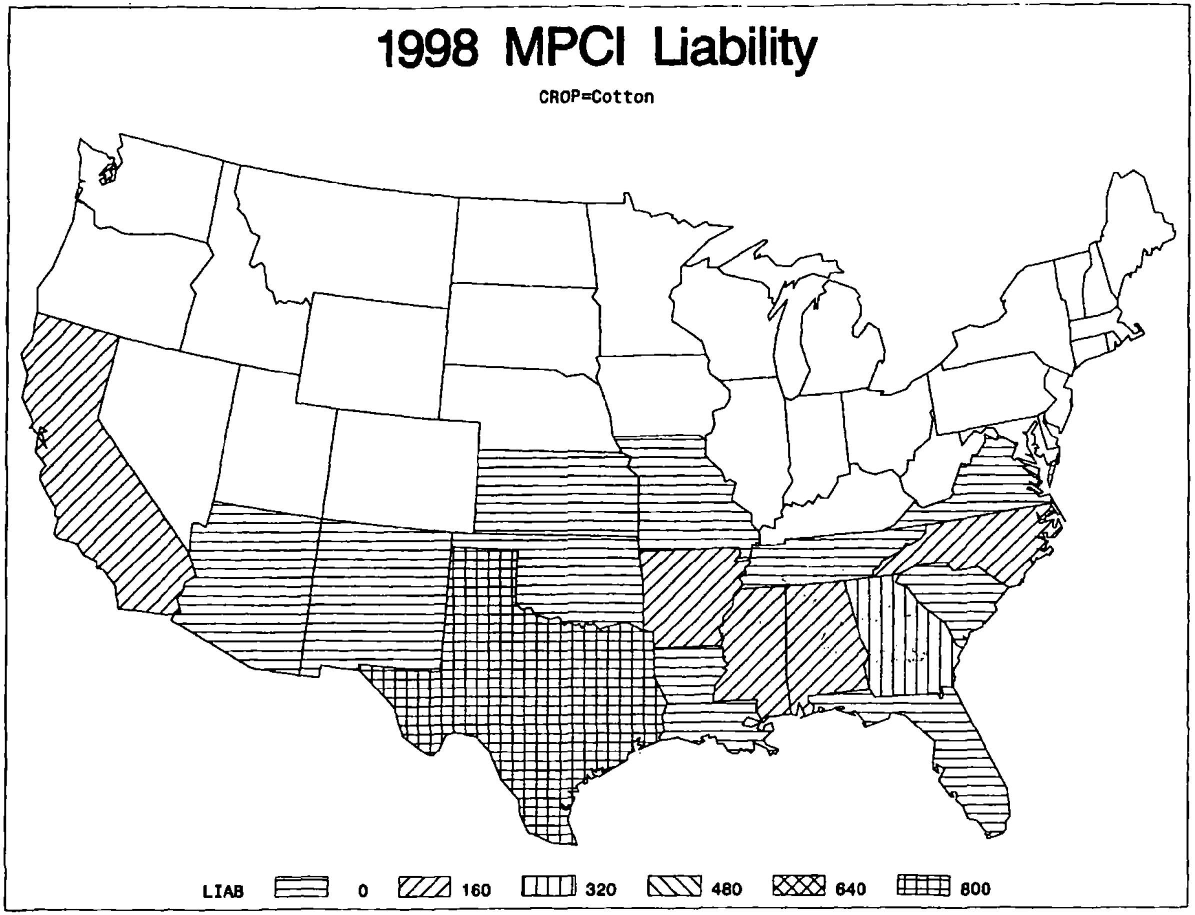


Exhibit 9

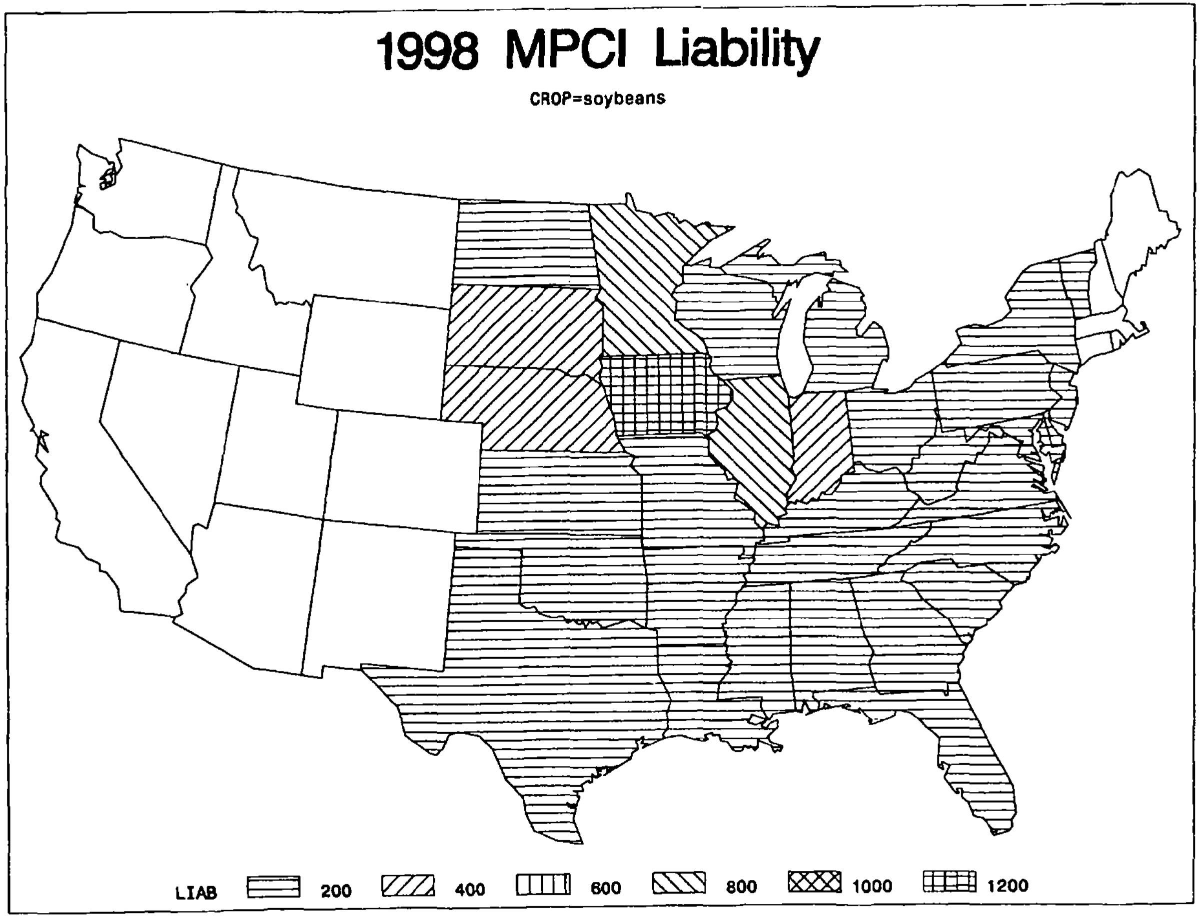

1998 MPCI Liability
CROP=soybeans
LIAB
200
400
600
800
1000
1200

Exhibit 10

1998 MPCI Liability

CROP=Wheat

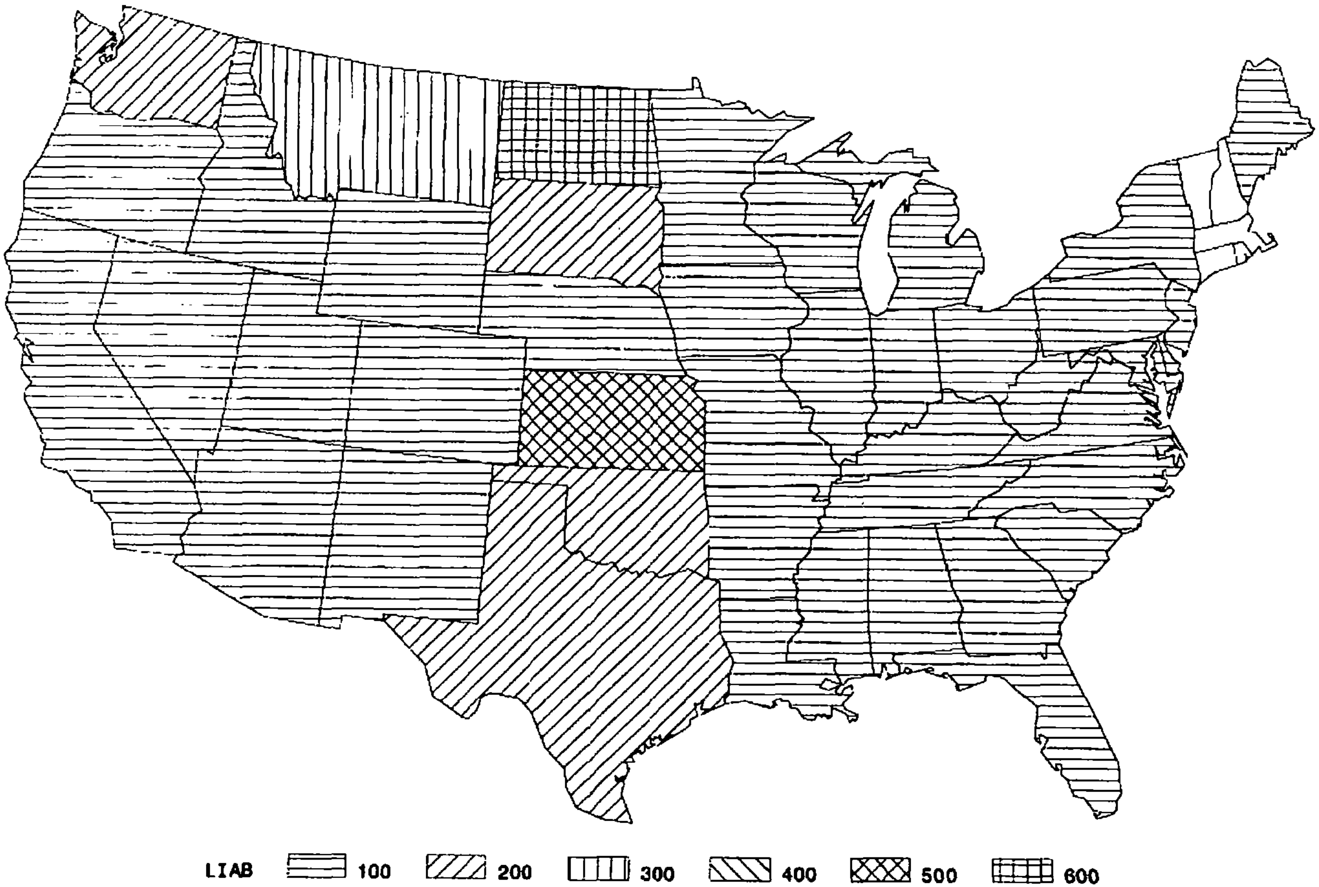


Exhibit 11

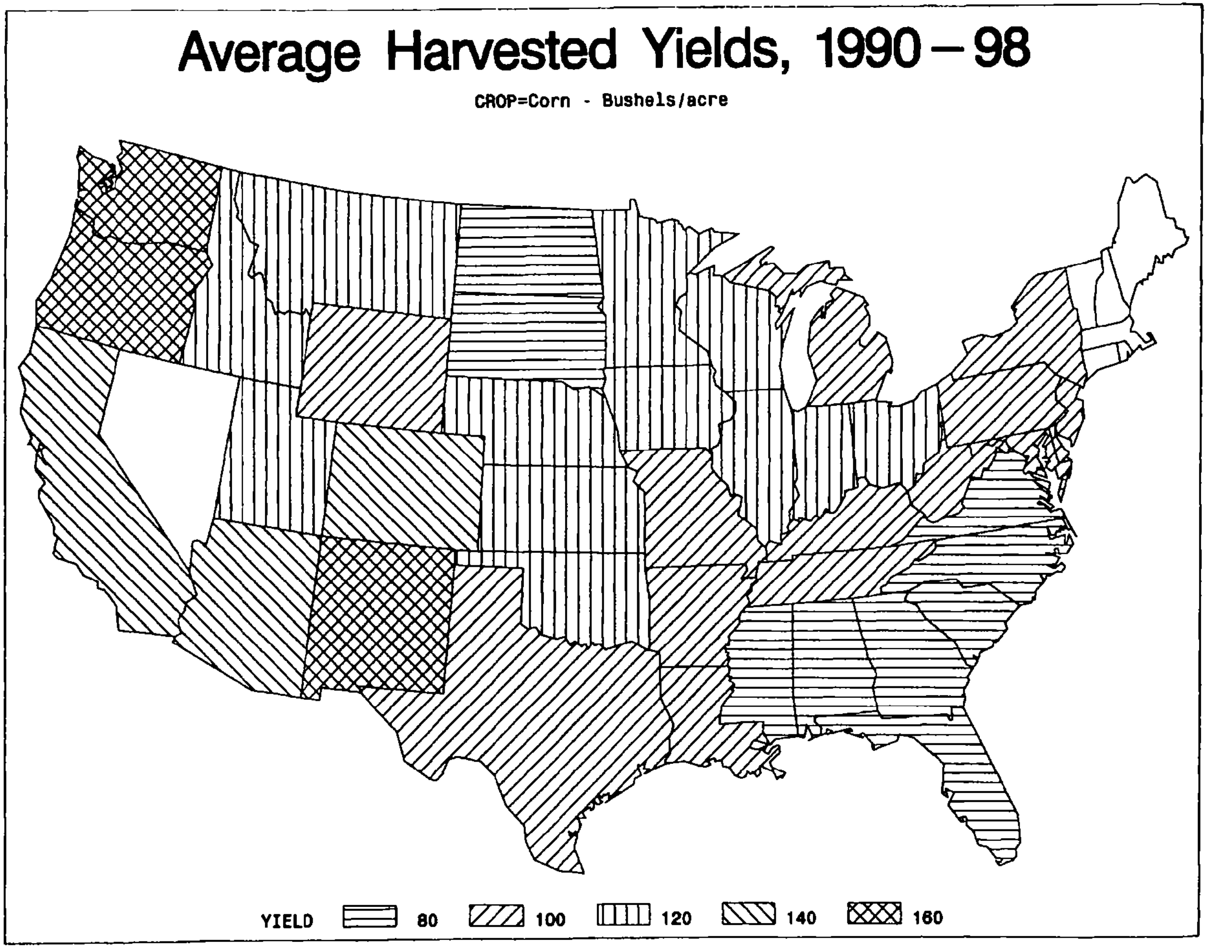

Average Harvested Yields, 1990 – 98
CROP=Corn - Bushels/acre
YIELD
80
100
120
140
160

Exhibit 12

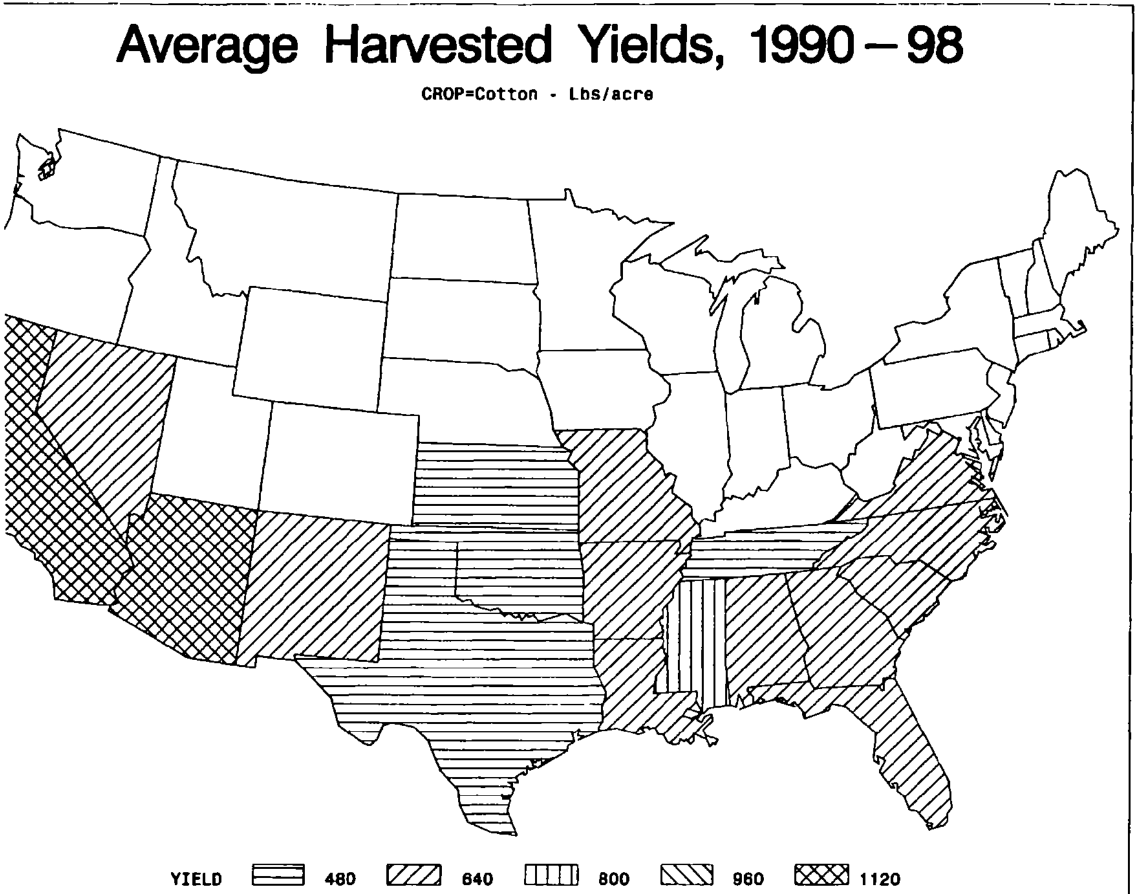

Average Harvested Yields, 1990 – 98
CROP=Cotton - Lbs/acre
YIELD
480
640
800
960
1120

Exhibit 13

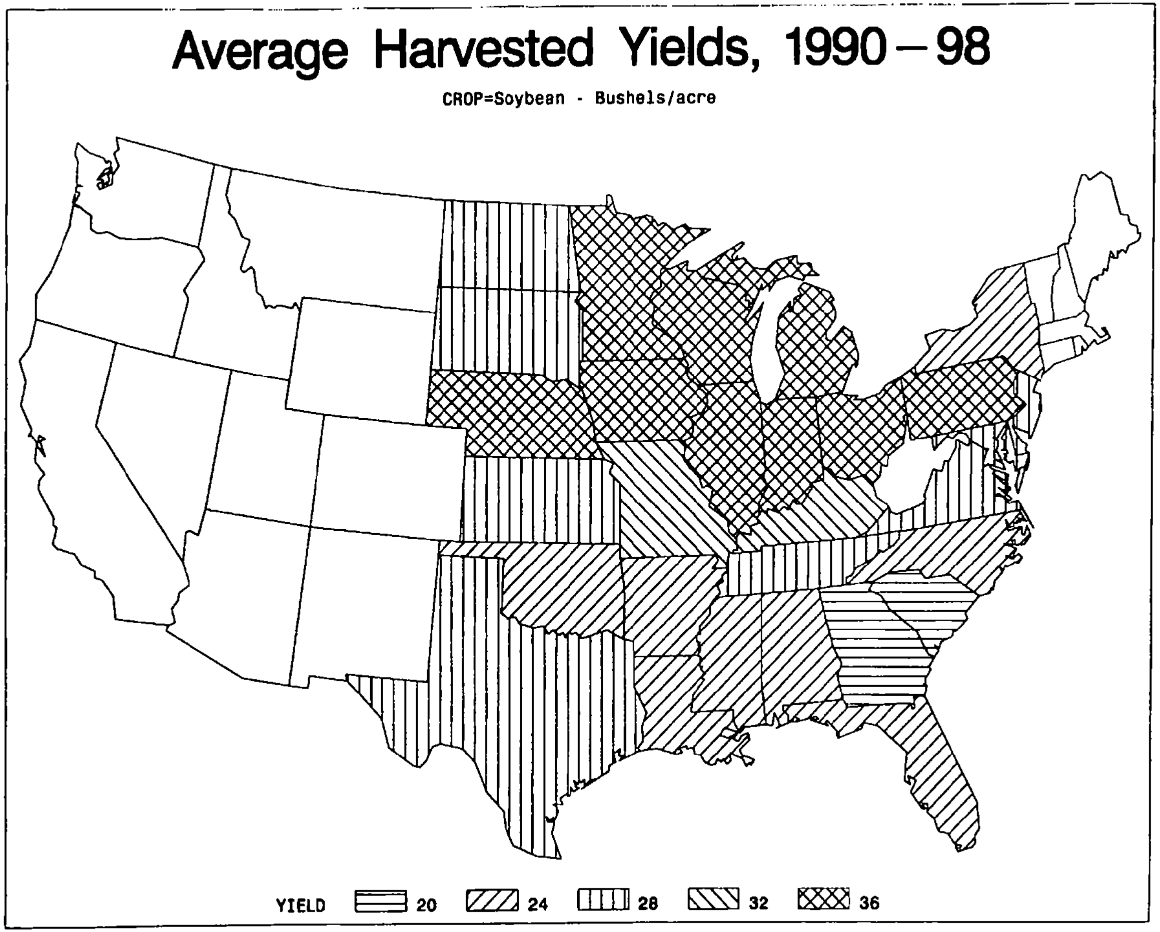

Average Harvested Yields, 1990 – 98
CROP=Soybean - Bushels/acre
YIELD
20
24
28
32
36

Exhibit 14

Average Harvested Yields, 1990 – 98

CROP=Wheat - Bushels/acre

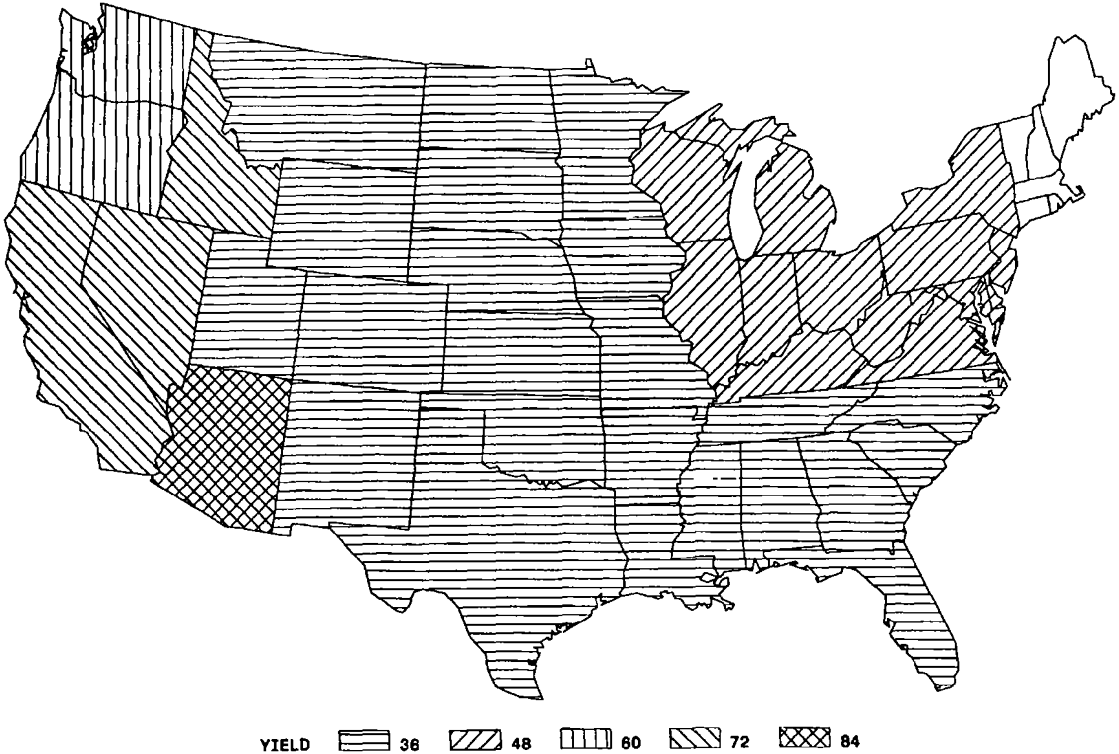


Exhibit 15

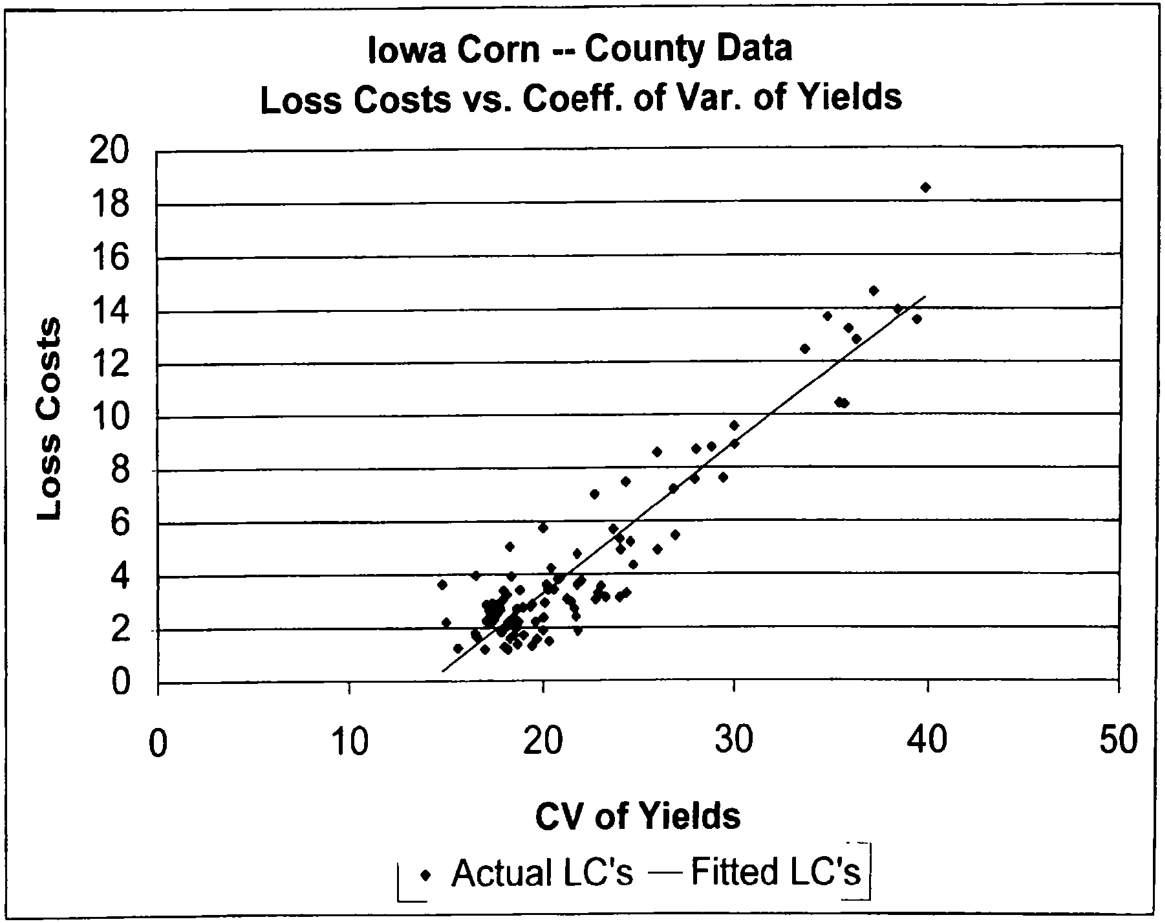
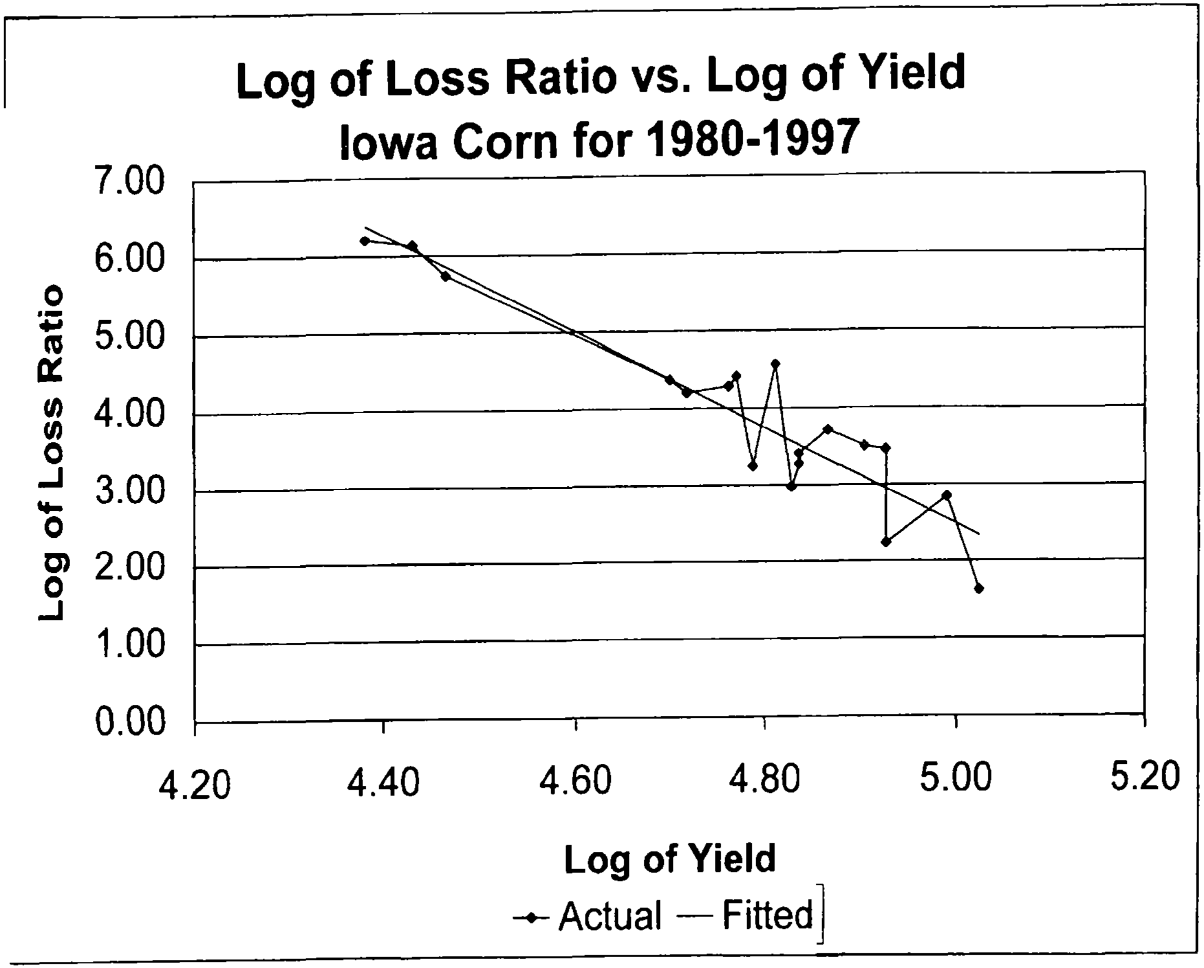

Iowa Corn -- County Data
Loss Costs vs. Coeff. of Var. of Yields
Loss Costs
CV of Yields
0
2
4
6
8
10
12
14
16
18
20
0
10
20
30
40
50
Actual LC's
Fitted LC's

Exhibit 16



Log of Loss Ratio vs. Log of Yield
Iowa Corn for 1980-1997
Log of Loss Ratio
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
4.20
4.40
4.60
4.80
5.00
5.20
Log of Yield
Actual
Fitted

Exhibit 17

Corn: Countrywide Yields over Time

Actual vs. Exponential Fit

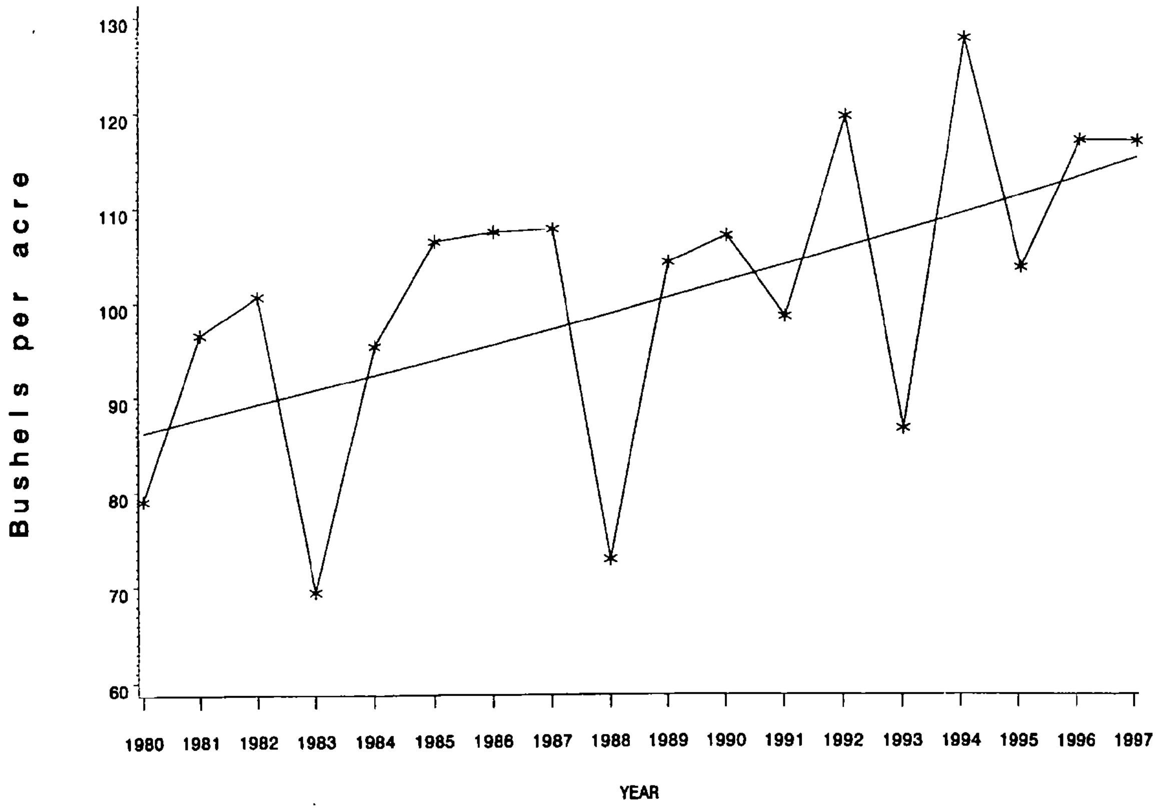


Exhibit 18

Corn: Heteroscedasticity test of Yield residuals

Note: the t statistic of the fitted slope is not significantly different from 0

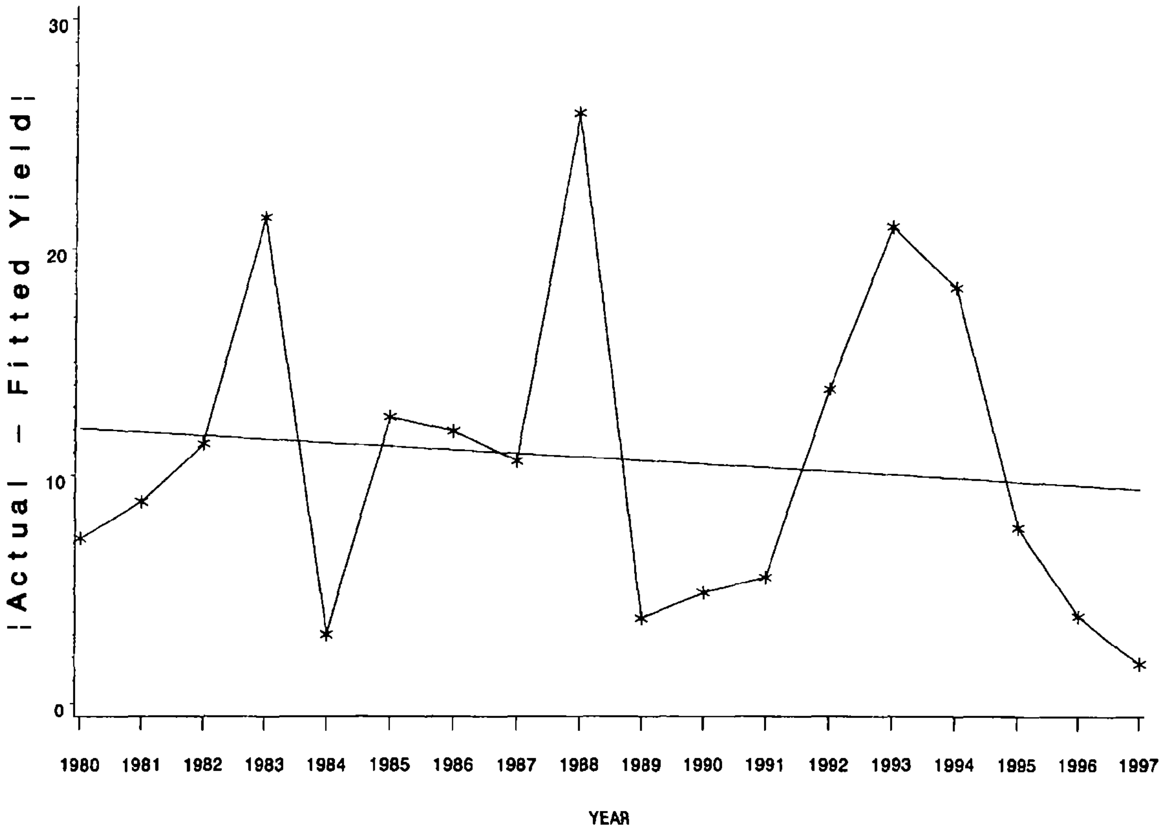


Exhibit 19

Corn: Countrywide Loss Costs over Time

Actual (stars) and Exponential Fit of Loss Costs vs. Yield

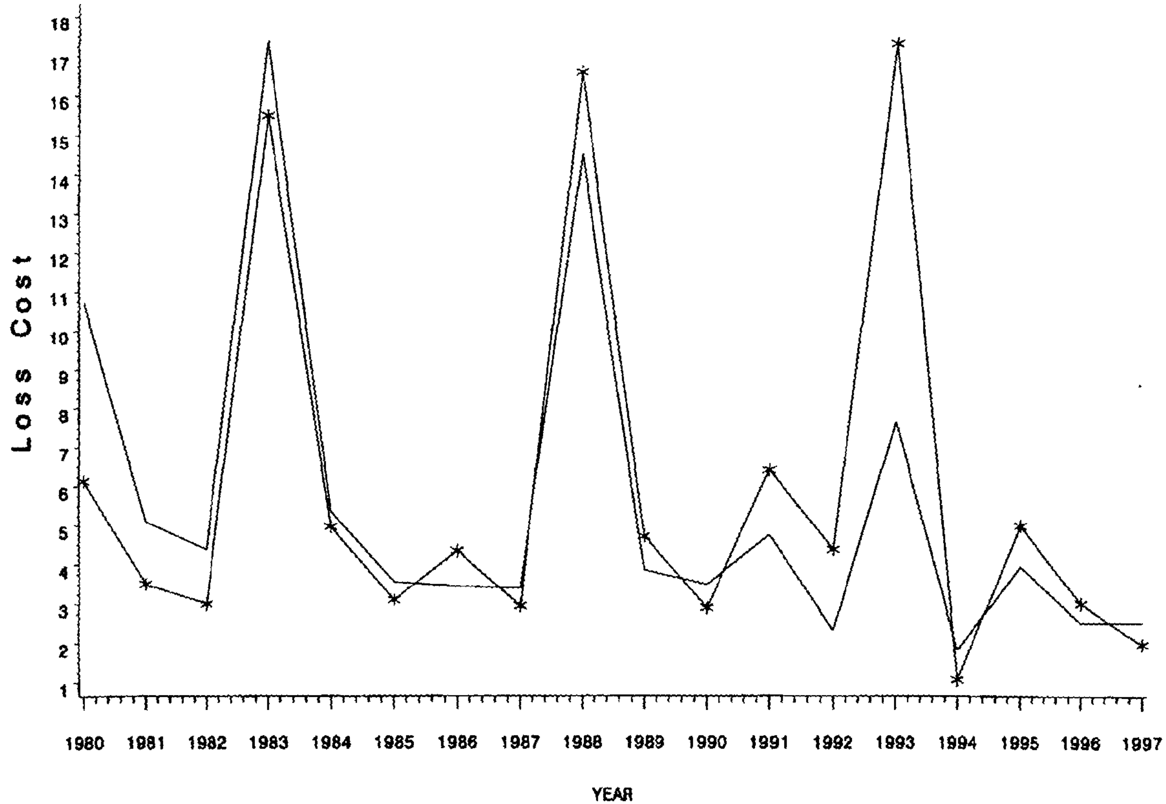


Exhibit 20

Corn: Heteroscedasticity test of Loss Cost residuals

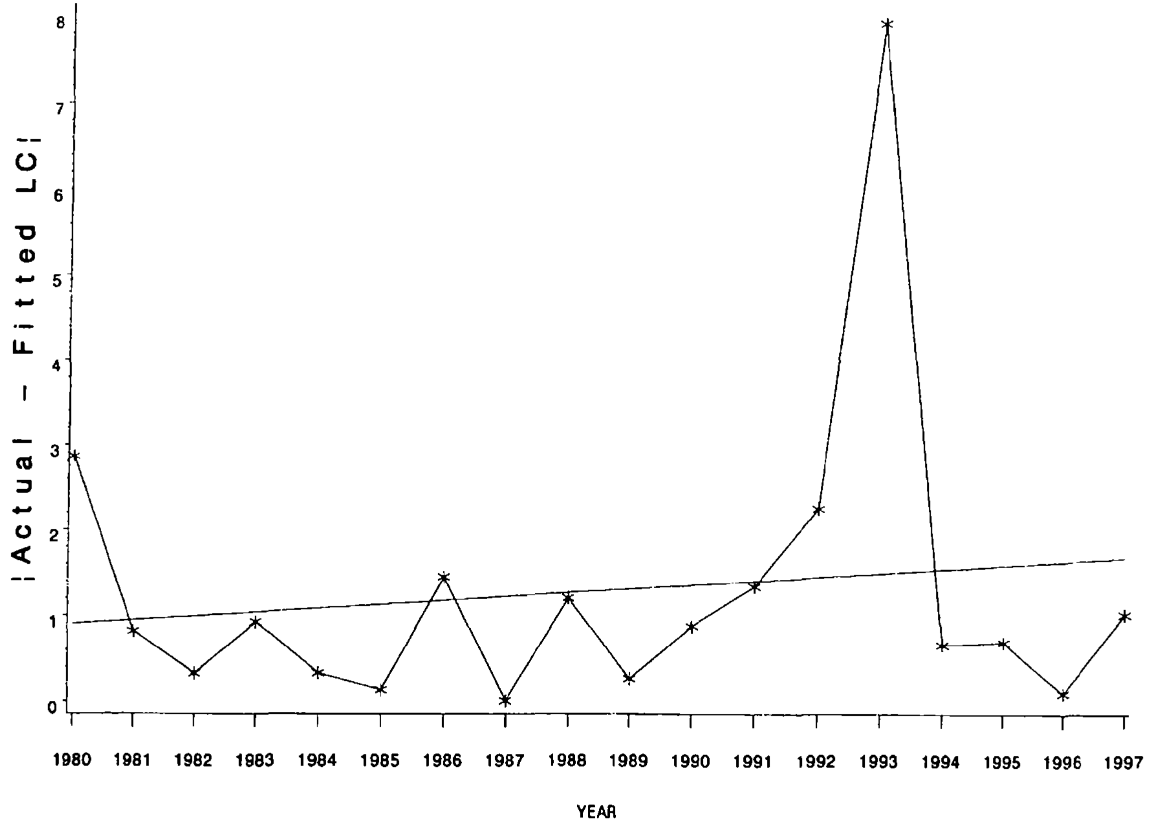
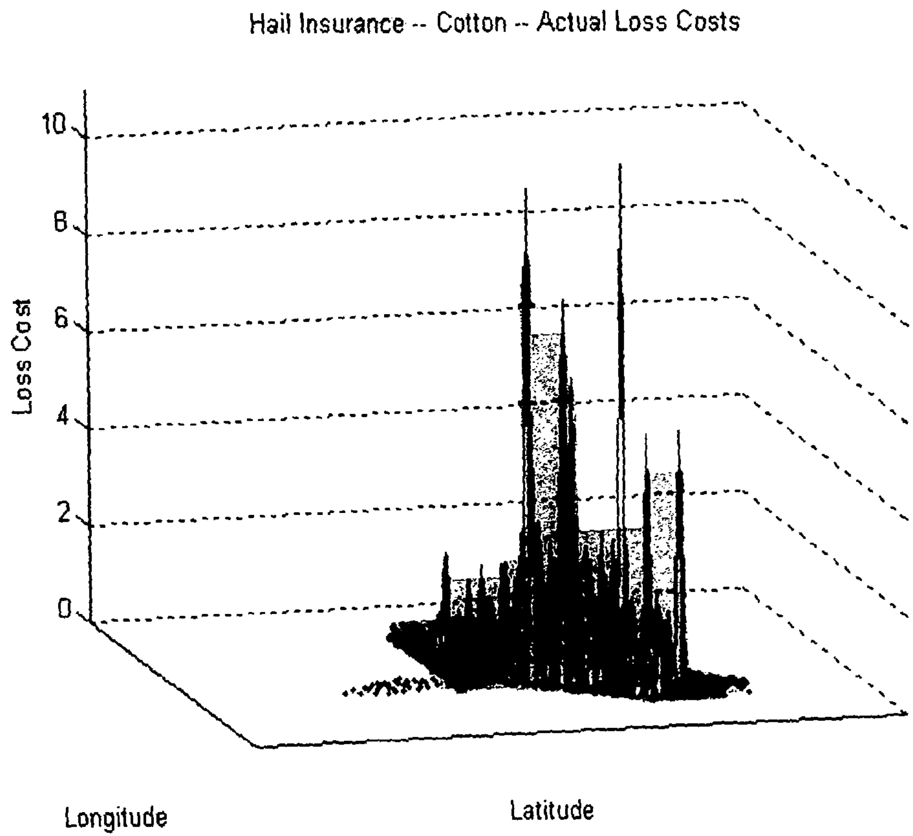


Exhibit 21



Hail Insurance -- Cotton -- Actual Loss Costs
10
8
6
4
2
0
Loss Cost
Longitude
Latitude

Exhibit 22

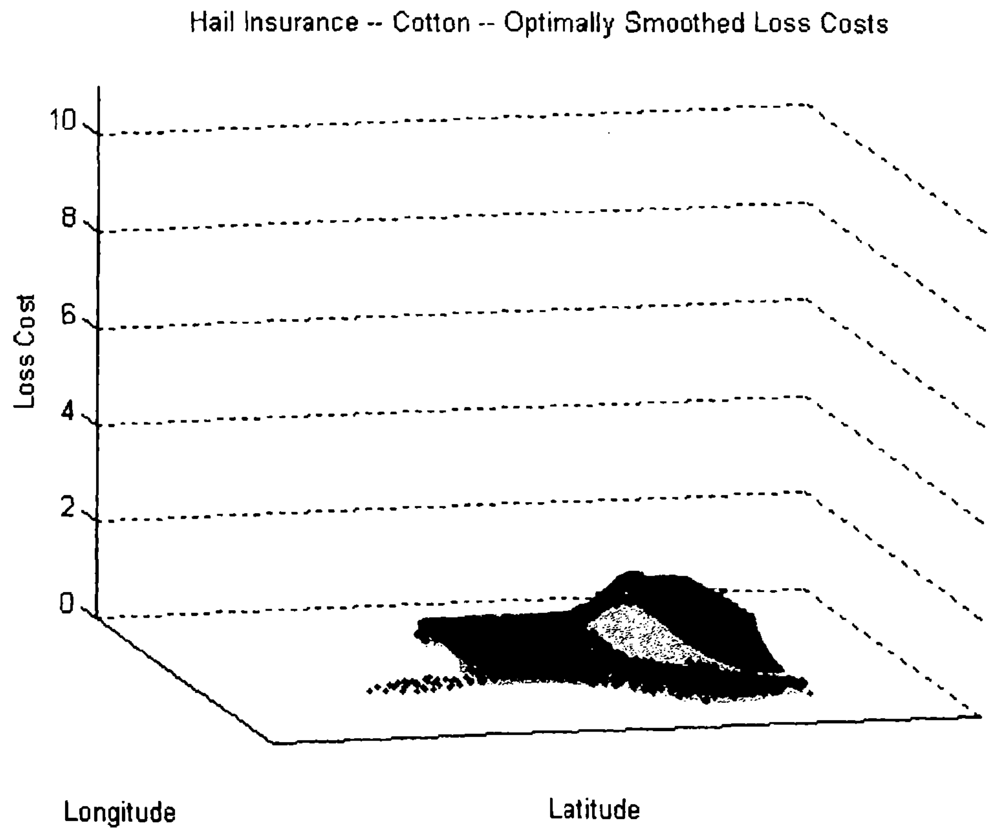

Hail Insurance -- Cotton -- Optimally Smoothed Loss Costs
Loss Cost
10
8
6
4
2
0
Longitude
Latitude

Exhibit 23

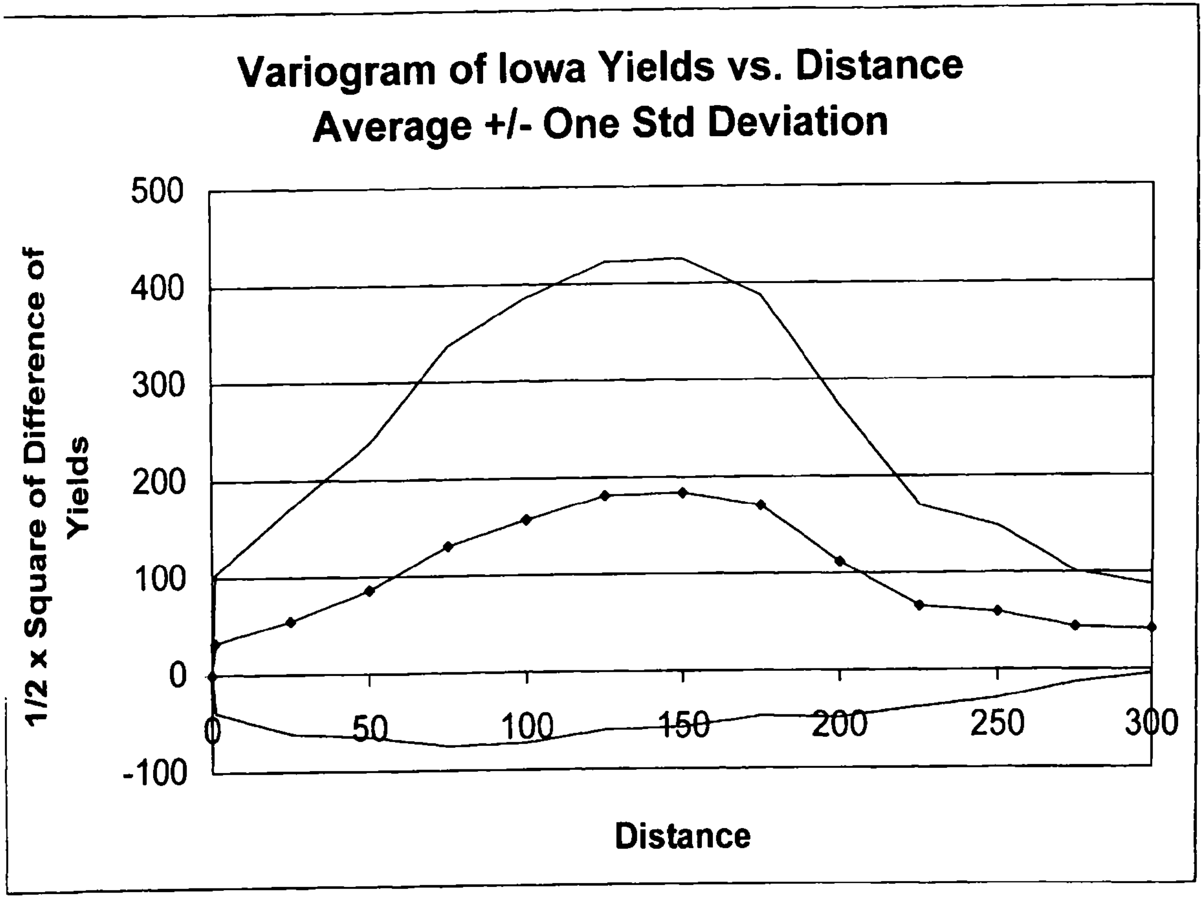

Variogram of Iowa Yields vs. Distance
Average +/- One Std Deviation
500
400
300
200
100
0
-100
1/2 x Square of Difference of Yields
0
50
100
150
200
250
300
Distance

Exhibit 24

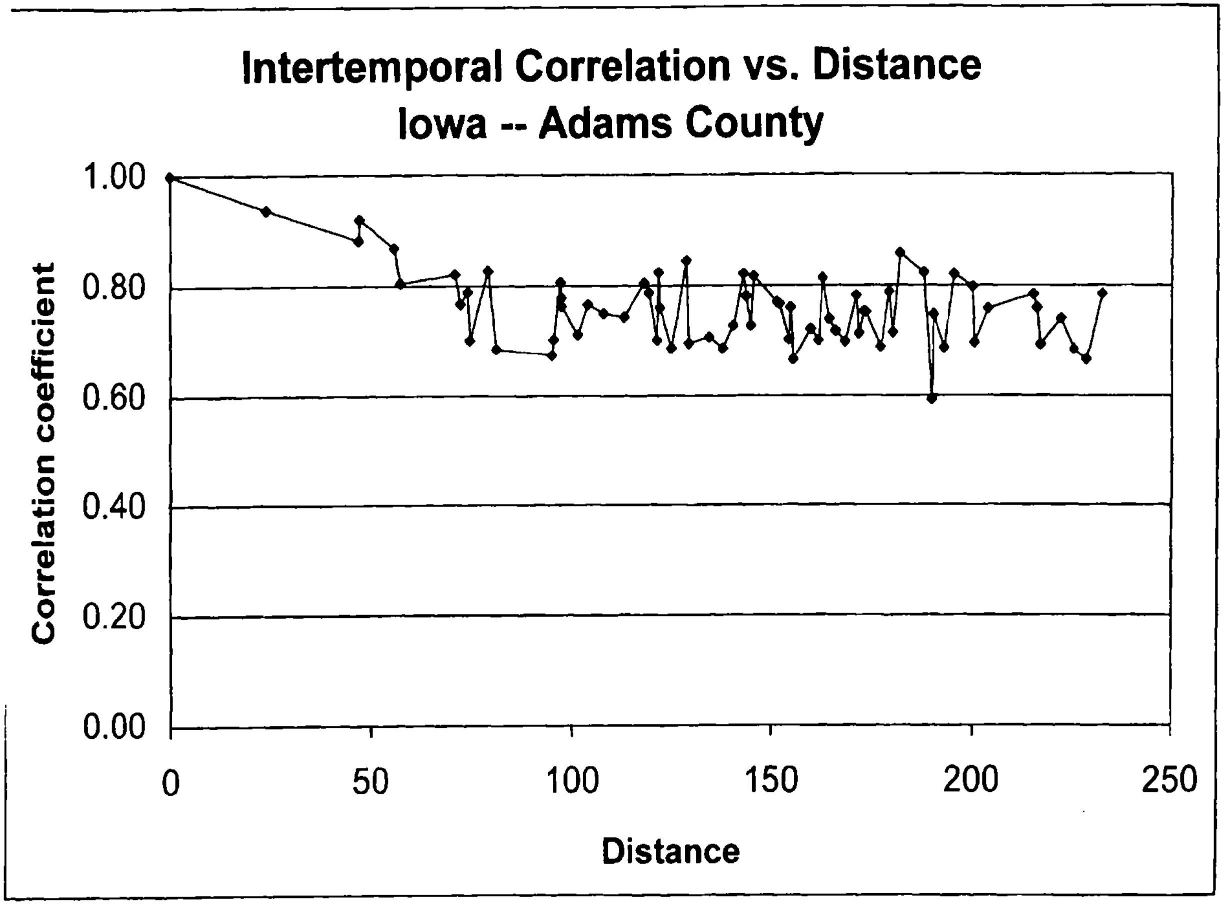

Intertemporal Correlation vs. Distance
Iowa -- Adams County
Correlation coefficient
1.00
0.80
0.60
0.40
0.20
0.00
0
50
100
150
200
250
Distance

Exhibit 25

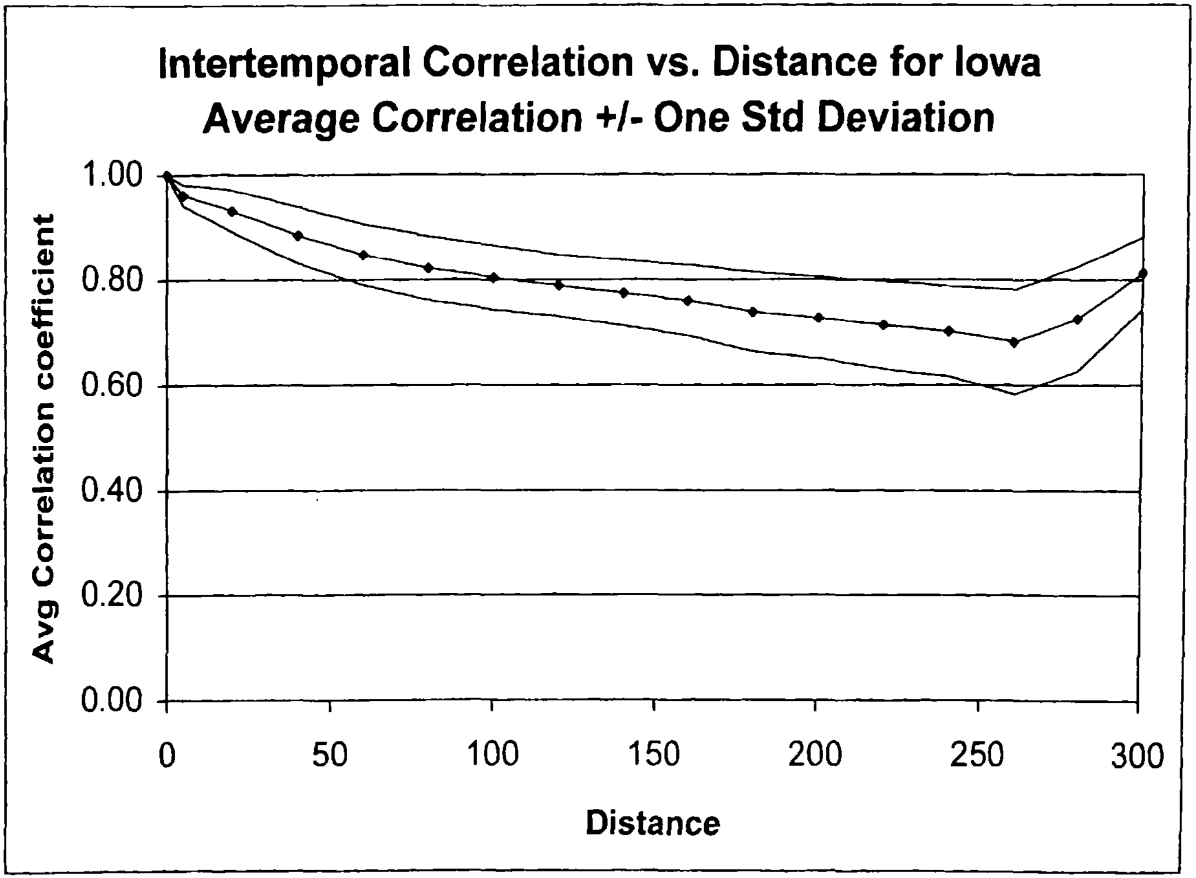

Intertemporal Correlation vs. Distance for Iowa
Average Correlation +/- One Std Deviation
Avg Correlation coefficient
1.00
0.80
0.60
0.40
0.20
0.00
0
50
100
150
200
250
300
Distance

Exhibit 26

Measurement of the Effect of Classification Factor Changes in Complex Rating Plans

David B. Schofield, FSA, MAAA

Abstract

Insurance companies strive to distinguish themselves from their competitors. One way of doing so is to refine the rating plan so that it more precisely estimates the appropriate rate for each risk. This refinement process adds complexity to the rating plan and in turn makes the measurement of changes to the individual components of the rating plan (the classifications) more difficult. In this paper, several different types of rating plans are analyzed. The rating plans range from simple plans, with either multiplicative or additive classification factors, to more complex rating plans, with mixtures of each of these types of classification factors. Methods are developed for measuring the effect of changes to classification factors on the overall rate.

1. INTRODUCTION

When a rating plan is structured such that all classification factors are multiplicative, the effect of a change in an individual classification factor on the overall rate is easy to measure. When an additive classification factor like an expense fee is included in the rating plan, measuring the effect becomes more complex. The focus of the paper is on the accurate measurement of the percent change associated with a classification factor when the rating plan contains both additive and multiplicative classification factors. The reader will be presented two different, though mathematically equivalent, methods of presenting the percentage changes in average classification factors.

Some terms are used throughout the text. In the interest of clarity, those terms are defined here:

Base rate – The dollar amount to which classification factors are applied to obtain the final rate.

Classification – A type of characteristic of the policyholder, the insured property, or type or level of coverage (e.g. Increased Limits, Deductibles, Model Year, Amount of Insurance, Town Class, etc.) that affects the final rate through the Classification factor.

Classification level – the specific value of a classification associated with a policy (e.g. a 100/300 Auto Bodily Injury Limit, a $500 deductible for Collision coverage).

Classification factor – a numeric quantity that adjusts the otherwise applicable rate to the level associated with the policy's particular classification level.

Rating plan – A mathematical model incorporating a base rate and classification factors in such a way as to produce an applicable rate.

This paper will begin with a review of the measurement of rating plan changes for four different rating models: a simple multiplicative model, a simple additive model, a simple combined additive and multiplicative model, and a more complex model with multiple additive and multiplicative classification factors. These models are represented algebraically as:

(1) $R = BM$

(2) $R = B + A$

(3) $R = BM + A$

(4) $R = B\mathbf{M} + \mathbf{A}$

Where

R = Average Rate
B = Average Base Rate
M = Average Multiplicative Classification Factor
A = Average Additive Classification Factor
$\mathbf{M}$ = The Product of All Average Multiplicative Classification Factors
$\mathbf{A}$ = The Sum of All Average Additive Classification Factors

Average classification factors are calculated using either exposures or premium that has been adjusted to remove the effect of the particular classification factor (See Appendix A for methods of calculating average classification factors).

In each of the four models, the goal will be to derive a set of multiplicative factors (f_i's), one for the base rate and each of the classifications, that, when applied to the present rate, R_0, give the proposed rate R_1, i.e.:

$$R_1 = R_0 \prod f_i \tag{1.1}$$

Equation 1.1 may also be represented as

$$R_1 = R_0\left(1 + \sum g_i\right) \tag{1.2}$$

Where the g_i's are the effects of the individual classification factors. The method for converting between the forms in Equations 1.1 and 1.2 will be shown in Section 3. Equations 1.1 and 1.2 represent the two different ways of looking at changes in classification factors. Each may be appropriate in different circumstances, as will be clear when we look at the heuristic examples associated with Model (1) and Model (2). In general, the Equation 1.1 seems more intuitive when the rating model contains all multiplicative classification factors, while Equation 1.2 seems more appropriate when there are predominantly additive classification factors. When the rating model is mixed (contains both additive and multiplicative classification factors), the actuary can choose the most appropriate method of representing the percent change.

2. MULTIPLICATIVE MODEL (Model 1)

Model (1) is the most basic of rating plans. It consists of a base rate and one multiplicative classification factor. Throughout this paper, the subscripts 0 and 1 represent current and proposed respectively. The current and proposed rating models are:

$$R_0 = B_0 M_0 \qquad (2.1)$$

$$R_1 = B_1 M_1 \qquad (2.2)$$

The goal is to find factors, f_B and f_M, representing the changes in average base rate and classification factor M, respectively, such that:

$$R_1 = R_0 f_B f_M \qquad (2.3)$$

Rearranging Equations 2.1 – 2.3, we find:

$$f_B f_M = \frac{B_1 M_1}{B_0 M_0} \qquad (2.4)$$

The following factors are selected:

$$f_B = \frac{B_1}{B_0} \qquad (2.5)$$

$$f_M = \frac{M_1}{M_0} \qquad (2.6)$$

This is the natural factorization of $f_B f_M$, the same factorization that most people would use without realizing assumptions are being made. Here, the factors are just (1 + change in classification factor). However, it should be kept in mind that it is not the only possible factorization. Consider, for instance,

$$f_B = \frac{2B_1}{B_0}; \quad f_M = \frac{M_1}{2M_0}$$

This is also a mathematically valid factorization of $f_B f_M$, though it makes little intuitive sense. This situation arises because we have one equation, Equation 2.3, with two unknowns, f_B and f_M. Additional assumptions are needed in order to restrict the possible factorizations of $f_B f_M$ to the one, intuitive, factorization we originally selected.

This paper will not detail the assumptions that are being made when factoring, but will instead use a common sense approach in factoring the more complex models.

Let's look at a numeric example. Suppose the current average base rate and average classification factor take on the following values, $B_0 = \$100$ and $M_0 = 1.0$. Also, suppose that each is being increased 10%, so that $B_1 = \$110$ and $M_1 = 1.1$. Using Equations 2.5 and 2.6, the multiplicative factors f_B and f_M are each 1.1. The overall change is 21.0%. The question is "What values of g_B and g_M should be selected to represent the percent change in base rate and classification factor M?" In our example, the base rate and the classification factor are increasing by the same percentage amount, so it makes sense to split the 21.0% evenly between those two components of the change and measure the change in each as 10.5%. So we have for Equations 1.1 and 1.2 respectively:

$$R_1 = R_0 f_B f_M = 100 \cdot (1.1)(1.1) = 121$$

$$R_1 = R_0 (1 + g_B + g_M) = 100 \cdot (1 + 0.105 + 0.105) = 121$$

This example is fairly simple. When the percent changes in the classification factors are different, determining appropriate values of g_B and g_M is more difficult. This subject will be discussed in further detail in Section 5, Multiple Additive and Multiplicative Model.

3. ADDITIVE MODEL (Model 2)

Model (2) is another simple rating plan. It consists of a base rate (B) and one additive classification factor (A). The current and proposed rating plans are:

$$R_0 = B_0 + A_0 \tag{3.1}$$

$$R_1 = B_1 + A_1 \tag{3.2}$$

Again, the goal is to find the factors representing the change in base rates and classification factor A, f_B and f_A, such that:

$$R_1 = R_0 f_B f_A \tag{3.3}$$

Rearranging Equations 3.1 – 3.3 we find:

$$f_B f_A = \frac{B_1 + A_1}{B_0 + A_0} \tag{3.4}$$

Factoring Equation 3.4 is a little more difficult than factoring Equation 2.4 in Model (1). Consider the example:

$$B_0 = 100; A_0 = 50$$

$$B_1 = 115; A_1 = 55$$

The overall change is 13.33%. There is a $15 change in the base rate and a five dollar change in the additive classification factor. How should $f_B f_A$ be factored? Two methods are readily apparent. These two methods are discussed in detail.

Additive Factoring Method I

Method I is the first of two methods that will be explored. While Method I makes some intuitive sense, it has several drawbacks that will be explored later. Method II is the author's preferred method of factoring and that method will be used primarily throughout the remainder of the paper. With Method I, we start with the percentage effect each of the dollar changes has on R_0. The percentage effects are:

$$\text{Change in base rate as a percent of present average rate} = \frac{B_1 - B_0}{R_0} = \frac{15}{150} = 10.0\%$$

$$\text{Change in classification factor as a percent of present average rate} = \frac{A_1 - A_0}{R_0} = \frac{5}{150} = 3.33\%$$

Multiplying the two factors together we get:

$$(1.1000)(1.0333) = 1.1366$$

Thus, these two factors overestimate the total change of 13.33%. The factors can be scaled to reach the 13.33% by way of the following:

$$1.1333 = 1.1366^{\alpha}$$

$$\alpha \cdot \ln 1.1366 = \ln 1.1333$$

$$\alpha = \frac{\ln 1.1333}{\ln 1.1366} = 0.9773$$

Once the scaling factor has been found, it can be applied to the individual percent changes:

$$f_B = (1.1000)^{0.9773} = 1.0976$$

$$f_A = (1.0333)^{0.9773} = 1.0325$$

The general form for calculating α is:

$$\alpha = \frac{\ln\left(\frac{R_0 + \Delta B + \Delta A}{R_0}\right)}{\ln\left(\frac{R_0 + \Delta B}{R_0}\right) + \ln\left(\frac{R_0 + \Delta A}{R_0}\right)} \quad (3.5)$$

Where

$$\Delta B = B_1 - B_0 ; \Delta A = A_1 - A_0$$

The final factors are:

$$f_B = \left(\frac{R_0 + \Delta B}{R_0} \right)^{\alpha} \tag{3.6}$$

$$f_A = \left(\frac{R_0 + \Delta A}{R_0} \right)^{\alpha} \tag{3.7}$$

Though Additive Method I certainly produces values in the range that we would expect, it has three drawbacks; the calculations are a little cumbersome, the use of the scaling factor is not intuitively appealing, and there is an interaction effect between classification factor changes. Additive Factoring Method II avoids these problems. The interaction effect will be made clear in Section 6, Interactions and Additive Factoring.

Additive Factoring Method II

Additive Method II takes a more direct approach. It uses exponential weighting to factor the overall percent change into f_B and f_A. Continuing with our previous example, we have a \$15 increase in the base rate and a five dollar increase in the additive classification factor. The Method II factorization is:

$$f_B f_A = 1.1333 = 1.1333^{\frac{15+5}{20}}$$

$$f_B f_A = (1.1333)^{\frac{15}{20}} (1.1333)^{\frac{5}{20}}$$

or

$$f_B = (1.1333)^{\frac{15}{20}} = 1.0984$$

$$f_A = (1.1333)^{\frac{5}{20}} = 1.0318$$

The general representation of this is:

$$f_B = \left(\frac{R_1}{R_0}\right)^{\frac{\Delta B}{\Delta B + \Delta A}} \quad (3.8)$$

$$f_A = \left(\frac{R_1}{R_0}\right)^{\frac{\Delta A}{\Delta B + \Delta A}} \quad (3.9)$$

when

$$\Delta B + \Delta A \neq 0$$

When $\Delta B + \Delta A = 0$, we get the following interesting result:

$$f_B = e^{\frac{\Delta B}{R_0}} \quad (3.10)$$

$$f_A = e^{\frac{\Delta A}{R_0}} \quad (3.11)$$

The derivation of this result is in Appendix B and can be generalized for use in the remainder of the models discussed in this paper.

Equations 3.8 and 3.9 are the heart of Method II. Conceptually, these can be written as

$$f_i = \left(\frac{\text{New Average Rate}}{\text{Old Average Rate}}\right)^{\frac{\text{Avg. Dollar Change in Rate Due to Change in Classification Factor i}}{\text{Average Dollar Change in Rate}}} \quad (3.12)$$

The Method II factorization is a much simpler calculation than Method I and gives results in the range we expect. No scaling factors are needed for this method. Again, another reason to prefer Method II over Method I will be discussed in Section 6, Interactions and

Additive Factoring. For the remainder of the paper, Additive Factoring Method II will be used.

We still need to find the solution in the form of Equation 1.2. Let:

$$1 + g_B + g_A = \frac{B_1 + A_1}{B_0 + A_0} \qquad (3.13)$$

$$1 + g_B + g_A = \frac{R_0 + \Delta B + \Delta A}{R_0} = 1 + \frac{\Delta B}{R_0} + \frac{\Delta A}{R_0}$$

so:

$$g_B = \frac{\Delta B}{R_0} \qquad (3.14)$$

$$g_A = \frac{\Delta A}{R_0} \qquad (3.15)$$

Again, in our example, we have

$$B_0 = 100; A_0 = 50$$

$$B_1 = 115; A_1 = 55$$

So,

$$g_B = \frac{15}{150} = 0.1000$$

$$g_A = \frac{5}{150} = 0.0333$$

We can also use the results in Equation 3.13 to find a conversion method between those factors shown in Equation 1.1 and those of Equation 1.2. From Equation 3.8 we have:

$$f_B = \left(\frac{R_1}{R_0}\right)^{\frac{\Delta B}{\Delta R}} \qquad (3.16)$$

Multiplying the exponent by a fancy form of one (R_0 / R_0) produces:

$$f_B = \left(\frac{R_1}{R_0}\right)^{\frac{\Delta B}{\Delta R}\cdot\frac{R_0}{R_0}} = \left(\frac{R_1}{R_0}\right)^{\frac{\Delta B}{R_0}\cdot\frac{R_0}{\Delta R}}$$

Substituting using Equation 3.14

$$f_B = \left(\frac{R_1}{R_0}\right)^{g_B\cdot\frac{R_0}{\Delta R}}$$

Solving for g_B gives:

$$g_B = \frac{\Delta R}{R_0}\cdot\frac{\ln f_B}{\ln\left(\frac{R_1}{R_0}\right)} \tag{3.17}$$

Under the Multiplicative model, calculating percentage changes using Equation 1.1 makes the most intuitive sense. In contrast, Equation 1.2 seems more appropriate under the Additive model. In a mixed model, the decision about the form of the percent change to use is less clear cut. Since there is a one to one correspondence between the two forms, it is up to the actuary to decide which method is most accurate in representing these changes.

4. SIMPLE ADDITIVE AND MULTIPLICATIVE MODEL (Model 3)

Model (3) is a simple combination of Model (1) and Model (2). Under Model (3), the current and proposed rating plans are:

$$R_0 = B_0 M_0 + A_0 \tag{4.1}$$

$$R_1 = B_1 M_1 + A_1 \tag{4.2}$$

The objective is to find a set of factors that accurately represent the multiplicative effect of changes to B, M and A. We want to find f_B, f_M and f_A that satisfy:

$$R_1 = R_0 f_B f_M f_A \tag{4.3}$$

Consider the following numeric example. Table 1 contains the information about the current and proposed rates needed to determine the individual effects of each classification factor change.

Table 1

Variable	Current	Proposed	Δ	%Δ
R	$200	$240	$40	20.00%
B	$100	$110	$10	10.00%
M	1.65	1.80	0.15	9.10%
A	$35	$42	$7	20.00%

Finding the factors is a two-step process. First, determine the effects of the two additive components (BM) and A, i.e., we find values for $f_B f_M$ and f_A. Next, the multiplicative sub-components of the (BM) component are partitioned into f_B and f_M. There is a total dollar change of $40. The change in A accounts for seven dollars of the total change; the changes in B and M account for the remaining $33. Measuring these changes individually results in the following percent changes by classification factor:

$$\text{Additive}: \frac{\Delta A}{R_0} = \frac{7}{200} = 0.035 \quad \text{or } 3.5\%$$

$$\text{Multiplicative}: \frac{\Delta(\text{BM})}{\text{R}_0} = \frac{33}{200} = 0.165 \quad \text{or } 16.5\%$$

$$\begin{aligned}\Delta(BM) &= B_1 M_1 - B_0 M_0 = (B_0 + \Delta B)(M_0 + \Delta M) - B_0 M_0 \\ &= B_0 \Delta M + \Delta B M_0 + \Delta B \Delta M \\ &= 100 \cdot 0.15 + 10 \cdot 1.65 + 10 \cdot 0.15 = 33\end{aligned}$$

The total change in R, of 20.0%, is factored into additive and multiplicative effects by using Method II factorization:

$$\frac{R_1}{R_0} = 1.2^1 = 1.2^{\frac{\Delta A + \Delta(BM)}{\Delta R}} = 1.2^{\frac{7+33}{40}}$$

We let

$$f_A = 1.2^{\frac{7}{40}} = 1.0324$$

$$f_B f_M = 1.2^{\frac{33}{40}} = 1.1623$$

If there were no additive classification factor in our model, $f_B f_M$ would be the same as in Model (1), the product of the two percentage changes associated with the base rate change and the multiplicative classification factor change:

$$\frac{B_1 M_1}{B_0 M_0} = 1.2$$

But we know the overall effect of the base rate and multiplicative classification factor change is 1.1623. So we factor this multiplicative part of our model just as was done in Model (1), then scale the factors to produce the overall effect of 1.1623. To determine the individual factors f_B and f_M while retaining the relative effect of the underlying classification factors, let:

$$f_B f_M = \left(\frac{B_1}{B_0} \cdot \frac{M_1}{M_0}\right)^{\alpha} \tag{4.4}$$

$$f_B f_M = \left(\frac{B_1}{B_0}\right)^{\alpha} \left(\frac{M_1}{M_0}\right)^{\alpha} \tag{4.5}$$

Factoring results in:

$$f_B = \left(\frac{B_1}{B_0}\right)^{\alpha} \tag{4.6}$$

$$f_M = \left(\frac{M_1}{M_0}\right)^{\alpha} \tag{4.7}$$

Solving Equation 4.4 for α :

$$\alpha = \frac{\ln(f_B f_M)}{\ln\left(\frac{B_1 M_1}{B_0 M_0}\right)} \tag{4.8}$$

$$\alpha = \frac{\ln(1.1623)}{\ln(1.2)}$$

$$\alpha = 0.8249$$

Plugging this result back into Equations 4.6 and 4.7 yields:

$$f_B = \left(\frac{B_1}{B_0}\right)^{\alpha} = \left(\frac{110}{100}\right)^{0.8249} = 1.0818$$

$$f_M = \left(\frac{M_1}{M_0}\right)^{\alpha} = \left(\frac{1.80}{1.65}\right)^{0.8249} = 1.0744$$

The effects of the changes in the base rate and classification factors on the overall rate are shown in Table 2.

Table 2

	Equation 1.1	Equation 1.2
Base Rate (B):	8.18%	8.63%
Classification Factor (M):	7.44%	7.87%
Classification Factor (A):	3.24%	3.50%

Model (3) has presented all the tools necessary to measure the effect of even very complex rating plans. The basic idea is to group whatever rating plan model you may have into a series of additive components (we know how to measure these from Model (2)), and then calculate the multiplicative effects within each of the additive components. Model (4) is a slightly more complex than Model (3), and shows how factors can be determined when there are multiple additive and multiplicative classification factors.

5. MULTIPLE ADDITIVE AND MULTIPLICATIVE MODEL (Model 4)

Model (4) expands on all that has been learned in the first three models. It is very similar to Model (3), but contains both multiple additive and multiplicative classification factors. In its simplest form, the model can be written as:

$$R = B\mathbf{M} + \mathbf{A}$$

M represents the product of all the multiplicative classification factors (other than the base rate) and **A** is the sum of the additive classification factors. So the present and proposed model can be represented as:

$$R_0 = \left(B_0 \prod M_{j0}\right) + \sum A_{k0} \tag{5.1}$$

$$R_1 = \left(B_1 \prod M_{j1}\right) + \sum A_{k1} \tag{5.2}$$

We want to find f_B, f_M and f_A, a set of factors that satisfy:

$$R_1 = R_0 \prod f_i \tag{5.3}$$

Consider the following numeric example. Table 3 contains the information about the current and proposed rates needed to determine the individual effects of each classification factor change.

Table 3

Variable	Current	Proposed	Δ	%Δ

R	$233.25	$279.80	$46.55	19.96%
B	$100.00	$110.00	$10.00	10.00%
M1	1.65	1.80	0.15	9.09%
M2	1.05	1.10	0.05	4.76%
A1	$35.00	$42.00	$7.00	20.00%
A2	$10.00	$8.00	($2.00)	-20.00%
A3	$15.00	$12.00	($3.00)	-20.00%

Table 3 is similar to the example shown in Section 4, Table 1. In Table 3, multiplicative classification factor M2 has been added as well as additive classification factors, A2 and A3. The base rate and other classification factors remain as they did in Table 1. Finding the factors is again a two-step process. First partition all the additive components and determine each of those effects. Then calculate the effects of the multiplicative sub-components of each (if any). The model for this rating plan is:

$$R = B \cdot M1 \cdot M2 + A1 + A2 + A3$$

We calculate $f_B f_{M1} f_{M2}$, f_{A1}, f_{A2} and f_{A3} using Additive Factoring Method II. First note:

$$\Delta A1 = 7$$
$$\Delta A2 = -2$$
$$\Delta A3 = -3$$
$$\Delta(B \cdot M1 \cdot M2) = \Delta R - (\Delta A1 + \Delta A2 + \Delta A3) = 46.55 - (7 - 2 - 3) = 44.55$$

So

$$f_B f_{M1} f_{M2} = (1.1996)^{\frac{44.55}{46.55}} = 1.1903$$
$$f_{A1} = (1.1996)^{\frac{7}{46.55}} = 1.0277$$
$$f_{A2} = (1.1996)^{\frac{-2}{46.55}} = 0.9922$$
$$f_{A3} = (1.1996)^{\frac{-3}{46.55}} = 0.9883$$

The scaling factor for $f_B f_{M1} f_{M2}$ is:

$$\alpha = \frac{\ln(f_B f_{M1} f_{M2})}{\ln\left(\frac{B_1 M1_1 M2_1}{B_0 M1_0 M2_0}\right)}$$

$$\alpha = \frac{\ln(1.1903)}{\ln(1.2571)}$$

$$\alpha = 0.7614$$

Giving us:

$$f_B = \left(\frac{B_1}{B_0}\right)^{\alpha} = \left(\frac{110}{100}\right)^{0.7614} = 1.0753$$

$$f_{M1} = \left(\frac{M1_1}{M1_0}\right)^{\alpha} = \left(\frac{1.80}{1.65}\right)^{0.7614} = 1.0685$$

$$f_{M2} = \left(\frac{M2_1}{M2_0}\right)^{\alpha} = \left(\frac{1.10}{1.05}\right)^{0.7614} = 1.0361$$

The effect of the changes in the base rate and classification factors on the overall rate is show in Table 4.

Table 4

	Equation 1.1	Equation 1.2
Base Rate (B):	7.53%	7.96%
Classification Factor (M1):	6.85%	7.27%
Classification Factor (M2):	3.61%	3.89%
Classification Factor (A1):	2.77%	3.00%
Classification Factor (A2):	-0.78%	-0.86%
Classification Factor (A3):	-1.17%	-1.29%

A comparison of these results with those of the example in Section 4, shows that the addition of classifications can have an impact on the measurement of the effect of the changes in the original classification factors.

6. INTERACTIONS AND ADDITIVE FACTORING

The major motivation for using Additive Factoring Method II, over Method I, is how each treats interaction effects between additive portions of the rating plan. For instance, suppose we have the following rating model:

$$R = B + A1 + A2 \tag{6.1}$$

Now consider the numeric example shown in Table 5. It contains data about the current rates but has different proposed rates for classifications A1 and A2.

Table 5

	Variable	Current	Proposed	Δ	%Δ
Scenario	R	$200.00	$230.00	$30.00	15.00%
1	B	100.00	110.00	10.00	10.00%
	A1	50.00	60.00	10.00	20.00%
	A2	50.00	60.00	10.00	20.00%
Scenario	R	$200.00	$230.00	$30.00	15.00%
2	B	100.00	110.00	10.00	10.00%
	A1	50.00	50.00	0.00	0.00%
	A2	50.00	70.00	20.00	40.00%

Under each scenario, the overall rate change and base rate change are 15.0% and 10.0% respectively. The total dollar change for classification factors A1 and A2 is constant over the two scenarios (i.e. $\Delta A1 + \Delta A2 = 20$). We have the following changes under each scenario:

Scenario 1: $\Delta R = \Delta B + \Delta A1 + \Delta A2 = 10 + 10 + 10 = 30$

Scenario 2: $\Delta R = \Delta B + \Delta A1 + \Delta A2 = 10 + 0 + 20 = 30$

We would expect that the measured effect of the change in the base rate would be the same under both scenarios. Under Additive Factoring Method I we have for the base rate change:

$$f_B = \left(\frac{R_0 + \Delta B}{R_0}\right)^{\alpha} = \left(\frac{200+10}{200}\right)^{\alpha} = 1.05^{\alpha}$$

$$\alpha = \frac{\ln\left(\frac{R_0 + \Delta B + \Delta A1 + \Delta A2}{R_0}\right)}{\ln\left(\frac{R_0 + \Delta B}{R_0}\right) + \ln\left(\frac{R_0 + \Delta A1}{R_0}\right) + \ln\left(\frac{R_0 + \Delta A2}{R_0}\right)}$$

Under Scenario 1 we have:

$$\alpha = \frac{\ln\left(\frac{230}{200}\right)}{\ln\left(\frac{210}{200}\right) + \ln\left(\frac{210}{200}\right) + \ln\left(\frac{210}{200}\right)} = \frac{0.1398}{0.0488 + 0.0488 + 0.0488} = 0.9549$$

$$f_B = 1.05^{0.9549} = 1.0477 \quad \text{or} \quad 4.77\%$$

Under Scenario 2:

$$\alpha = \frac{\ln\left(\frac{230}{200}\right)}{\ln\left(\frac{210}{200}\right) + \ln\left(\frac{200}{200}\right) + \ln\left(\frac{220}{200}\right)} = \frac{0.1398}{0.0488 + 0.0000 + 0.0953} = 0.9702$$

$$f_B = 1.05^{0.9702} = 1.0485 \quad \text{or} \quad 4.85\%$$

This difference between the results shown under these scenarios is one reason for preferring Additive Factoring Method II. Under Method II, each scenario produces:

$$f_B = \left(\frac{R_1}{R_0}\right)^{\frac{\Delta B}{\Delta R}} = \left(\frac{230}{200}\right)^{\frac{10}{30}} = 1.0477 \quad \text{or} \quad 4.77\%$$

This is, of course, the same result we obtained under Scenario 1 above (because all the dollar changes in classification factors were the same under that scenario). The value of f_B will remain constant for all $\Delta A1$ and $\Delta A2$, provided $\Delta A1 + \Delta A2 = 20$. It is this invariance property that makes Additive Factoring Method II favorable to Method I.

7. MULTIPLE EFFECTS OF A SINGLE CLASSIFICATION

Suppose, a rating plan can be represented by the following model:

$$R = B \cdot M1 + B \cdot M2 + A1 \tag{7.1}$$

Here the base rate has effects in two of the components. We still want one value for f_B. How can this be accomplished? The answer is to simply calculate the effect of the base rate change in each of the components and multiply them together. Consider the heuristic example shown in Table 6.

Table 6

Variable	Current	Proposed	Δ	%Δ
R	$250.00	$313.00	$63.00	25.20%
B	$100.00	$110.00	$10.00	10.00%
M1	1.00	1.10	0.10	10.00%
M2	1.00	1.20	0.20	20.00%
A1	$50.00	$60.00	$10.00	20.00%

First find the factors for the components, $f_B f_{M1}$, $f_B f_{M2}$ and f_{A1}.

$$\Delta(B \cdot M1) = B_1 M1_1 - B_0 M1_0 = 121 - 100 = 21$$
$$\Delta(B \cdot M2) = B_1 M2_1 - B_0 M2_0 = 132 - 100 = 32$$

$$\Delta A1 = 60 - 50 = 10$$

Using Additive Factoring Method II, we get:

$$f_B f_{M1} = \left(\frac{313}{250}\right)^{\frac{21}{63}} = 1.0778.$$

$$f_B f_{M2} = \left(\frac{313}{250}\right)^{\frac{32}{63}} = 1.1209$$

$$f_{A1} = \left(\frac{313}{250}\right)^{\frac{10}{63}} = 1.0363$$

Factoring $f_B f_{M1}$:

$$f_B f_{M1} = \left(\frac{B_1}{B_0} \cdot \frac{M1_1}{M1_0}\right)^{\alpha} = \left(\frac{110}{100} \cdot \frac{1.1}{1.0}\right)^{\alpha} = 1.21^{\alpha}$$

$$\alpha = \frac{\ln(f_B f_{M1})}{\ln\left(\frac{B_1 M1_1}{B_0 M1_0}\right)} = \frac{\ln(1.0778)}{\ln(1.21)} = 0.3930$$

$$f_B = \left(\frac{110}{100}\right)^{0.3930} = 1.0382$$

$$f_{M1} = \left(\frac{1.1}{1.0}\right)^{0.3930} = 1.0382$$

Next, factor $f_B f_{M2}$:

$$f_B f_{M2} = \left(\frac{B_1}{B_0} \cdot \frac{M2_1}{M2_0}\right)^{\alpha} = \left(\frac{110}{100} \cdot \frac{1.2}{1.0}\right)^{\alpha} = 1.32^{\alpha}$$

$$\alpha = \frac{\ln(f_B f_{M2})}{\ln\left(\frac{B_1 M2_1}{B_0 M2_0}\right)} = \frac{\ln(1.1209)}{\ln(1.32)} = 0.4111$$

$$f_B = \left(\frac{110}{100}\right)^{0.4111} = 1.0400$$

$$f_{M2} = \left(\frac{1.2}{1.0}\right)^{0.4111} = 1.0778$$

To calculate the total effect of f_B:

$$f_B = (1.0382)(1.0400) = 1.0797$$

The effects of each of the classification factor changes are shown in Table 7.

Table 7

	Equation 1.1	Equation 1.2
Base Rate	7.97%	8.60%
Classification Factor (M1):	3.82%	4.20%
Classification Factor (M2):	7.78%	8.40%
Classification Factor (A1):	3.63%	4.00%

8. PROCESS SUMMARY

The methodology has now been developed to measure the effect of changes in classification factors (CF's) where the form of the rating plan can be represented as:

$$R = \sum_{i=1}^{n} \prod_{j=1}^{m_i} CF_{ij} \qquad CF_{ij} \in (\mathrm{B}, \mathrm{M}_1, \ldots, \mathrm{M}_p, \mathrm{A}_1, \ldots, \mathrm{A}_q) \qquad (8.1)$$

Where there are p multiplicative and q additive classifications; n represents the number of additive components; m_i is the number of multiplicative sub-components within additive component i. Note that each of the four original models can be represented using Equation 8.1. In order to calculate the effects of the classification factor changes, the following steps should be followed:

1) Determine the mathematical representation of the rating plan and put it in the form of Equation 8.1.

2) Calculate the average classification factor change (Appendix A).

3) Determine the effects of each of the additive components of the rating plan (Section 3, Additive Factoring Method II).

4) Determine the effects for each set of multiplicative sub-components (Section 4).

5) Combine the effects of classifications that are represented in more than one additive component (Section 7).

6) Transform the result into either Equation 1.1 or Equation 1.2 format if necessary.

The result of this process is the accurate partition of the total rate change into the effects of the changes in each of the classification factors.

9. CONCLUSION

Measuring the effect of classification factor changes can be difficult in anything other than a simple rating plan. This paper has shown methods for accurately measuring the effects of individual classification factor changes in the context of more complex rating plans. The rating models included combinations of both additive and multiplicative classification factors. It was also shown that the percent change for the classification factors within a rating plan could be viewed from a factor point of view (Equation 1.1) or as an additive percent (Equation 1.2). The techniques described in this paper can of course be used in models other than those shown.

Appendix A

Methods for Calculating the Average Classification Factor

An average classification factor is the weighted average of the factors for the individual levels of the classification. There are two different weighting schemes used in calculating this weighted average, exposures and adjusted premium at current rate level. Using exposures as weights assumes that the classification's factors are uncorrelated with the factors of the other classifications. If the actuary does not wish to make this assumption, then adjusted premium should be used. The adjustment removes the effect of the particular classification on the premium. If unadjusted premiums are used as weights, the effect of the classification factor is, in effect, doubled. The actuary must also decide which of either written or earned weights is appropriate; this will generally depend on the type of data being used for the analysis, Accident Year, Calendar Year or Policy Year data (of course, this is all conditioned on the actuary having the appropriate data available). Table A.1 provides the data for the first example of the calculation of an average classification (Limit) factor using each of exposures and adjusted premium.

Table A.1

Limit of Liability	Limit Factor	Exposures	Premium	Adjusted Premium
1,000	1.0	1,000	100,000	100,000
2,000	1.4	800	123,200	88,000
5,000	1.6	500	112,000	70,000
10,000	1.8	200	72,000	40,000
Total		2,500	407,200	298,000

For each Limit of Liability we calculate (assuming all multiplicative classification factors):

$$\text{Adjusted Premium} = \frac{\text{Premium}}{\text{Limit Factor}}$$

Using exposures as weights yields an average classification factor (in this case, Limit factor) of:

$$\begin{aligned}\text{Average Limit Factor} &= \frac{\sum_{\text{all limits}} (\text{Limit Factor}) \cdot (\text{Exposures})}{\sum_{\text{all limits}} \text{Exposures}} \\ &= \frac{3280}{2500} \\ &= 1.312\end{aligned}$$

When the limit factor is correlated with other classification factors, adjusted premiums are used as weights, yielding:

$$\begin{aligned}\text{Average Limit Factor} &= \frac{\sum_{\text{all limits}} (\text{Limit Factor}) \cdot (\text{Adjusted Premium})}{\sum_{\text{all limits}} \text{Adjusted Premium}} \\ &= \frac{\sum_{\text{all limits}} \text{Premium}}{\sum_{\text{all limits}} \text{Adjusted Premium}} \\ &= \frac{407{,}200}{298{,}000} \\ &= 1.366\end{aligned}$$

Care should be taken when choosing which premium to use and the adjustments to be made. As in the above example, when the rating plan has all multiplicative classification factors, the premium used can be total premium; the adjusted premium has the effect of the classification removed. Suppose that the premium in the example shown in Table A.1 includes a fixed expense fee of $20 per exposure. Table A.2 gives the data to show the effect of this difference in rating plans on the premium to be used as a weighting variable.

Table A.2

Limit of Liability	Limit Factor	Exposures	Expense Fees	Premium with Exp. Fees	Premium without Exp. Fees	Adjusted Premium
1,000	1.0	1,000	20,000	100,000	80,000	80,000
2,000	1.4	800	16,000	123,200	107,200	76,571
5,000	1.6	500	10,000	112,000	102,000	63,750
10,000	1.8	200	4,000	72,000	68,000	37,778
Total		2,500	50,000	407,200	357,200	258,099

Here the Adjusted Premium is calculated as:

$$\text{Adjusted Premium} = \frac{\text{Premium without Expense Fees}}{\text{Limit Factor}}$$

The average classification factor (Limit Factor in this case) is calculated as:

$$\begin{aligned}
\text{Averge Limit Factor} &= \frac{\sum_{\text{all limits}} (\text{Limit Factor}) \cdot (\text{Adjusted Premium})}{\sum_{\text{all limits}} \text{Adjusted Premium}} \\
&= \frac{\sum_{\text{all limits}} \text{Premium without Expense Fees}}{\sum_{\text{all limits}} \text{Adjusted Premium}} \\
&= \frac{357{,}200}{258{,}099} \\
&= 1.384
\end{aligned}$$

When using premium as a weight, it is important to use only that portion of the premium to which the classification factor is being applied; then adjust that premium to remove the effects of the classification factor. The ratio of the applicable premium to adjusted premium is the average classification factor.

Appendix B

Derivation of Percent Changes for Additive Components When $\Delta R = 0$

A revenue neutral rate change ($\Delta R = 0$) must be handled differently than a non-zero change when using Additive Factoring Method II. The general form of Additive Factoring Method II, given additive classification factor A and $\Delta R = R_1 - R_0$, is:

$$f_A = \left(\frac{R_1}{R_0}\right)^{\frac{\Delta A}{\Delta R}}$$

This can be rewritten as:

$$f_A = \left(1 + \frac{\Delta R}{R_0}\right)^{\frac{\Delta A}{\Delta R}\cdot\frac{R_0}{R_0}}$$

$$= \left(1 + \frac{\Delta R}{R_0}\right)^{\frac{R_0}{\Delta R}\cdot\frac{\Delta A}{R_0}}$$

Taking the limit as $\Delta R \to 0$:

$$f_A = \lim_{\Delta R \to 0}\left(1 + \frac{\Delta R}{R_0}\right)^{\frac{R_0}{\Delta R}\cdot\frac{\Delta A}{R_0}}$$

$$= \left(\lim_{\Delta R \to 0}\left(1 + \frac{\Delta R}{R_0}\right)^{\frac{R_0}{\Delta R}}\right)^{\frac{\Delta A}{R_0}}$$

$$= e^{\frac{\Delta A}{R_0}} \quad \therefore$$

Results of the International Survey on Ratemaking Principles and Methods Used in Other Countries

Gregory S. Wilson, FCAS, MAAA,
on the behalf of the CAS Ratemaking Committe

Results of the International Survey on Ratemaking Principles and Methods Used in Other Countries

Gregory S. Wilson, FCAS, MAAA

The CAS Committee on Ratemaking is responsible for furthering and disseminating ratemaking theory and principles, and identifying topics for research and discussion. The Committee sought information about ratemaking internationally and specifically about actuarial participation in this area. It was felt that there are common concerns and problems with the insurance industry around the world, and that sharing information and jointly collaborating in some areas of research would prove beneficial to all that participate.

In 1996 the Committee, under the guidance of Pamela Sealand Reale, FCAS and Steven M. Visner, FCAS, designed a survey that was sent to the president of the local actuarial society, as indicated in the international directory maintained by the International Actuarial Association. The survey is attached as Exhibit 1. An initial test survey was sent to six countries, four of which responded. After reviewing the initial results, the survey was sent to an additional thirty-five countries. Thirteen countries responded, bringing the total to seventeen. The following table identifies the countries that responded:

	Responses Received	
Test Group		
Australia	Mexio	
China	South Africa	
Full Mailing		
Norway	India	Portugal
Switzerland	Finland	Netherlands
Japan	Czech Republic	Ireland
New Zealand	Austria	Pakistan
United Kingdom		

Much of the information obtained from the survey was presented at the 1997 ratemaking seminar by Pam and Steve. Following is an itemization of the questions and answers from the survey, as well as an overall summary and future plans:

1. How is property and casualty and/or general insurance sold?

Private insurance companies sell Property/Casualty insurance in the majority of countries responding to the survey. Insurance in some instances is sold by the government in the following six countries:

Switzerland	"Fire insurance for buildings is compulsory in some states (cantons)."
New Zealand	Accident compensation and limited coverage for earthquake
Australia	Compulsory third party insurance is sold as a separate contract, usually by a government insurer.
South Africa	Third party liability auto insurance is funded by a fuel levy.
India	All non-life insurance is nationalized.
Pakistan	Risks pertaining to government or government owned organizations must be insured with the National Insurance Corporation, which is owned by the Federal Government of Pakistan.

2. Who determines the property and casualty rates?

Rates are determined by private companies without government approval in most countries. Government approval is generally required for all lines in Japan, India, South Africa and China.

3. What roles do actuaries/non-actuaries play in developing the rates?

This question was a free form response in the original survey. Those responses varied from little actuarial involvement in South Africa to significant involvement in Australia. Many of the companies responding to the full mailing have actuaries and non-actuaries sharing various roles. There was no actuarial involvement in New Zealand, India and Pakistan, where the non-life actuarial profession was either non-existent or in its infancy.

The following comments were received:

Switzerland	"Actuaries are usually not involved in data gathering, but actuaries say which data should be gathered and which data are needed for ratemaking; actuaries are also responsible for design of the statistics to be done for calculation purposes."
New Zealand	Practical actuarial experience in general insurance is limited. Actuaries are involved with the Accident Compensation Organization financial planning, rate setting process. They have also been involved with the Earthquake Commission.
Norway	Every non-life insurance company is required by law to have an appointed actuary. One of the responsibilities of the appointed actuary is to set premiums and thus set the rates.
Japan	These functions are not necessarily done exclusively by admitted actuaries, but those who engage in such operations are practically expected to have some actuarial knowledge. Actuaries are involved with Accident Compensation Organization financial planning, rate setting process. They are also involved with the earthquake commission.
India	The non-life business has been nationalized and is written by the General Insurance Corporation of India through its four subsidiaries. No actuaries work in this company at the present time. Netherlands

"Most actuaries in non-life (including econometricians acting in the non-life actuarial field) have an advisory role within their companies. The final responsibility for the actual commercial rating is taken by management."

4a. Are there any published statements of principles or generally accepted actuarial methods for property and casualty insurance ratemaking?

No country has comprehensive statements of Principles or Generally Accepted Methodologies. Australia has standards that are limited to the New South Wales Compulsory Third Party Liability coverage, Mexico has standards, which apply to personal lines, as well as lines with excess agreements, and Portugal has standards in its Bonus-Malus system.

5. How are the risk classification systems developed? (Different responses for different types of insurance may be appropriate.)

Risk classification systems in most countries are developed by private companies with no governmental approval. Those countries where governmental approval of rates is required also have requirements for governmental approval of classification systems.

6a. Are there any published statements or principles of generally accepted actuarial methods for property and casualty risk classification systems?

There are generally no published statements or principles for property and casualty risk classification systems.

6b. If the answer to 6a. is "Yes;" please identify and attach a copy:

There were no responses to this question.

7. What is the primary source of data used in ratemaking?

In most countries, private companies use their own experience, supplemented by industry data where available. In Japan and New Zealand, governmental statistics are used for some lines.

The following additional comments were received:

Switzerland	Beside industry and company statistics, general economic data is used. Note that there is no access to personal data about policyholders because of legal data protection.
Austria	Industry data is from the Austrian Association of Insurance Companies.
New Zealand	Also use overseas experience for commercial ratemaking. Government statistics are used particularly in health and re accidents.(classified as work-related, motor vehicle related or other) and earthquake losses.
China	Industry data is from the Insurance Institute of the Republic of China.
Mexico	Data used depends on the size of the insurance company. Larger companies use their own experience, smaller companies use the Mexican Association of Insurance Companies (AMIS) data.
India	The Tariff Advisory Committee gathers statistics and advises the nationalized insurer on the rates.
Netherlands	"In foregoing years, ratemaking advice (net risk premium level) was given by CVS (statistical bureau). But now, due to further tariff-segmentation and differentiation more and more large companies do their own research, using mostly their own portfolio data."

Pakistan	For tariff lines, premium and claim data filed with the Insurance Association of Pakistan (IAP) are used. For non-tariff lines, companies rely on their own experience supplemented by guidance from professional reinsurers. Commercial Property: Fire, Auto and Marine use IAP data, while other property rates are made from insurer's own experience.

8. What publications are typically used to disseminate research on ratemaking issues?

There are a variety of publications used in the different countries. Most of the European respondents use the ASTIN Bulletin, along with the Scandinavian Actuarial Journal and bulletins published by their own societies. The majority of the remaining countries rely primarily on bulletins published by their own societies, with the exception of South Africa that relies on overseas publications, if at all.

9. Are you aware of any research on ratemaking issues done in your country that would be of interest to CAS members?

Only six of the countries that answered this question are doing research that they feel would benefit the CAS. The following two were specific about the research being done:

Japan	Methodologies of Earthquake Risk Evaluation Applied to Earthquake Insurance—how to reflect catastrophe losses in ratemaking for insurance policies covering natural disaster such as earthquake, flood, etc.
United Kingdom	"Statistical Motor Rating: Making Effective Use of Your Data," Brockman & Wright, Journal of the Institute of Actuaries, 119, III. Pricing in the London Market, General Insurance Study Group (GISG) 2995 convention.

10. Do actuaries in your country: (a) Utilize research published by the CAS? (b) If so, how and what? (c) Utilize the CAS Ratemaking Principles or the ASB Standards of Practice? (d) If so, how and what?

10a&b. Five countries utilize research published by the CAS and four are not sure. Japan and the Czech Republic use Foundations of Casualty Actuarial Science, Austria and the Czech Republic use the Forum, Finland uses various CAS publications and the United Kingdom uses the Thomas Mack chain ladder method.

10c&d. Only Japan reports that they utilize any CAS principles or ASB Standards of Practice. They use the Statement of Principles Regarding Ratemaking, as well as the ASB Standards of Practice #12 on Risk Classification and #13 on Trending Procedures.

11. Does a rating bureau or similar organization exist in your country? If yes, please describe its function.

Nine countries report having a rating bureau or similar organization. Many of the European respondents report that the influence of the rating bureaus seems to be declining due to recent legislation regarding competition. Following are the specific responses:

Finland	"The central association of the Finnish insurance companies has a bureau called Insurance Data LTD, which basically is a rating bureau. However, recently the role of Insurance Data has diminished. The role of Insurance Data is "strongest" in Workers' Comp."
New Zealand	"Insurance Council of New Zealand does this as one of its services to members."

Norway	"Association of Norwegian Insurance Companies: they represent its members and take care of their interests in relations with the authorities and other institutions. They also gather data on premiums and claims by lines of business and data on causes for fire."
Japan	(1) "Computation of and filing rates with the Minister of Finance. (2) Evaluation of loss experience and risks, and research and study of other matters relevant to rates. (3) Research and study of the policy conditions and various underwriting terms. (4) Research and study of the matters concerning valuation of the objects of insurance. (5) Compilation of insurance statistics. (6) Research and study of the matters concerning damage and its prevention and mitigation."
Australia	"'Insurance Statistics Australia' goes some way toward fulfilling the functions of an independent rating bureau."
Mexico	The Mexican Association of Insurance Companies (AMIS) sets certain procedures or standards to determine rates.
India	"There is a Tariff Advisory Committee who collects the statistical data and advises the General Insurance Corporation of India on the rates for various classes of non-life business."
Netherlands	"In the Netherlands, we have CVS, the Center for Insurance Statistics, which gathers all relevant industry data for most non-life classes of business. With the aid of actuarial committees, they analyze all the data and report about risk premium rating system that are applicable. However, due to European legislation as regards competition, more and more research on rating transfers to companies and the influence of CVS seems to lessen. Notice that non-life insurance in Holland is a totally free market with high level of competition."
Ireland	For Private Auto, statistics are assembled for all insurers (processed by U.K. ABI).

Pakistan	"Insurance Association of Pakistan. All reputable insurance companies (other than National Insurance Corp.) belong to the IAP. The IAP sets minimum tariffs for Fire and Allied perils, Motor, Marine and Workers' Compensation. Member companies must charge at least the minimum tariff. Consequential loss (business interruption), engineering, burglary and some other lines are non-tariff, so companies can set their own rates. All rates (whether tariff or non-tariff) must be filed with government. Prior approval is not needed but government has the right to object."

12. Are there any areas where joint research, related to insurance pricing, would be beneficial between actuaries of your country and the CAS?

The following responses were received:

Switzerland	European and American actuaries could only profit by having closer contacts. Research in actuarial science has a long tradition in Switzerland. Some favorite areas of study are credibility theory, mathematical statistics, extreme value theory, modem financial mathematics, calculation of total claim distribution, concepts on risk adjusted capital.
Austria	Credibility rating in non-life; risk classification in car insurance.
Finland	Personal auto, personal home.
Japan	"Joint research for selecting risk classification factors by different types of insurance policies."
Czech Republic	Mathematical modeling for dynamic solvency testing.
Portugal	Workers' Compensation, Auto Insurance.
Netherlands	"Dutch non-life actuarial profession is relatively very young and small. Most research is done at the University of Amsterdam (Goovuerts, Kaas, Dannenburg) which is reported in Astin Colloquia."
United Kingdom	Rating of Excess of Loss Reinsurance Long tail/Liability business Catastrophes

13a. Describe the process for admission to the actuarial society in your country, along with the ratemaking topics the actuary needs to master for admission.

There is a wide range of requirements for admission to the actuarial society of the various countries, ranging from exams and university training in the Netherlands and United Kingdom to no specific requirements in Portugal. Following are the specific responses by country:

Switzerland	"Ordinary Member: admission by the board of the Swiss Association of Actuaries (SAA), conditions: (i) university degree in actuarial science, OR university degree in mathematics and 3 years of experience in actuarial duties, OR experience as staff-member in actuarial duties (at least 10 years) AND in any case: (ii) recommendation by 2 experienced members of the SAA Actuary SAA: (e.g. "full member" in the sense of the IFAA) admission by the board of the Swiss Association of Actuaries (SAA), conditions: (i) ordinary member of the SAA (ii) examination (in preparation actually)"
Austria	"A full degree in mathematics and an additional education in insurance mathematics OR a full degree in insurance mathematics and three years of practical working experience."
Finland	"First you have to be graduated from a university with main degree in mathematics. Besides that you must have at least one year working experience on actuarial matters. On top of that, qualified actuaries must pass a series of postgraduate examinations (which on the average take 2-3 years to accomplish)."
New Zealand	"FFA, FIA, FSA, FIAA--effectively eligibility to be a Fellow of the NZ Society requires qualification as a Fellow of an overseas actuarial examining body. Most such bodies now include some knowledge of general insurance including ratemaking in their examination systems."

Norway	"The degree of cand. scient. from the University with both life and non-life topics and a thesis relevant to insurance mathematics, is required (corresponds to a Masters degree). A written application describing your background along with a recommendation from 2 members, is also required. Must be elected at an ordinary meeting."
Japan	"Passing the examination of eight subjects (six subjects in basic course and two subjects in advanced course) is the prerequisite to become a fellowship actuary of the IAJ. One of the two subjects in the advanced course deals with practical methods of insurance business operations including ratemaking principles and methods."
India	There are no society actuaries currently involved in ratemaking for general insurance. Hence, it does not appear that there is property/casualty representation in their society.
Portugal	"We launched a program to the formation (?) of non-life actuaries: risk theory, statistics and probability, numerical methods and simulation, regression models, statistical time series, claims reserving, credibility theory, a priori tarification (?), posteriori tarification (?), reinsurance. The program is not yet compulsory but many members are attending it."
Netherlands	"Rules have changed recently. Full membership can only be obtained by actuaries, graduated at University OR finishing a long-lasting special course outside University, combined with a post-university course (2 years)."
Ireland	"Any full member of an actuarial association in any E. C. country who works in Ireland in an actuarial capacity can apply to become a fellow member of the society."
Pakistan	"The Pakistan Society of Actuaries relies on the examination system of the UK Institute of Actuaries and the US Society of Actuaries. The UK Institute has a compulsory exam on general insurance. The SOA's system is well known to the CAS. The PSOA hopes eventually to conduct examinations for its own Fellowship."

United Kingdom	"Members admitted as students only if they have a mathematicaleducation (e.g. mathematics degree). There are then 9 exams to pass. The 3 most relevant exams are subject C (Statistics), G (General Insurance), as well as a Fellowship paper."

13b. Are there non-members of your actuarial society who would also be referred to as actuaries? If so, what qualifications do they have?

Many of the respondents exclusively use the title actuary to mean a member of the society. However, there are several that call a variety of non-members actuaries. These may include students, teachers and economists. Following are the responses received from each country:

Switzerland	"The legal requirement for Pension Schemes Experts is to pass successfully the examination organized by the SAA; this examination is open to any applicant. Pension Scheme Experts (PVE) is a title recognized by federal legislation."
Norway	"The actuarial title is not protected by law in Norway any longer. What we call an actuary is a person with a degree of cand. scient. with what we call an actuarial competence as described under 13a above. Therefore, there are actuaries that are not members of the actuarial society."
Portugal	Usually they are teachers of actuarial science not working outside of universities.
Netherlands	"In non-life, we have a number of econometricians, working in the actuarial field (and even sometimes acting as the certifying actuary), which are affiliate members of the Institute of Actuaries."
Ireland	"The society has associate members who do not engage in actuarial work in Ireland, or who are students of the Institute or Faculty in the U. K."

14. What are currently the most significant ratemaking issues within your country?

Following are the responses:

Switzerland	Generalized linear models; credibility; total claim distribution; extreme value theory; financial mathematics.
Austria	Car insurance
Finland	"With the decreasing significance of the rating bureau, companies must learn to stand on their own feet; ratemaking in motor third-party liability."
New Zealand	Catastrophe; human rights acts issues (discrimination on age/race/disability); distribution.
Norway	Workers' Compensation.
Japan	"Most reasonable ratemaking systems not contradictory to Anti-Monopoly Law under the current rate deregulation environments."
Netherlands	"The trend to more differentiation in several classes of business. The impact of privatization of significant parts of social insurance."
Ireland	Level of competition; introduction of new low cost direct insurers; self-insurance for commercial risks.
Pakistan	Rate adequacy in auto insurance; ratemaking of property and casualty companies is primitive. There is a shortage of actuaries due to complacency of companies.
United Kingdom	Neural network technology Integrating technical rates with market rates, using price elasticity models.

SUMMARY

The results of the survey clearly indicate a significant amount of interest in casualty actuarial topics outside the United States and Canada. While some countries have knowledge of CAS activities, there are a large number of countries where there is a lack of familiarity with the CAS.

There was clearly a great deal of interest throughout the international actuarial community in sharing information with the CAS in the area of ratemaking. The Committee on Ratemaking intends to work within the CAS and with ASTIN to encourage the sharing of ratemaking research internationally.

Using Resampling Techniques to Measure the Effectiveness of Providers in Workers' Compensation Insurance

David B. Speights, Ph.D.,
Terry J. Woodfield, Ph.D., and Deborah Becker

David B. Speights is a research statistician at HNC Insurance Solutions. He holds a doctorate in Biostatistics from the University of California, Los Angeles. His current research is in the area of individual claims reserving, resampling techniques, and neural networks for right-censored data.

Terry J. Woodfield is Chief Statistician at HNC Insurance Solutions. He holds a Ph.D. in statistics from Texas A&M University. His research interests include statistical forecasting, statistical computing, and applications of advanced statistical modeling techniques to problems in insurance, finance, and retail sales.

Deborah Becker is a Product Manager at HNC Insurance Solutions. She has over 20 years of experience in the property/casualty insurance industry. Her research for the past three years has been provider profiling for the workers' compensation industry.

HNC Insurance Solutions is a business unit of HNC Software Inc. and was formed in the 1998 merger of CompReview Inc. and Risk Data Corporation. HNC Insurance Solutions is an insurance information service and statistical research company which creates and markets solutions for the insurance industry.

Abstract

We use resampling techniques to analyze the impact of providers on workers' compensation costs taking into consideration inherent differences in claim populations between providers. Resampling techniques provide a nonparametric determination of a statistic's distribution and a measure of effectiveness that is not sensitive to deviations from the assumptions underlying most parametric statistical procedures. These techniques are applied to a subset of an extensive nationwide database of workers' compensation claims to demonstrate the methods.

1 Introduction

A major cost saving method for the workers' compensation industry is to refer injured workers' to the most cost-effective provider. Often, inefficient providers dramatically increase costs by over treating or performing ineffective treatments which reduce the quality of care, prolong the length of disability, and increase both the potential for litigation and permanent disability. The goal of treatment should be to return the worker to suitable gainful employment as soon as possible, reducing costs and increasing patient satisfaction. It is important to be able to compare medical costs and indemnity costs between providers when determining the cost outcome of a workers' compensation claim.

When comparing providers, considerations must be made for the differences in claim populations represented by each provider. For example, we may be interested in comparing the total claim cost for patients served by providers A and B. If provider A services a large number of severe injuries, and provider B services no severe injuries, then we will most likely conclude that provider B is less expensive than provider A, even if the two providers are equivalent. Therefore, it is difficult to identify the provider with the lowest costs without accounting for inherent differences in characteristics indicative of claim severity. Throughout this paper, we refer to characteristics indicative of claim severity as comorbidity factors. In section 3, we describe techniques to risk adjust the data so that the comparisons are based on "like to like" factors. The type of risk adjustment used in this paper is also known as normalization.

Bootstrap resampling is a relatively new statistical technique that allows for nonparametric or semiparametric estimates of a statistic's distribution. Traditionally, statistical methods sought to determine analytically the distribution of a statistic. For example, the asymptotic and small sample distribution of statistics needed to compare population means or variances is well known. However, these distributional properties are often rooted in unrealistic assumptions about the population. In Section 2 we give a brief introduction to the idea of resampling. The idea is very straightforward and is applicable to a wide variety of situations. In addition, bootstrap techniques allow us to form complicated statistics that would normally have asymptotic and small sample properties that are difficult if not impossible to derive.

We apply the basic resampling concept for comparing the distribution of ultimate claim cost on two populations while adjusting for inherent differences in claim severity. The statistical methods are an extension of the methods presented in Efron & Tibshirani (1986) to deal with normalized populations. We show two different techniques for comparing two populations while adjusting for claim characteristics that are indicative of the severity of a claim. In addition, we form complicated statistics for comparing the two populations that would normally have asymptotic distributions that are difficult to obtain.

In this paper, two examples using data from HNC Insurance Solutions' Provider Compare ® database are given to demonstrate how the methods can be applied to comparing claim costs between providers. First, it has become standard to refer injured workers to a provider network. It makes sense to compare providers not in a network to providers in a network. The outcome of profiling providers in and out of networks is outlined in section 4.1. The second example compares the total claim costs of one provider to a group of providers while accounting for the differences induced by 13 separate claim characteristics. This example can be found in section 4.2.

2 Introduction to Resampling Methods

In this section we give an introduction to resampling methodology. Resampling is a simple technique that was developed to serve two basic purposes. First, resampling provides a departure from the rigid assumptions that underlie many statistical procedures. Like many nonparametric and semi-parametric methods, bootstrap resampling provides a framework that is not constrained by assumptions on the data and error distributions. Second, resampling provides a framework for estimating the distribution of very complex statistics. Many times, a procedure is developed for estimating model parameters, but the distribution of the estimate is either too difficult to derive or requires unrealistic assumptions. Resampling techniques provide a straight-forward method for determining the distribution of any statistic.

Define our data as a sample of size n, $X_1, ..., X_n$, where X can represent a vector or a scalar. Assume the data arises from an unknown distribution function F. Based on the data, it is typically of interest to estimate a population parameter. We can usually denote a population parameter as a function of the distribution function, $\Theta = \Theta(F)$. For example, the population mean is defined as

$$\Theta(F) = \int u dF(u). \tag{2.1}$$

Analogously, we can define a corresponding estimate of that parameter as $\Theta(\hat{F}) = \Theta(X_1, ..., X_n) = \hat{\Theta}$, where $\hat{F}$ represents the empirical distribution function. For an estimate of the population mean, we would use the sample mean

$$\Theta(\hat{F}) = \int u d\hat{F}(u) = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{2.2}$$

Often times, the distribution of $\Theta(\hat{F})$ is difficult if not impossible to obtain. In these situations, we can use repeated samples from the original data set to obtain the distribution of $\Theta(\hat{F})$.

Let $X_1^{(*)}, ..., X_n^{(*)}$ represent a simple random sample taken with replacement from the original data $X_1, ..., X_n$. Using the data set $X_1^{(*)}, ..., X_n^{(*)}$ we can obtain an estimate of the population parameter Θ with $\Theta(X_1^{(*)}, ..., X_n^{(*)})$. The estimate of Θ using this procedure is a single bootstrap estimate and $X_1^{(*)}, ..., X_n^{(*)}$ is known as a bootstrap sample.

In order to obtain an estimate for the distribution of $\Theta(\hat{F})$, we must take repeated bootstrap samples. Denote the k^{th} bootstrap sample by $X_1^{(k)}, ..., X_n^{(k)}$ and the corresponding estimate of the population parameter Θ by $\Theta(\hat{F}^{(k)}) = \hat{\Theta}^{(k)}$. Repeat the procedure and obtain B bootstrap samples. From the B samples we obtain a set of estimates for the parameter Θ, $\{\hat{\Theta}^{(1)}, ..., \hat{\Theta}^{(B)}\}$. The distribution of the parameter estimate $\hat{\Theta}$ can be estimated with

$$\hat{F}_{\hat{\Theta}}(x) = \frac{1}{B}\sum_{k=1}^{B} I(\hat{\Theta}^{(k)} \leq x), \tag{2.3}$$

where $I()$ is the indicator function defined as

$$I(A) = \left\{ \begin{array}{ll} 1 & \text{A is true} \\ 0 & \text{A is false} \end{array} \right\}. \tag{2.4}$$

The estimate for the distribution of $\hat{\Theta}$ in equation 2.3 allows us to obtain the mean and standard deviation of our statistic as well as any other relevant measures. The parameter estimate is often taken to be the mean of this distribution. A p-value for testing the hypothesis $H_0 : \Theta = \Theta_0$ versus a two-sided alternative can be obtained as

$$p = 2\min(\hat{F}_{\hat{\Theta}}(\Theta_0), 1 - \hat{F}_{\hat{\Theta}}(\Theta_0)). \tag{2.5}$$

Another way to test this hypothesis is by constructing a 95% confidence interval

$$[\hat{F}_{\hat{\Theta}}^{-1}(0.025), \hat{F}_{\hat{\Theta}}^{-1}(0.975)]. \tag{2.6}$$

If this interval contains the point Θ_0, we would not reject $H_0 : \Theta = \Theta_0$. The two methods are equivalent if the distribution of $\hat{\Theta}$ is symmetric.

The introduction to bootstrap resampling methods given in this section is not meant to be exhaustive. We are simply providing the foundation of resampling techniques so that we may develop methods for comparing providers while controlling for comorbidity factors. For further reference to this topic, consult Efron & Tibshirani (1986), Efron (1982), and Efron & Tibshirani (1993) . For an insurance application see Derrig, Ostaszewski & Rempala (1998).

3 Resampling Techniques for Comparing Two Populations

In this section, we present two applications of the general bootstrap technique for comparing two populations in the presence of covariates. The methods presented can be used in many analysis situations, but we restrict our attention to comparing the effectiveness of providers in lowering the cost of workers' compensation insurance claims.

Assume there are two distinct populations we are interested in comparing. Let the cost of a claim, C, from the two populations have distribution functions $F_1(c)$ and $F_2(c)$ respectively. Define $Z = 1$ if we are in population one and $Z = 2$ if we are in population two. We can rewrite the distribution of claim costs conditional on Z as $F(c|z = 1) = F_1(c)$ and $F(c|z = 2) = F_2(c)$. Further assume that there is a set of extraneous variables in these populations that influence the ultimate claim cost. Denote this set of claim characteristics by $X = \{X_1, ..., X_n\}$. We present two methods to compare the distribution of claim costs in the two populations while removing the effects of the extraneous variables.

3.1 Method 1: Normalized Comparisons for Two Populations

The first method of comparison assumes very little about the structure of the data, but requires X to consist of categorical variables exclusively. Techniques for normalizing populations are used to account for the differences in the distribution of X. The first step in doing normalized comparisons is to write the distribution of claim costs for population one adjusted for the distribution of X in population two. This distribution is written

$$F^{(2)}(C|Z = 1) = \sum_{i=1}^{R} F(C|Z = 1, X = x_i)P(X = x_i|Z = 2). \tag{3.1}$$

where R is the number of different possible values the vector X can represent. If we have 5 categorical variables with 3 levels each, then there will be $R = 3^5 = 243$ possible combinations. Due to this limitation, the methods in this section are limited to only a few covariates. The superscript (2) is added to the distribution function to indicate that it has been normalized to the distribution of X in population two. The corresponding distribution function for population two is

$$F(C|Z = 2) = \sum_{i=1}^{R} F(C|Z = 2, X = x_i)P(X = x_i|Z = 2). \tag{3.2}$$

3.1.1 Comparing Two Population Means

A simple comparison that can be made on the two populations is a comparison of means. We would compare

$$\mu_1^{(2)} = \int uF^{(2)}(du|Z = 1) \tag{3.3}$$

to

$$\mu_2 = \int uF(du|Z = 2). \tag{3.4}$$

A comparison of the two populations can be made using the bootstrap distribution of the statistic $\Theta(F) = \Theta = \mu_1^{(2)} - \mu_2$.

Assume that we have a sample of data from population one, $(C_1, X_1)_i, i = 1, ..., n_1$, and a sample of data from population two, $(C_2, X_2)_i, i = 1, ..., n_2$. To make the comparison of these populations we would resample from each population B times. Let the k^{th} bootstrap sample be denoted as $(C_1, X_1)_i^{(k)}, i = 1, ..., n_1$ and $(C_2, X_2)_i^{(k)}, i = 1, ..., n_2$. To obtain an estimate of Θ we would need to estimate $F(C|Z = 1, X = x_i)$, $F(C|Z = 2, X = x_i)$, and $P(X = x_i|Z = 2)$. Estimates of each of these functions can be obtained with

$$\hat{F}^{(k)}(u|Z = 1, X = x) = \frac{\sum_{i=1}^{n_1} I(C_{1i}^{(k)} \le u, X_{1i}^{(k)} = x)}{\sum_{i=1}^{n_1} I(X_{1i}^{(k)} = x)}, \tag{3.5}$$

$$\hat{F}^{(k)}(u|Z = 2, X = x) = \frac{\sum_{i=1}^{n_2} I(C_{2i}^{(k)} \le u, X_{2i}^{(k)} = x)}{\sum_{i=1}^{n_2} I(X_{2i}^{(k)} = x)}, \tag{3.6}$$

and

$$\hat{P}^{(k)}(X = x|Z = 2) = \frac{1}{n_2} \sum_{i=1}^{n_2} I(X_{2i}^{(k)} = x). \tag{3.7}$$

For each bootstrap sample, the estimates in equations 3.5, 3.6, and 3.7 can be used to obtain an estimate of $\hat{\Theta}^{(k)}$ with

$$\begin{aligned} \hat{\Theta}^{(k)} \;\; = \;\; & \textstyle\sum_{i=1}^{R} \int u\hat{F}^{(k)}(du|Z = 1, X = x_i)\hat{P}^{(k)}(X = x_i|Z = 2) \\ & - \textstyle\sum_{i=1}^{R} \int u\hat{F}^{(k)}(du|Z = 2, X = x_i)\hat{P}^{(k)}(X = x_i|Z = 2). \end{aligned} \tag{3.8}$$

This equation reduces to

$$\hat{\Theta}^{(k)} = \sum_{j=1}^{R} \bar{C}_{1j} \frac{n_{2j}}{n_2} - \bar{C}_2, \tag{3.9}$$

where $\bar{C}_{1j}$ is the mean cost of sample one for the j^{th} category of the vector X and n_{2j} is the number of observations in sample 2 for the j^{th} category of the vector X.

Using the B bootstrap estimates of $\hat{\Theta}^{(k)}$, we can obtain the distribution of $\hat{\Theta}$ using equation 2.3. In addition, confidence intervals and tests of hypothesis can be constructed using the methods described at the end of Section 2.

3.1.2 Comparing Two Populations Percentiles

Since the bootstrap procedure is very flexible with respect to the form of the statistic, we can estimate the distribution of statistics that may otherwise be very difficult to estimate. One example is found when comparing two population percentiles. For example, we may be interested in comparing normalized distribution functions from equations 3.1 and 3.2 for a percentile p. The statistic for this comparison is

$$\Theta(F) = (F^{(2)})^{-1}(p|Z = 1) - F^{-1}(p|Z = 2) \qquad (3.10)$$

The k^{th} bootstrap estimate of $\Theta = \Theta(F)$ is obtained from the equations

$$\hat{F}^{(2k)}(C|Z = 1) = \sum_{i=1}^{R} \hat{F}^{(k)}(C|Z = 1, X = x_i)\hat{P}^{(k)}(X = x_i|Z = 2), \qquad (3.11)$$

and

$$\hat{F}^{(k)}(C|Z = 2) = \sum_{i=1}^{R} \hat{F}^{(k)}(C|Z = 2, X = x_i)\hat{P}^{(k)}(X = x_i|Z = 2), \qquad (3.12)$$

where the estimates indicated on the right-hand side of equations 3.11 and 3.12 are found from equations 3.5, 3.6, and 3.7. Combining equations 3.11 and 3.12 the k^{th} bootstrap estimate of Θ is

$$\hat{\Theta}^{(k)} = (\hat{F}^{(2)})^{-1}(p|Z = 1) - \hat{F}^{-1}(p|Z = 2). \qquad (3.13)$$

Using the B bootstrap estimates of $\hat{\Theta}^{(k)}$, we can obtain the distribution of $\hat{\Theta}$ using equation 2.3. With this distribution, confidence intervals and tests of hypothesis can be constructed using the methods described at the end of Section 2.

3.2 Method 2: Bootstrapping Linear Regression

In this section, methods for bootstrapping in a linear regression model are used to control for both continuous and categorical variables within the covariate vector X. We only provide an overview of the topic, for a more detailed description see Freedman (1981) and Freedman & Peters (1984). The techniques used in this section assume that the log of the claim cost in the population follows a linear model

$$\log(c) = \alpha + \gamma I(Z = 2) + X'\beta + \epsilon = R'\eta + \epsilon, \tag{3.14}$$

where ϵ is random error term with distribution function F, $R = (1, I(Z = 2), X')'$ is a $p + 2$ dimensional vector of covariates, and $\eta = (\alpha, \gamma, \beta')'$ is a $p + 2$ dimensional vector of parameters.

We can estimate η with the standard least-squares estimate. Since we do not want to disturb the distribution of X in each population, we resample from the set of residuals $e_i = \log(c_i) - R_i'\hat{\eta}, i = 1, ..., n$. Denote the k^{th} bootstrap sample of the residuals with $\{e_1^{(k)}, ..., e_n^{(k)}\}$. The corresponding bootstrap sample of $c_i, i = 1, ..., n$ is found with $\{\exp(R_1'\hat{\eta}+ e_1^{(k)}), ..., \exp(R_n'\hat{\eta}+e_n^{(k)})\}$. Using this setup we allow the distribution of X to remain constant and we reconstruct the bootstrapped values of c_i from the residuals.

With B bootstrap sample obtained from the above procedure, an estimate of η from the k^{th} bootstrap sample is the least squares estimate from the regression of $\{\exp(R_1'\hat{\eta} + e_1^{(k)}), ..., \exp(R_n'\hat{\eta} + e_n^{(k)})\}$ on $\{R_1, ..., R_n\}$. Denote the k^{th} bootstrap estimate of η as $\hat{\eta}^{(k)}$. From the B bootstrap estimates of η we can estimate the distribution of $\hat{\eta}$ using equation 2.3. Confidence intervals and tests of hypothesis can be constructed using the methods described at the end of Section 2.

4 Comparing Providers with Resampling Methods

To demonstrate the use of the preceding resampling methods, we give two examples that analyze subsets of HNC Insurance Solutions' Provider Compare ® Database. The first example compares the quantiles of providers in one network to all other providers while controlling for claim severity. The second example uses bootstrap regression techniques to compare one provider to all other providers while controlling for 13 variables.

4.1 Example 1: Comparing Quantiles

In order to measure the effectiveness of a network of providers, we compare the median, seventy-fifth percentile, and ninety-fifth percentile to all other providers outside of the network. To control for the severity of a claim, a grouping of ICD9 code and NCCI injury type are used. We singled out one network of providers from the rest of the providers and measured the effectiveness of that network in lowering the total cost of a workers compensation claim. We refer to the providers in the network of interest as "Network A" and the remaining providers as "Other Providers"

The distribution of claim costs in the "Other Providers" group in each sample generated in the bootstrap process is normalized to the distribution of claim severity in "Network A" as outlined in section 3.1.2. We computed the median, 75^{th} percentile, and 95^{th} percentile from the distribution determined by each bootstrap sample. Figures 1, 2, and 3 show the bootstrap distributions for the median, 75^{th} percentile, and 95^{th} percentile respectively.

The upper left-hand corners of Figures 1, 2, and 3 show a histogram representing the distribution of the median, 75^{th} percentile, and 95^{th} percentile respectively as calculated from each bootstrap sample of the claim costs in the "Other Providers" group. The upper right-hand corners of each figure demonstrate the same statistic as calculated on the "Network A" providers. The lower-left hand corner of Figures 1, 2, and 3 shows the bootstrap distribution of the difference between the "Network A" group and the "Other Providers" group for the median, 75^{th} percentile, and 95^{th} percentile respectively.

From the graphs shown in Figures 1, 2, and 3, we can conclude that the providers in Network A have significantly lower median claim costs. However, upon closer inspection, the difference of 75^{th} and 95^{th} percentiles are not significant at the 0.05 Type I error rate level. Looking at the distributions, we can see a trend of the two populations towards one another as we approach the 95^{th} percentile. This finding implies that Network A may not be as effective on the more severe claims.

4.2 Example 2: Bootstrap Regression Techniques for Comparing One Provider to All Other Providers

This example is based on specific client feedback about a suspicious provider. We will identify the suspicious provider as Z. Using data from our workers' compensation provider database, lost time claims where provider Z is listed as the primary provider have mean indemnity costs of $10,317. The combined sample without provider Z produces a mean indemnity cost of $7,228. The unadjusted estimate of the increase in indemnity costs associated with provider Z is 100%(10317-7228)/7228=42.7%.

A model for the natural logarithm of total indemnity cost regressed against thirteen claim characteristics and the provider Z dummy variable (1 if provider Z, 0 otherwise) was identified through a standard variable selection/model building process. Predictor variables included body part, nature of accident, cause of accident, industry class code, age, gender, injury type, and a reduction of ICD9 code that indicates body region and injury severity through a ten level variable. We used a bootstrap regression technique as outlined in section 3.2 to compute the distribution of the parameter estimate associated with provider Z. Figure 4 shows a histogram representing the bootstrap distribution of the parameter estimate.

Using the bootstrap hypothesis testing strategy developed in this paper in section 2, we determined that the one-sided p-value for a test of no significant difference between provider Z and all other providers in the database was approximately 0.15. The mean increase in indemnity cost associated with provider Z is 54.5%, and the median increase is 53.8%. The unadjusted statistics suggest that provider Z produces 42.7% higher indemnity outcomes than the remaining providers as a group. With $p=0.15$ we have only marginal evidence of an effect and would not reject our null hypothesis using traditional significance levels of 0.05, and 0.10. Nonetheless, the bootstrap results provide more compelling evidence that provider Z is worth watching compared to standard, unadjusted statistics.

5 Examples Using Simulated Data

Since the methods demonstrated in the last section were applied to proprietary data, we will show additional examples in this section that utilize a randomly generated data set that can be found in Appendix A. The data set was generated from the linear regression equation

$$Y = 5 + 1.0Z + 1.5G + \epsilon, \tag{5.1}$$

where Z is a random deviate from a normal distribution with mean 5 and variance 1 and represents a continuous covariate, G is a Bernoulli random variable with the probability of success equal to 0.5, and ϵ is a random error term with mean 0 and variance 1. The random variables Z, G, and ϵ were generated independently. In addition a categorical variable, C, was generated from a binomial distribution with two trials and a probability of success equal to 0.5. The variable C was generated independently of all variables. A listing of the data can be found in Appendix A. This equation basically represents two straight line equations between Y and Z with a slope of one and an additive error term. The Equation for Group 0 is Y=5+Z and the equation for Group 1 is Y=6.5+Z.

5.1 Example 1: Comparing the Medians of Two Groups

Using the same techniques applied to the data sets in Section 4, we compare the medians of the two groups defined by G. We use C in this example as the normalization variable. We used 500 bootstrap samples to compare the medians of group 0 and group 1. Table 1 represents the sample and bootstrap estimates of the medians normalized for C. In addition, the standard deviation of the bootstrap estimates is also given. Figure 5 shows a histogram of the difference in medians for each bootstrap sample. From the histogram we can conclude that there is a statistically significant difference between Group 0 and Group 1 since no bootstrap sample had a difference less than or equal to zero. This is consistent with the model given by equation 5.1 that was used in generating the sample. The mean difference is 11.576-9.901=1.675 which compares to the difference in means of 1.5 represented in equation 5.1.

Table 1: Results from Simulated Data

Group	Raw Sample Mean	Bootstrap Mean	Bootstrap Std. Dev.
0	9.893	9.901	0.232
1	11.659	11.576	0.137

5.2 Example 2: Bootstrapping Regression

Following the techniques presented in section 3.2, we fit the regression equation

$$E(Y|G,C,Z) = \beta_0 + \beta_1 G + \beta_2 I(C=0) + \beta_3 I(C=1) + \beta_4 Z \tag{5.2}$$

to the simulated data listed in Appendix A. Table 2 shows the estimates from the initial model fit to the whole data set. We took 200 bootstrap samples from the residuals to generate the bootstrap samples as describes in Section 3.2. For consistency, we present a histogram of the bootstrap estimates for β_1 in figure 6, the estimated effect of the binary variable G. From the histogram, we can conclude that there is a statistically significant difference in Y between Group 0 and Group 1 since none of the bootstrap estimates were less than or equal to zero. This histogram is similar to the results presented in Section 4.2 on provider, where G=1 represents a provider with higher cost claims and G=0 refers to a provider with lower cost claims. The mean bootstrap estimate of β_1 is 1.577 and the standard deviation is 0.181. This is in the same range to the least squares estimates that are shown in Table 2. To make stronger statements on how these quantities relate, a simulation study would be needed.

Table 2: Regression Results from Simulated Data

Parameter	Estimate	Std. Err.
β_0	4.984	0.366
β_1	1.625	0.142
β_2	-0.049	0.202
β_3	0.083	0.178
β_4	0.985	0.069

6 Conclusions

We have demonstrated how bootstrap techniques are applicable to comparing providers in workers compensation insurance. The methods outlined in this paper provide a powerful set of tools for assessing the effectiveness of individual providers or a group of providers. The examples presented in section 3 are only a few methods that can be constructed using bootstrap techniques. Since the distribution of any statistic is attainable using resampling methods, it is possible to construct a wide range of meaningful tests about our populations.

From the examples presented, we can determine the effectiveness of subsets of providers while controlling for various comorbidity factors. While the methods presented here are the most effective methods for retrospective study, a word of caution is in order. If we are unable to represent all of the comorbidity variables, it is possible to show a difference between providers that can be explained from unaccounted for comorbidity variables. For example, suppose factor A is present in provider X 20% of the time and factor A is present in provider Y 60% of the time. If the presence of factor A is associated with higher claim costs, then factor A is a comorbidity factor. Even if provider X and Y perform equally, if we do not take factor A into account in the analysis, we will likely show that provider X is less expensive than provider Y. This example shows the importance of considering all comorbidity factors in the analysis.

One way to avoid incorrect assessment in the analysis of providers is to randomly assign claims to providers to create a balance of the claim characteristics between the two samples. This type of study design requires careful planning and execution. Since random assignment is often impractical or costly, the methods in this paper should be used to best account for the differences that may exist.

A Appendix

The following data set was randomly generated using the methods described in section 5.

Y	G	C	Z	Y	G	C	Z	Y	G	C	Z	Y	G	C	Z	Y	G	C	Z
6.58	0	0	2.64	7.65	0	0	3.73	8.10	0	0	5.56	8.25	0	0	3.14	8.39	0	0	4.79
8.61	0	0	2.91	8.63	0	0	2 13	9.33	0	0	5.15	9.38	0	0	4.20	9.41	0	0	4.20
9.42	0	0	4.07	9.42	0	0	4.13	9.48	0	0	5.86	9.56	0	0	4.38	9.58	0	0	4.83
9.68	0	0	4.99	9.79	0	0	4.69	9.89	0	0	4.20	9.94	0	0	5.17	9.95	0	0	5.15
10.04	0	0	4.45	10.10	0	0	4.85	10.21	0	0	5.38	10.46	0	0	6.32	10.70	0	0	6.24
10.84	0	0	5.56	11.06	0	0	6.83	11.21	0	0	6.10	11.42	0	0	6.00	12.10	0	0	5.86
12.13	0	0	7.10	12.98	0	0	7.28	6.30	0	1	3.08	6.33	0	1	3.42	6.38	0	1	3.54
6.41	0	1	3.13	6.77	0	1	3.87	7.93	0	1	4.71	8.08	0	1	3.44	8.32	0	1	5.17
8.34	0	1	5.68	8.59	0	1	4.15	8.72	0	1	3.25	8.89	0	1	4.80	8.91	0	1	3.66
8.93	0	1	5.30	9.04	0	1	3.59	9.05	0	1	4.74	9.17	0	1	4.76	9.22	0	1	5.41
9.47	0	1	4.34	9 50	0	1	5.95	9.52	0	1	5.02	9.53	0	1	5.28	9.68	0	1	5.05
9.73	0	1	5.30	9.75	0	1	4.18	9.82	0	1	5.35	9.86	0	1	4.22	9.94	0	1	5.22
10.03	0	1	5.22	10.16	0	1	6.27	10 18	0	1	5.54	10.22	0	1	4.85	10.28	0	1	5.63
10.40	0	1	5.59	10.46	0	1	5.24	10.46	0	1	5.09	10.63	0	1	5.10	10.65	0	1	5.50
10.70	0	1	5 84	10.70	0	1	3.45	10.72	0	1	4.18	10.79	0	1	5.10	10.83	0	1	5.59
10.95	0	1	6.10	11.02	0	1	5.67	11.08	0	1	5.44	11.21	0	1	5.13	11.26	0	1	4.43
11.29	0	1	6.70	11.38	0	1	4.86	11.61	0	1	6.56	11.70	0	1	5.79	11.73	0	1	6.85
12 00	0	1	4.76	12.28	0	1	5.95	12.31	0	1	5.16	12.45	0	1	6.05	12.78	0	1	6.28
13.00	0	1	6.16	7.69	0	2	3.81	8.22	0	2	3.59	8.52	0	2	3.98	8.53	0	2	3.32
8.63	0	2	4.70	8.67	0	2	3.37	8.71	0	2	4.07	8.90	0	2	5.08	9.05	0	2	4.81
9.14	0	2	3.85	9.38	0	2	3.84	9.81	0	2	5.14	9.88	0	2	4.71	9.90	0	2	4.81
10.17	0	2	4.29	10.18	0	2	5.34	10.22	0	2	4.70	10.30	0	2	5.00	10.66	0	2	5.93
10.69	0	2	6.90	10.74	0	2	5.14	11.79	0	2	5.96	11.79	0	2	4.90	12.72	0	2	6.94
9.00	1	0	4.08	9.86	1	0	3.73	9.97	1	0	4.64	10.10	1	0	3.49	10.44	1	0	3.65
10.47	1	0	4.46	10.49	1	0	5.19	10.62	1	0	4.34	10.89	1	0	5.50	11.32	1	0	4.43
11.66	1	0	4.81	11 68	1	0	4.77	11.97	1	0	5.33	12.21	1	0	5.28	12.30	1	0	5.76
12.31	1	0	5.47	12.49	1	0	6.54	13.05	1	0	6.55	13.51	1	0	5.85	14.55	1	0	5.99
9.03	1	1	3.61	9.25	1	1	2.99	9.39	1	1	4.85	9.47	1	1	3.96	9.52	1	1	4.96
9.99	1	1	2.94	10.44	1	1	3.94	10.58	1	1	3.51	10.74	1	1	5.53	10.75	1	1	4.24
10.80	1	1	5 45	11.06	1	1	4.68	11.09	1	1	5.11	11.22	1	1	3.32	11.26	1	1	5.40
11.26	1	1	4.22	11.29	1	1	3.43	11.33	1	1	4.00	11.59	1	1	4.37	11.67	1	1	5.46
11.68	1	1	4.78	11.75	1	1	5.64	11.97	1	1	5.60	11.99	1	1	5.01	11.99	1	1	3.98
12.05	1	1	5.41	12.08	1	1	6.01	12.12	1	1	5.92	12.13	1	1	6.84	12.42	1	1	5.40
12.50	1	1	5.03	12.50	1	1	6.12	12.62	1	1	6.66	12.76	1	1	5.67	12.83	1	1	4.39
12.95	1	1	6.55	13.03	1	1	4.81	13.24	1	1	6.46	13 53	1	1	7.10	14.20	1	1	7.17
14.39	1	1	6.53	14.46	1	1	7.94	14.67	1	1	5.44	15.38	1	1	6.56	9.78	1	2	3.80
9.85	1	2	5.14	9.93	1	2	4.89	10.23	1	2	4.57	10.57	1	2	4.09	10.79	1	2	4.31
10.79	1	2	4.91	10.96	1	2	3.88	11.05	1	2	6.15	11.13	1	2	5 16	11.19	1	2	4.76
11.32	1	2	5.35	11.43	1	2	4.49	11.78	1	2	4.69	11.98	1	2	4.55	11.99	1	2	5.84
12.13	1	2	5.29	12.15	1	2	4.24	12.25	1	2	5.34	12.44	1	2	5.50	13.70	1	2	5.35

References

Derrig, R., Ostaszewski, K. & Rempala, G. (1998), 'Applications of Resampling Methods in Actuarial Practice', *CAS Forum, Summer*.

Efron, B. (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans*, CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38, Philadelphia: Society for Industrial and Applied Mathematics.

Efron, B. & Tibshirani, R. (1986), 'Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy', *Statistical Science* **1**, 54–77.

Efron, B. & Tibshirani, R. (1993), *An Introduction to the Bootstrap*, Chapman and Hall.

Freedman, D. (1981), 'Bootstrapping Regression Models', *Annals of Statistics* **9**, 1218–1228.

Freedman, D. & Peters, S. (1984), 'Bootstrapping a Regression Equation: Some Empirical Results', *Journal of the American Statistical Association* **79**, 97–106.

Figure 1: Bootstrap distribution of the Median Claim Cost in Network A and in Other Providers and Bootstrap Distribution of the Difference of Medians

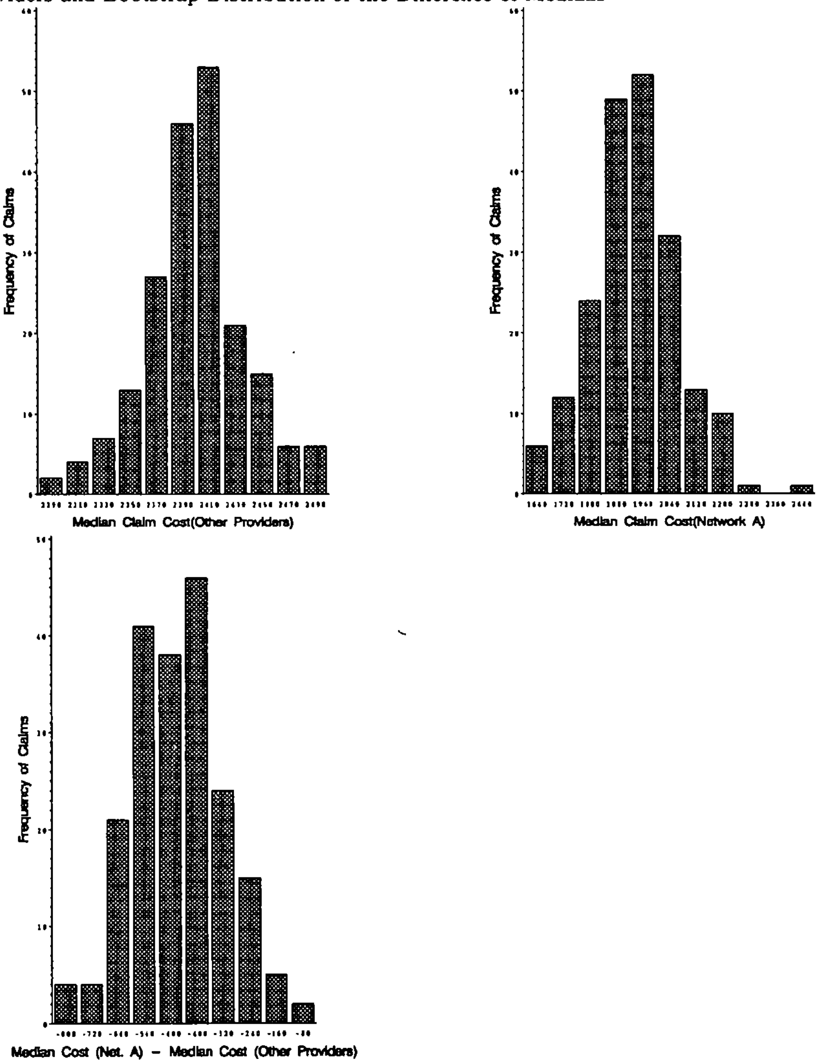


Figure 2: Bootstrap distribution of the 75th Percentile of Claim Costs in Network A and in Other Providers and Bootstrap Distribution of the Difference of 75th Percentiles

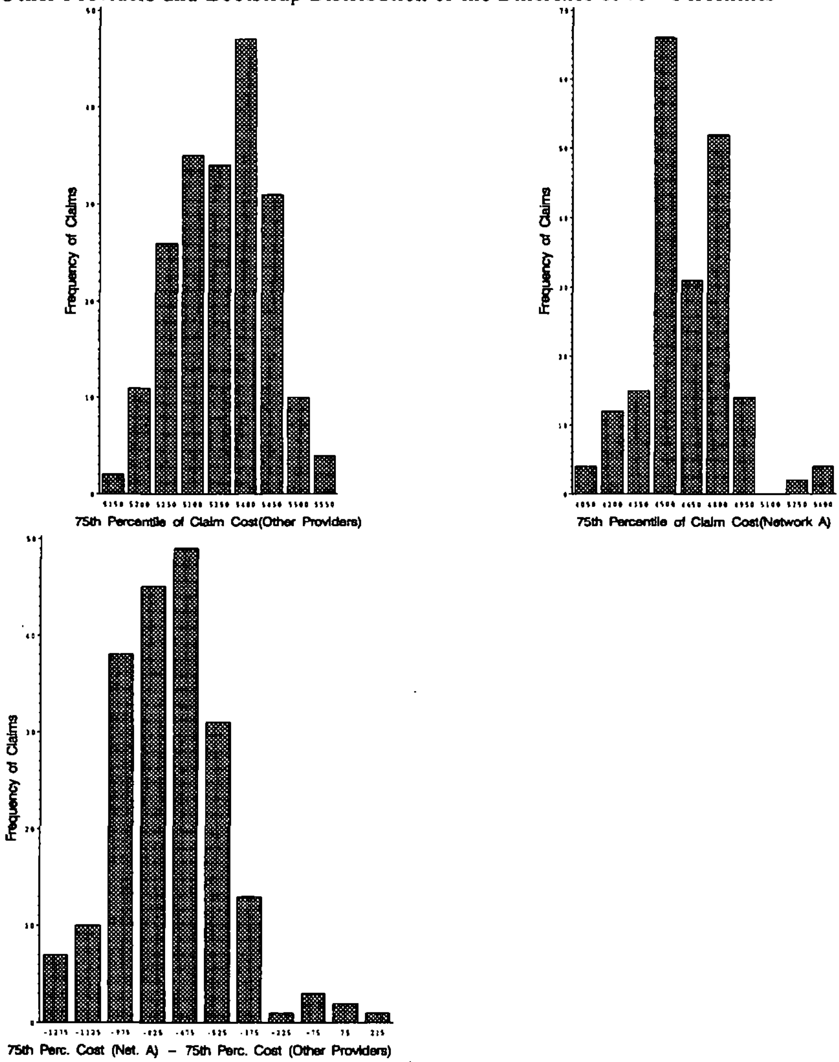


Figure 3: Bootstrap distribution of the 95th Percentile of Claim Costs in Network A and in Other Providers and Bootstrap Distribution of the Difference of 95th Percentiles

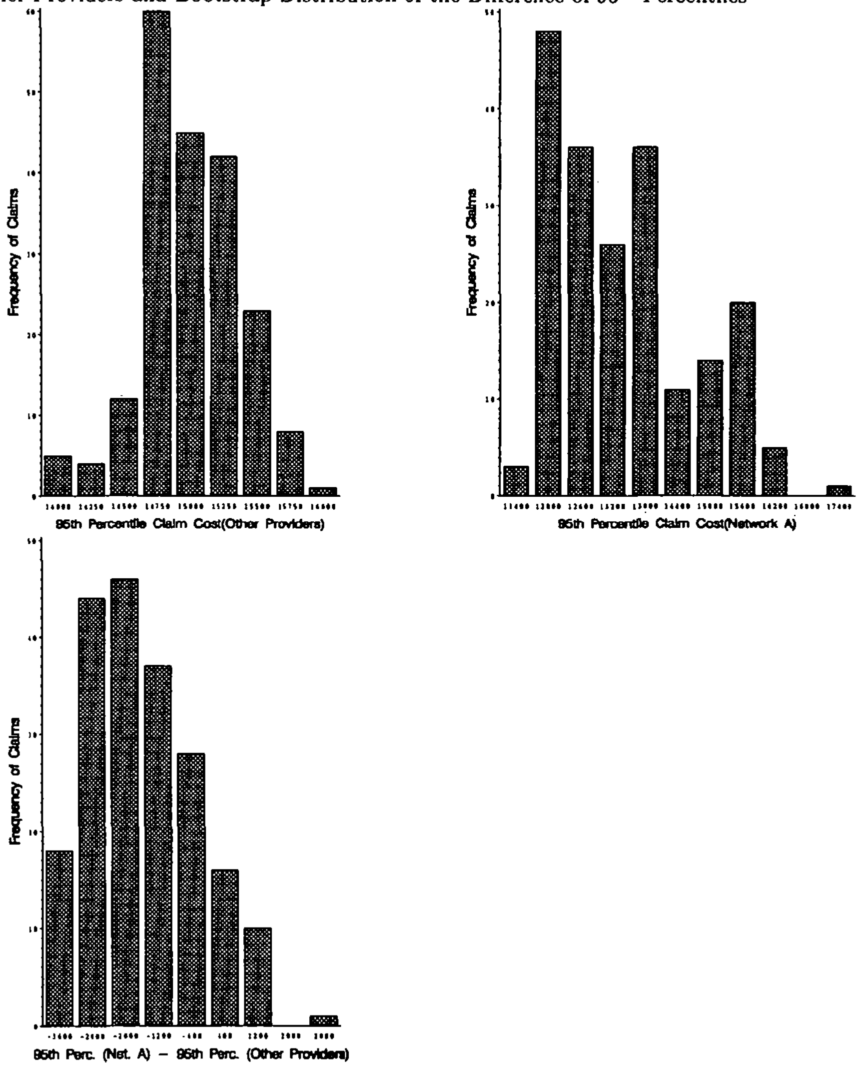


Figure 4: Bootstrap distribution of Regression Coefficient Measuring the Relative Change in Log-Cost for Provider Z versus all other Providers

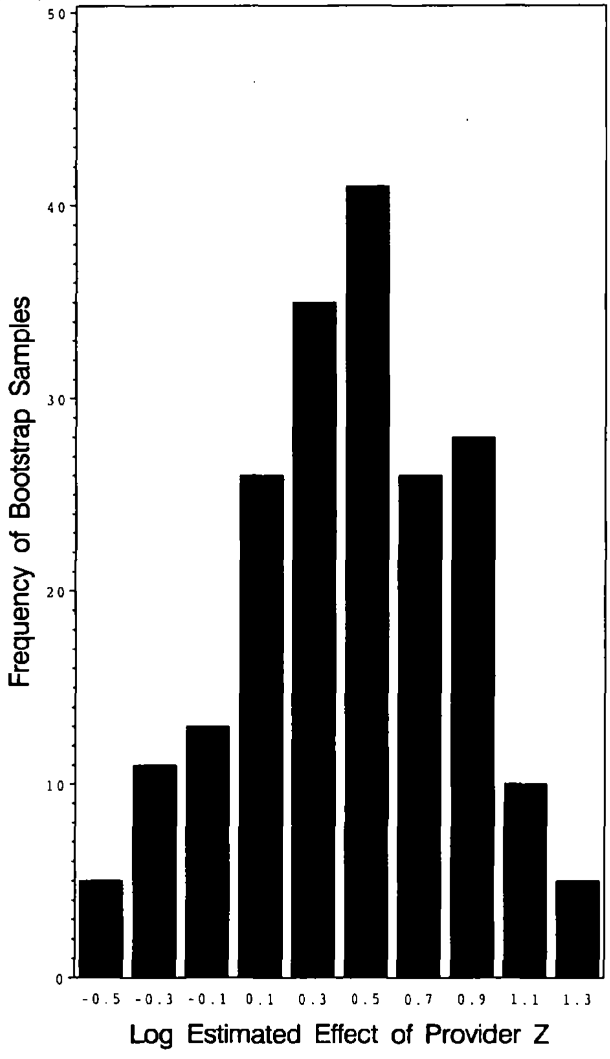


Figure 5: Bootstrap distribution for the Difference of Medians for Group=0 and Group=1

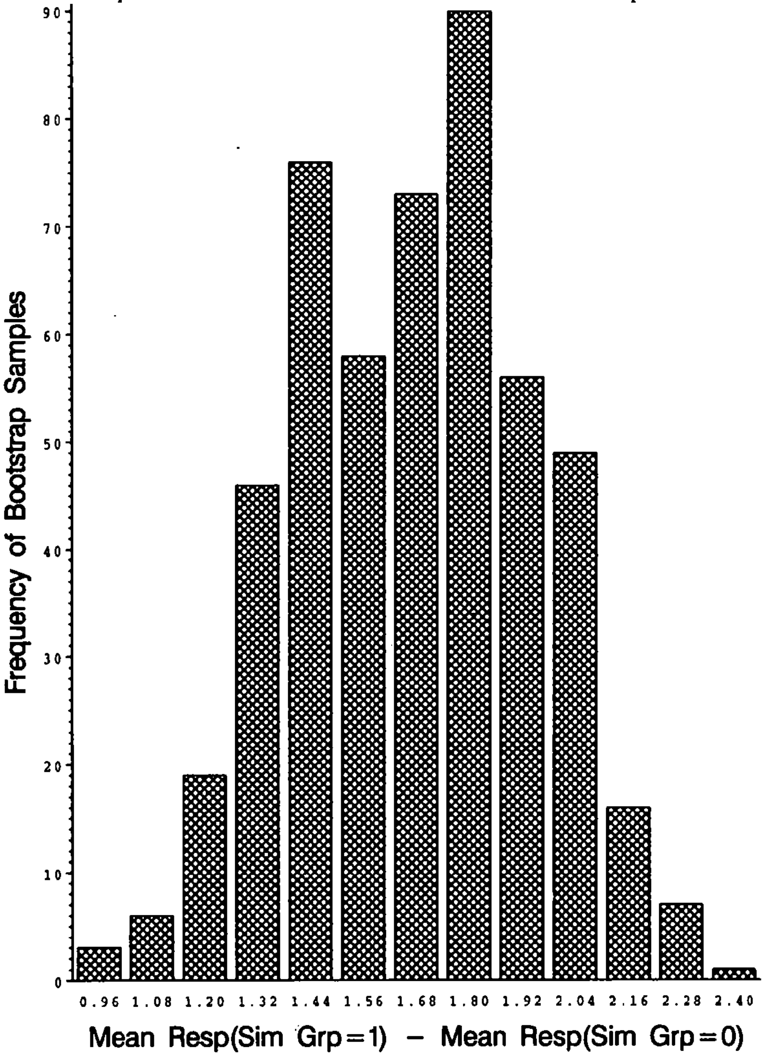
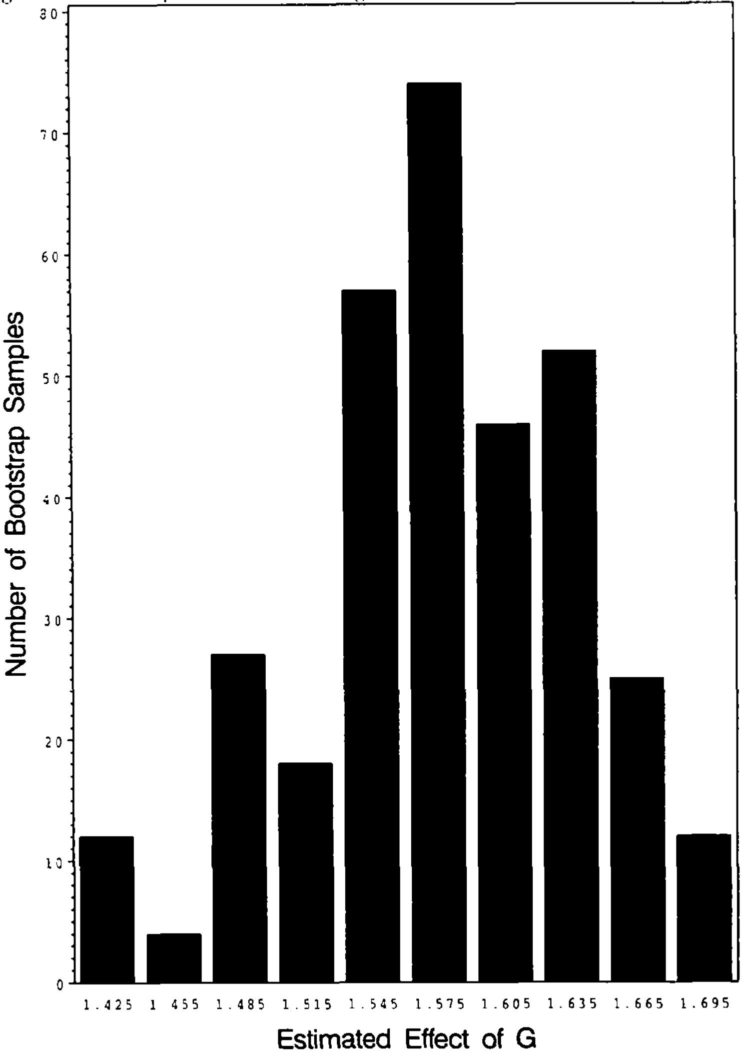


Figure 6: Bootstrap distribution of Regression Coefficient for Simulated Data



The Benefits and Challenges of Profiling Providers in Workers Compensation

Nancy R. Treitel, FCAS, MAAA, ARM, CPU,
Miriam Perkins, ACAS, MAAA, and
Bart Margoshes, M.D.

The Benefits and Challenges of Profiling Providers in Workers Compensation

Abstract

This paper provides a general overview of ways in which provider profiling can be used in developing, maintaining, and evaluating workers compensation managed care programs. It discusses some of the practical challenges that actuaries face in actually developing such profiles and using them. Specifics covered include determining the types of statistics one might want to review, creating the appropriate database needed to do the analysis, and adjusting and segmenting data so that differences in the types of claims providers handle are taken into account. Provider profiling in WC is relatively new to the managed care world. This intent of this paper is to introduce actuaries to provider profiling.

Authors: Nancy R. Treitel, FCAS, Liberty Mutual Group
Miriam Perkins, ACAS, Liberty Mutual Group
Dr. Bart Margoshes, M.D., Liberty Mutual Group

1. Introduction

Due to the significant rise in medical inflation in the late 1980's and early 1990's in workers compensation, cost containment strategies patterned after the group health model were developed[i]. These strategies included utilization management, medical case management, and discounted provider networks[ii]. In theory, providers would be willing to accept a discounted fee for an increase in patient referral volume. Providers were chosen based on anecdotal information from the insurers, a willingness by the provider to treat workers compensation patients, and an agreement to accept a discount. Objective criteria based on treatment outcomes were not part of the selection process. As this model has matured and information systems have become more sophisticated, insurers and managed care companies are beginning to analyze provider performance and outcomes. The result of this capability is that insurers are beginning to select providers who objectively provide consistent high quality, cost effective care and are forming provider networks that are more limited in size but more focused on outcomes.

Profiling providers in the workers compensation system has some significant challenges and differs from the group health system. There are three major areas of divergence. In group health PPO models, a primary care provider is selected by the patient based on the group health contract. This primary care provider acts as a gatekeeper for care. In workers compensation, the selection of the provider is made by the employer (directed), the employee (nondirected), or a combination of the two. The issue of choice is based on the individual state's workers compensation laws. Another area of divergence is the fact that group health outcomes are generally based on discrete, time limited episodes of care. In contrast, in workers compensation, an episode lasts the life of the claim which can be years long. Finally, group health outcomes are limited to medical care only. Since both the medical and indemnity costs are the responsibility of the carrier in workers compensation, a provider's performance (effectiveness) must be based on both these costs.

This paper provides a general overview of ways in which provider profiling can be used in developing, maintaining, and evaluating workers compensation managed care networks. It will then discuss some of the practical challenges that actuaries face in developing these profiles. Specifics covered include determining the types of outcomes one might want to review, creating the appropriate database needed to do the analysis, and adjusting and segmenting data so that differences in the types of claims providers handle are taken into account. Since provider profiling in workers compensation is relatively new to the managed care world, this paper is intended to introduce actuaries to provider profiling and to stimulate discussion as to ways to enhance techniques in place today.

The benefits of provider profiling are plentiful. These include fine tuning provider network composition, providing a feedback tool for quality improvement, training and education, and improving the quality and cost effectiveness of the care delivered to the injured worker. The challenges, however, one faces in building a provider profiling system are substantial.

2. The Concept of Provider Profiling

The basic concept is to use an insurer's existing data repositories to rank individual providers or provider groups relative to one another. This concept is akin to the idea of experience rating, with insurers trying to determine which subset of providers are better than average by reviewing historical experience to the extent credible. Insurers also may look at other risk characteristics (e.g. location of providers to customers, quality of provider's staff) to supplement experience based information similar to what is done in the risk selection process when deciding which employers to insure. Experience rating and risk selection across employers enables insurers to rank employers and enables the insurance company to charge a price commensurate with expected cost. In the case of provider profiling, the process includes:

- selecting and maintaining superior networks and panel lists of providers (analogous to determining the quality of an insurer's book of business);
- rewarding better than average providers (analogous to charging insureds different prices based on experience);
- using the findings in quality assurance and provider education (analogous to risk management and loss prevention);
- using the findings with customers to encourage channeling and to differentiate one managed care program from another; and
- maintaining a competitive edge.

3. Data Considerations

A large practical problem in the process of provider profiling is securing relevant data for the task. In the group health environment, studies done are largely based on clinical information[iii]. Typically, a single type of illness is reviewed (e.g. myocardial infarctions), and clinical data covering single episodes of treatment are gathered. Because medical insurance covers calendar year periods of time and because people may change health insurers and health care providers over time, group health studies find it difficult to track patients and their corresponding medical treatment over long periods of time.

In workers compensation, medical coverage is provided for the life of the injury. Indemnity benefits are also often tied to the duration of the injury. Workers compensation, therefore, has the benefit of having information that tracks results by claimant (e.g. patient) over time, something not available in the group health world. All is not ideal, however. While information is captured over time in workers compensation, it is often difficult to determine exactly when an episode ended. Claim closure date is often used as a proxy but that date depends on more than just the medical condition of a claimant. In addition, clinical data at the provider or medical bill level is not something that has historically been tracked by claimant in workers compensation. In order to profile providers in workers compensation, specialized databases need to be created.

The primary databases available for workers compensation are those developed in conjunction with statistical agents/rating organizations to support the ratemaking process, i.e. those based on unit report data, financial aggregate data, or the sample of detail claim information (a.k.a. DCI data) collected by statistical agents/rating organizations from insurers. In addition, companies typically have reserving databases with workers compensation data used in analyses done to determine ultimate liabilities for the company and estimates of incurred but not reported losses (IBNR).

The above mentioned databases are designed for ratemaking and reserving and do not contain information pertaining to medical case management. Instead, medical management information may reside on a variety of systems not traditionally accessed by actuaries. Two of the first challenges actuaries face when trying to profile providers are 1) developing the list of desired items to measure (see next section of paper) and 2) determining what data exists at the provider level so the list can be narrowed down to those items that can realistically be measured. In some instances, determining what data is available may mean going to the managed care vendor used by the insurance company (e.g. PPO network administrator or medical bill review company). In other cases, a company may already retain the detailed medical information in internal systems. Once finding out the data that is available, the actuary will likely need to request that a special database be built for provider profiling: one that contains provider transaction level detail, as well as medical and indemnity loss outcomes on each claim.

4. Performance Measurements

Given the complexity of the medical encounter in workers compensation, it is a challenge to select those provider performance outcomes that have relevance to quality and cost effectiveness. In order to narrow the list of potential outcomes it is important to determine at a high level what the insurer values as indicating cost effective, quality care. In this paper, we focus on treatment quality, return to work outcomes, medical outcomes and patient satisfaction. Another factor in determining what outcomes to measure is the ability to obtain the data and the data integrity. Since several disparate databases contain the relevant information, this decision process can be formidable.

Treatment Quality

There is no single statistic that perfectly measures the quality of treatment provided by providers. Possible measurements for treatment quality include relapse rates (reopened claims), litigation rates, closure ratios, the percentage of cases referred to nurse case managers, referral rates to a specialist, use of diagnostic studies, average number of visits, use of prescriptions, and degree of documentation of the medical record. A brief discussion of some of these indicators follows.

The provider often plays a pivotal role in determining whether or not a claimant returns to work. High levels of claims reopening may be indicative of claimants returning to work too soon. This is not a desirable outcome. Since the provider is not solely responsible for the return to work outcome, this measure, like most, is not a perfect measure of quality. Additionally, carriers may not always track reopenings in such a way that the information is usable in developing provider profiles.

Managed care programs which encourage providers to get involved early in a claim and to play an active role in facilitating a patient's return to work are thought to have an influence in an injured worker's decision to litigate or not. The theory is that injured workers receiving high quality medical care will be less apt to litigate. Thus high levels of litigation would reflect poorly on a provider while low levels of litigation would reflect well. How highly correlated litigation rates are to provider quality is open to debate. Some may argue that litigation is driven more by employer-employee relations. Even if one reaches agreement that providers play a role in litigation, companies may find it a challenge to accurately measure litigation rate. Insurers often have statistics on their own use of attorneys but information on claimant's defense attorneys is often incomplete.

Higher closure ratios are better than lower closer ratios when profiling providers. High closure ratios indicate earlier return to work and efficient medical care. As long as they are not accompanied by higher relapse rates, high closure rates are considered to be a favorable measure of quality. Of course, in addition to the provider, the injured worker, employer, and insurer play a role in the injured worker's return to work so again, it is not a perfect measure. On the positive side, this statistic is often one that an insurer can easily measure.

Some indicators more consistently represent positive behavior than others. Other indicators may be hard to interpret. For example, high levels of nurse case management on claims handled by a provider may be a good thing or a bad thing depending on the reason for the nurse's involvement. Higher levels of this indicator could mean that the insurer was compelled to bring the nurse in due to the claim not being resolved as anticipated, or in contrast, it could reflect a proactive doctor who involves the nurse by choice to facilitate the injured worker's return to work. Referral to a specialist could be a positive result if it brings experience and knowledge into the process at an early point, or a negative result if the case has deteriorated and higher intensity care is needed. Well documented provider records are an indicator of quality but may present some practical problems in incorporating into a formal ranking procedure.

Return to Work Outcomes

Possible measurements for return to work outcomes include the number of temporary total disability days (TTD days), the ratio of temporary partial disability dollars to temporary total disability dollars, the indemnity cost per claim, the total cost per claim, the percentage of return to modified duty claims, the percent of all claims that are lost time, and the percent of claims where return to work is within published guidelines.

These measures all relate to the success in getting the injured worker back to work. Lower TTD days are desirable. Light duty is encouraged and a higher ratio of temporary partial disability days to temporary total disability is a sign that the provider is making use of light duty programs (again, many others also play a role in returning a worker to light duty). The percentage of claimants returning to modified duty is another measure along these lines. Companies typically track different types of payments being made (e.g. temporary total, temporary partial) and the length of disability for which these payments are being made within their claim systems. However, some programming may be needed to turn the information into usable output. Adjustments may also be needed to account for waiting periods. How to deal with lump sum settlements when trying to estimate the duration of disability is yet another challenge.

Total indemnity costs (possibly excluding fatal and permanent total claims, due to their low frequency and high severity) are a sign of dollars spent but may be heavily influenced by factors other than provider treatment. Assuming each provider has a similar profile of claims, low levels of claims that are lost time may indicate that providers are getting claimants back to work within the waiting period (state statutes typically have a 3 or 7 day waiting period before indemnity benefits get paid). Both these statistics are generally easily accessible within company systems.

Measuring the success of providers in returning injured workers back to work within specified guidelines is desirable but may not be practical. First one needs to have guidelines in place. There are several sets of guidelines commercially available for a price (e.g. Presley Reed, Milliman & Robertson). Then one needs to consistently track how a provider does relative to the guideline on a claim by claim basis. This requires keeping sufficient detail to know that a claim's injury type is consistent with that of the guidelines. The guidelines also often give a range of disability days rather than a single point estimate so one needs to decide how to deal with the ranges.

Medical Outcomes

Possible measurements for medical outcomes include medical cost per lost time claim, medical on medical only claims, the percent of cases referred to physical therapy, and the average number of office visits.

Potential concerns with medical outcome indicators include the fact that high dollars on a medical only claim could underlie a situation where a lost time claim was avoided, perhaps by the physician authorizing an employee's return to work under a modified duty program. The avoidance of an indemnity loss is a positive outcome, yet it would increase the physician's average severity on medical only claims. An offsetting result could be a decrease in claim severity on the indemnity side, if the physician's behavior reduces time lost from work in general. Thus, the measurements are not independent and a positive behavior would not necessarily result in a positive result across all measurements. The calculation of an overall score should take these interdependencies into account, through the weights assigned to the measurements. Also of concern when measuring medical outcomes is the difference in medical treatment costs between communities. A state may need to be subdivided into regions reflecting medical cost differences, and the results normalized across regions before being combined. This type of procedure is often used in pricing health insurance.

Detailed clinical outcomes that are diagnostic specific, such as pain relief after medication administration, are not practical at this time due to data related issues (e.g. much of information needed exists in doctor's reports but is not captured electronically today) but are an area of significant potential.

Satisfaction

Satisfaction can be measured using patient satisfaction surveys, customer (employer) satisfaction surveys, and retention of the patient in the network for the first 30 days. The response rate on patient surveys may be low unless the patient is comfortable that their anonymity will be preserved.

Overall

The measurements can use closed claim data only or data on open as well as closed claims. Measurements involving dollars can use paid or incurred loss amounts. There are benefits and drawbacks to each. Closed claim data is the final result on a claim, and as such is not subject to question or manipulation. However, depending on the length of the experience period used, closed claims may include only non-serious cases. This would mean the exclusion of results on the more difficult claims. Paid loss data does not include the subjective element of case reserves over which the doctor is not in control. However, paid amounts to date may be minor and do not contain information on the best estimate of the ultimate claim amount.

Dealing with outliers also needs to be addressed. Averages can be heavily influenced by a small number of very large claims. In some cases, it may be more appropriate to use medians or to only include data that is within a certain percentile (perhaps from the 10^{th} to 90^{th} percentile) when computing the average.

Once measurements have been selected and the underlying data to be used is determined, weights need to be selected in order to combine the measurements for an overall score. Weights are needed within each of the four areas and across the areas. At the beginning of a provider profiling program, selection of the weights may be heavily judgmental, with the greatest weight assigned to measurements with the most direct link to ultimate costs. For example, under the return to work category, more weight would be given to the median indemnity cost than to the percentage of claims with lost time. In the overall rating, the most weight might be given to medical outcomes and return to work outcomes since these are the ones most directly tied to costs. Over time as experience becomes available, the correlation of the measurements with aggregate ultimate loss cost outcomes should be tested. For example, did a group of doctors that scored well on temporary total disability days also have low average claim severity once the claim matured? Did doctors with a high number of office visits have higher average claim severity? The relative weights assigned to the measurements should be adjusted based on the results. Measurements that do not prove to be good predictors of loss costs should be eliminated.

5. Data Considerations Revisited

Once the statistics to be measured are determined, the actual data needs to be obtained. Provider information and claimant data can typically be linked by a common claim number. Medical bill detail and provider information can typically be linked together by a common provider number which can then be linked to claim level data (bill level data may not always contain claim numbers).

Some preliminary questions still need to be addressed when developing and using the database before provider results can be compared. First, one must decide how to uniquely identify providers or provider groups. This may be a challenge. Names and addresses can be used but if the spelling is not the same from one bill to the next, it may be hard to link all claims together. Using provider tax identification number (TIN) is an option, but there may be multiple providers paid under the same TIN. Depending on company systems, one may find it necessary to manually review provider lists to figure out what approach works best.

Second, one needs to decide how to identify the primary care physician on a claim. A single doctor or clinic needs to be designated as primarily "responsible" for the case's outcome. This is analogous to the gatekeeper in the group health system Should this be a specific type of provider such as an occupational physician. Should this be the doctor with the greatest billed amount? Should it relate to the number of office visits? Should it be the specialist, if the case was referred?

Third, one must decide which doctors or clinics will be among those to be evaluated: those with very few claims will not have credible experience. Should there be a minimum number of claims above which the doctor's experience will be used in full? Given that random fluctuations can affect conclusions, should partial credibility be introduced? Obtaining a sufficient sample size may be a stumbling block even for large writers of workers compensation given the number of providers that exist in a state.

Fourth, one must decide how to deal with claims at different maturities across providers (or how to ensure that the data is all at a similar age). At a minimum, a check should be done to see if there is wide variation in the age of claims. If there is, one can limit the data used to common accident periods and valuation periods. Alternatively, one can look at individual claims a specific number of days after the claim has occurred (similar to data reporting for NCCI's detail claim information). Development factors might also be used though their use is not without problems. The volume of data available may play a role in deciding how best to address this issue if it exists.

6. Adjusting for Case Mix Differences

An important consideration when comparing average costs between doctors is the type of claims being handled. Case mix differences will be present due to physician specialties, the types of occupational injuries that predominate in an area, and random variation in the severity of injuries being treated. Many of the performance measurements will be distorted unless an adjustment is made for case mix. The provider's results must be normalized for case mix differences. This is necessary even if we are only profiling a subset of providers (e.g. physical therapists, occupational health clinics).

The process to normalize the data begins with defining the injury categories. This can be done using any number of methods and degree of detail. Using body part and injury type codes is one way to categorize claims. Other options include using broader injury groupings or using ICD9 codes.

The relative severity of a body part/injury type (or other chosen segmentation) then needs to be determined. One possible methodology is the Bailey Simon method[iv], an interative technique that determines the proper relativities when a general category can be grouped in more than one manner. This method has been used in auto ratemaking to produce the proper relativities by territory and class of driver. In provider profiling we may use it to do the same for body part and injury type, allowing us to assign high, medium, or low severity to body part/injury type combinations. Some body part/injury type combinations are extremely severe and unusual, and as "outliers" are unsuitable to include in the measurement of provider results. These should be excluded from the calculation of the severity index.

Once injury categories have been determined, the data can be normalized for case mix differences. One may determine if an individual provider's case mix is significantly different from the overall population using the chi square test, which tests the hypothesis that the sample has the same probability density function as the overall population. In health insurance, adjustments are made for case mix (differing levels of patient risk) using one of two techniques. In the first, indirect standardization, expected results are determined from the overall population and then applied to the provider's case mix. These expected results using the provider case mix are then compared to actual results. In the second, direct standardization, provider results are applied to a standard case mix. The latter approach produces results that best compare relative performance. Adjusting the provider profiling data for case mix differences removes the penalty that would otherwise be given to a provider who sees a disproportionate number of high severity injury types. Appendix A contains a simplified example of a normalization technique.

7. Ranking Providers

The actual ranking of providers is the easy part of provider profiling, once one has gotten the required data and computed the desired statistics with adjustments for case mix as needed. Provider performance relative to their peer group may be determined by seeing where the provider falls in the distribution of results. This can be done for individual statistics with the four categories discussed above. If scores are assigned to each statistic and weights given to each, the ranking can also be done by category (quality, return to work outcomes, medical outcomes, and satisfaction), or it can be done overall. One may determine if an individual provider's results are statistically different from the overall population using the t test, which tests the hypothesis that the sample has the same mean as the overall population. The number of standard deviations from the mean measures the degree of departure in the results. Using the standard deviation also allows for consistency of results from one measurement period to the next.

Once providers have been ranked, we can use the information in a variety of ways:

- selecting and maintaining superior networks and panel lists
- rewarding better than average providers
- use in quality assurance and provider education
- use with customers to encourage channeling and to differentiate one managed care program from another
- maintaining a competitive edge

8. Summary

Workers compensation managed care programs have grown extensively in the 1990s. The use of preferred provider networks is now commonplace as a cost containment strategy. How effective are these networks and which doctors deliver the best outcomes are often asked questions. Provider profiling can be used to answer these important questions and to give companies a competitive edge. Techniques to profile providers in workers compensation are in their infancy. As highlighted in this paper, the benefits of provider profiling are plentiful, but there are many challenges to overcome before implementing a provider profiling system.

Appendix A

Normalization Technique
Adjusting for Case Mix Differences

Data as Reported:

	Simple Claims		**Complex Claims**		**All Claims**	
Provider	Number	Avg $	Number	Avg $	Number	Avg $
A	90	600	10	6000	100	1140
B	70	500	30	5000	100	1850
Total	160	556	40	5250	200	1495
% Total	**80%**		**20%**		**100%**	

If we look at the average cost per case for All Claims in the above chart, provider A appears to outperform provider B since provider A's average cost is $1,140 versus provider B's average cost of $1,850. However, when we look by type of claim, provider A's average costs are actually higher than provider B's average costs. The results for All Claims reflects differences in the mix of simple and complex claims between providers.

How should we adjust overall results to remove distortions due to differences in case mix? One method is to recompute the averages using the distribution of claims for all providers combined. This is done in the chart below.

Information Normalized to Reflect the Same Mix of Claims for All Providers:

	Simple Claims		**Complex Claims**		**All Claims**	
Provider	%	Avg $	%	Avg $	%	Avg $
A	80%	600	20%	6000	100%	1680
B	80%	500	20%	5000	100%	1400
% Total	**80%**	**556**	**20%**	**5250**	**100%**	**1495**

After adjusting for case mix, provider A looks worse than provider B in all cases.

[i] Vincenzino, JV., "Health care cost: Market forces and reform." Statistical Bulletin Metropolitan Insurance Company 76(1):29-35. 1995.

[ii] Margoshes, Bart, " Disability Management and Occupational Health", OCCUPATIONAL MEDICINE: State of the Art Reviews, MANAGED CARE, Volume 13, Number 4, October-December, 1998, pp. 693-704.

[iii] Iezzoni, Lisa I., "Risk Adjustment for Measuring Healthcare Outcomes," Second Edition, Health Administrator Press, 1997.

[iv] Bailey, Robert A., "Insurance Rates with Minimum Bias," PCAS L, 1963, pp. 4-14.

A Note on Decomposing the Difference of Two Ratios

Daniel R. Corro

A Note on Decomposing the Difference of Two Ratios

Dan Corro
Director of Claims Research
National Council on Compensation Insurance, Inc.

August, 1999

Abstract

Consider a ratio statistic (e.g. the mean) built from observations assigned into classes. An example would be losses=L, claim counts=C, and exposures=E each aggregated by rating class with the applicable statistic being either case severity=L/C or case frequency=C/E. The note discusses comparing two observed values for such a statistic. The difference is expressed as a sum of two components. One component measures the change due to the change in class mix. The other measures the change "holding the class mix constant". It is shown that a T-test on each component can assess whether it represents a nonzero difference. A simple numeric example is presented and an Appendix provides a SAS routine to perform the calculations.

Introduction

One of the ongoing assignments of the claims research department at the National Council on Compensation Insurance, Inc. (NCCI) is to monitor the experience of states which have introduced reforms in their workers compensation (WC) systems. These state specific post reform monitor (PRM) reports analyze WC costs by breaking them down into their frequency and severity components. That experience is compared with a benchmark determined from the pooled experience of a collection of other states, selected from among states that did not undergo major changes in their WC systems over the time frame of the study. The time frame of the study is, in turn, broken down into two subsets, corresponding to the pre- and post-reform experience of the PRM state. This approach leads to a number of two-way comparisons: PRM state vs. benchmark states; pre- vs. post-reform time period.

A primary PRM data source is the NCCI unit record system (URS), used for WC ratemaking and experience rating. A key feature of URS data is its capture according to the job classification system used for pricing WC insurance. When comparing frequency and severity, over time or between jurisdictions, it is obviously important to be able to account for differences in exposure mix. This short note describes the technique developed specifically for the PRMs but which clearly has general application.

The Decomposition

The idea comes from simple arithmetic. Consider any paired comparison, indexed by $j \in \{1,2\}$, in which a "numerator" $N_j = \sum_i n_{ji}$ and "denominator" $D_j = \sum_i d_{ji}$ are determined by summing over a common set of disjoint classes, indexed by i.

The difference of the ratios $\rho_j = N_j / D_j$ can be decomposed as:

$$\rho_2 - \rho_1 = \alpha + \beta \quad \textit{where, letting } r_{ji} = n_{ji} / d_{ji},$$

$$\alpha = \sum_i r_{1i}\left(\frac{d_{2i}}{D_2} - \frac{d_{1i}}{D_1}\right) \quad and \quad \beta = \sum_i (r_{2i} - r_{1i})\left(\frac{d_{2i}}{D_2}\right).$$

The component α is referred to as the *class mix component* and β, which is a weighted sum of ratio differences by class, as the *matched difference component*. Observe that when the two denominators share a common class decomposition, i.e. $d_{2i} = a \cdot d_{1i}$ for some fixed number a, then $\alpha = 0$. By the same token, if the two ratios are the same for all classes ($r_{2i} = r_{1i}$ for all i) then $\beta = 0$.

Observe that the matched difference component

$$\beta = \sum_i \left(\frac{n_{2i}}{d_{2i}}\right) \cdot \left(\frac{d_{2i}}{D_2}\right) - \sum_i r_{1i} \cdot \left(\frac{d_{2i}}{D_2}\right) = \rho_2 - \sum_i r_{1i}\left(\frac{d_{2i}}{D_2}\right) = \rho_2 - \hat{\rho}_1$$

may be interpreted as the difference between ρ_2 and the result of reweighting the observed ratios $\{r_{1j}\}$, which yielded the first ratio ρ_1, to match ρ_2's denominator distribution $\{d_{2i}\}$.

Of course, the ratios ρ_1 and ρ_2 can be regarded as weighted means. Indeed, we will regard the d_{ji} both as individual observed "denominators" as well as weights. It is natural to consider testing whether the difference of means $\rho_2 - \rho_1 = \alpha + \beta$ is significantly different from 0. The usual test for this is the conventional T-test of mean difference. In its customary formulation, however, that test is not suited for weighted observations. For example, the SAS TTEST procedure does not support a weight variable, even though the SAS package is most accommodating of weighted data. It is well known, however, that the customary T-test of mean difference is a special case of the OLS regression calculation. Indeed, the coefficient parameters are routinely tested for significance using a T-test. The Appendix illustrates, using SAS, a simple and general way to test whether the difference $\rho_2 - \rho_1$ of weighted means is significantly different from 0, via weighted OLS.

In this regard, consider the set of matched pairs $\{(r_{1i}, r_{2i}) \mid 1 \le i \le n\}$ in which the paired observations are assigned weights according to the second denominator distribution $\{d_{2i}\}$. Those familiar with what the SAS documentation refers to as a "matched T-test" to determine whether the ratios are different, will note that it is in fact the matched difference component $\beta = \rho_2 - \hat{\rho}_1$ that is being tested. To see this, first recall that, unlike the conventional T-test of mean difference in PROC TTEST, SAS recommends the use of PROC MEANS to perform a matched T-test. The SAS PROC MEANS directly accommodates weighted data (although one has to choose a weight, here we chose the $\{d_{2i}\}$). The idea is to consider the set of matched differences $\{x_i = r_{2i} - r_{1i} \mid 1 \le i \le n\}$ to determine whether its mean is different from 0. The value being tested is thus the weighted mean:

$$\sum_i (x_i) \cdot \left(\frac{d_{2i}}{D_2}\right) = \sum_i (r_{2i} - r_{1i}) \cdot \left(\frac{d_{2i}}{D_2}\right) = \beta \text{, as claimed.}$$

A more formal statement of our result (which allows for nonnegative weights, rather than the strictly positive weights demanded by the $r_{ji} = {}^{n_{ji}}\!/_{d_{ji}}$ formulation) is provided below:

***Proposition**: Given any ordered set of 2N nonnegative real numbers:*

$$\{r_{ji}, d_{ji} \mid j = 1,2; 1 \le i \le N\}$$

Set

$$D_j = \sum_i d_{ji},\ R_j = \sum_i r_{ji} \quad \text{and} \quad \rho_j = \sum r_{ji}\left(\frac{d_{ji}}{D_j}\right)$$

and assume $D_j > 0$ *and* $R_j > 0,\ j = 1,2.$

Define

$$\alpha = \sum_i r_{1i}\left({d_{2i}}\big/{D_2} - {d_{1i}}\big/{D_1}\right) \quad and \quad \beta = \sum_i (r_{2i} - r_{1i})\left({d_{2i}}\big/{D_2}\right).$$

α is referred to as the **class mix component** *and β as the* **matched difference component.**

Then:

(i) $\rho_2 - \rho_1 = \alpha + \beta$.

(ii) *An appropriate test of the hypothesis $\alpha = 0$ is a matched T–test on the pairs* $\{(\frac{d_{1i}}{D_1}, \frac{d_{2i}}{D_2}) \mid 1 \le i \le N\}$ *using $\{r_{1i}\}$ as weights.*

(iii) *An appropriate test of the hypothesis $\beta = 0$ is a matched T–test on the pairs* $\{(r_{1i}, r_{2i}) \mid 1 \le i \le N\}$ *using $\{d_{2i}\}$ as weights.*

Proof: Everything follows directly from earlier remarks except (ii) on testing the hypothesis $\alpha = 0$. In that regard, set

$$\hat{r}_{2i} = d_{2i}, \hat{r}_{1i} = d_{1i}, \hat{d}_{2i} = r_{1i} \; and \; \hat{d}_{1i} = r_{2i}$$

and apply the established part of the proposition to the hatted numbers, noting that:

$$\alpha = R_1\hat{\beta}.$$

Since a T-test is unaffected by multiplication by a positive constant, the result follows from (iii) as applied to $\hat{\beta}$. This establishes the proposition.

An (unmatched) T-test for the difference of means $\rho_2 - \rho_1$ involves $2N - 2 = 2(N - 1)$ degrees of freedom, which the proposition suggests can be split evenly between the two matched T-tests for α and β, each involving $N - 1$ degrees of freedom.

A Numeric Example

This note concludes with a simple numeric example, designed to illustrate the calculation as well as the need to account for class mix when making comparisons. Think of the r-values as a cost measure (frequency or severity) and the d-values as the exposure base (payrolls or cases). The Group 1 data is meant to suggest a starting point situation in which much of the exposure lies in high cost classes (I and J). This changes into the Group 2 situation with most of the exposure assumed to move into the lower cost classes (A, B and C). The cost within class is fairly similar between the two groups but note that for every class, the Group 2 cost equals or exceeds that for Group1. The Appendix provides a SASLOG and listing of the routine used to make the calculations.

Data Table				
	Group 1		Group 2	
Class	r_1	d_1	r_2	d_2
A	95	2	98	30
B	100	2	100	30
C	105	1	106	20
D	295	5	298	2
E	300	10	300	5
F	305	10	308	2
G	310	10	310	2
H	495	10	500	1
I	500	25	505	5
J	505	25	505	3

Decomposition of Ratio Difference				
Component	Value	T-Test	T-Value	P-Value
α	-254.15	Matched	-1.9456	0.0836
β	1.52	Matched	2.9135	0.0172
ρ_2-ρ_1	159.32 - 411.95 = - 252.63	Unmatched	-4.409	0.0003

All of the decline in overall mean cost from Group 1 to Group 2 is attributed to the change in class mix component α. The matched difference component β works in the opposite direction, due to the higher class costs for Group 2. Observe that the overall difference is the most statistically significant finding, as measured by the lowest P-value. It is interesting to note that the dominating component numerically, the change in the class mix, is of marginal statistical significance. On the other hand, the numerically smaller matched difference component reflects a statistically significant increase in the by class costs.

APPENDIX

SASLOG

```
*****************************************************************;
 373         DATA ONE;
 374         INPUT  R1 D1 R2 D2;
 375         CARDS;

 NOTE: The data set WORK.ONE has 10 observations and 4 variables.
 NOTE: The DATA statement used 3248K.

 386         ;
 387         PROC PRINT   DATA=ONE;

 NOTE: The PROCEDURE PRINT printed page 1.
 NOTE: The PROCEDURE PRINT used 3381K.

 388         PROC SUMMARY DATA=ONE;
 389         VAR D1 D2 R1 R2; ;
 390         OUTPUT OUT = SUMM SUM = SD1 SD2 SR1 SR2;
 391         *DEFINE DIFFERENCES;

 NOTE: The data set WORK.SUMM has 1 observations and 6 variables.
 NOTE: The PROCEDURE SUMMARY used 3521K.

 392         DATA ONE;
 393         SET ONE;
 394         KEEP R1 R2 D1 D2 R2_R1 D2_D1;
 395         RETAIN SD1 SD2 SR1 SR2;
 396         IF _N_ = 1 THEN SET SUMM;
 397         D2_D1 = (D2/SD2 - D1/SD1)*SR1;
 398         R2_R1 = R2 - R1;

 NOTE: The data set WORK.ONE has 10 observations and 6 variables.
 NOTE: The DATA statement used 3558K.

 399         PROC MEANS DATA=ONE;
 400         VAR R1;
 401         WEIGHT D1;
 402         TITLE2 'WEIGHTED MEAN OF R1, WEIGHT D1';

 NOTE: The PROCEDURE MEANS printed page 2.
 NOTE: The PROCEDURE MEANS used 3572K.

 403         PROC MEANS DATA=ONE MEAN STDERR T PRT;
 404         VAR R2 R2_R1;
 405         WEIGHT D2;
 406         TITLE2 'WEIGHTED MEAN OF R2 AND MATCHED T-TEST,WEIGHT D2';

 NOTE: The standard error of the mean is computed as sqrt( weighted
sample variance / sum of weights ).
 NOTE: The PROCEDURE MEANS printed page 3.
 NOTE: The PROCEDURE MEANS used 3572K.

 407         PROC MEANS DATA=ONE MEAN STDERR T PRT;
```

```
        VAR D2_D1;
        WEIGHT R1;
        TITLE2 '  MATCHED WEIGHTED T-TEST D2 - D1 MEANS WEIGHT R1';
13 The SAS System
11:51 Monday, August 23, 1999

        *SET UP TO DO WEIGHTED TTEST OF R1-R2 USING OLS';

NOTE: The standard error of the mean is computed as sqrt( weighted
sample variance / sum of weights ).
NOTE: The PROCEDURE MEANS printed page 4.
NOTE: The PROCEDURE MEANS used 3572K.

        DATA TWO;SET ONE;
        KEEP R D C;
        R = R1;D = D1;C = 0;OUTPUT;
        R = R2;D = D2;C = 1;OUTPUT;

NOTE: The data set WORK.TWO has 20 observations and 3 variables.
NOTE: The DATA statement used 3572K.

        PROC REG DATA=TWO;
        MODEL R = C;
        WEIGHT D;
        TITLE2 'UNMATCHED WEIGHTED T-TEST USING OLS';
NOTE: 20 observations read.
NOTE: 20 observations used in computations.
NOTE: The PROCEDURE REG printed page 5.
NOTE: The PROCEDURE REG used 4071K.

NOTE: The SAS session used 4071K.
NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
```

OUTPUT LISTING

```
Page 1:
OBS      R1     D1      R2     D2
  1       95     2       98    30
  2      100     2      100    30
  3      105     1      106    20
  4      295     5      298     2
  5      300    10      300     5
  6      305    10      308     2
  7      310    10      310     2
  8      495    10      500     1
  9      500    25      505     5
 10      505    25      505     3

Page 2: WEIGHTED MEAN OF R1, WEIGHT D1
 Analysis Variable : R1
  N          Mean          Std Dev        Minimum        Maximum
 -----------------------------------------------------------------
 10   411.9500000   391.7262011    95.0000000   505.0000000
 -----------------------------------------------------------------

Page 3: WEIGHTED MEAN OF R2 AND MATCHED T-TEST,WEIGHT D2
 Variable            Mean      Std Error              T   Prob>³T³
 -------------------------------------------------------------------
 R2          159.3200000    41.8235681      3.8093354    0.0042
 R2_R1         1.5200000     0.5217066      2.9135150    0.0172
 -------------------------------------------------------------------

Page 4: MATCHED WEIGHTED T-TEST D2 - D1 MEANS WEIGHT R1
 Analysis Variable : D2_D1
        Mean        Std Error              T    Prob>³T³
 ---------------------------------------------------------
 -254.1500000   130.6297110    -1.9455758    0.0836
 ---------------------------------------------------------

Page 5:UNMATCHED WEIGHTED T-TEST USING OLS
 Model: MODEL1
 Dependent Variable: R
 Analysis of Variance
                               Sum of           Mean
 Source             DF        Squares         Square        F Value          Prob>F
 Model               1    3191095.845    3191095.845         19.436          0.0003
 Error              18     2955334.51   164185.25056
 C Total            19    6146430.355

     Root MSE         405.19779        R-square          0.5192
     Dep Mean         285.63500        Adj R-sq          0.4925
     C.V.             141.85859

 Parameter Estimates
                        Parameter        Standard      T for H0:
 Variable  DF           Estimate           Error     Parameter=0      Prob > ³T³
 INTERCEP   1         411.950000     40.51977919          10.167          0.0001
 C          1        -252.630000     57.30362127          -4.409          0.0003
```

Dynamic Financial Models of Property-Casualty Insurers

The CAS Committee on
Dynamic Financial Analysis

DYNAMIC FINANCIAL MODELS OF PROPERTY-CASUALTY INSURERS

Prepared by
the Dynamic Financial Analysis Committee
of the Casualty Actuarial Society

Charles C. Emma, Chairman

Donald F. Behan
Roger W. Bovard
Richard Derrig
Owen M. Gleeson
Philip E. Heckman
Gerald S. Kirschner
Pierre Lepage
Eduardo P. Marchena
Glenn G. Meyers
Raymond S. Nichols
Marc B. Pearl
Mark R. Shapland
William R. Van Ark
Thomas V. Warthen
Peter G. Wick

January, 2000

TABLE OF CONTENTS

Chapter

Appendices

CHAPTER 1[1]
PURPOSE

The purpose of this paper is to discuss and provide guidance on the important issues and considerations that confront actuaries when designing, building or selecting dynamic financial models of property-casualty risks. The Casualty Actuarial Society's Dynamic Financial Analysis Committee has prepared it as a part of the Society's ongoing educational efforts on issues affecting actuaries responsible for the strategic and dynamic financial analysis of insurers.[2]

This paper should not be interpreted as placing requirements on actuaries or the models used by actuaries. Such requirements have been and will continue to be addressed by the Actuarial Standards Board.

---

1 This document is an update to "Dynamic Financial Models of Property-Casualty Insurers" prepared by the Subcommittee on Dynamic Financial Models of the Casualty Actuarial Society's Valuation and Financial Analysis Committee on September 22, 1995.

2 Other sources of information regarding dynamic financial models are included in Appendix A.

CHAPTER 2
INTRODUCTION AND BACKGROUND

What is Dynamic Financial Analysis?

One of the early works related to dynamic financial analysis comes from Jay W. Forrester in Industrial Dynamics. He defines it as ". . . a way of studying the behavior of industrial systems to show how policies, decisions, structure, and delays are interrelated to influence growth and stability. It integrates the separate functional areas of management -- marketing, investment, research, personnel, production, accounting, etc. Each of these functions is reduced to a common basis by recognizing that any economic or corporate activity consists of flows of money, orders, materials, personnel, and capital equipment. These five flows are integrated by an information network."[3]

Models are the key tools in dynamic financial analysis. Such models are ". . . a systematic way to express our wealth of descriptive knowledge about industrial activity. The model tells us how the behavior of the system results from the interactions of its component parts."[4]

For insurance applications[5], the underlying system differs from an industrial one in the degree to which the functioning of the system manifests itself as pure flows of cash. Additionally, the many processes that can affect the amount and timing of the insurance cash flows are complex: some are stochastic, some allow for varying degrees of management control, and some may be imposed as constraints by either the marketplace or external regulatory entities.

This paper discusses the use of dynamic financial analysis as it applies to insurance. In this context, a systematic approach to analyzing all the major flows of cash is key.

---

3 MIT Press, 1961, p. vii.

4 Ibid.

5 Throughout this paper, the application of dynamic financial models to insurers is discussed. These models are equally useful for captives, risk retention groups, self-insurance pools and large self-insureds, as well as conventional insurers. Dynamic financial models are used in other financial sectors as well, e.g., the banking and investment industries.

The Actuary's Changing Role

Historically, casualty actuaries have primarily focused on rates and loss and loss adjustment expense reserves. Since 1980, property-casualty actuaries have had increasing responsibility to provide opinions on the loss and loss adjustment expense reserves of property-casualty insurance companies in the U.S.

In more recent years, regulatory and competitive pressures, as well as the desire for a broader understanding of the insurance process, have led and continue to lead to expansion of the casualty actuary's role. To meet the demands of this expanded role, actuaries now need a more complete understanding of insurance company cash flows: both assets and liabilities and their associated risks as well as their interrelationships.

This broader role will also increase the number of situations where the actuary must function in an interdisciplinary setting, communicating with the other major functional specialists of a company: those in investments, underwriting, claims, accounting and finance. This will bring new challenges in that it is likely that what is "normal" in terms of language or quantitative measures for the individual specialties may need to be described or measured differently for purposes of the dynamic financial model. However, if done effectively, this interdisciplinary communication network among specialists, and ultimately the company's management, can be one of the most valuable end results of building a dynamic financial model.

Why Use Dynamic Financial Models?

Dynamic financial models generally reflect the interplay between assets and liabilities and the resultant risks to income and cash flows. The explicit recognition of all of the insurer's operations gives dynamic financial models the power to illustrate the links between strategies and results. Therefore, these models make a unique contribution to the analytical tools available for performing financial analysis.

Dynamic financial models are valuable in effectively dealing with the complex interrelationships of variables relevant to an insurer's future results. Uncertainty related to contingent events occurring during the potentially long delay between writing a property-casualty insurance policy and the payment of claims against that policy make it difficult or impossible to evaluate strategies and decisions without explicit consideration of their effects on the flows of funds. Indeed, the results of management decisions or the effects of outside forces may be counter-intuitive. Use of a dynamic financial model can provide the insights necessary to clarify situations such as these.[6]

The explicit consideration of time delays, alternative outcomes of contingent events and interrelationships between different aspects of an insurers'operations gives dynamic financial models a unique role in helping management to identify profit opportunities, avoid negative outcomes and encourage investment in the company. Such explicit considerations can also assist both management and regulators in identifying and understanding problems early, before they grow to crisis size. Furthermore, in the event that problems do arise, these considerations can assist regulators in distinguishing short-term problems that do not warrant intervention from long-term problems that require action.

[6] These types of situations may be most common when a company changes strategy, either entering or exiting a line of business that has different characteristics from its existing book. A dynamic financial model can provide insight into the changing mix of the company's cash flows, both assets and liabilities, and the timing of profit recognition in published financial statements.

CHAPTER 3
USES, USERS AND RESOURCES

The design and/or selection of a dynamic financial model will depend heavily upon the question(s) to be addressed, the users of the model and its expected results, and the available resources.[7] Also, an effective design and selection process will solicit the expertise of a company's various major functional units. In and of itself, this communication network can be a significant benefit to company management as it will tend to reduce instances where actions might be taken that have not considered the ramifications to all areas of the company.

Uses

Dynamic financial models have a variety of applications, including:

- Determination of the value of an insurance company or a block of policies to a potential buyer or seller;
- Assessment of how an insurer might fare in a range of future economic, competitive, and regulatory environments;
- Strategic planning, including asset-liability management, claims management and settlement strategy, tax planning, reinsurance planning and costing, and market strategy;
- Tactical decision-making, including product pricing;
- Capital adequacy and capital allocation decisions;
- Liquidity analysis;
- Identification of the kinds of risks that most threaten the solvency of the insurer; and
- Support for company discussions with rating agencies.

[7] These considerations, along with the others identified in this report, are summarized in Appendix B.

The application will be a key determinant of many of the model's requirements. Examples include:

- The complexity of the model should reflect the question(s) being investigated. For example, if modeling long term capital needs, the underwriting experience may not need to be modeled by state and by coverage;
- The model output should reflect an appropriate time horizon and accounting basis. For example, if the question(s) being addressed only require statutory results, there may not be a need to include a GAAP module;
- The application may determine whether a deterministic or stochastic model[8] is more appropriate. This decision in turn will greatly affect the resources and data needed, the model structure, and the form that output will take. As an example, if the goal is to develop probability distributions of results, then an actuary will be more likely to use a stochastic model.

The development of a model must balance the need for complexity with usability and cost issues. Indeed, if a model is too complex, it could add unnecessarily to the development time and cost, it could mask errors in the model, and it could make results from the model harder to interpret. On the other hand, if a model is too simple, it may miss an important source of variation in results, it may not answer one or more of the questions being addressed, and it may lead to incorrect conclusions and actions.

Users

Users of dynamic financial models include insurers that employ such models as tools for tactical and strategic decision-making, including pricing decisions. Other users of the results of dynamic financial models can include regulators, reinsurers, investment bankers, financial intermediaries, institutional investors, securities rating organizations, and financial analysts.

The intended users' needs are the primary consideration in designing and selecting the model. The type of model used and its structure depend on users and their needs. As an example, regulators may focus mainly on the insurer in total. Company management may focus on the total corporation as well as individual products.

---

[8] A stochastic model will reflect the uncertainty in a company's estimated cash flows by treating one or more components of the cash flows as random variables from specified probability distributions. A deterministic model will treat all estimated cash flows as though they are certain.

As a practical matter, the model design should also take into consideration the expertise of the end user. At one end of the spectrum, it may be that a model with a limited number of user-specified scenario options and input variables provides the best fit to the user's needs. At the other end of the spectrum, a user may want the control and flexibility to address almost any situation. In the former case, if the user is applying the model almost like a "black box," it becomes more important to have a plan of periodic review and update to the internal workings of the model. Otherwise, the user may continue producing results when in fact the model's parameters have become outdated.

Resources

The choice of dynamic financial model will depend on the available resources:

- people available for system design and programming;
- data from which to derive assumptions and with which to initialize the model;
- money available to purchase an existing software package; and
- computer architecture.

Detailed dynamic financial models require a significant investment of time for research to determine assumptions, for validation of results, and for maintenance to keep the model's logic current and to revise assumptions in light of new data. Such models also require a significant expenditure of time to interpret the results.

The purpose of the analysis and the level of detail of the projections often determine the choice of computer architecture. A simple spreadsheet might be appropriate if the purpose of the study is to highlight the effects on financial results of one particular risk, such as adverse development of loss reserves. At the other extreme, complex, report-generating software with a user-friendly front-end and efficient coding of the detailed calculations might be appropriate if the model is intended to cope with a wide range of different problems and be used by a wide range of users.

CHAPTER 4
TYPES OF MODELS

In many disciplines, mathematical models have become important tools in the study of the behavior of systems. The systems are diverse, almost unlimited, ranging from the biology of the human body to the weather to insurance. For any of these applications, the type of mathematical model employed will depend upon the nature of the system, how broadly the system is defined, and what needs to be learned about the system's behavior. Considerations for building a model include:

- the extent to which the system is described by stochastic versus deterministic mathematical relationships;
- the length of time horizons if predictions of future behavior are important;
- the ability of the system to adapt to changing environments; and
- the nature of the system's interrelationships with external factors or other systems.

These considerations, and the extent to which a model must emulate all facets of the real system, will determine how simple or sophisticated a model must be.

In the context of property-casualty insurance, dynamic financial models will incorporate different features depending on the application and the types of risks that must be measured. The extent that certain features are emphasized will determine what might be called the "type" of model (i.e., is it primarily stochastic or deterministic; does it include feedback loops, etc.). However, different models may include any or all of these features to different degrees. Therefore, the spectrum of types of models can be viewed as a continuum rather than a collection of discrete categories. At one end of the spectrum, sophisticated models may incorporate many features, emulating an entire company and most of its interrelationships. At the other end of the spectrum, simpler models may incorporate few of these features and may be designed for specific narrowly defined problems. A key consideration in the design of a dynamic financial model is its ability to evaluate the material sources of risk for the problem at hand.

Primary Modeling Considerations

Stochastic vs. Deterministic

If material random fluctuations in a variable are significant for a particular application, then stochastic features can be added to a model. Random fluctuations around projected losses, for example, may be incorporated into a model by introducing probability distributions about loss costs or loss ratios, by modeling the collective risk process, or by modeling the underlying claim settlement process.

A simple model of the collective risk process may assume probability distributions for the frequency and severity of losses. A more complicated model of the collective risk process may include estimates of parameter uncertainty for frequency and size-of-loss, and may include a number of different kinds of losses, each with its own frequency and size-of-loss assumptions. A model of the underlying claim settlement process may be a multi-state Markov chain model or some other appropriate model.

Identifying and modeling the interactions among variables is important when either stochastic or deterministic variables are used. However, when assumptions are stochastically generated, a model that does not reflect these interactions may generate scenarios that are meaningless. At best, the results of such models would be difficult to interpret.

Time Horizon

The time frame for the analysis is an important consideration in the choice and design of a dynamic financial model. For example, the choice of time frame may reflect whether the model includes only the run-off of current business, a going concern for some stated period, or a going concern in a long-range projection valuation.

In addition to the time horizon of the model, the model also reflects a choice about the length of time intervals under study. While annual time intervals may be appropriate for some purposes, quarterly or even monthly time intervals might be appropriate for other purposes. The user must consider the procedures to be carried out over various time frames and their suitability for the use of the model. For example, a model might generate cash flows on a monthly basis, but test statutory solvency on an annual basis.

Feedback Loops/Adapting to Change

Dynamic financial models may employ feedback loops (automatic conditional decisions) which are algorithms that make calculations for each modeled time period dependent on values calculated for earlier periods. Feedback loops provide for reactions to specific conditions. For example, if a given scenario shows a loss ratio that is unacceptably high for a certain line of business, then the model could assume that rate level and other underwriting decisions will be made by management to mitigate the unacceptable results.

Models without feedback loops may be under-determined, showing excessive income under favorable scenarios and excessive loss under unfavorable scenarios. Models with feedback loops, however, may be over determined, showing little risk regardless of the scenario because the model builder often assumes that management will respond quickly to increased risk with appropriate strategic or operational responses. The issues of feedback loops and strategy specification are closely related.

Interrelationships with External Systems

The insurance process is subject to constraints imposed by the choice of available investments, underwriting commitments, laws and regulations, rating agencies, and income tax laws. Comprehensive models, for example, a model designed to determine the value of an insurance company, will reflect all or most of these constraints. Less comprehensive models, for example, a model designed to price a specific product, may be appropriate, however, for specific applications.

Other Modeling Considerations

Generalized vs. Tailor-Made

Generalized models usually permit the user to specify several different types of insurance products or lines of business and a range of different investments. Other models are often tailor-made, such as one that addresses the unique characteristics of a company or one developed for a situation in which a simple model is sufficient.

If a generalized model is used, it is important to consider whether results may be distorted by features inconsistent with a particular application or because a characteristic of the particular company is not addressed. For example, if a general-purpose model is used for an insurer that plans to invest only in bonds and cash equivalents, the model does not need to include a strategy that involves investment in other assets. If it does, care should be taken so that the ramifications of that logic do not distort the projections.

Logic vs. Input

There are always tradeoffs between the coding of logic versus the selection of parameters. Dynamic financial models differ in the choices the developers make about which assumptions will be represented by variables and which will be fixed by the software. Also, the user will be able to determine the values of certain variables used by the model, whereas others will have been pre-set by the developer. The mix between input and logic will be determined in part by the users of the model (both the operator and the decision-maker). Models with extremely large numbers of variables can be daunting to use and difficult to interpret, while models with too many decisions incorporated into logic may not be flexible enough.

In selecting or building a dynamic financial model, decisions must also be made about the level of detail to be captured. For example, some choices include the detail of the insurance coverage (by broad product group, statutory line of business, individual form, etc.), the factual context (including the level of detail about accounting and tax rules), and the precision with which strategies are defined.

Strategies are inevitably a part of the logic of a model. The strategies incorporated in the model should be reasonably consistent with its purpose. Some models allow the user to build in explicit recognition of management strategies. Other models assume certain strategies, even to the extent of letting presumptions about strategies affect the architecture or design of the model.

Relationship between Parent and Subsidiaries

Parents and subsidiaries have a number of different effects on an enterprise. A consolidated model of the entire organization can be developed, or the existence of the parent and subsidiaries might simply show up as assumptions about flows of funds, tax calculations, and income. A model may explicitly reflect a range of scenarios regarding the availability of or drain on surplus due to external influences. Alternatively, each entity may be modeled separately, with output from one model serving as input for other models.

CHAPTER 5
PROPERTY CASUALTY RISKS AND MODELING CONSIDERATIONS

Evaluation of risk is the focus of dynamic financial models. The relative importance of each type of risk will determine the detail of assumptions and analyses built into any model. Ultimately, a model must provide a quantitative evaluation of risk in terms of its effects on the amount and timing of flows of cash. This chapter describes the risks affecting the property-casualty insurance business and addresses the related modeling considerations.

Property-casualty insurance risks can be divided into many categories.[9] In this paper, we will follow the definitions originated by the Committee on Valuation and Related Matters of the Society of Actuaries and will discuss these risks in the following four categories:

> **Asset Risk** – The risk that the amount or timing of items of cash flow connected with assets will differ from expectations or assumptions as of the valuation date for reasons other than a change in interest rates.
>
> **Obligation Risk** – The risk that the amount or timing of items of cash flow connected with the obligations[10] considered will differ from expectations or assumptions for reasons other than a change in interest rates.
>
> **Interest Rate Risk** – The risk that the amount or timing of items of cash flow connected with assets or obligations will differ from expectations or assumptions because of changes in interest rates.
>
> **Mismanagement Risks** – Uncertainty from taking incorrect or fraudulent actions in light of the available information.

As do many discussions of insurance risks, this paper will focus on the first three of these risks. At present, measuring Mismanagement Risk is beyond the scope of most actuarial analysis.

---

[9] For example, the NAIC's risk-based capital formula divides risk into 5 categories: asset risk, credit risk, reserve risk, premium risk, and off balance sheet risk (e.g., growth).

[10] Any tangible or intangible commitment by, requirement of, or liability of a plan or an insurer that can reduce receipts or generate disbursements.

Asset Risk

Asset risk encompasses uncertainty regarding:

- Default rates;
- Future market value of equity assets; and
- Liquidity of assets.

In addition to these inherent asset risks, model builders should take care to look beyond the general description of the various asset classes to make sure that all relevant risk characteristics are incorporated in the model. This precaution increases in importance as capital markets develop a greater range of non-equity investments that have many of the risk characteristics of equity investments.

Appropriate data and methods are critical to the development of ranges of assumptions to reflect asset risk in the projected performance of the insurer. Historical data developed for investment managers is readily available, including time series of default rates of various classes of assets as a function of age.

Dynamic financial models can be used to estimate the effects of these risks alone on the projected performance of the insurer and can also be used to estimate the interrelationships between these risks and other risks. In modeling, asset risks may be assumed to correlate with inflation or some other variable or to be autoregressive.

Obligation Risk

Obligation risk encompasses:

- **Reserve Risk** - the risk that the actual cost of losses for obligations incurred before the valuation date will differ from expectations or assumptions for reasons other than a change in interest rates;
- **Premium Risk** - the risk that premium for future obligations will differ from expectations or assumptions;
- **Loss Projection risk** - the uncertainty regarding assumptions (other than interest rates) about future loss costs (including LAE);

- **Catastrophe risk** - the uncertainty regarding the costs of natural disasters and other catastrophes;
- **Reinsurance Risk** - the uncertainty regarding the cost, value, availability and collectibility of reinsurance; and
- **Expense Risk** - the risk that expenses and taxes will differ from those projected.

Dynamic financial models can be used to estimate the effects of these risks individually on the projected performance of the insurer and to evaluate the interrelationships between these risks and other risks.

Reserve risk may be a function of:

- Inflation in claim costs (other than that related to interest rates);
- The legal environment in which claims will be resolved, including the environment in which claims are pursued by policyholders or third parties;
- The possibility of a breakdown in some basic premise underlying the reserves for a particular coverage (such as has occurred with environmental impairment liability);
- Past patterns of pricing adequacy which affect case reserves or financial reserves;
- Corporate culture, training, and incentives that affect the payment of claims or the adequacy of case reserves;
- Currency fluctuations which affect the costs of losses when expressed in local currency;
- The randomness of the claims process itself;[11] and
- Incompleteness of databases.

Premium risk may be a function of:

- Competitive pressures that do not allow the insurer to achieve assumed levels of exposure and/or rate adequacy;
- Regulatory intervention that restrains premium increases or decreases or requires business to be underwritten that would not be underwritten in the absence of such intervention;

[11] The randomness of the claims process itself can be studied by modeling the patterns of loss development or by more detailed analysis of the claims process. Inevitably, however, data for such models always include the effects of other factors affecting the claims process.

- Premiums for involuntary business underwritten at premium rates and in volumes that differ from assumptions;
- Retrospective premiums or dividends that differ from assumptions; and
- Amounts collectible from agents that differ from assumptions.

Loss projection risk is a function of the factors that affect reserve risk and also of the uncertainty regarding:

- Unanticipated changes in loss costs and exposures from the historical experience period;
- Loss costs for the mix of new policies being underwritten, including the effect of adverse selection; and
- Loss adjustment practices in the future that may differ from those in the past.

Catastrophe risk can be considered a component of loss projection risk. It is a function of:

- The coverages being written;
- The concentration of insured values in specific geographic areas or legal jurisdictions; and
- Uncertainty regarding the frequency, severity, and nature of catastrophic events.

Computerized models of the damage arising out of certain types of catastrophes are available and may be of value in determining assumptions about the probabilities and sizes of catastrophic losses. Output from these catastrophe models may be used in a variety of ways. A link between models could be constructed to feed catastrophe simulations directly into the Dynamic Financial Analysis (DFA) model. Alternatively, the output could be used as an input table to a DFA model to generate catastrophe risk scenarios. Further, the output could be analyzed to obtain values to parameterize catastrophe risk within the DFA model.

Reinsurance risk is a function of changes in the price and availability of desired reinsurance, and of uncertainty regarding the collectibility of reinsurance recoverables arising from the financial condition of the reinsurer or ambiguity about the coverage provided. Reinsurance risk exists in each of the four obligation risks identified thus far. In many models, projections are made on a net of reinsurance basis. Such projections incorporate implicit assumptions regarding reinsurance risks, whereas projections made on a gross of reinsurance basis require explicit instructions regarding the reinsurance mechanism. Reinsurance risk recognizes how reinsurance responds under stress, such as a large catastrophe or other strain on collectibility, aggregates, reinstatements and other reinsurance parameters.

Expense risks, those associated with expenses (other than loss adjustment expenses) and taxes, include uncertainty regarding:

- Contingent commissions to agents;
- Marginal expenses of adding new business;
- Overhead costs, including the risk that overhead costs will be changed by regulatory intervention, and the risk that there may be periods of changing premium during which overhead costs will not change in proportion to premium;
- Assigned risk overburdens, second injury funds and other assessments;
- Policyholder dividends; and
- Federal and local income taxes, both in interpreting the current Internal Revenue Code and in anticipating changes to the code.

These lists of uncertainties regarding the major components of obligation risk are illustrative. Other factors may also affect obligation risk.

Interest Rate Risk

Interest rate risk is the risk of financial loss caused by changes in future interest rates. It encompasses:

- The risk of a change in the economic value of asset cash flows caused by changes in interest rates – this includes cash inflows such as those from bonds, mortgages, real estate, and dividends from equity investments; and
- The risk of a change in the economic value of obligation cash flows caused by changes in interest rates – this includes both cash outflows (such as those related to loss reserves) and cash inflows (such as expected future premium receipts).

A dynamic financial model is an important tool in measuring the financial effects of these components of interest rate risk, both individually and in combination. The model's ability to measure interest rate effects on all cash flows - cash inflows, cash outflows and net cash flow - will also enable a company to develop management strategies that mitigate the potential adverse financial effects related to interest rate changes.

Asset Cash Flows – Asset cash flows may be fixed or may change in response to interest rate changes. If cash flows are fixed (e.g.. some types of bonds) an increase in interest rates produces a reduction in market value and possibly a reduction in earnings if conditions force the insurer to sell the bond in the high interest rate environment (see disinvestment risk below). If the cash flows are interest sensitive (e.g., a bond with fixed payments but having a prepayment option), then both the timing and amount of the flows may change in response to an interest rate change. For example, a bond that has a prepayment option would tend to be called in times of declining interest rates. In this situation, the borrower would prepay the bond in order to take advantage of more favorable borrowing costs elsewhere. On the opposite side of the transaction, the insurer would realize an adverse economic impact in the loss of future investment income from the higher yielding asset after reinvestment at the lower prevailing rates. The same effects can occur when cash flows are not fixed as in these examples, unless the variable cash flows change in concert with interest rate changes (such as with debt with interest linked to a market index).

Cash flows from other assets may also be fixed or interest sensitive. Generally, the sophistication with which the effects of interest rate risk on assets need to be modeled is directly related to the asset's importance to the insurer. For most property casualty insurers, more effort would be made to appropriately model the effects of interest rate changes on bonds than on real estate and equities.

Obligation Cash Flows – Obligation cash flows may also be fixed or may change in response to interest rate changes.

By far, the largest obligation cash outflows for property-casualty insurers are payments for losses and loss adjustment expenses. The degree to which interest rate risk is an issue and the degree to which these cash flows are fixed or interest sensitive will vary by line. At one end of the spectrum, if the cash flow for losses incurred prior to the valuation date is fixed relative to interest rates (i.e., excluding reserve risk), then a decrease in interest rates would produce an adverse financial impact (measured on an economic basis). To the extent that the loss payments are interest sensitive, the economic impact will be reduced, provided that they move in the same direction that interest rates move. Generally, interest rate risk will be more significant for the longer tail lines of business because of the longer duration of the cash flows.

On the premium side, an increase in interest rates could produce a decrease in future premium cash inflows to the extent that insurance companies in the marketplace rely on investment income to maintain overall profitability. Other components of underwriting income could also show varying degrees of sensitivity to interest rate changes.

Again, the needed degree of effort and sophistication applied to modeling the effects of interest rate changes on each component of the obligation cash flows will depend on the relative importance of each component. This will vary in each situation according to the specific characteristics of the insurance operations being modeled.

Net Cash Flows – Differences in timing and amount between cash inflows and cash outflows produce risks and opportunities with respect to the potential financial loss associated with interest rate changes. The risks include reinvestment risk when cash inflows exceed outflows and disinvestment risk when cash outflows exceed inflows. Opportunities exist to the extent that these risks can be mitigated by managing cash inflows and cash outflows in such a way that the economic value of the net cash flow is immunized, to some extent, from changes in interest rates. The degree of immunization may be limited by the available choice of investments if the optimal asset cash flow is not produced by any readily available asset. Such differences could arise from the interaction of economic factors with assets or liabilities.

Reinvestment and disinvestment risks are components of interest rate risk that arise when differences in the timing and amount of cash inflows and outflows cause the insurer's net cash flow in a period to be substantially different from zero.

Reinvestment risk relates to the uncertainty regarding investment returns that will be available upon the reinvestment of excess cash flow related to proceeds from investments. If interest rates have decreased, then the excess cash flow will have to be reinvested at rates below those on the existing or maturing assets.

Disinvestment risk arises when fixed-income assets must be sold prior to maturity to meet cash flow needs, typically because the net cash flow is negative absent the sale of these assets. If interest rates have increased, then the market value of these assets has decreased and they will be sold at a relative loss.

Interest rate risk includes the portion of market value uncertainty due only to changes in interest rates. The portion of market value uncertainty related to changes in perceived credit or default risk is a component of Asset Risk. Also, the reinvestment rate assumption in a dynamic financial model determines both reinvestment risk and disinvestment risk for fixed-income assets.

Consequently, the reinvestment rate can have a significant impact on the results of a dynamic financial model.

CHAPTER 6
POTENTIAL DANGERS/PITFALLS INHERENT IN THE MODELING PROCESS

Once the risks to be incorporated in the model have been identified and the model built, there are a number of dangers inherent in the modeling process to consider, including:

- The range of scenarios may not reflect the user's intent;
- The model may be incorrectly or incompletely specified for the intended purpose; and
- The model may quickly become obsolete if it is not adaptable to change.

Importance of Scenario Testing and Selection of Assumptions

For a particular application, proper use of a model depends on the selection of appropriate scenarios[12] to evaluate and the development of consistent assumptions within each scenario, which, in turn, will influence the data and methods used to provide assumptions for understanding the projected performance of the insurer. Scenarios permit links between assumptions for various parts of the model. For example, a high interest rate scenario might include assumptions of high bond yields, low common stock values with high dividends, high inflation in medical costs, and a low level of unemployment.

Scenarios provide a useful tool for determining the implications of risks on the projected performance of an insurer. Observing the results for a variety of scenarios yields information about the company's response to risk. Careful selection of scenarios is essential.

Often times, the scenarios to be studied will be specified by company management. There may also be times when scenarios are specified by external sources. For example, the Canadian regulations provide general guidance on the choice of scenarios. By whatever means, the range of scenarios is selected, its choice will impact the results that the model produces. It may be appropriate to observe the model under scenarios other than those specified by regulators or management to adequately understand the implications of the scenarios that were specified.

---

[12] A scenario is a description (set of assumptions) of a group of variables (such as interest rates or combined ratios) that can reasonably be expected to impact an insurance enterprise. The description of the group of variables constitutes the environment within which the insurance enterprise will operate.

When the range of scenarios has been selected using only retrospective tests as a guide, the model may be prone to under-determination. For example, the danger that the probability distributions in a stochastic model are incorrectly specified can be reduced by choosing probability distributions that have greater uncertainty (dispersion and skewness) than historical data.

Model Specification and Validation

A model that is incorrectly or insufficiently specified will fail in its intended purpose and could lead to costly mistakes. To reduce this danger, model validation is crucial, i.e., matching the model to the insurer's own history over some period of time. A well-specified model will reasonably reproduce past actual results. Actual results varying from projections may not be an indication of a poor model. Rather, it is generally appropriate to investigate such differences and reconcile the model's results with the actual results. This process of reconciliation may identify weaknesses in the model, or clarify ways in which the enterprise's activities departed from what would have been reasonably expected (e.g., writing more, rather than less, unprofitable business to cover up poor experience).

Keeping the Model Relevant[13]

Work does not end once a model is built. Change is constant and a model must keep pace with this change to stay relevant. Examples of continuing change include:

Proliferation of Insurance Products: Although regulation and custom tend to slow the creation of insurance products by entrepreneurs, changes in the markets served by insurance enterprises constantly press for new products and services. Dynamic financial models may need to be refined to adapt to these changes.

Competitive Pressures: In the past, pressures were perceived to arise from competition at the point of sale of the insurance product. Since at least 1970, competitive pressure has increasingly come to mean competition at the point that capital is being raised. Dynamic financial models are playing an increasingly visible role in corporate decisions regarding purchases and sales of business units, means to tap capital markets, and trade-off between

[13] The following subsection was adopted from R. Blanchard, Actuarial Digest, Volume 15, No. 5, Oct./Nov. 1996: "A Mechanics Perspective to Model Building".

trend is expected to continue.

Innovation in Assets: Recent innovations in asset design make it difficult to understand the riskiness of many investments by looking at their financial designations for accounting purposes. For example, some bonds have the risks of stock investments or mortgages, and mortgages are often backed by a wide range of securities. Existing accounting classifications may be misleading to tabulate information about assets for input into dynamic financial models.

New types of asset classes are emerging, some with purposes other than purely generation of investment returns. For example, some assets, such as catastrophe futures, can hedge risks undertaken by the insurer's underwriting activities. More innovation can be expected, along with the need to model these kinds of investments.

Regulatory, Accounting, and Tax Requirements: Dynamic financial models may need to be revised from time to time to reflect the latest developments in regulation. Such changes may be as simple as adding a set of calculations, or they may require modeling of the corporate response to the impact of the regulations (e.g., a shift in marketing or investment strategy to accommodate surplus constraints of risk-based capital). Projections of cash flow may react to changes in these constraints differently from projections of statutory results. Dynamic financial models with feedback loops may react differently from static models.

In a changing environment, to keep a model from rapidly becoming a dinosaur, it should be designed with change in mind. A structured model comprised of smaller interrelated program modules will tend to be much more adaptable than one big monster.

CHAPTER 7
MEASURING RESULTS

In order to be an effective analytical tool, a dynamic financial model should be capable of producing various types of output, both financial and analytical. Financial summaries could range from high level, e.g., the overall company balance sheet and income statement, to detailed financial statements at the level that a company would manage and plan its business strategies. Analytical output could include various statistics including graphic representations such as plots of results in a risk vs. reward format (e.g., the asset/liability efficient frontier). A comprehensive model would also generally be capable of producing these results under various bases of accounting.

In addition to the appropriate output summaries, a model must also be designed to maintain whatever additional detail might be needed – either at still lower levels of detail or at intermediate calculation points – that would be needed in order to analyze and interpret output. This "drill down" capability is crucial to successfully reconcile model output with expectations or to diagnose those situations where the model output appears either counter-intuitive or even unreasonable.

- **Basis of Accounting** - Comprehensive dynamic financial models will usually include accounts on at least four bases simultaneously: cash (or economic), statutory, GAAP, and tax.[14] This is the only way to reflect the details of the interrelationships among constraints imposed by investment opportunities, underwriting commitments, laws and regulations, generally accepted accounting principles, and income tax laws. However, less comprehensive models may be appropriate depending on the use.

- **Interpreting Output/Drill Down** – Proper interpretation of output is possibly the most important aspect of using a dynamic financial model. The danger of inappropriate interpretation can be reduced by communicating the possibly limited extent of variation among modeled scenarios in comparison to the potential range of variation in the year to year results of the insurer's operations.

[14] Financial reporting, and therefore modeling, may be more complex for international users.

Additionally, developing conceptual interpretations of model results is crucial to communicating these results. This can be a challenge and may entail an intensive drill down through model output in order to identify major "drivers" of the results. Because the volume of output data generated by a detailed model can be overwhelming, the task will be made easier if the model design includes drill down and diagnostic capabilities on its output. These may include:[15]

- Expectation and distribution of selected output variables;
- Identification and categorization of scenarios that resulted in extreme values;
- Determination of explanatory variables relative to selected output variables (e.g., regression techniques);
- Evaluation of decision rules, reinsurance programs, etc. relative to selected output variables (e.g., "on/off" switches); and
- "Good vs. bad" analyses (e.g., risk vs. reward types comparisons).

The results of the model could suggest that either one or more assumptions are incorrect (in which case the assumptions will likely be revised before results are presented) or that the insurer's strategies could be improved. As an example of the latter, the results of the model may suggest that the insurer may be particularly at peril due to one or more sources of risk.

[15] Adapted from "Dynamic Financial Modeling - Issues and Approaches," Thomas V. Warthen III and David B. Sommer; CAS Forum, Spring, 1996.

APPENDIX A
BIBLIOGRAPHY

Adelberger, Otto L., "SIMULFIN: A General Financial Simulation Model of the Firm for Teaching and Research Purposes," Journal of Financial Education, Vol. 3, 1974, pp. 96-104.

Arvan, Lanny and Leon N. Moses, "A Model of the Firm in Time and Space," Journal of Economic Dynamics and Control, Vol. 9, No. 1, 1985, pp. 77-100.

Chalke, Shane, "Heuristic Modeling," The Chalke Perspective, 3rd Quarter, 1993.

Cho, Dongsae, "The Impact of Risk Management Decisions on Firm Value: Gordon's Growth Model Approach," Journal of Risk and Insurance, Vol. 55, No. 1, 1988, pp. 118-131.

Clouse, Maclyn L., "A Simulation Model to Link the Economic Environment with Corporate Financial Decisions," Journal of Financial Education, Vol. 8, 1979, pp. 94-96.

Elliott, J. Walter, "Forecasting and Analysis of Corporate Financial Performance with an Economic Model of the Firm," Journal of Financial and Quantitative Analysis, Vol. 7, No. 2, 1972, pp. 1499-1526.

Forrester, Jay W., Industrial Dynamics, MIT Press, 1961.

Francis, Jack Clark and Dexter R. Rowell, "A Simultaneous Equation Model of the Firm for Financial Analysis and Planning," Financial Management, Vol. 7, No. 1, 1978, pp. 29-44.

Herendeen, James B., "A Financial Model of the Regulated Firm," Southern Economic Journal, Vol. 42, No. 2, 1975, pp. 279-284.

MacMinn, Richard D. and John D. Martin, "Uncertainty, The Fisher Model, and Corporate Financial Theory," Research in Finance, Vol. 7, 1988, pp. 227-264.

Myers, S. C., "A Simple Model of Firm Behavior Under Regulation and Uncertainty," Bell Journal of Economics, Vol. 4, No. 1, 1973, pp. 304-315.

Prisman, Eliezer Z., Myron B. Slovin and Marie E. Sushka, "A General Model of the Banking Firm Under Conditions of Monopoly, Uncertainty, and Recourse," Journal of Monetary Economics, Vol. 17, No. 2, 1986, pp. 293-304.

Taggart, Robert A., Jr., "A Model of Corporate Financing Decisions," Journal of Finance, Vol. 32, No. 5, 1977, pp. 1467-1484.

Walker, Kenton B. and Lawrence A. McClelland, "Management Forecasts and Statistical Prediction Model Forecasts in Corporate Budgeting," Journal of Accounting Research, Vol. 29, No. 2, 1991, pp. 371-381.

APPENDIX B
CHECKLIST OF CONSIDERATIONS[16]

1. Is the model appropriate for the intended use?

2. Are the model and related communications appropriate for the expected users of its results?

3. Can the model be developed, purchased, maintained and/or used within the personnel, time, hardware, software and budget resources available?

4. Does the model contain input, output and processing regarding each of the risks to be evaluated in appropriate detail? Are the available historical data regarding these risks sufficient to use to derive the assumptions needed by the model? These risks include:

   - Asset risk
   - Obligation risk
     - Reserve risk
     - Premium risk
     - Loss projection risk
     - Catastrophe risk
     - Reinsurance risk
     - Expense risk
   - Interest rate risk

5. Is the range of scenarios broad enough to reasonably address the questions at hand?

6. Is the model specification accurate and appropriately complete?

7. Are the measures used to summarize and interpret the range of results reasonable for the application?

[16] This is an abbreviated list of considerations. A more comprehensive list is contained in the CAS Handbook for Dynamic Financial Analysis.

8. Have the limitations of the model and range of scenarios been communicated clearly to reduce the risk of misinterpretation?

9. Is a generalized model reasonable for the application or would a tailor-made model better address specific issues?

10. Does the model have a reasonable balance between input assumptions and hard-coded logic?

11. Is the model's time horizon appropriate to the application?

12. Are the accounting bases upon which the model makes forecasts of appropriate breadth to the application?

13. Does the model provide sufficient detail (input and output) with respect to interactions with parents, subsidiaries and affiliates?

14. Will the value of the model results be enhanced enough by the presence of feedback loops (automatic conditional decisions) to warrant a model with such features?

15. Is a deterministic or stochastic model better suited for the application?