CASUALTY ACTUARIAL SOCIETY FORUM

Winter 2007
Including the Ratemaking Call Papers

CASUALTY ACTUARIAL SOCIETY
ORGANIZED 1914

Printed for the Society by
United Book Press
Baltimore, Maryland

NOTICE

The Casualty Actuarial Society is not responsible for statements or opinions expressed in the papers in this publication. These papers have not been reviewed by the CAS Committee on Review of Papers.

The Casualty Actuarial Society *Forum*

Winter 2007 Edition

Including the 2007 Ratemaking Call Papers

To CAS Members:

This is the Winter 2007 Edition of the Casualty Actuarial Society *Forum*. It contains nine Ratemaking Call Papers, one CAS Research Working Party Report, and four additional papers.

The *Forum* is a nonrefereed journal printed by the Casualty Actuarial Society. The CAS is not responsible for statements or opinions expressed in the papers in this publication. See the Guide for Submission to CAS *Forum* (http://www.casact.org/about/index.cfm?fa=forum) for more information.

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 must submit papers in conformance with the research paper template established by the CAS. See the CAS Web Site at http://www.casact.org/research/index.cfm?fa=template for formatting instructions. Papers not conforming to the template will not be accepted.
2. Authors should submit a camera-ready original paper and two copies.
3. Authors should not number their pages.
4. All exhibits, tables, charts, and graphs should be in original format and camera-ready.
5. Authors should format graphs, tables, exhibits, and text in solid black and white. Avoid using color or gray-shaded graphs, tables, or exhibits as the printing process will produce muddy and moire patterns.
6. 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,

Glenn M. Walker, CAS *Forum* Chairperson

The 2007 CAS Ratemaking Call Papers
Presented at the
2007 CAS Ratemaking Seminar
March 8–9, 2007
Hyatt Regency Atlanta
Atlanta, Georgia

The Winter 2007 Edition of the CAS *Forum* is a cooperative effort between the Committee for the CAS *Forum* and the Committee on Ratemaking.

The CAS Committee on Ratemaiking present for discussion nine papers prepared in response to their 2007 call for papers.

This *Forum* includes papers that will be discussed by the authors at the 2007 CAS Ratemaking Seminar, March 8–9, 2007, in Atlanta, Georgia.

2007 Committee on Ratemaking

John J. Lewandowski, *Chairperson*

Lee M. Bowron
William M. Carpenter
Donald L. Closter
Kiera Elizabeth Doster
Charles E. Gegax
RenBin Guo
Keith D. Holler
Eric J. Johnson
Todd W. Lehmann
Neal Marev Leibowitz
Pierre Lepage
Taylan Matkap
Dennis T. McNeese
Jin Park
David A. Smith
Jane C. Taylor
Natalie Vishnevsky
Benjamin A. Walden
Jonathan White

2007 CAS Ratemaking Call Papers

Research Working Party Report

Additional Papers

Catastrophes and Workers Compensation Ratemaking

Tom Daley, ACAS, MAAA

Abstract: The CAS Statement of Principles Regarding Property and Casualty Insurance Ratemaking states that "Consideration should be given to the impact of catastrophes on the experience, and procedures should be developed to include an allowance for the catastrophe exposure in the rate."

For the first time in many years, NCCI has modified the methodology used to determine a state's overall average loss cost or rate level indication for workers compensation. The aggregate ratemaking methodology was modified specifically to handle two general categories of large events for which workers compensation exposure exists. They are: a) large individual claims, and b) catastrophic events related to the perils of industrial accidents, earthquake, and terrorism. NCCI actuaries worked with a well-known modeling firm to determine provisions for catastrophic events on a state basis.

This paper describes the new methodology NCCI has developed, implemented, and filed in many of its states. It discusses in detail how the traditional areas of aggregate ratemaking were modified: loss development, the tail factor, trend, selection of loss limits by state, and application of excess provisions.

The paper also documents for the first time in CAS literature how computer modeling was applied in workers compensation to determine a loss cost by state. Consideration was given to the protection of proprietary trade secrets of the EQECAT modeling firm, with whom NCCI partnered.

Keywords: workers compensation; NCCI ratemaking; NCCI loss cost filings; catastrophic events; large losses; TRIA.

1. INTRODUCTION

For the first time in many years, NCCI has modified the methodology used to determine a state's overall average loss cost or rate level indication for workers compensation insurance. The aggregate ratemaking methodology was modified specifically to handle two general categories of large events for which workers compensation exposure exists. They are: a) large individual claims, and b) catastrophic events related to the perils of industrial accidents, earthquake, and terrorism.

This paper describes the new methodology NCCI has developed, implemented, and filed in many of its states. It discusses how the traditional methods for aggregate ratemaking were modified, as well as how advanced modeling techniques were used to quantify loss cost provisions by state for those perils. The large loss ratemaking procedure can be described as one that uses reported losses capped at a given dollar threshold and adds a provision for expected losses excess of this threshold. The details underlying the specifics of the approach and the decision making process are documented in the pages that follow.

1.1 Research Context

The focus of this research is two-fold. It addresses the use of modeling outside of personal lines, as well as providing an update on current workers compensation ratemaking methods. Current CAS literature that addresses some of the same issues include "Workers Compensation Ratemaking" by Sholom Feldblum, and "Issues in the Regulatory Acceptance of Computer Modeling for Property Insurance Ratemaking" by Rade Musulin.

1.2 Objective

This paper updates the CAS literature on workers compensation ratemaking techniques, with particular attention to recent modifications in the NCCI ratemaking methods for handling large claims and very large events. To address its absence in the current CAS literature, this paper also discusses the use of computer modeling in workers compensation ratemaking. Class ratemaking considerations will not be addressed in this paper, as the considerations of large losses on class relativities is currently being reviewed at NCCI.

1.3 Outline

The remainder of the paper proceeds as follows. Section 2 will discuss the reasons and impetus for the changes made, the thought process NCCI followed, the research approach and results, and specific aggregate ratemaking methodology changes. Section 3 documents the modeling approach used for several catastrophic perils, and how the modeling of large events was used to estimate loss costs.

2. BACKGROUND AND METHODS

Prior to the 1970s, the workers compensation rates promulgated by NCCI included a 1-cent catastrophe provision in every rate. This provision was eventually removed from ratemaking.

The events of September 11, 2001, which caused the greatest insured loss in property-casualty history to date (that may or may not have been exceeded by Hurricane Katrina), brought into focus the potential that large events may have on workers compensation. Previous NCCI estimates of the insured loss for the workers compensation line of insurance due to the events of September 11th range from $1.3 – $2.0 billion on a direct (of reinsurance) basis. As with so many other lines of insurance and perspectives about risk-taking, the events of that day created a compelling reason to take a fresh look at how workers compensation ratemaking could fund such large, infrequent events prospectively. It was clear that funding large events is an issue in workers compensation, and is no longer just an issue confined to personal and commercial property insurance.

2.1 Overview of the Methodology Change

NCCI revised its aggregate ratemaking approach in 2004 to more closely resemble basic limits ratemaking. Limited losses are calculated by subtracting the actual loss dollars for each claim that are excess of a given dollar threshold from the aggregate unlimited losses for a state. Next, the limited aggregate losses are multiplied by limited loss development factors (discussed later) to obtain ultimate limited losses. Trends (loss ratio and severity) are then calculated using these limited losses and benefit changes are applied (to the limited base of losses).

Finally, the trended ultimate limited losses are divided by a factor (1 -XS), where XS is the Excess Ratio (described later) for the appropriate dollar threshold, resulting in total projected ultimate losses, for use in ratemaking.

In the sections below the details of the large loss procedure will be described and how the different aspects of the ratemaking process and the overall rate filing are affected. The terms "limited" and "capped" will be used interchangeably.

2.2 How Were Large Events Handled in the Past?

Historically, NCCI actuaries occasionally encountered one or more large individual or multi-claim occurrences in past loss cost and rate filings that impacted a state's overall loss cost or rate level indication. The methods of handling these claims varied from state to state. Treatments in filings of historical experience that included large claims or occurrences in a filing included the following ad hoc approaches:

- Making no adjustment to the reported experience for the state
- Selecting a longer experience period (for example, three policy years in lieu of two

years)

- Allowing the large claim(s) to remain in the base losses, without applying loss development factors to the specific large losses
- Removing the large claim completely from the experience period, without building back any excess provision

Similar decisions were made for loss development and loss ratio trend selections. It made sense for NCCI to develop an approach that was standardized and uniformly applicable across its states.

2.3 Goals and Objectives

The goal of this research was to develop an aggregate ratemaking methodology, which would provide long-term adequacy of loss costs, rates, and rating values while recognizing the need for rate stability, particularly at a state level. It also aided in standardizing the methodology for handling individual large claims in aggregate ratemaking.

2.4 Defining a Large Event

Beginning in 2002, NCCI began working with EQECAT, a division of ABS Consulting. EQECAT is a modeling firm that has performed modeling for the California Earthquake Authority, a large earthquake pool, and has performed modeling extensively used in windstorm filings. The perils EQECAT modeled specifically for NCCI included the following:

- Terrorism
- Earthquake
- Catastrophic Industrial Accidents

Naturally, only injuries and losses resulting from the simulated events that related to workers compensation were a priority from NCCI's perspective.

It soon became clear to NCCI actuaries that the most practical approach for treating large catastrophic events in a ratemaking context was to exclude entirely from the NCCI ratemaking data any actual catastrophic events that occurred in the past due to these perils. The reasons for doing this included:

1. Actual catastrophic events of this nature that impact the workers compensation line of insurance have rarely occurred. Thus, they would not be predictive by their nature.

2. Actual catastrophic events would create volatility for a state's loss cost structure.

3. Direct carriers cannot put per-claim or per-occurrence limits on workers compensation policies. Therefore, events such as these would not be able to be excluded from workers compensation coverage without statutory actions by legislators.

4. Reporting and aggregating information from such large events would create difficult data reporting issues, conceivably involving multiple employers with multiple claims involving multiple insurance carriers.

5. Very few catastrophic events of this nature ever occurred, and thus, it was easy to remove data for these perils from NCCI's historical databases, provided the loss limitation dollar amount chosen was significantly large.

6. State of the art modeling techniques could be used to better estimate the cost of large events directly caused by one of the named perils.

After much discussion both internally at NCCI, and with external parties including carrier representatives and regulatory authorities, NCCI selected a threshold of $50 million for the specific perils of terrorism, earthquake, or catastrophic multi-claim occurrences. This threshold applies per occurrence, across all states for which claims arise from a single occurrence.

The entire ground-up amount of losses generated from a catastrophic multi-claim event is removed from the ratemaking data, not just the portion excess of $50 million. The loss costs derived from the modeling for the named perils include the cost of the first $50 million layer, as well as the excess.

NCCI removes the catastrophic occurrences first, and then caps individual claims secondly. The $50 million limit applies to individual claimant large losses that occur in workers compensation, but a more stringent limiting approach is applied. Large individual claims are treated state-specifically and the loss limitations are applied based on 1) the size of the state, and 2) the maturity of the claim. This procedure is described in more detail under the section entitled "Selecting a Threshold by State".

2.4.1 Capturing the Detail on Large Individual Claims and Events

For use in workers compensation ratemaking, NCCI collects the Policy Year Call (#3) and Calendar-Accident Year Call (#5), amongst other calls. The data calls are due by April 1 each year, and provide a year-end snapshot of twenty individual years of cumulative data and certain aggregate data on prior years. NCCI collects the data by carrier and by state, and it is reconciled to each carrier's Annual Statement. Because this data is reported on a summarized basis, large individual claims are not identified.

A review of the other databases NCCI maintains showed that a new call would be required to provide the information needed to implement the new large loss procedure.

NCCI designed Call #31, considering input from NCCI's Actuarial Committee and Data Collection Procedures Subcommittee, to capture detail on large individual claims greater than $500,000, and multi-claim occurrences from large catastrophic events. Extraordinary loss events that may involve multiple insurance lines of business, states, or data collection organizations are synchronized with the already existing catastrophe numbering system administered by the Insurance Services Office (ISO) for the property casualty industry. Consideration was given to having a large occurrence data call, but was not pursued due to practical considerations of the data providers.

A copy of Call #31 (i.e. Large Loss and Catastrophe Call) is included in Appendix C.

2.4.2 Selecting a Threshold by State

In order to perform the large loss limitation procedure in aggregate ratemaking, a threshold is needed at which individual claims will be limited.

Thresholds are state-specific. They were initially calculated based on a given state's on-leveled and developed experience period Designated Statistical Reporting (DSR) level premium from the previous year's filing. The initial dollar threshold is calculated as one percent of this premium figure—after all currently approved expense provisions have been removed—rounded to the nearest one million dollars. As an example, in a full rate state, this would mean standard premium at DSR level less all expenses multiplied by 0.01. This includes all policy (or accident) years in the experience period used in the most recent previous filing.

Essentially, a large individual claim is defined as one for which the impact of the claim under the prior methodology would result in an overall average statewide loss cost level change of at least one percent. Depending on the state, two or three years of experience will generally be used for the experience period. The advantages of this approach are that loss limitation thresholds:

1. Reflect the actual loss volume in each state,
2. Are inflation sensitive,

3. Temper the impact that one large claim may have on the overall statewide loss cost level indication, and
4. Install a standardized approach across states

As will be described in a later section, a lower threshold results in more claims being limited (and losses removed), but also results in a greater expected excess factor being applied. Conversely, if a larger threshold was selected, fewer losses are limited and removed from ratemaking, but the magnitude of the expected excess factor is smaller. NCCI had considered a two percent of DSR pure premium threshold, but after considering the two thresholds and observing the hypothetical results of previous loss cost filings for many states under both thresholds, a one percent threshold was chosen. One of the main reasons for selecting the one percent threshold was to provide stability.

2.5 Limited Loss Development

Historically, NCCI workers compensation aggregate ratemaking was based on using unlimited loss development factors applied to unlimited losses. The new methodology revised the loss development procedure to use limited loss development factors and apply them to limited base losses from the state's experience period. Thus, the ultimate losses derived are limited to a given threshold, analogous to the concept of basic limits losses, commonly found in other property - casualty lines of insurance. In other lines of insurance, the insured makes the decision as to how much coverage to purchase, and increased limit factors are computed and applied to derive the proper loss estimate for the limit sold on the policy.

The important difference that separates the workers compensation line of business from those other lines of insurance is that the benefits the coverage provides is based on statutory provisions, and essentially workers compensation provides unlimited medical benefits. In some jurisdictions, wage replacement benefits are also unlimited as to their duration. (One exception to that general statement is employers' liability coverage, with a basic limit of $100,000, and the employer has the option to purchase higher limits if desired.) Therefore, the unique coverage differences that workers compensation presents for NCCI actuaries is that the limited ultimate losses must be brought to an unlimited ultimate basis. This is addressed by the application of the excess ratio, which will be discussed in a later section of the paper.

NCCI computes loss development factors separately for indemnity and medical benefits. The large claims that are subject to loss limitation almost always have both an indemnity and medical component. Therefore, by limiting individual claims, a procedure had to be determined for capping the two components. The procedure NCCI uses to cap individual claims is discussed in a later section of this paper.

A difficult hurdle the NCCI actuaries had in implementing a new methodology based on limited loss development was how to handle the workers compensation tail factor, which is a 19th report to ultimate factor based on incurred losses including IBNR. In addition to the many well-documented, difficult challenges that exist estimating the tail factor in workers compensation was the challenge of answering the question, "How does one cap a bulk reserve?" A subsequent section of this paper is devoted to the details underlying the modifications made to the NCCI tail factor methodology.

2.5.1 De-Trending Loss Thresholds for Loss Development

The maturity of the claim is considered in the loss limitation that is applied. This is achieved through a process NCCI calls de-trending. De-trending is a procedure that progressively reduces the thresholds in historical periods to remove the distortion inflation has on loss development triangles. A detailed example may be found in the Appendix. Thresholds are de-trended each year by the corresponding change in the annual state-specific CPS wage index. This procedure was chosen for the following reasons:

1. State-specific wage changes will reflect indemnity inflation, and, through actual testing, provided a very reasonable proxy for medical inflation over a long period,
2. For consistency, as annual state-specific wage information is already used in other areas of the filing such as the wage adjustment used in loss ratio trend calculations, and
3. The medical CPI commonly used to approximate medical inflation is only available on a countrywide or regional basis rather than a state-specific basis.

NCCI performed actual data testing of the differences that would result in thresholds based on de-trend factors using annual medical CPI percentage changes in lieu of de-trend factors using CPS wage changes. The overall differences in loss cost level indications that resulted by state between the two de-trending approaches tested were hardly discernable. Thus, it was not clear that the countrywide medical CPI would better represent state-specific medical inflation than the state-specific CPS wage index.

Another very important, yet subtle, point to clarify is that the de-trending percentage does not represent, nor was intended to quantify, the total loss severity trend that occurred from year to year. It represents an inflationary amount to recognize the change in the average nominal costs of a claim over time. A loss severity trend in workers compensation measures much more than inflation. It measures changes such as the following:

- Changes in the utilization of benefits such as longer or shorter claim durations, or the propensity of claimants to return to work sooner or later than in the past,
- Changes in medical utilization, such as increased usage of more expensive treatments, medical procedures, pharmaceuticals with no generic equivalents, etc.,
- Changes to a state's administration of its workers compensation system, which may increase or reduce adjudication delays, alter dispute resolution processes, increase or decrease attorney involvement, etc.

If the de-trend percentage selected was the total loss severity trend that was incurred (which is very difficult to isolate and quantify), then it would be difficult for NCCI to accurately forecast loss costs. The historical data (adjusted for de-trending) used for loss development and to forecast the trend would be adjusted in such a way that the projected loss costs would be inaccurate by some implicit amount. Actuaries at NCCI tested two possible indices for de-trending, namely CPI inflation and changes in total claim severity. By developing simple models, it was demonstrated that the de-trending index should be based on inflation because it produces more predictive loss development factors than using claim severity for the de-trending index.

Using the simple models, the actuaries separately tested the impact on loss development factors and resulting ultimate losses of de-trending the cap using both an inflation index and a severity index. De-trending by an inflation index preserved the value of the age-to-age link ratios when average claim size is increasing due to inflation, which is what one would expect. When severity increases due to changes in claim duration, using inflation to de-trend preserved the value of the age-to-age link ratios for early reports, but link ratios for later reports would need to be adjusted to reflect lengthening durations. The alternative de-trending index, claim severity, resulted in distorted age-to-age link ratios at every age, making the resulting ultimate losses less predictive. In conclusion, the resulting ultimate losses were more predictive using an inflation index to de-trend large loss thresholds.

The initial state-specific thresholds were rounded to the nearest million for the policy year effective period when the new methodology was first implemented. For example, if the experience period DSR pure premium volume is $525M, a 1.0% threshold would imply a (rounded) large loss limitation of $5.0 million for the midpoint of the rate effective period. The rate effective period is also known as the "base year". The thresholds for each of the years prior to the effective period are not rounded.

Because NCCI actuaries develop a range of indications using policy year and calendar/accident year data, NCCI must de-trend large loss thresholds applicable to both sets of data. NCCI calculates the accident year de-trended thresholds first, and then calculates the de-trended policy year thresholds second. This is accomplished by weighting together two adjacent calendar/accident year thresholds using the state-specific distributions of premium writings by month. The reason for de-trending accident year thresholds first is that the CPS wage changes are on a calendar year basis, which is a better match with calendar/accident year data. A detailed example of the de-trending approach used may be found in Appendix A.

Once de-trended thresholds are computed for individual years, they are fixed at those dollar amounts going forward. In this way, the limited loss development factors will not vary from year to year due to revisions to the thresholds. In subsequent loss cost filings, the base year threshold will be trended forward utilizing actual CPS wages to the extent possible and then projected CPS wage changes. For example, if the AY 2006 threshold is $5,000,000, the newly calculated AY 2007 threshold will be $5,000,000 multiplied by the expected 2006-2007 CPS wage change.

In the future if a state grows or shrinks such that the threshold seems too high or low, NCCI may consider recalibrating the threshold at that time. Thresholds in years subsequent to the base year will not be rounded.

NCCI uses the same threshold and excess ratio for loss cost level indications based on paid and "paid+case" losses. Since large losses are reported to NCCI only for those claims with "paid+case" loss amounts greater than $500,000, the minimum de-trended threshold used in a state is $500,000, despite the fact that de-trending could generate a lower threshold.

Due to the size of DSR pure premium in the states of Florida and Illinois, and hence, the very large indicated threshold, the large loss procedure was not filed in those jurisdictions.

2.5.2 Applying the Loss Limitations to Individual Claims

In workers compensation ratemaking, losses are separately analyzed by type of benefit; namely, indemnity and medical losses. This impacts the method one chooses to limit a large claim. Further complicating loss limitation is that the traditional chain-ladder loss development techniques project ultimate losses using cumulative paid losses as the base (i.e. "paid" methods), as well as cumulative paid losses plus case reserve amounts (i.e.

"paid+case" methods).

In a given state, the NCCI actuaries review a range of indications based on both "paid" and "paid+case" methodologies. Therefore, capping large claims was more challenging than expected. After reviewing several loss limitation possibilities, the decision was made to use a methodology that limited payments first, followed by limiting the case reserves. The capping would be applied to individual claims within the experience period as well as within the historical loss development triangles. The myriad of other options considered by NCCI for capping claims is not included in this paper for sake of brevity.

NCCI uses proportional capping to allocate limited claim amounts. Limited loss amounts for claims above the threshold will be allocated to layers and to indemnity and medical in the proportion that their values contribute to the total value of the claim and the threshold. NCCI limits paid losses first, then limits the case reserves until the per claim threshold is reached. The remaining excess losses are subtracted from the aggregate unlimited losses in order to calculate limited losses for use in ratemaking. In order to understand the mechanics of how claims are limited, the following hypothetical illustrative examples are included:

Illustration 1. For claims that have pierced the threshold on a "paid" basis; State threshold = $1M:

UNLIMITED LOSSES ($Millions)	Paid	Case	Total
Indemnity	0.4	0.6	1.0
Medical	4.8	2.2	7.0
Total	5.2	2.8	8.0

In this situation, the resultant limited amounts are as follows:

LIMITED LOSSES ($Millions)	Paid	Case	Total
Indemnity	0.077	0	0.077
Medical	0.923	0	0.923
Total	1.0	0	1.0

The formula for deriving the limited paid amounts for indemnity and medical is:

(Indemnity paid/total paid) x threshold = (0.4 / 5.2) x 1.0 = 0.077

(Medical paid/total paid) x threshold = (4.8 / 5.2) x 1.0 = 0.923

Illustration 2: A claim that has not pierced the threshold on "paid" basis, but has pierced the threshold on a "paid+case" basis; State threshold = $1M:

UNLIMITED LOSSES ($Millions)	Paid	Case	Total
Indemnity	0.1	0.8	0.9
Medical	0.3	6.8	7.1
Total	0.4	7.6	8.0

In this situation, the resultant limited amounts are as follows:

LIMITED LOSSES ($Millions)	Paid	Case	Total
Indemnity	0.1	0.063	0.163
Medical	0.3	0.537	0.837
Total	0.4	0.6	1.0

In Illustration 2, the limited paid amounts are identical to the unlimited paid amounts. The "remainder of threshold" is computed as follows:

"remainder of threshold" = (threshold – total paid) = (1.0 – 0.4) = 0.6

The formula for limited case reserve amounts for indemnity and medical:

(Indemnity reserve/total reserve) x "remainder of threshold" = (0.8 / 7.6) x 0.6 = 0.063

(Medical reserve/total reserve) x "remainder of threshold" = (6.8 / 7.6) x 0.6 = 0.537

It is possible to have negative development on a limited basis for individual claims when the uncapped claim value increases. Usually this results from a shift in the proportion of paid losses and/or case reserves between indemnity and medical claim benefits from one evaluation to the next. This is simply a situation to be aware of, and should not significantly impact the limited loss development factors.

2.5.3 Tail Factor Adjustment

A limited tail factor (referred to as a capped tail factor in the terminology that is being introduced in this section) is needed to properly develop capped "paid" and "paid+case" losses to an ultimate basis. The previous NCCI tail methodology generates uncapped (i.e., unlimited) tail factors. Because claims with accident dates prior to 1984 are not reported on Call #31 (Large Loss and Catastrophe Call), it is not possible to adjust the state uncapped tail to a capped tail by removing the effect of losses excess of the state threshold. In order to convert the uncapped "paid+case" tail factor to a capped "paid+case" tail factor, we use a tail adjustment.

In general terms, the tail adjustment considers the relationship between a countrywide capped "paid+case" tail factor and a countrywide uncapped "paid+case" tail factor, and applies that relationship to individual state uncapped "paid+case" tail factors to generate state-specific capped "paid+case" tail factors.

First, a countrywide capped tail factor $CLDF_T$ is derived for the threshold T from countrywide uncapped tail factors, countrywide excess tail factors, and countrywide excess ratios, using the formula:

$$CLDF_T = \frac{1 - XS_T}{\left[\dfrac{1}{ULDF} - \dfrac{XS_T}{ELDF_T}\right]} \qquad (2.1)$$

Where,

$CLDF_T$ = Capped "paid+case" tail factor, 19th - to - ultimate, for threshold T

$ULDF$ = Uncapped "paid+case" tail factor, 19th - to - ultimate

XS_T = Excess ratio for threshold T, i.e., the ratio of losses excess of T to total losses at an ultimate report.

$ELDF_T$ = Excess "paid+case" tail factor, 19th - to -ultimate, for threshold T

All of the above factors are on a countrywide basis for medical and indemnity benefits combined, across all injury types. Thresholds are de-trended to the 19th prior report.

The numerator of the right hand side of (2.1), $1\text{-}XS_T$, is the proportion of total ultimate losses that are below the dollar threshold T. The denominator is the proportion of total ultimate losses below the threshold T reported at 19 years of maturity. To see this, note that $1/ULDF$ is the proportion of total unlimited losses reported at 19 years, and $XS_T/ELDF_T$ is the proportion of total losses that are excess losses reported at 19 years. The difference is the proportion of total losses less than the threshold reported at 19 years. The ratio of the numerator and denominator is the loss development factor. The adjustment factor F_T is

$$F_T = \frac{CLDF_T - 1}{ULDF - 1} \tag{2.2}$$

where $CLDF_T$ and $ULDF$ are as described above. The state capped tail factor is derived as follows:

$$SCLDF_T = 1 + F_T(SULDF - 1) \tag{2.3}$$

Where,

$SCLDF_T$ = State-specific capped "paid+case" tail factor, 19th - to - ultimate, for threshold T

$SULDF$ = State-specific uncapped "paid+case" tail factor, 19th - to - ultimate.

The state-specific uncapped "paid+case" tail factor, $SULDF$, is the state uncapped incurred (including IBNR) tail factor times the ratio of uncapped incurred (including IBNR) at 19th report to uncapped "paid+case" at 19th report. This is computed separately for medical and indemnity losses.

In practice, the factor F_T is applied to the uncapped medical and indemnity "paid+case" tail factors separately, to produce separate capped "paid+case" medical and indemnity tail factors.

An additional step is necessary to convert to a state-specific paid tail factor on a capped basis. The state-specific capped "paid+case" tail factor, $SCLDF_T$, is divided by the ratio of capped "paid" losses to capped "paid+case" losses at 19th report, separately for medical and indemnity losses. The de-trended dollar thresholds are used in the calculations of the "paid" to "paid+case" ratio for each state.

Unlimited "paid+case" tail factors, SULDF, will not be adjusted (i.e. reduced) if the unlimited "paid+case" tail factor is less than or equal to 1.000.

NCCI used Reinsurance Association of America (RAA) data [2] to calculate countrywide excess loss development factors (ELDFs). Data is submitted to RAA by reinsurers on an accident year de-trended basis. The RAA excess loss development factors are available only for combined "paid+case" losses (not "paid" losses) for five attachment point ranges (in thousands of dollars: $1-150, $151-350, $351-1500, $1500-4000, $4001 and greater) through an 18th report. NCCI fit curves through average period-to-period development factors for the lowest four ranges to extrapolate 19th-to-ultimate tail factors for each of the ranges (reported development for the highest range was deemed too volatile to provide a reliable base for extrapolation). A curve was fit through these four tail factors to extrapolate tail factors for higher attachment points. RAA produces excess loss development data every two years, which will allow NCCI to update the underlying factors periodically.

NCCI class ratemaking is generally not impacted by the new procedure at this time. NCCI concluded that it is appropriate for class ratemaking development factors to be unlimited for the following reasons: 1) The excess ratios are computed using the same 5th-to-ultimate factors as are applied to the loss dollars on serious claims, and 2) The first through fifth report link ratios derived from the Workers Compensation Statistical Plan data are currently based on unlimited losses. This is an area that is being explored in class ratemaking research.

An alternative considered was to apply the tail adjustment only to medical tail factors since it is believed that most development after 19th report, especially for large claims, occurs on medical rather than indemnity. One reason why this procedure was not followed is that adjustments to indemnity factors are usually small, since the uncapped indemnity factor is generally small, so most of the impact of the tail adjustment is to the medical tail factor. For states whose indemnity tail factor is large, it is likely that large loss development occurs on indemnity claims as well as medical, in which case it is appropriate to adjust indemnity tail factors.

A future consideration might be to incorporate state-specific excess ratios and unlimited "paid + case" tail factors, which are inputs into the tail adjustment calculation, in lieu of countrywide excess ratios and tail factors.

2.6 Application of the Excess Ratios

Adjusted per claim excess ratios will be used in calculating unlimited ultimate losses from limited ultimate losses. Excess losses are defined as the sum of the excess portion of claims above a given per claim threshold. NCCI produces proposed excess ratios with each loss cost or rate filing.

The excess ratio, XS_T, for a given threshold T, is defined as:

$$XS_T = \frac{\text{Expected Excess Losses Above Threshold } T}{\text{Expected Total Unlimited Losses}} \tag{2.4}$$

The ratio of excess losses to total unlimited losses is at an ultimate value. The excess ratio applied in the large loss procedure is on a per claim basis and varies by state as well as by threshold. This differs from an excess loss factor as excess loss factors are on a per occurrence basis, and also may include a provision for expenses.

Excess ratios are <u>not</u> adjusted when applied to different experience period years for purposes of calculating experience period loss ratios for ratemaking or for trend calculations. Therefore, in a given filing, the same excess ratio is applied to each year in the experience period. This is due to the fact that the dollar thresholds applicable to historical years are de-trended. By de-trending the threshold in the loss development and trend calculations, the proportion of losses above the threshold is preserved. Consider the following simple example. If a state's threshold is \$5.0M in 2005, and that corresponds to a 2.0% excess ratio, then a \$4.8M threshold in 2004 would also correspond to a 2.0% excess ratio, assuming that the 1.042 (1.042 = \$5.0M/\$4.8M) change in threshold values is solely due to inflation and correctly measures the actual rate of claim inflation in the state.

The adjusted, per claim excess ratio is applied as a factor, 1/ (1 -XS), to limited ultimate losses that have been on-leveled and trended to the midpoint of the proposed filing effective period. Similarly, the excess ratio applied has also been trended to the midpoint of the proposed filing effective period. Each policy period in the experience period has the same 1/ (1-XS) factor applied to both indemnity and medical losses, since the size-of-loss distributions are on a combined indemnity and medical basis. The excess ratios for aggregate ratemaking are a weighted average across hazard groups using expected losses as weights, and are based on the values contained in the state's latest approved filing.

2.7 Loss Ratio Trend

Indicated exponential loss ratio and severity trends, as well as econometric trends, are based on the losses that are derived from the large loss procedure. That is, trend indications are based on ultimate limited losses, where the limit is determined using the same de-trended thresholds by year as those used for loss development. This is consistent with the general approach that the ratemaking analysis is done on a limited basis, and is consistent with the fact that the excess ratio used in the filing implicitly contains inflationary trend over time.

2.8 Defense, Cost Containment and Adjusting and Other Expenses (formerly Loss Adjustment Expenses)

No changes to the calculation of Loss Adjustment Expense (LAE) factors were made as a result of using the aggregate large loss limitation procedure. This is a potential area of future study.

2.9 Summary of Filing Results for the Large Loss Methodology

NCCI filed the new large loss procedure for the first time in the filing season with effective dates from October 1, 2004 through July 1, 2005. The new procedure was filed in 32 states and it coincided with NCCI's revised excess loss factor procedure. Most state regulatory officials were satisfied with the implementation of NCCI's new methodology and its long-term advantages, and NCCI staff tracked results for each state on both an "unlimited" basis (i.e. the previous methodology) and the newly filed large loss procedure.

In the implementation year of the large loss procedure, the overall limited rate/loss cost level change was the same as would have been filed using the prior unlimited loss procedure when averaged across the 32 states where it was filed by NCCI. The indicated loss cost/ rate level change approved across individual NCCI states ranged from 0.973 to 1.028, indicating that the difference between the new methodology and the previous one, even at the extreme ends of the spectrum, were relatively modest and generally symmetric around 1.00.

In summary, as of May, 2006, the large loss methodology was adopted in 30 of the 32 states where it was filed. Colorado and Virginia have not adopted the change in methodology, and it has not been filed in Nevada, Illinois, or Florida.

3. THE USE OF CATASTROPHE MODELING IN WORKERS COMPENSATION

A secondary, but very important, goal of this paper is to discuss how modeling was used to derive loss cost provisions for catastrophic events due to terrorism, earthquake, and industrial accidents. In late 2002, NCCI filed Item B-1383, which was a national item filing proposing new loss cost/rate provisions by state for events that result from acts of foreign terrorism. This filing was designed to align with conditions of the Terrorism Risk Insurance Act (TRIA) passed by Congress in 2002.

In 2004, NCCI filed Item B-1393, which was a national item filing proposing new loss cost/rate provisions by state for events that result from the following perils: acts of domestic terrorism, earthquake (and tsunami, in certain states), and catastrophic industrial accidents.

Almost all states approved the voluntary loss cost and assigned risk rate provisions that NCCI filed, and many workers compensation insurers now apply these values to payroll in hundreds of dollars to determine the premium it generates. As part of Item B-1393, this premium is applied after standard premium is determined, and is not subject to any other modifications including, but not limited to, premium discounts, experience rating, retrospective rating, and schedule rating. It is an additive amount applied in the calculation of a policy's estimated annual premium initially charged to an employer, which is subject to a final audit when payroll is finalized at policy expiration.

3.1 Definition of the Perils

Terrorism, earthquakes, and catastrophic industrial accidents can result in losses of extraordinary magnitude for workers compensation. While the exposure is real, the absence of a large event in recent history within the data means that the current loss costs and rates do not provide for this type of exposure. NCCI's new approach is to exclude losses resulting from these major catastrophes once a provision for their exposure is contained in the loss costs and rates. The threshold for each of these exposures is $50 million. The modeling results described below assume that all events exceeding $50 million of loss for workers compensation would be removed from ratemaking on a first-dollar basis.

For purposes of the modeling, the following definitions apply:

• ***Acts of Foreign Terrorism:*** All acts of terrorism within the scope of TRIA with aggregate workers compensation losses in excess of $50 million. This is defined as:

a. Any act that is violent or dangerous to human life, property, or infrastructure; and

b. The act has been committed by an individual or individuals acting on behalf of any foreign person or foreign interest, as part of an effort to coerce the civilian population of the United States or to influence the policy or affect the conduct of the U.S. Government by coercion

• ***Domestic Terrorism:*** All acts of terrorism outside the scope of TRIA with aggregate workers compensation losses in excess of $50 million.

• ***Earthquake:*** The shaking and vibration at the surface of the earth resulting from underground movement along a fault plane or volcanic activity where the aggregate workers compensation losses from the single event are in excess of $50 million.

• ***Catastrophic Industrial Accident:*** Any single event other than an act of terrorism or an earthquake resulting in workers compensation losses in excess of $50 million.

Note that for workers compensation, obligations to pay benefits are dictated by state law, and exclusions of these perils are not possible without statutory changes. Because TRIA has a unique mechanism for triggering federal reinsurance, separate statistical codes were created to capture premium credits or debits reported to NCCI for the Foreign Terrorism catastrophe provision and the catastrophe provision covering the other three perils, commonly referred to as DTEC (Domestic Terrorism, Earthquake, and Catastrophic Industrial Accidents).

3.2 Overview of the Approach to Determining Loss Costs Using Modeling

Beginning in 2002, NCCI began working with EQECAT, a division of ABS Consulting. EQECAT is a modeling firm that performed modeling for the California Earthquake Authority, a large earthquake pool, and performed modeling extensively used in windstorm filings. Serving the global property and casualty industry, EQECAT is known as a technical leader and innovator in the development of analysis tools and methodologies to quantify insured exposure to natural and man-made catastrophic risk. EQECAT developed three models for NCCI. These models address the potential exposure to workers compensation for terrorism, earthquake, and catastrophic industrial accidents. The models are described in detail in the following sections.

The framework of determining loss costs/rates using the modeling can best be described in the following manner:

1. Events are simulated for specific states using qualitatively defined thresholds. Some events modeled may actually result in no losses. The qualitative thresholds used by peril were:

- Large industrial accidents likely to cause at least two worker fatalities or at least ten worker hospitalizations,
- Terrorist attacks with the potential to cause at least $25M in workers compensation

losses according to the magnitude of physical event, and

- All possible earthquakes are modeled

2. Expected Annual Losses (EAL) were calculated for every state and peril analyzed. These losses were obtained using the casualty counts generated from the simulated events and by using state-specific benefit payments by injury type by state provided by NCCI.

3. Using the loss exceedance distribution underlying the EAL estimates, NCCI actuaries remove from the distribution events that do not exceed the selected dollar threshold of $50M. See Appendix B for more explanation.

4. The modified EAL was divided by the number of full-time-equivalent (FTE) employees and divided by the annual wage per employee (based on Current Population Survey or CPS) to derive a pure loss cost per $100 of payroll.

5. This was computed by peril and summed to determine the catastrophic (DTEC) provision. (Note: the foreign terrorism provision was computed similarly except for a final adjustment to remove the portion of losses from events that exceeded the federal backstop provided under TRIA.)

3.3 Modeling the Three Perils: Terrorism, Industrial Accidents, and Earthquake

Separate EQECAT models have been utilized to provide estimates of the risks to workers compensation insurers due to the following perils:

- Terrorism events
- Industrial accidents
- Earthquake ground shaking

All three models consist of the following primary components:

- Definition of the portfolio exposures
- Definition of the peril hazards
- Definition of the casualty vulnerability
- Calculation of loss due to casualty

Each of the above components is described separately below.

3.4 Portfolio Exposures Within the Models

The location, number, and types of employees are needed to characterize the risk exposures to all three perils listed above. Business information databases were used to obtain the addresses of businesses and the estimated number of employees assigned to each location. For the perils of terrorism events and industrial accidents, the exposures were aggregated to the census block level (typically a city block). This aggregation level was suitable for terrorist events and industrial accidents that span hundreds of meters. Since the definition of seismic hazard data is rather refined, the exposure data at each work site were used.

The number of workers at each aggregate level (census block or work site) was prorated to approximately account for part-time workers, workers absent for various reasons, and the self-employed. The workers were then grouped into five NCCI industry groupings: Manufacturing, Contracting, Office & Clerical, Goods & Services, and All Others. Certain government classifications not covered by workers compensation were excluded.

In addition to the employee information, required exposure data for the earthquake peril include information on the buildings where the employees are located. Building information consists of the structure type and age.

Furthermore, the number of employees used for the earthquake peril was defined for four different work shifts:

- Day shift
- Swing shift
- Night shift
- Weekends and holidays

Since the number of casualties vary depending on the time of the day and day of the week when the earthquake strikes, it is necessary to determine the number of employees for the different work shifts. The day shift accounts for most of the workers compensation exposure.

The definition of exposure by work shift was only performed for the earthquake peril. Earthquakes are natural disasters and can occur at any time in a random manner. Therefore, it is considered important to "average" the losses from all possible outcomes. Conversely, terrorism events and industrial accidents can be considered to occur most likely during the day shifts when there are more people and activities. Terrorism events are planned to inflict maximum casualties, and industrial accidents are more prone to occur during the peak hours of activities.

3.5 Peril Hazards Within the Models

3.5.1 Peril Hazards for Terrorism Events

EQECAT assembled data on the insurers' exposure and subjected that exposure to a large number of simulated terrorist events. These simulated terrorist events consist of three primary elements:

1. Weapon types
2. Target selection
3. Frequencies of weapon attacks

A brief description of each element follows.

<u>1. Weapon Types</u>

Specific weapons were selected from the range of known or hypothesized terrorist weapons. The selection process considered weapons that have been previously employed, weapons that could cause large numbers of casualties, or weapons that would be more readily available. In some cases a "likely" or "practical" weapon size (or quantity of agent) was selected; in other cases, a range of weapon sizes was selected, in part, to reflect standard quantities that might be available. Some of the selected modes of attack are listed below.

a) Blast/Explosion
b) Chemical
c) Biological
d) Radiological
e) Other

<u>2. Target Selection</u>

A target is the location of a terrorist attack and, in the model, represents the locus of a casualty footprint. An inventory of targets having the following characteristics was created such as:

- Tall buildings—10 stories and higher
- Government buildings—with a large number of employees or serving a critical or sensitive nature (e.g., FBI office).

- Airports—major
- Ports—major
- Military bases—U.S. armed forces
- Prominent locations—capitol buildings, major amusement parks, etc.
- Nuclear power plants—operational
- Railroads, railroad yards and stations—freight lines for railroad cars carrying chemicals
- Dams—large ones near urban areas
- Chemical facilities—emphasizes those with chlorine and ammonia on site

Nuclear power plants, dams, and chemical facilities receive only specific casualty footprints. Other locations are assigned more than one type of terrorist weapon. Some footprints have no specific target but are distributed at regular intervals throughout the urban area. This spreads out the effect to a larger population in the urban area. Mobile release anthrax is not located at any target but located in the general downtown area in major metropolitan areas.

<u>3. Frequency of Weapon Attack</u>

The relative likelihood of a type of attack occurring at a target location is represented by an assigned (annual) frequency. The significance of an attack's frequency is in its relationship to other attacks. Attack frequency is based on the following considerations:

- Availability of weapons
- Attractiveness of target
- Relative attractiveness of the region to other regions based on various theories

For footprints that are atmospheric releases of chemical, biological, and radiological agents, wind direction affects the assigned frequency. The frequency for each wind direction is weighted by the likelihood of the wind blowing in that direction based on historical wind speed and direction measurements for the region.

3.5.2 Peril Hazards for Catastrophic Industrial Accidents

Industrial accidents are characterized by the following elements:

- Facilities where industrial accidents occur

- Accident types
- Frequencies of accidents

Facilities

Facilities capable of large industrial accidents resulting in casualties above a threshold were identified from several public and commercial data sources. The facilities considered as potential sources for large industrial accidents are identified below:

- Refineries
- Chemical plants (oil, gas, petrochemical, etc.)
- Water utilities
- Power utilities
- Other manufacturing plants

Accident Types

Depending on ,the peril, the atmospheric conditions, the plant configuration and location, etc., the footprint of an accident could reach beyond the plant boundaries and affect workers in adjacent facilities and beyond. The perils considered in the study were broadly classified into three categories: chemical releases, large explosions, and all other accidents.

Chemical Releases: Chemicals considered included chlorine, anhydrous ammonia, and other nonspecific chemicals. A range of potential atmospheric releases of chemicals was considered in the analysis. The range encompassed an upper quantity represented by the total amount of chemical stored on site and, in some cases, identified in the facility's Risk Management Program submittal as the worst-case scenario, and a lower release quantity representing the minimum release quantity that could produce consequences to meet the threshold definition of large industrial accidents. A continuous range of release quantities was considered within the range.

All of the scenarios considered were modeled probabilistically and included the likelihood of the releases and their consequences as described above.

Large Explosions: Explosion simulation software is used to estimate blast pressures and consequences of the explosion in terms of casualties. These footprints were varied probabilistically to simulate the variability in the effects of an explosion. The size of the explosion varied by facility. The largest explosions were modeled to occur at oil refineries, where a significant potential for explosions exists.

All Other Accidents: In addition to the above accident types, a smaller event was considered at all modeled facilities to simulate all other industrial accidents such as fires, explosions, confined space accidents, structure and component collapse, and all other random accidents that meet the threshold damage criteria of large industrial accidents.

Frequencies of Accidents: The frequencies of occurrence of large industrial accidents in each of the modeled states were derived based on historical fatality and injury data. Frequencies of extreme events, which are very large and very rare, were based on ABS Consulting expert opinion and historical data.

The relative likelihood of the three categories of perils simulated in the analysis was derived from historical data and varies by state.

3.5.3 Peril Hazards for Earthquakes

Regional Hazard

The calculation of annualized losses requires a probabilistic representation of the location, frequency, and anticipated ground shaking of all earthquakes that can be expected to occur in the region. The characterization of the location and frequency of earthquakes comprise what is commonly known as a seismotectonic model.

One component of the seismic hazard model is the source zonation. Source zonation entails identifying potential seismogenic sources that can affect the site. These sources can either be faults or diffuse zones of seismic activity, commonly referred to as area sources and background seismicity. Each source zone represents a fault or area in which earthquakes are expected to be uniformly distributed with respect to location and size. Background seismicity is distinguished from an area source by the way that earthquake locations are treated. Earthquakes associated with background seismicity are allowed to have recurrence frequencies that smoothly vary over a region. Both area sources and background seismicity can include large earthquakes and are intended to model areas containing hidden or unknown faults or known faults, which are too numerous to be modeled individually. Earthquake source zones are identified from information on the geology, tectonics, and historical seismicity of the region.

The seismic hazard model also integrates the recurrence frequency of earthquakes. For each of the earthquake source zones, an earthquake recurrence relationship is developed. For area sources and background seismicity, this relationship is developed using an appropriate earthquake catalog, which is a listing of historically recorded or documented earthquakes. The catalog is analyzed for completeness by determining the time period over which all earthquakes of a given magnitude are believed to have been reported.

Magnitudes are converted to a consistent magnitude measure (e.g., moment magnitude, MW) for use with the strong-shaking attenuation relationships (described in the next section) and for the determination of earthquake recurrence relationships.

Faults are modeled by a characteristic earthquake model or a Gutenberg-Richter recurrence relationship, or both, depending upon the available geologic information. The characteristic earthquake model assumes that earthquakes of about the same magnitude occur at quasi-periodic intervals on the fault. The characteristic recurrence relationship is consistent with paleoseismic and historical earthquake data on individual faults. For most faults, the recurrence relationships are constrained to be consistent with known geologic deformation along the fault, since there are usually very few historical earthquakes from which to develop a reliable earthquake recurrence relationship.

The maximum magnitude for each earthquake source zone is estimated from the published literature, from comparisons with similar tectonic regimes, from historical seismicity, and from the dimensions of mapped faults.

The seismic hazard model simulates approximately 2,000,000 stochastic events across the United States.

Site Hazard Severity

Attenuation relationships are used to predict the expected amplitude of ground shaking at a site of interest knowing an earthquake's magnitude and the distance from the fault to the site. The ground shaking is characterized by one or more ground-shaking parameters, the most notable of which are peak ground acceleration (PGA), response-spectral acceleration (Sa), and Modified Mercalli intensity (MMI). These predictions are made for a uniform soil condition. Attenuation relationships are chosen to correspond as closely as possible to the tectonic environment of the region, since regional differences in earthquake source characteristics, crustal propagation properties, and site-response characteristics are known to have a significant effect on the observed ground shaking.

Soil amplification factors are used to modify the ground-shaking parameter calculated for a uniform soil condition for the specific soil conditions at the site of interest. These factors are different for each ground-shaking parameter. They are defined in terms of one or more site categories (or classes), each representing a specific set of site-response characteristics.

Soil categories are defined in terms of simple qualitative or quantitative site descriptions, such as surface geology and shear-wave velocity (the speed at which seismic waves travel through the soil deposit, a measure of the strength of the deposit).

The effect of local soil conditions within each individual zip code was taken into account. In general, soft soil sites will experience higher earthquake motions than firm soil or rock sites for comparable locations relative to the earthquake fault rupture zone, thereby increasing the likelihood of damage to buildings on soft soil for a given earthquake.

3.6 Casualty Vulnerability

Casualty vulnerability establishes the casualty levels to various peril event magnitudes. While the casualty vulnerability for terrorism events and industrial accidents are rather similar, the casualty vulnerability for earthquakes is established rather differently.

3.6.1 Casualty Vulnerability for Terrorism Events

The casualty footprint of a weapon is a measure of the physical distribution of the intensity of the agent as it spreads out from its initial target. The effects of each type of weapon will vary with the size of the weapon, with atmospheric conditions, and in some cases with local terrain. If detailed knowledge is available, a correspondingly detailed simulation of the effects is possible, but it would be time-consuming to perform. In a large-scale nationwide analysis with millions of simulated events, where local atmospheric and terrain are only generally known, a simpler, more generalized simulation is necessary. The simplifications necessary to efficiently model footprints of weapons effects are described below.

For conventional blast loading, blast simulation software is used to estimate casualties in various urban settings where the geometry and height of the buildings are varied. The results of these detailed simulations are used to develop simplified blast attenuation functions that vary with distance and with the general terrain.

For conventional blast loading, the footprint is defined as a decreasing function of distance from the source of the blast. The casualties for nuclear blast can be estimated on the basis of empirical data resulting from wartime and nuclear test experience. Casualties are assumed to be a function of distance from ground zero with the source located either at ground level or at a relatively low altitude. A simplified, conservative casualty footprint was created to encompass the range of conditions that could exist. Long-term radiation effects were not considered.

The casualty effects for aircraft impact are very much dependent upon the details of the event, so much so that only a simple, conservative footprint can be employed. A simplifying assumption is made that the extent of the footprint is a function of the height of the building.

For chemical, biological, and radiological agent releases, a plume is formed that is influenced by atmospheric conditions and by the terrain. The footprint of the cumulative dose that is deposited by a plume over time was calculated using the simulation software, MIDAS-AT™ (Meteorological Information and Dispersion Assessment System—Anti-Terrorism™). Terrain conditions were assumed to be "rough" to conservatively approximate a general urban terrain. The wind direction was assumed to be unchanging. The plume footprint was calculated for low, medium, and high wind speeds and for three different atmospheric turbulence conditions. Any of the footprints could then be oriented in each of eight compass directions. Most of the footprints were truncated after an elapsed time of about two hours to account for successful evacuation.

Casualties due to dam failure are approximated using simple hydraulic relationships and assumptions made about the terrain over which the water will flow. The resulting footprint varies as depth of water (and casualty) decreases with distance away from the dam.

The analysis methodology is to apply a casualty footprint to an assigned target and to calculate the extent of casualties to the covered workers within the footprint. For chemical, biological, and radiological footprints, the dose to each employee is calculated, and a conversion is made to the degree of casualty (outpatient treatment, minor/temporary disability, major/permanent disability, and death). Degree of casualty is then converted to loss based upon the average costs by injury type provided by NCCI. The average costs provided vary by state.

3.6.2 Casualty Vulnerability for Industrial Accidents

As discussed earlier in Section 3.2, three accident types were considered in the Industrial Accidents study: chemical releases, large explosions, and all other accidents. The latter category includes a variety of accidents that are localized in nature and affect workers in a small perimeter, the size of a building. These smaller scale accidents were simulated as small blasts.

The methodology used to model chemical releases and blasts is the same as in the

---

™ ABS Consulting Trade Mark

terrorism model described earlier.

3.6.3 Casualty Vulnerability for Earthquakes

Workers' casualties due to earthquakes are directly correlated to the damage extent incurred by the buildings in which they work. Therefore, casualties due to earthquakes are estimated in two sequential stages:

- Estimation of building damage
- Estimation of worker casualties based on the building damage

<u>Building Damage at the Workplaces</u>

Individual building vulnerability functions, that is, the probability of building damage given a level of ground shaking at the site, depends of the structure type, the age of construction, and the building height. Vulnerability functions account for variability by assigning a probability distribution bounded by 0% and 100% with a prescribed mean value and standard deviation. The vulnerability functions were based on historical damage data and insurance claims data—including the analysis of over 50,000 claims from the Northridge and other earthquakes.

The probability distributions of ground shaking at the site and vulnerability functions are combined to estimate the probability of building damage for each earthquake event. The probability of damage at the site level is also combined probabilistically, accounting for correlation in ground shaking between zip codes and in damage level between the same and different structure types within and between zip codes.

Note that considerable randomness exists in earthquake damage patterns where randomness denotes the irreducible variability associated with the earthquake event. Randomness as characterized by the following parameters:

- Ground shaking
- Damage to the average structure of a given class at a given level of ground shaking
- Each structure's seismic vulnerability relative to the average structure of its class

Modeling uncertainty, the lack of knowledge in characterizing each element of the model, is statistically combined with randomness and correlation to estimate overall variability in damage and loss to the entire portfolio.

<u>Casualties Due to Building Damage</u>

Workers' casualty data resulting from earthquakes is very scarce in the United States. EQECAT is constantly using data from the most recent earthquakes worldwide to update its casualty functions, which correlate building damage to casualties. Because of differences in building design codes and construction practices, data from earthquakes outside the U.S. are adapted to local U.S. conditions. This adaptation takes into consideration building damage, the state, and its resulting casualties.

EQECAT's proprietary workers compensation casualty rate functions are defined for four injury types: death, severe/major, minor/light, and medical-only.

3.7 Calculation of Loss Due to Casualties

Average costs by injury type were provided by NCCI and used in calculating losses due to workers' casualties. The same average costs were applied to all three perils.

Earthquake exposures were defined for different work shifts. The number of casualties by work shift for each work site and earthquake event is estimated prior to the application of the average costs.

3.7.1 Calculation of Loss Due to Tsunami

Although all coastal states on the West Coast are prone to tsunamis, only Alaska was analyzed for this peril.

Alaska has a higher worker rate near the shore in inundetable zones and its coastline is in close proximity to the subduction zone capable of triggering tsunamis. In addition, in remote locations of Alaska, workers compensation extends coverage after the employee leaves the immediate worksite. Other states such as Oregon and Hawaii can benefit from a warning advantage that would reduce the impact of tsunamis generated far away.

A simplified model was formulated to estimate workers compensation loss due to tsunami inundation. This model is based on tsunami modeling developed for Japan, which makes use of historical data to derive a relationship between earthquake moment magnitude (Mw), distance from the earthquake rupture to the shore, and direct or indirect exposure to the wave to determine the run-up height of a tsunami wave. The quantity of historical data needed to develop such a relationship is not available for Alaska; however, the model adopts the Japanese method where the detailed physics of the wave are not being calculated.

<u>Injury Rate</u>

Casualties due to tsunami run-up are estimated by assuming a simple relationship between depth of inundation and the likelihood of being in one of four NCCI injury classes (medical-only, minor permanent partial /temporary disability, major permanent partial /permanent disability, and death). There is scarce data available and the conditions under which the casualties occur is extremely variable. For this simplified approach, the injury relationships were subjected to the 1964 Mega-Thrust earthquake and the relationships calibrated to produce roughly the casualties suffered in the event.

Earthquake Modeling

The source of tsunami in Alaska is limited to the lengthy subduction zone that lies along the undersea trench that stretches from about Seward to the tip of the Aleutians. This subduction zone produces earthquake magnitudes estimated to be as large as Mw 9.2. Only the larger magnitude events have a potential for causing tsunami. For this analysis, magnitudes down to Mw 7.7 were considered.

Based on the geometry of the subduction zone adopted from the USGS, ruptures of magnitudes between Mw 9.2 and Mw 7.7 were placed along the length of the trench. The frequency of each event, as a function of magnitude, was derived from an analysis of the earthquake catalog for the region.

For each earthquake rupture, the surface distance between any location on the rupture plane and each near-shore business location was calculated.

Tsunami Analysis

The computations were performed for each earthquake rupture and for each site. Given the magnitude of the rupture and the distance from the ruptures to the site, the simplified equation estimates the run-up height. The difference between the elevation above sea level and the run-up height determines the depth of inundation.

Inundation depth is then used to determine the percentage of employees who are in each injury category. From the number of employees at the location, the total casualty cost is estimated using the mean costs for each injury category. The cost is multiplied by the event frequency, and aggregated by NCCI occupation class and by county.

The losses from earthquake shaking and tsunami were combined through summation. This conservative treatment neglects the potential for overlap in casualties caused by shaking and by tsunami.

3.8 Deriving Loss Costs from the Modeling

As described earlier, Expected Annual Losses (EAL) were calculated for every state and peril analyzed. These losses were obtained using the casualty counts generated from the simulated events and by using state-specific benefit payments by injury type by state provided by NCCI. The losses do include self-insured employers.

Using the loss exceedance distribution underlying the EAL estimates, NCCI actuaries remove from the distribution events that do not exceed the selected dollar threshold of $50 million. See Appendix B for a detailed hypothetical illustration of this process.

The modified EAL was divided by the number of full-time-equivalent (FTE) employees and divided by the annual wage per employee based on Current Population Survey (CPS) to derive a pure loss cost per $100 of payroll. Note the number of employees also includes self-insured employers.

This was computed by peril and summed to determine the catastrophic (DTEC) loss cost/rate provision. (Note: the foreign terrorism provision was computed similarly except for a final adjustment to remove the portion of losses from events that exceeded the federal backstop provided under TRIA.)

3.9 Other Insights from the Modeling

The relative magnitude of different catastrophes varies based on the time horizon. In workers compensation, for shorter time horizons, industrial accidents are expected to generate the largest expected losses. However, for very long term horizons, earthquakes generate the largest expected losses.

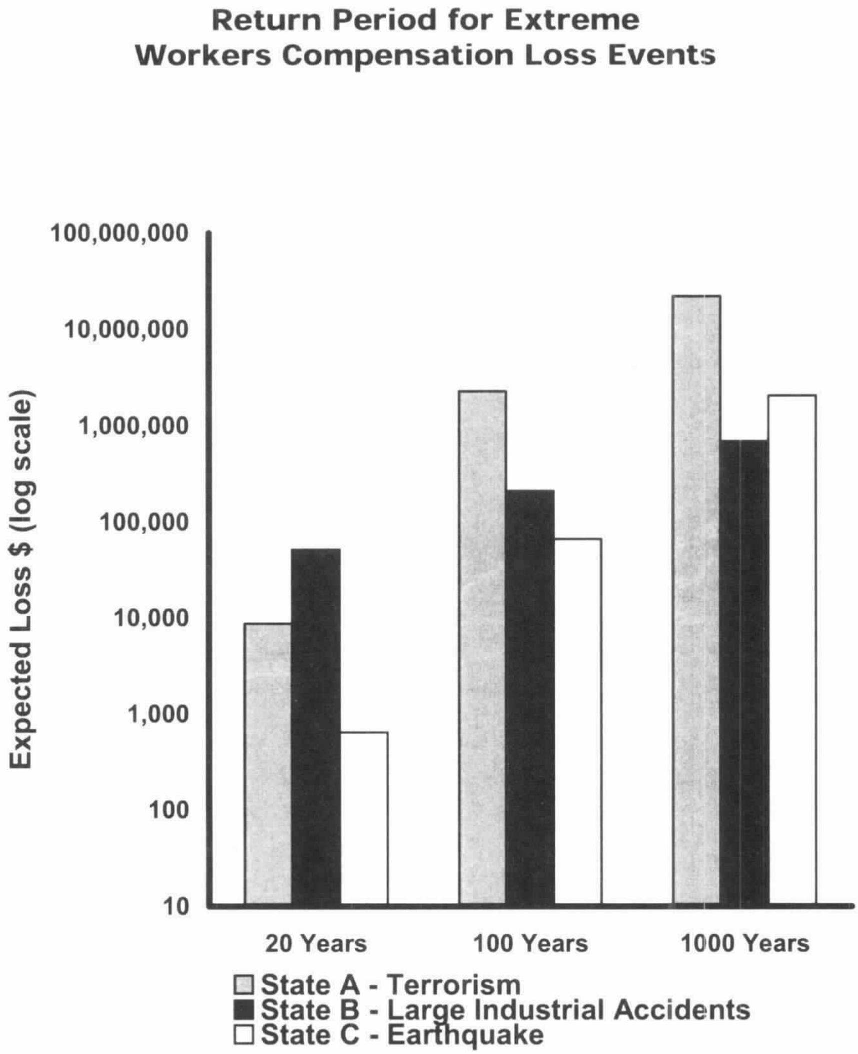
When talking about the relative length of time horizons, one references the return period. The return period for extreme events is defined as the expected length of time between occurrences. It is an approximate measure of frequency per unit of time.

The following bar chart shows that the relative magnitude of perils varies based on the time horizon. Here, we look at the three different perils in three different states: industrial accidents, terrorism, and earthquake. The focus is on the different time horizons.

The first observation is that regardless of peril, over longer time horizons, the expected loss amounts increase. This is because the very largest catastrophic events dominate the calculation of expected loss, despite their very low return period. This is generally true across all states modeled, and all three perils.

The second observation is based on comparing the expected losses by peril relative to each other based on the different time horizons. The following chart shows the differences at a 20-year, 100-year, and 1000-year return period. When taking a 20-year time horizon, industrial accidents are expected to be the largest events in terms of generating expected losses, and earthquake is the lowest. When a 100-year time frame is viewed, industrial accidents rank second, and terrorism events are first. At a time horizon of 500 years or more, terrorism events rank first, followed by earthquakes, and then industrial accidents.

Note the following chart shows results for three different states. This is still representative of the pattern that would likely result in a single state over the same time horizons, but the model results underlying this analysis did not include a single state modeled for all three perils.


Return Period for Extreme
Workers Compensation Loss Events
Expected Loss $ (log scale)
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10
20 Years
100 Years
1000 Years
State A - Terrorism
State B - Large Industrial Accidents
State C - Earthquake

3.10 The Pros and Cons of Using Catastrophe Modeling in Workers Compensation

Catastrophic events are a low frequency occurrence with very high severity, and cannot be adequately addressed through standard actuarial techniques to quantify risk. The data on such events is limited to a very small number of historical events — often without an event having been observed in a state.

Used in conjunction with the actual historical data, stochastic simulations were used in the modeling to provide additional data points. Repeat simulations of an event provide a broader perspective of the possible outcomes. Variations of parameters are also modeled and result in a comprehensive stochastic event set.

Modeling is being used extensively in the insurance and reinsurance industries. State regulators are scrutinizing the models, more fully understanding how they operate, and asking better questions to learn more and more. Over time, there has been a wider acceptance of catastrophe modeling by regulatory officials.

As for disadvantages, there are several parameters with varying levels of uncertainty involved in each of the hazard, vulnerability, casualty, and loss modules which are integrated in these complex models. These uncertainties lead to differences between models and raise questions among regulators who have to determine the validity of these tools which are becoming increasingly used in rate making.

3.11 Possible Future Enhancements to the Catastrophe Modeling

The catastrophe models rely heavily on underlying databases which contain information on the different parameters used in the analysis. To the extent that the refinement and quality of these databases increases, the result may be a reduction in the margin of uncertainty in the final results.

An enhancement to the workers compensation models described earlier would result if a database containing the employment data at each business location and for each work shift were updated regularly. This would improve the estimates of the numbers of workplace injuries and the subsequent modeled loss estimates resulting from events emanating from the perils of terrorism, earthquake, and catastrophic industrial accidents.

Some other examples of information or databases which might improve the estimation of the workers compensation loss estimates follows, organized by peril.

Earthquake Peril

A more refined soil database would be a possible enhancement if used in the earthquake model for workers compensation. It could allow for better estimation of the site amplification of the ground motion, which in turn is used to calculate the building damage, and hence, the resulting casualties among its occupants.

Also, building structure information, if more accurately defined, would allow for the use of a more fine-tuned building vulnerability function. In the absence of such information, assumptions are generally made based on information that could possibly be dated.

The casualty rate functions allow the estimation of the casualties by injury type in different building structures. These functions are developed from limited earthquake casualty data and as more data is collected from future occurrences, loss estimates could be improved as the estimation of casualties improved.

Catastrophic Industrial Accidents and Terrorism Perils

The potential for extreme industrial events needs to be constantly reviewed based on safety regulations and their enforcement, emergency planning, and medical emergency care. These conditions may vary greatly over time and across facilities. This type of information directly impacts the frequency assumption underlying the loss cost. As this information becomes more refined, one should be better able to target the frequency assumption.

Other areas of possible enhancement include obtaining more refined information on the potential target sites. In particular, those sites storing toxic chemicals need to be constantly updated as some plants open or close or change their product lines. The nature and quantities of the toxic chemicals need also be kept current.

For terrorism, the statements above apply with respect to the potential target sites. Also, event frequencies need to be regularly evaluated based on current conditions and the possible threats they may generate. The frequency assumption, as always, is very important to determining the appropriate loss cost levels for all perils.

3.12 Using Models Outside the Actuary's Expertise

The author relied upon the expertise of other NCCI actuaries, whose work product has been described in parts of the modeling discussion presented. Such information has been documented in accordance with ASOP No. 38.

The NCCI actuaries relied upon simulation models supplied by EQECAT for calculating expected losses due to the earthquake perils. The accuracy of these models heavily depends upon the accuracy of seismological and engineering assumptions included.

The NCCI actuaries also relied upon simulation models supplied by EQECAT for calculating expected losses due to the perils of terrorism and catastrophic industrial accidents. The models produce estimated losses due to physical, chemical, and biological terrorist acts. They also produce estimated losses due to chemical releases and explosions at industrial plants, and both perils include the input and opinions from experts in related fields and experts at ABS Consulting. The accuracy of these models heavily depends upon the accuracy of meteorological, engineering, and expert claim frequency assumptions.

4. CONCLUSIONS

This paper documents several important changes that have been implemented in the aggregate ratemaking process used to determine indicated workers compensation loss cost and rate changes by state. The changes NCCI implemented support the long-term goals of adequacy and stability of loss costs and rates based on the explicit consideration of how to treat large events consistently from state to state in the ratemaking methodology.

This paper also serves to document for the first time in CAS literature how computer modeling was used in workers compensation.

Acknowledgment

The author acknowledges the work of Ia Hauck, John Robertson, John Deacon, and numerous other NCCI staff members, whose tireless efforts allowed NCCI to complete the aggregate ratemaking research and testing that allowed for the implementation of the new methodology. The author also acknowledges the contributions to this paper by Jon Evans and Barry Lipton, whose work with EQECAT allowed NCCI the ability to produce the loss cost provisions for catastrophic events, and Omar Khemici, Andrew Cowell, and Ken Campbell of EQECAT, whose insights and work on performing the modeling was the basis for many of the results shown in the paper. I would also like to thank EQECAT, a division of ABS Consulting, for allowing me to disclose an overview of their modeling techniques, much of which is proprietary.

Appendix A – Example of De-Trending Procedure

State X - Effective 9/1/2004

Calculation of Base Threshold (Using information from latest approved filing):

Experience Period of Latest Approved Filing	1PY / 1AY
On-leveled, Developed Premium for PY 2001	247,605,878
On-leveled, Developed Premium for AY 2002	240,782,386
Experience Period On-leveled, Developed Premium	488,388,264
Factor to Remove Expenses	1.000
Experience Period On-leveled, Developed Premium Excluding Expenses	488,388,264
1% of the Total Experience Period Premium	4,883,883
Threshold for the Base Year	5,000,000
Midpoint of the Proposed Filing Policy Period (Base Year)	8/13/2005

Calculation of De-trended Thresholds:

(1) Midpoint	(2) CY	(3) Actual CY CPS Wage	(4) Change in Col (2)	(5) Year	(6) AY Threshold	(7) PY Threshold
7/1/1984	**1984**	294.17	1.016	**1984**	2,346,511	2,360,628
7/1/1985	**1985**	298.84	1.010	**1985**	2,384,055	2,393,019
7/1/1986	**1986**	301.72	1.064	**1986**	2,407,896	2,465,839
7/1/1987	**1987**	320.92	1.021	**1987**	2,562,001	2,582,231
7/1/1988	**1988**	327.57	1.063	**1988**	2,615,803	2,677,766
7/1/1989	**1989**	348.30	1.024	**1989**	2,780,599	2,805,691
7/1/1990	**1990**	356.51	1.054	**1990**	2,847,333	2,905,145
7/1/1991	**1991**	375.76	1.101	**1991**	3,001,089	3,115,058
7/1/1992	**1992**	413.85	1.004	**1992**	3,304,199	3,309,169
7/1/1993	**1993**	415.55	1.020	**1993**	3,317,416	3,342,363
7/1/1994	**1994**	423.89	1.030	**1994**	3,383,764	3,421,933
7/1/1995	**1995**	436.46	1.064	**1995**	3,485,277	3,569,147
7/1/1996	**1996**	464.18	1.039	**1996**	3,708,335	3,762,714
7/1/1997	**1997**	482.45	1.021	**1997**	3,852,960	3,883,383
7/1/1998	**1998**	492.61	1.047	**1998**	3,933,872	4,003,391
7/1/1999	**1999**	515.60	1.044	**1999**	4,118,764	4,186,905
7/1/2000	**2000**	538.48	1.049	**2000**	4,299,990	4,379,213
7/1/2001	**2001**	564.63	1.020	**2001**	4,510,689	4,544,609
7/1/2002	**2002**	576.17	1.014	**2002**	4,600,903	4,625,122
7/1/2003	**2003**	584.52	1.026	**2003**	4,665,316	
7/1/2004	**2004**	599.66	1.040	**2004**	4,786,614	
7/1/2005	**2005**	623.80	1.038	**2005**	**4,978,079**	
7/1/2006	**2006**	647.54				

Appendix B – Loss Exceedance Curves and the Catastrophic Event Threshold

Loss exceedance curves are a standard output format from catastrophe models. Table B.1 shows a hypothetical example of output from a catastrophe model. For illustration purposes only 4 points on the loss exceedance curve are shown in Table B.1. Typically, loss exceedance curves will consist of at least several hundred points. The curve is usually represented by loss amounts sorted in descending order along with associated probabilities of exceedance. The probability of exceedance of a given loss amount is the probability that at least one event causing at least as much loss as that loss amount will occur in a single year. The loss exceedance curve is assumed to result from an underlying collective risk model with a Poisson frequency distribution. Based on this assumption, frequencies (exceedance and incremental), return periods, and the severity density can be derived easily.

Table B.1

Hypothetical Example of Various Components of Common Representations of Loss Exceedance Curves

Event Loss	Probability of Exceedance	Frequency of Exceedance	Return Period	Incremental Frequency	Severity Distribution	Severity Density
[1]	[2]	[3]	[4]	[5]	[6]	[7]
=Model Output	=Model Output	= - ln (1-[2])	= 1 / [3]	= Difference [3]	=100%-{Shift[3]/ Total [5]}	= Difference [6]
1,000,000,000	0.1998%	0.002	500	0.002	100%	1%
100,000,000	0.9950%	0.010	100	0.008	99%	4%
10,000,000	9.5163%	0.100	10	0.090	95%	45%
1,000,000	18.1269%	0.200	5	0.100	50%	50%
Total	18.1269%			0.200		

For NCCI's large loss procedure, catastrophic losses from events exceeding $50 million dollars are completely excluded from experience used for aggregate ratemaking. A corresponding provision based on catastrophe model results is added to loss costs. Although the catastrophe model assumptions may be designed to only contemplate events likely to cause a large loss, this is only a qualitative threshold. Actual model output will include some events that when simulated with various stochastic assumptions happen to generate a small loss or even no loss at all.

Table B.2 shows the quantitative exclusion of losses exceeding $50 million, on both excess and ground-up bases, from the exceedance curve in Table B.1. Expected values were calculated for the two types of exclusions. Column (12) is used in the derivation of the catastrophe provisions. Note the excess exclusion shown in column (11) of Table B.2 is not used to calculate the replacement provision for catastrophic events in the large loss procedure. However, this is the type of calculation that would be applicable if events greater than $50 million were simply capped, as is done in the large loss procedure with large individual claims exceeding the state's per claim threshold.

Table B.2

Exclusion of Losses Excess of $50 Million Event Losses From Table B.1

$50m Excess	>$50m Ground-up	Expected Ground-up	Expected $50m Excess	Expected >$50m Ground-up
[8]	[9]	[10]	[11]	[12]
= Max(0, [1] 50m)	= if([1] > 50m , [1] ,0)	= [5] x [1]	= [5] x [8]	= [5] x [9]
950,000,000	1,000,000,000	2,000,000	1,900,000	2,000,000
50,000,000	100,000,000	800,000	400,000	800,000
0	0	900,000	0	0
0	0	100,000	0	0
Total		3,800,000	2,300,000	2,800,000

Illustration 3. NCCI 's formula for the calculation of one catastrophe peril's pure loss cost:

Catastrophe Pure Loss Cost (per $100 limited payroll) =

100 x Catastrophe Expected Losses / (# Workers x Limited Average Annual Wage)

So, if the loss exceedance curve in Tables B.1 and B.2 were based on a modeling assumption of 1,000,000 workers and the average annual wage was $40,000 the provision for the excluded large event losses would be:

100 x $2,800,000 / (1,000,000 x $40,000) = 0.007

For the DTEC provision, a similar provision would be computed for the other perils and added to the 0.007. The sum would then be multiplied by a factor to account for loss based expenses (or fully loaded expenses in administered pricing jurisdictions) and then rounded to the nearest penny to produce an additive provision for loss costs/rates.

Appendix C – NCCI Call #31, Large Loss and Catastrophe Call

CALL #31

NATIONAL COUNCIL ON COMPENSATION INSURANCE, INC.
LARGE LOSS AND CATASTROPHE CALL
VALUED AS OF DECEMBER 31, 2003

CARRIER/CARRIER GROUP____CARRIER CODE NUMBER__________________________
SUBMITTED BY _________________________TITLE ___________________TELEPHONE NO____________DATE SUBMITTED ________________

Claim Number (1)	Policy Number (2)	NCCI Catastrophe Number (3)	Exposure State Code (4)	Market Type Code (5)	Policy Effective Date (6)	Accident Date (7)	Claim Status Code (8)	Accumulated Paid Losses		Case Outstanding		Defense and Cost Containment Expense	
								Indemnity (9)	Medical (10)	Indemnity (11)	Medical (12)	Accumulated Paid (13)	Case Outstanding (14)

Market Type Code:
3 - Voluntary (not Large Deductible)
2 - Large Deductible
0 - Assigned Risk (not Large Deductible)

Claim Status Code:
0 - Open
1 - Closed
2 - Reopened

5. REFERENCES

[1] Historical Loss Development Study, 2003 Edition, Reinsurance Association of America, 2003.

Abbreviations and notations

AY, accident year

BLS – Bureau of Labor Statistics

CAS, Casualty Actuarial Society

$CLDF_T$- Capped "paid + case" tail factor, 19th to ultimate, for threshold T.

CPI-consumer price index

CPS-Current Population Survey

CY-calendar year

DSR-Designated Statistical Reporting level of NCCI

DTEC-Domestic Terrorism, Earthquake, and Catastrophic Industrial Accident provision

EAL-expected annual loss

$ELDF_T$- Excess "paid+case" tail factor, 19th to ultimate, for threshold T.

EQECAT- modeling company, a division of ABS Consulting Group

F_T- Factor to apply to state-specific *ULDF* to get state-specific $CLDF_T$ for threshold T.

FTE- full-time equivalents

ISO-Insurance Services Office

LAE-loss adjustment expense

M- $millions

MMI- Modified Mercalli intensity

MIDAS-AT- Meteorological Information and Dispersion Assessment System—Anti-Terrorism™

MW- moment magnitude

NCCI- National Council on Compensation Insurance, Inc.

OSHA- Occupational Safety and Hazard Administration

PGA-peak ground acceleration

PY-policy year

RAA-Reinsurance Association of America

Sa-response-spectral acceleration

$SCLDF_T$- State-specific capped "paid+case" tail factor, 19th to ultimate, for threshold T.

$SULDF_T$- State-specific uncapped "paid+case" tail factor, 19th to ultimate, for threshold T.

TRIA -Terrorism Risk Insurance Act of 2002

ULDF - Uncapped "paid+case" tail factor, 19th to ultimate

US – United States

USGS – United States Geological Survey

WCSP- NCCI 's Workers Compensation Statistical Plan

XS_T- Per Claim adjusted excess ratio at threshold T

Biography of the Author

Tom Daley is Director and Actuary at NCCI, Inc. He is currently responsible for both applied research and production duties in class ratemaking for all NCCI states, and handling state actuary loss cost and rate filing duties in several other states. He has a B.S. degree in Mathematics from the Pennsylvania State University. He is an Associate of the CAS and a Member of the American Academy of Actuaries.

Forecasting Workers Compensation Severities and Frequency Using the Kalman Filter

Jonathan Evans and Frank Schmid

Motivation. Estimating trend rates of growth of severity and frequency is crucial to workers compensation ratemaking. Such trend rates can be estimated using unobserved components models and structural time series models. These two types of models derive from parsimonious and transparent data-generating processes and, in the case of structural time series models, allow the researcher to incorporate economically meaningful explanatory variables into a time series framework. When specified in state-space form, unobserved components models and structural time series models become available to the Kalman filter estimation technique. The Kalman filter explicitly accounts for possible measurement errors in the observed severity and frequency data.
Model. Structural time series models, which nest unobserved components models, are applied to state-level time series data for (on-leveled and wage-adjusted) indemnity and medical severities, and for frequency. Parameter estimates, hypothesis tests, and growth forecasts are provided for by the software package STAMP. STAMP is especially designed for estimating unobserved components and structural time series models.
Results. NCCI developed a production process that employs unobserved components and structural time series models to state-level data of indemnity and medical severities, and frequency. Trend growth forecasts generated with such models were presented in state advisory forums and served as a consideration in rate filings.
Conclusions. NCCI's experience with Kalman-filtered estimation of trend rates during the policy year 2006 rate-filing season was encouraging. NCCI anticipates continued use of unobserved components models and structural time series models in future rate filings.
Availability. STAMP is an easy-to-use windows-driven software package that runs on the GiveWin platform. STAMP and GiveWin are available from Timberlake Consultants Ltd.
Keywords. Workers compensation, trend growth rates, Kalman filter, unobserved components model, structural time series models, state-space modeling

1. INTRODUCTION

Forecasting frequency and severity is crucial to workers compensation ratemaking. Such forecasting is performed using time series models, which are models that account for the time dependence in the observed data. In many time series models, this time dependence is modeled as a (potentially rather complex) autoregressive structure, as is the case in ARIMA (Auto-Regressive Integrated Moving Average) or ARMA (Auto-Regressive Moving Average) models. To many, such autoregressive structures appear mechanistic. In search for more transparent and parsimonious representations of the underlying data-generating processes, unobserved components (UC) models and, as an extension to UC models, structural time series (STS) models have been developed. In UC models the quantities of most interest are not directly observed and must be estimated using both empirical data and estimates of underlying statistical parameters (sometimes called *hyperparameters* in this context). STS models are linear combinations of UC models for the time series of interest and standard linear regression models including explanatory variables that are exogenous to the time series of interest. For example an STS model for stock market prices might combine a UC random walk with drift for the logarithm of stock prices and a standard linear regression for changes in the logarithm of stock market prices against recent changes in interest rates.

Forecasting is a signal extraction and signal extrapolation exercise. Signal extraction is the process of filtering out measurement errors from empirical data. Measurement errors include the total impact from all sources of *noise*, deviations of the empirical data from the underlying signal that do not affect the expected values of future observations (such as medical or indemnity severities, and frequency). In forecasting, the signal is the quantity of interest, because it is the signal that determines the expected values of future observations. Specifically, it is the objective of a forecasting model to elicit from historical observations the process that generates the unobservable signal. Because the forecasting model replicates the data-generating process of the signal (instead of fitting historical observations), the quality of these models cannot be judged by the (in-sample) fit to the observed data, as gauged, for instance, by the R^2. In fact, good fit to heretofore observed data harbors the risk of overfitting. Such overfitting implies that the (in-sample) fits and (out-of-sample) forecasts may not center on the signal, thus giving rise to potentially large forecasting errors.

As an example, consider a game of dice, where each die has six faces, the number of spots ranging from 1 to 6. In any toss of a pair of dice, the expected value of the outcome is 7. This expected value is the signal, which manifests itself as the mean outcome as the number of tosses goes to infinity. The difference between the observations and the signal is noise. The signal offers an unbiased forecast for any future toss. Thus, among all possible forecasting models, the one that simply produces this time-invariant signal as its forecast has the lowest expected root mean squared error. Yet, this model offers the worst in-sample fit possible, as the model has no explanatory power with regards to the variation of the outcome around the expected value. Not surprisingly, a least-squares regression of the 36 possible outcomes on the time-invariant signal reveals an R^2 equal to zero.

The risk of overfitting awards parsimony a critical role in time series modeling. UC models are conducive to such parsimonious modeling as the underlying data-generating process is highly transparent. UC models, and their extension, STS models, can be written in state-space form (defined in section 2.2), which makes these models available to the Kalman filter estimation technique. The Kalman filter has been developed in engineering as a signal extraction algorithm and, as such, recommends itself for estimating forecasting models. In fact, the Kalman filter is an estimation technique that explicitly accounts for possible measurement errors in the reported data.

NCCI estimates UC and STS models using the software packages STAMP and SsfPack. SsfPack is a collection of functions for state-space modeling, including maximum likelihood-estimation (MLE) and Kalman-filtered estimation and smoothing. This package runs in two alternative environments: the programming language Ox (Koopman et al. [11]) and the platform S-Plus (Zivot et al.[17]). We use SsfPack within Ox Professional on the GiveWin platform. STAMP (Koopman et al. [10]) also runs on the GiveWin platform. Ox Professional, GiveWin, and STAMP are distributed by Timberlake Consultants Ltd. S-Plus is a commercial platform available from Insightful Corporation. The models presented here were estimated using STAMP. Due to the complexity of code development, practical implementation of the Kalman filter in actuarial applications generally requires the acquisition of a preexisting specialized statistical software package from an external vendor.

NCCI developed a production process that employs unobserved components and structural time series models for indemnity and medical severities and for frequency for more than 30 U.S. states. Trend growth forecasts derived from these models were presented in state advisory forums and served as a consideration in rate filings.

1.1 Research Context

The material in this paper falls under CAS Research Categories II.G.12 Actuarial Applications and Methodologies/Ratemaking/Trend and Loss Development and III.H.15 Financial and Statistical Models/Statistical Models and Models/Time Series. Econometric models for actuarial trends have been dealt with in Hartwig et al.[5], Lommele and Sturgis.[12], McGuinness[13], and Van Slyke[15]. Credibility adjusted trending has been discussed in Venter[16]. None of these sources utilize the Kalman filter.

1.2 Objective

Economic support for actuarial trending of workers compensation losses at NCCI currently includes UC and STS models for forecasting (on-leveled and wage-adjusted) medical and indemnity severities, and frequency (number of claims, divided by on-leveled and wage-adjusted premium). (Wage-adjusting brings past exposure, as gauged by payroll, up to current wage levels; on-leveling brings past loss experience up to current benefit levels.) This paper describes current practice at NCCI of estimating such models using the Kalman filter. In addition to this set of three single-equation models, NCCI operates a Bayesian five-equation state-space forecasting model for severities, frequency, and the corresponding loss ratios—this multi-equation model, which accounts for add-up constraints and contemporaneous (cross-equation) covariances, is estimated using the Metropolis-Hastings algorithm. This paper is written from the perspective of actuarial researchers using preexisting statistical software packages from external vendors and does not include algorithmic details for statistical methods.

1.3 Outline

In Section 2.1, we describe the data-generating processes that underlie UC and STS models, and we put these models in state-space form. In section 2.2, we discuss the Kalman filter estimation technique and show how ML estimates for the moments are obtained from the Kalman filter output. The authors caution that readers need not completely understand the formulaic details in section 2.2 to understand the rest of this paper. Section 3 describes an implementation of UC and STS models in indemnity and medical severities and frequency forecasting. Section 4 concludes.

2. BACKGROUND AND MODELS

2.1 Unobserved Components and Structural Time Series Models

STS models are linear combinations of UC models and standard linear regression models. We start out by describing the data-generating processes of UC models and then expand these models to the STS framework.

UC models derive from the concept of Gaussian *innovations*, as exemplified in Brownian motion. Unlike noise, innovations propagate forward in time and affect the expected values of future observations. In their most basic (and, hence, most restrictive) form, these models postulate that innovations to the (unobserved) signal of a given (observable) variable are draws from the normal distribution. Put differently, the signal in question follows a random walk. Let y_t be the variable and θ_t the signal, then we can write the *local level model* in Equations 2.1.1 through 2.1.3 as follows:

$$y_t = \theta_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0,\sigma_\varepsilon^2),\ \ t = 1,\ldots,T \tag{2.1.1}$$

$$\theta_t = \mu_t \tag{2.1.2}$$

$$\mu_t = \mu_{t-1} + \eta_t, \qquad \eta_t \sim N(0, \sigma_\eta^2) \qquad (2.1.3)$$

The local level model is written in state-space form. The variable μ_t is the only state variable and, by definition, unobservable. The variable μ_t describes the time t *level* of the unobservable signal θ_t and is subject to the Guassian innovation η_t. The observed dependent variable y_t is the sum of this signal and Gaussian noise ε_t. Inserting Equation (2.1.2) into Equation (2.1.1) delivers the *measurement equation*. Equation (2.1.3) is the *transition equation*, which describes the trajectory of the state variable μ_t.

Local level models apply when the signal follows a random walk. Variables that follow random walks exhibit high degrees of persistence, as all innovations are permanent. An example of such a highly persistent variable is the rate of CPI (Consumer Price Index) inflation (see, for instance, Koopman et al.[11] and Green[4]).

Signals may exhibit drift. If this drift is stochastic, we obtain the *local linear model*, which is described in Equations 2.1.4 through 2.1.7:

$$y_t = \theta_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2),\ \ t = 1, \ldots, T \qquad (2.1.4)$$

$$\theta_t = \mu_t \qquad (2.1.5)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \qquad \eta_t \sim N(0, \sigma_\eta^2) \qquad (2.1.6)$$

$$\beta_t = \beta_{t-1} + \zeta_t, \qquad \zeta_t \sim N(0, \sigma_\zeta^2) \qquad (2.1.7)$$

The state variable μ_t indicates the *level* of the signal, and the state variable β_t describes the *slope* (or, synonymously, drift) of the signal. As with the level, the slope is governed by a Gaussian, permanent innovation ζ_t. Because there are two state variables in the local linear model, there are two transition equations, which are Equations 2.1.6 and 2.1.7.

An example of a variable the trajectory of which may be described using a local linear model is the logarithmic stock market total-return index, where β_t indicates the expected log return (or, equivalently, the drift in the logarithmic stock price). For $\sigma_\zeta^2 = 0$, the slope is non-stochastic. In the stock market example, non-stochastic drift implies constant expected return.

A third model of interest follows from the local level model by means of integration. Assume that the CPI rate of inflation indeed describes a random walk, as empirical studies indicate, and measure the rate of inflation by the first difference in the logarithmic price level. In this case then, the logarithmic price level follows an integrated random walk. For the ***integrated random walk*** (which sometimes is called ***smooth trend***) model, we can write in Equations 2.1.8 through 2.1.11:

$$y_t = \theta_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \; t = 1, \ldots, T \tag{2.1.8}$$

$$\theta_t = \mu_t \tag{2.1.9}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1}, \tag{2.1.10}$$

$$\beta_t = \beta_{t-1} + \zeta_t, \qquad \zeta_t \sim N(0, \sigma_\zeta^2) \tag{2.1.11}$$

The integrated random walk model results from the local linear model for $\sigma_\eta^2 = 0$.

The described types of UC models rest on parsimonious data-generating processes, which makes them appealing for signal-extraction purposes. On the other hand, these models are not cognizant of economic, causal relations that may exist between the dependent variable in question, y_t, and a vector of variables of economic activity, $(x_{1,t}, x_{2,t}, \ldots, x_{n,t})$. UC models can be expanded to STS models by adding a standard regression component, thus enabling such models to account for pertinent economic relations. When expanding the most general UC model—the local linear model—to an STS model, we can write in Equations 2.1.12 through 2.1.16:

$$y_t = \theta_t + \gamma_t \cdot x_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{2.1.12}$$

$$\theta_t = \mu_t \tag{2.1.13}$$

$$\gamma_t = \gamma_{t-1} + \upsilon_t, \qquad \upsilon_t \sim N(0, \sigma_\upsilon^2) \tag{2.1.14}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \tag{2.1.15}$$

$$\beta_t = \beta_{t-1} + \zeta_t, \qquad \zeta_t \sim N(0, \sigma_\zeta^2) \tag{2.1.16}$$

The STS model above has one explanatory variable, x_t. The regression parameter of this variable, γ_t, follows a random walk. Alternatively, this parameter may be specified as stationary ($\gamma_t = \bar{\gamma} + \upsilon_t$) or time-invariant ($\sigma_\upsilon^2 = 0$). At NCCI, when employing time-variant parameters, we estimate such STS models using SsfPack. The software package STAMP (we use version 6.21) can handle currently only time-invariant parameters. For short data series, time invariance in γ_t is an appropriate constraint to avoid over-parametrization. Such time-invariance presumes that the variable x_t is measured without error and that the economic relation depicted in the above measurement equation is time-invariant—these are standard assumptions in ordinary linear regression models.

In the next section, we describe the Kalman filter technique for estimating the state variables, and the accompanying ML estimation of the moments. Further, we discuss the relation between the Kalman filter and the Bühlmann credibility criterion.

2.2 The Kalman Filter

The Kalman filter was invented in 1960 by Rudolf Kalman (Kalman [7]) and saw almost immediate application in real-time signal processing for spacecraft. Up to the present, the Kalman filter is widely used in various aspects of aerospace operations, such as radar. The filter acts on an observed time series by removing an estimate of measurement noise. Thus, the filtered series represents an estimate for the underlying process of the signal, that is, the observed variable, purged of noise. The Kalman filter introduces into time series modeling the fundamental statistical philosophy that real-world observations are only shadows of ideal Platonic forms (Plato [14]).

The Kalman filter works in the context of time series models expressed in state-space form. The state-space form specifies a transition vector equation (Equation 2.2.1) for unobserved state variables of interest, and an associated measurement vector equation (Equation 2.2.2) for the observed series (Harvey[6] and Durbin and Koopman[3]). The transition equations describe the transition of the state variables from state t to state $t+1$. The measurement equations describe the relations between the signals and the state variables and, at the same time, account for measurement noise as the difference between the observed variables and the respective signals.

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t,\ \mathrm{E}(\eta_t) = 0,\ \mathrm{Var}(\eta_t) = Q_t \tag{2.2.1}$$

$$y_t = Z_t \alpha_t + d_t + \varepsilon_t, \mathrm{E}(\varepsilon_t) = 0,\ \mathrm{Var}(\varepsilon_t) = H_t \tag{2.2.2}$$

The vector α_t contains the unobservable state variables. The matrices T_t, R_t, and Z_t, as well as the parameter vectors c_t, and d_t are assumed to be non-stochastic; typically, these variables and parameters are known and may (but need not) be time-invariant. In engineering applications, the matrices of the variances of the innovations and the measurement errors, Q_t and H_t, respectively, are often determined by actual physical calibration with instruments. In financial and economic analyses, these moments are estimated using the ML approach. This ML estimation can easily be obtained through a decomposition of the prediction error of the Kalman filter (Kim and Nelson[8]). The Kalman filter is presented in Equations 2.2.3 through 2.2.7.

$$a_{t|t-1} = T_t a_{t-1} + c_t \tag{2.2.3}$$

$$P_{t|t-1} = T_t P_{t-1} T_t' + R_t Q_t R_t' \tag{2.2.4}$$

$$a_t = a_{t|t-1} + P_{t|t-1} Z_t' F_t^{-1} (y_t - Z_t a_{t|t-1} - d_t) \tag{2.2.5}$$

$$P_t = P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1} \tag{2.2.6}$$

$$F_t = Z_t P_{t|t-1} Z_t' + H_t \tag{2.2.7}$$

For initial values, it is assumed that $P_{2|1}$ is very large and that $a_{2|1} = 0$.

The coefficients $a_{t|t-1}$ and a_t represent estimates for α_t before and after y_t is observed, respectively.

The exists an analogy between the Kalman filter and the Bühlmann credibility criterion (Venter[18]); to make the analogy more apparent, assume $Z_t = 1$ and $d_t = 0$. Equation 2.2.5 contains the Bühlmann credibility-like term:

$$\frac{P_{t|t-1}}{P_{t|t-1} + H_t}$$

If y_t , $a_{t|t-1}$, and a_t are interpreted as the indication, the complement of credibility, and the credibility-weighted estimate, respectively, then Equation 2.5 is effectively a Bühlmann credibility estimate since $P_{t|t-1}$ and H_t can be interpreted as estimates of the variances of $a_{t|t-1}$ and y_t, respectively.

Note that the Kalman filter only estimates the series of underlying states α_t, given the observed series y_t and assumed values for the variance parameters contained in Q_t and H_t. These variance parameters must still be estimated via ML. In general, likelihood functions for time series models based on prior estimates of observations conditional on all previous observations can be stated as in Equation 2.2.8, where θ represents the parameter values:

$$L(y;\theta) = \prod_{t=1}^{T} p(y_t \mid y_{t-1},\dots,y_1) \tag{2.2.8}$$

The Kalman filter estimates can be used to derive prior means and variances of not yet observed points, conditional on the previous observations. Since the actual observations the are conditionally normally distributed, the log-likelihood function can be written as Equation 2.2.9, where N is the number of scalar components of y_t:

$$l(y;\theta) = -\frac{NT}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log|F_t| - v_t{}'F_t^{-1}v_t \tag{2.2.9}$$

$$v_t = y_t - Z_t a_t - d_t, \tag{2.2.10}$$

The Kalman filter works reasonably well even for some time series shorter than 30 points, although the filtered series may behave erratically on the first few data points.

The filtered estimates are predictions for the time-t vector of state variables, based on information available at time $t-1$. In the context of economics, these filtered estimates may be interpreted as expectations for time t that economic agents formed based on the information available at time $t-1$. Thus, the time-t filtered estimates of the rate of inflation may serve as a gauge of inflation expectations (Koopman et al. [10]).

Typically, the researcher is not interested in time-t estimates for the state vector that are based on information available only at time $t-1$. As in standard regression analysis, the researcher looks for time-t estimates of the state vector that use all information available as at the end of the data series, that is, as at time $T \geq t$. Such estimates can be obtained through a backward moment-smoothing algorithm (Bryson and Ho [1], de Jong [2], Kohn and Ansley [9]). This algorithm is presented in Equations 2.2.11 through 2.2.16, where the initial conditions are set as $N_T = r_T = 0$:

$$K_t = T_t P_{t|t-1} Z_t ' F_t^{-1} \tag{2.2.11}$$

$$e_t = F_t^{-1} v_t - K_t ' r_t \tag{2.2.12}$$

$$r_{t-1} = Z_t ' e_t + T_t ' r_t \tag{2.2.13}$$

$$D_t = F_t^{-1} + K_t ' N_t K_t \tag{2.2.14}$$

$$L_t = T_t - K_t Z_t \tag{2.2.15}$$

$$N_{t-1} = Z_t ' F_t^{-1} Z_t + L_t ' N_t L_t \tag{2.2.16}$$

Equations 2.2.17 through 2.2.20 present the moment-smoothed estimates of the stochastic elements and their associated variances:

$$\mathrm{E}[\varepsilon_t \mid \{y_1,\ldots,y_T\}] = H_t e_t \tag{2.2.17}$$

$$\mathrm{Var}[\varepsilon_t \mid \{y_1,\ldots,y_T\}] = H_t D_t H_t \tag{2.2.18}$$

$$\mathrm{E}[\eta_t \mid \{y_1,\ldots,y_T\}] = R_t ' Q_t R_t r_t \tag{2.2.19}$$

$$\mathrm{Var}[\eta_t \mid \{y_1,\ldots,y_T\}] = R_t ' Q_t R_t N_t R_t ' Q_t R_t \tag{2.2.20}$$

3. IMPLEMENTATION

We now demonstrate how to apply UC and STS models to state-level series of (on-leveled and wage-adjusted) indemnity and medical severities and of frequency (number of claims, divided by on-leveled and wage-adjusted premium). The objective is to forecast the growth factor $1+g_{T,T+3}$ for the indemnity and medical severities that applies to the 3-year period between the last observed period T and the future period $T+3$. (The number of years may not be an integer; for instance, the time interval may range from T to $T+3+\varepsilon$, $0<\varepsilon<1$, in which case the applicable growth factor reads $1+g_{T,T+3+\varepsilon}$.)

There are several routes to arriving at such a 3-year growth factor. One approach is to estimate directly 3-year rates of growth $\hat{g}_{T,T+3}$ from (successive and non-overlapping) 3-year time periods. This method requires long data series (as the number of data points is, at maximum, one-third of the number of annual observations) and, hence, is not an option at NCCI. An alternative route is to estimate annual rates of growth and then tally up the annual forecasts for the time periods $T+1$, $T+2$, and $T+3$ in order to obtain the 3-year rate of growth from T to $T+3$. Tallying up forecast rates of growth is not straightforward as these forecasts are random variables and annual compounding involves nonlinear transformations. For instance, let $\hat{g}_{T+1}$, $\hat{g}_{T+2}$, and $\hat{g}_{T+3}$ be the forecasts for the annual rates of growth and calculate the forecast for the 3-year growth rate by means of compounding: $\hat{g}_{T,T+3}=(1+\hat{g}_{T+1})\cdot(1+\hat{g}_{T+2})\cdot(1+\hat{g}_{T+3})-1$. In this case then, if the 3 annual forecasts $\hat{g}_{T+i}$ are unbiased forecasts of the actual annual rates of growth g_{T+f} $(f=1,2,3)$, $\hat{g}_{T,T+3}$ is a biased forecast for the pertinent future 3-year rate of growth $g_{T,T+3}=(1+g_{T+1})\cdot(1+g_{T+2})\cdot(1+g_{T+3})-1$.

We arrive at our forecast for the growth rate $g_{T,T+3}$ by means of estimating and tallying up logarithmic rates of growth—that is, first differences in natural logarithms. We choose this approach because, here, our interest is to estimate the geometric mean of the (continuously compounded) annual rates of growth rather than the arithmetic mean. Logarithmic rates of growth are additive; thus we can write: $\hat{g}^{\log}_{T,T+3}=\hat{g}^{\log}_{T+1}+\hat{g}^{\log}_{T+2}+\hat{g}^{\log}_{T+3}$. (When there is the fraction ε of an incomplete fourth year, then the multi-year growth rate amounts to $\hat{g}^{\log}_{T,T+3+\varepsilon}=\hat{g}^{\log}_{T+1}+\hat{g}^{\log}_{T+2}+\hat{g}^{\log}_{T+3}+\varepsilon\cdot\hat{g}^{\log}_{T+4}$.) This additivity property implies that the sum of the annual forecast growth rates is indeed an unbiased estimator of the multi-year logarithmic rate of growth. By means of invoking normality, it is possible to calculate a standard error for $\hat{g}^{\log}_{T,T+3}$ from the variances of the annual growth rates $\hat{g}^{\log}_{T+1}$, $\hat{g}^{\log}_{T+2}$, and $\hat{g}^{\log}_{T+3}$. These standard errors then enable us to compute confidence bounds around $\hat{g}^{\log}_{T,T+3}$. This provides valuable information, rarely if ever available from traditional actuarial trend analyses, about the uncertainty of trend estimates.

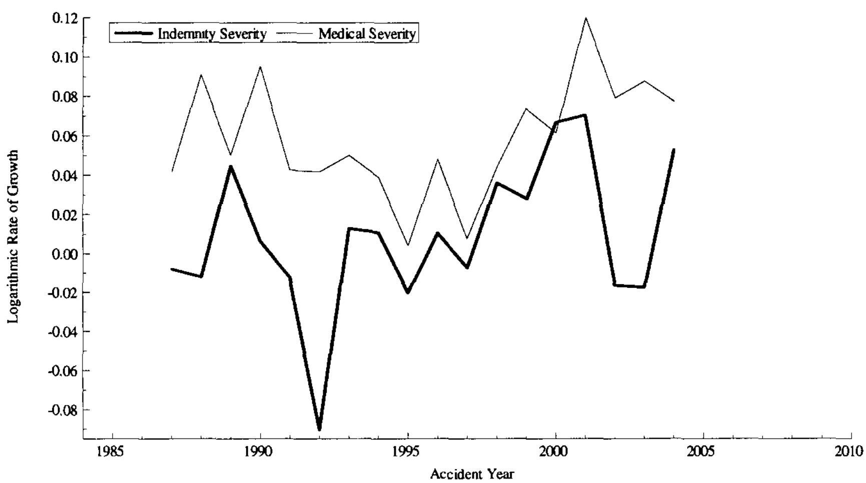
The pertinent severity series are on a "paid" basis. The severity and frequency data are from an anonymous U.S. state and refer to the policy year 2006 rate-filing season. These data series range from 1986 through 2004, thus affording 18 annual growth rate observations. The model estimates presented below are for illustration purposes only and are not necessarily identical to the estimates used in the rate-filing for the anonymous state in question.

According to the NBER (National Bureau of Economic Research, www.nber.org), there are two economic recessions that fall into the analyzed 1986-2004 period. Both of these recessions lasted for 8 months, as measured from peak to trough. The 1990/91 recession lasted from July 1990 to March 1991, and the 2000 recession lasted from March to November. This fluctuation in economic activity is potentially important for frequency. For instance, it can be shown that the growth rate of BLS (Bureau of Labor Statistics, www.bls.gov) on-the-job injury rates correlate with the change in the rate of unemployment. Similarly, it is common for NCCI states that the growth rate of frequency correlates with the change in the state-level rate of unemployment.

Chart 1 shows for the anonymous state in question the log growth rates for the on-leveled and wage-adjusted indemnity and medical severities. Chart 2 exhibits the log growth rate of frequency, along with the first difference in the percentage rate of unemployment (which has been divided by 10 in this exhibition, for scaling purposes).

Chart 1

Logarithmic Growth Rates of Indemnity and Medical Severities, State-Level Data, Accident Years 1987-2004



Visual inspection of Charts 1 and 2 suggests that the rates of growth of the indemnity and medical severities and of frequency follow random walks—thus, the local level model applies. A *Unit-root process* includes an autoregressive AR(1) coefficient of unity. We do not employ unit-root tests, such as Dickey-Fuller (see, for instance, Greene [4]). This is because, for short time series, these tests have little power, that is, the test are deficient in their ability to reject the null hypothesis of the presence of a unit root. What complicates matters for the frequency series is that this variable is not a pure unit-root process but instead is the sum of a unit-root process and a cyclical (that is, business cycle) component. For instance, as Chart 2 shows, the two recessions seem to have depressed the growth rate of frequency, although the drop during the 1990/1991 recession appears to have been permanent (instead of cyclical).

Chart 2

Logarithmic Growth Rate of Frequency and First Difference in Rate of Unemployment, State-Level Data, Accident Years 1987-2004

Note: The Rate of Unemployment was measured in percent; for scaling purposes, the first difference was divided by 10 (in this exhibition only).

Table 1 displays the regression results for the local-level UC (severities) and STS (frequency) models. This table shows the final state variable only—the level μ_T. The t-statistic displayed alongside μ_T pertains to this final, time T variable only. Put differently, the *t*-statistic for μ_T does not afford statistical inference for μ_t in prior periods ($t = 1,...,T-1$). For medical severity, we can reject the null hypotheses of zero growth as at time T (the time of the last observation). Most interestingly, we can reject the null hypothesis that there is no business cycle influence on the growth rate of frequency. Specifically, an increase in the rate of unemployment by 1 percentage point (for instance, from 4 percent to 5 percent) depresses the (logarithmic) rate of growth of frequency by 1.76 percentage points. This finding supports the commonly held view that when the labor market softens, the least productive workers (which, frequently, are the last hired and thus least experienced) are the first to be laid off—such layoffs leaves the remaining pool of employed workers more experienced, on average. Note that, for the purpose of forecasting, the lack of statistical significance of "baseline" growth (as at time T) in indemnity severity and in frequency is irrelevant.

Table 1

Regression Results for Growth Rates of Indemnity and Medical Severities, and Frequency

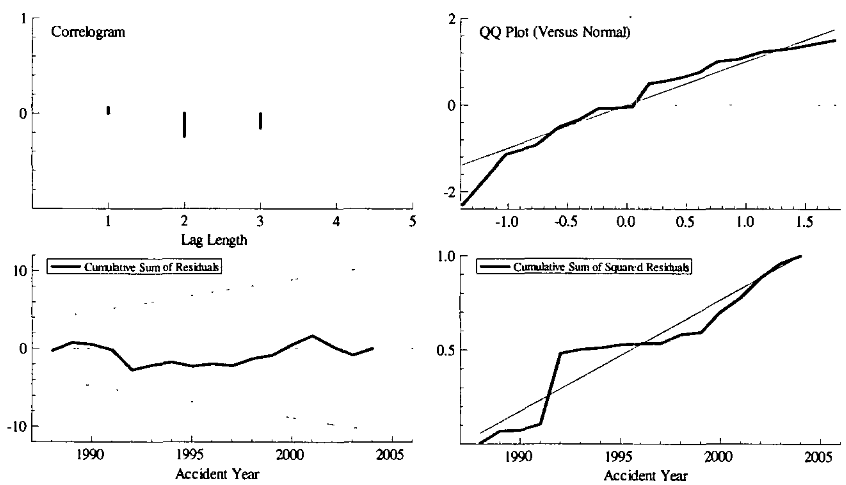
	Panel A: Indemnity Severity			
Variable	Coefficient	RMSE	*t*-statistic	Q-Ratio
μ_T (Level)	0.020474	0.015641	1.309	0.0455
Log Likelihood	54.3136			
	Panel B: Medical Severity			
Variable	Coefficient	RMSE	*t*-statistic	Q-Ratio
μ_T (Level)	0.081500	0.014593	5.585	0.4060
Log Likelihood	59.8715			
	Panel C: Frequency			
Variable	Coefficient	RMSE	*t*-statistic	Q-Ratio
μ_T (Level)	0.0023129	0.0045367	0.50983	1.000
Unemployment	-0.017635	0.0074340	-2.3722	---
Log Likelihood	52.8787			

The Q-ratio in the rightmost column of Table 1 is the ratio of the ML-estimated variance of the innovation in level ($\hat{\sigma}_\eta^2$) to the ML-estimated variance of the measurement error ($\hat{\sigma}_\varepsilon^2$). For both severities and for frequency, this Q-ratio is positive. A positive Q-ratio indicates that the level μ_t is time-variant or, equivalently, that the rate of growth of the severity in question (for frequency, this holds net of the cyclical influence) is non-stationary, as hypothesized.

As mentioned, traditional measures of goodness of fit are of limited use for forecasting models. This impediment puts the emphasis on regression diagnostics. Chart 4 shows for the indemnity severity UC model four diagnostic plots for the measurement error. The left-hand side plot in the top panel of Chart 3 presents autocorrelations in the residuals at lag lengths 1 through 3. These autocorrelations appear to be small on this plot, thus lending support to the assumption that the measurement errors are independently distributed. On the right-hand side of the top panel, there is a QQ-plot. This QQ-plot indicates that there are no fat tails. Specifically, there is neither statistically significant skewness (which equals 0.8208) nor statistically significant excess kurtosis (which measures 0.1698). The bottom panel of Chart 2 displays the cumulative sum of residuals (left) and cumulative sum of squared residuals (right). The cumulative residuals signify no discernible positive serial correlation as these sums are well within the error cone. The cumulative sum of squared residuals indicates no material heteroskedasticity, thus suggesting that the assumption of a time-invariant variance of the measurement error is adequate. The corresponding residual diagnostics for the medical severity UC model and the frequency STS model are displayed in Charts 5 and 6. Here again, there is no statistically significant skewness (0.3505 and -0.4684, respectively) or excess kurtosis (-0.8213 and -0.5680, respectively).

Chart 3

Regression Diagnostics (Local Level UC Model) for the Log Growth Rate of Indemnity Severity



A means of testing the performance of the forecasting model is to generate forecasts for a holdout period and then compare these forecasts with the actual, known observations. We estimate the two severity models for the time period 1987 through 2001 (that is, periods $t = 1,...,T-3$), assigning the years 2002 through 2004 (periods $T-2$ through T) to the holdout window. Then, we generate multi-step logarithmic annual growth rate forecasts for this holdout window from the shortened ($t = 1,...,T-3$) time series. Multi-step forecasts, by definition, do not incorporate information that arrives during the holdout period; for instance, the forecast for T does not incorporate information that becomes available during periods $T-2$ or $T-1$. The concept of multi-step forecasting agrees with the actual forecasting problem at hand. However, there is one important difference between the holdout forecasting exercise and the actual forecasting situation. When we employ STS models in forecasting for the periods $T+1$ through $T+3$, we have to feed to the model for the first difference in the rate of unemployment historical observations or, if such observations have not yet become available, forecasts. In the holdout forecasting exercise, the historical observations for the state-level rate of unemployment in periods $T-2$ through T are available. Although using historical forecasts rather than historical observations in the holdout forecasting exercise would remedy this problem, the exercise would still be three observations short of the data series available in the actual forecasting situation.

Chart 4

Regression Diagnostics (Local Level UC Model) for the Log Growth Rate of Medical Severity

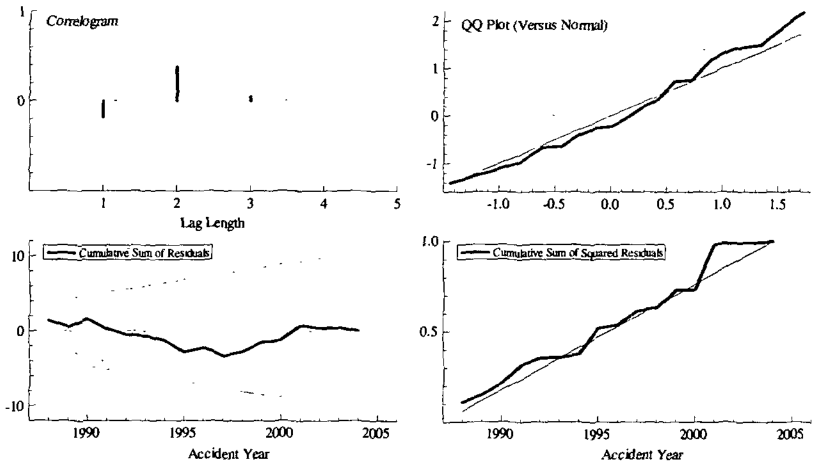
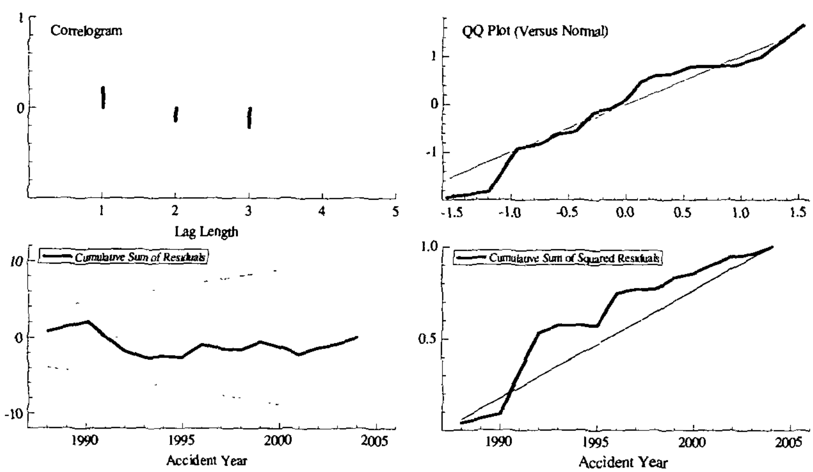


Chart 5

Regression Diagnostics (Local Level STS Model) for the Log Growth Rate of Frequency



In the actual forecasting situation, the accuracy of the forecasts for the frequency rate of growth is dependent on the accuracy of the forecasts for the rate of unemployment. (Note that our model does not account for the stochastic nature of the unemployment rate forecasts.) Most importantly, an STS model will forecast recession-related dips in the growth rate of frequency (as exhibited in the historical data of Chart 2) only if the pertinent forecasts for rate of unemployment describe such a recession. Unfortunately, economic recessions are next to impossible to forecast, because, if they were predictable, they would not occur as the Federal Reserve (or, possibly Congress, when it comes to fiscal policy) would act in a timely manner to prevent them.

Chart 6 and Chart 7 exhibit for the mentioned 3-year holdout window multi-step forecasts for the annual logarithmic growth rates of the indemnity and medical severities. Note that the displayed forecasts need to be multiplied by 100 to obtain percentage rates of growth. The confidence bounds around the forecasts range over 2 RMSE (root mean squared errors)—these confidence intervals are comparatively wide, which is due to the small number observations (14 by count). The forecasts of interest—those for the year $T+3$ (2004)—are quite accurate.

Chart 6

Holdout-Window Forecasts (Local Level UC Model) for the Growth Rate of Indemnity Severity

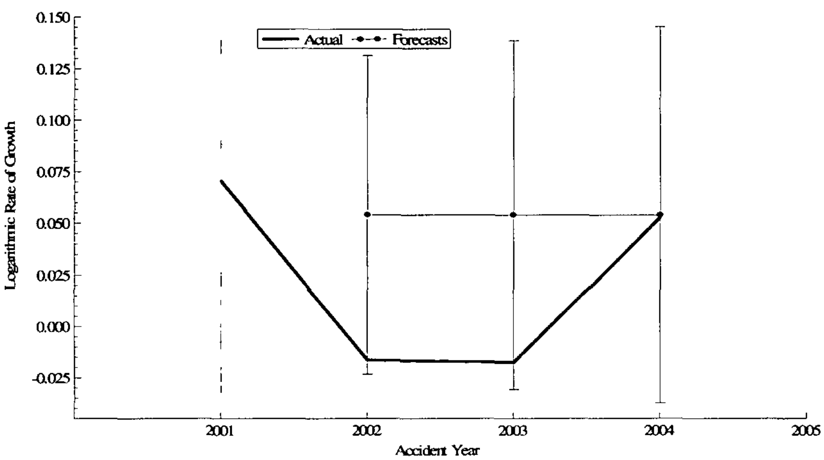
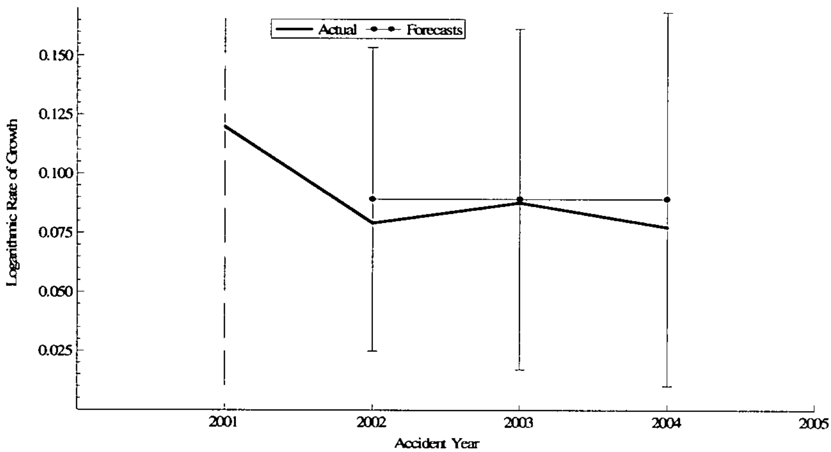


Chart 8 exhibits the hold-out window forecasts for the growth rate of frequency. The forecast of interest—the one for the year $T+3$ (2004)—clearly falls short of the observed value. Apparently, the regression coefficient that gauges influence of the first difference in the rate of unemployment underestimated the effect on the growth rate of frequency of the pronounced drop in the rate of unemployment rate during the economic recovery 2002-2004. It bears to mention that the shortening of the period of observation for this hold-out window exercise leaves the remaining period of observations (1987-2001) with only one economic recovery (the one following the 1990/91 recession); thus, it comes at now surprise that the regression coefficient in question is poorly estimated. Yet, it bears to mention that even in this shortened time period of only 15 observations, the STS model forecast for the year $T+3$ (2004) beats the benchmark forecast (the random walk), which equals -7.78 percent.

Chart 7

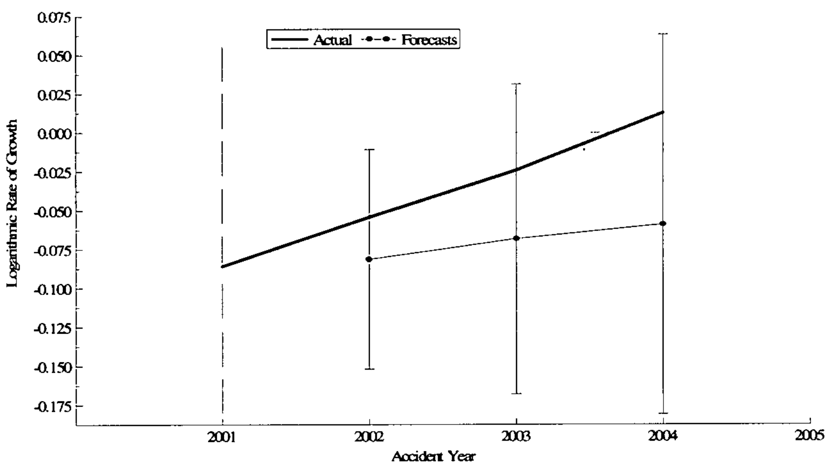
Holdout-Window Forecasts (Local Level UC Model) for the Growth Rate of Medical Severity



Charts 9 through 11 exhibit forecasts for the annual logarithmic rates of growth of the indemnity and medical severities and for frequency for the periods $T+1$ through $T+3$, based on the full ($t=1,\ldots,T$) model. The top panels of these charts exhibit recent observations (at the dashed vertical bar and to the left of it) and forecasts with error bars ranging over 2 RMSE (solid vertical bars). Remember that the forecasts for frequency (unlike those for the severities) are linear combinations of the trend rate of growth (μ_T) and a business cycle component. The bottom panels of Charts 9 through 11 display the estimated (at the dashed vertical bar and to the left of it) and forecast values for the level μ_t. Note that these forecasts for the level—the annual trend rate of growth of the indemnity severity, medical severity, or frequency—are equal to the respective final state vector μ_T for all $T+f$, $f=1,...,\infty$. Again, the presented forecasts need to be multiplied by 100 to obtain percentage rates of growth.

Chart 8

Holdout-Window Forecasts (Local Level STS Model) for the Growth Rate of Frequency



As mentioned, the sought-after growth factor $1+g_{T,T+3}$ is the exponentiated sum of the 3 forecasts for the annual logarithmic rates of growth $\hat{g}^{\log}_{T+f}$ $(f=1,2,3)$.

Finally, it is of interest how the Kalman filter technique compares with less sophisticated approaches to forecasting logarithmic rates of growth from past realizations. For simplicity, we focus on medical severity as this variable is, unlike frequency, not (hypothesized to be) subject to the business cycle.

In general, if a time series is stationary (here, is the sum of a constant and a Gaussian error term), then the mean of the series renders unbiased forecasts for any future value of this series. Although any past value of the series renders such an unbiased forecast, the more past realizations are averaged over, the lower is the expected RMSE of this forecast.

Hence, if the log rate of growth of indemnity severity were stationary (which it is not, we hypothesize), then taking the mean over all available historical observations is desirable when forecasting this variable.

Chart 9

Forecasts (Local Level UC Model) for the Log Growth Rate of Indemnity Severity

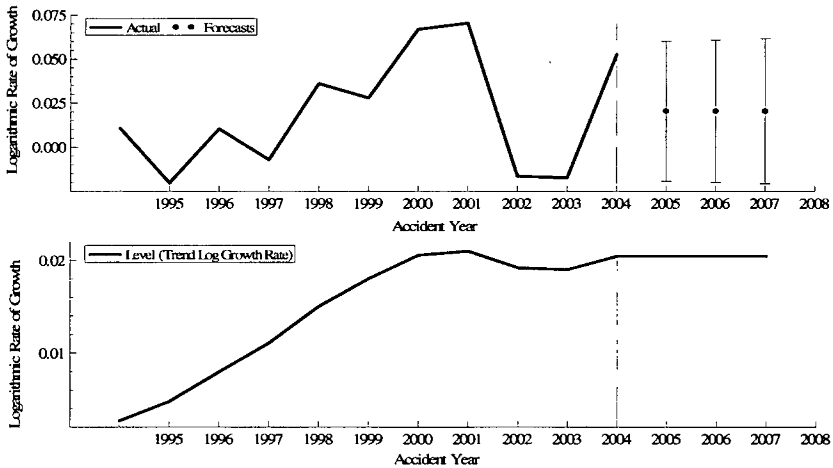


Chart 10

Forecasts (Local Level UC Model) for the Log Growth Rate of Medical Severity

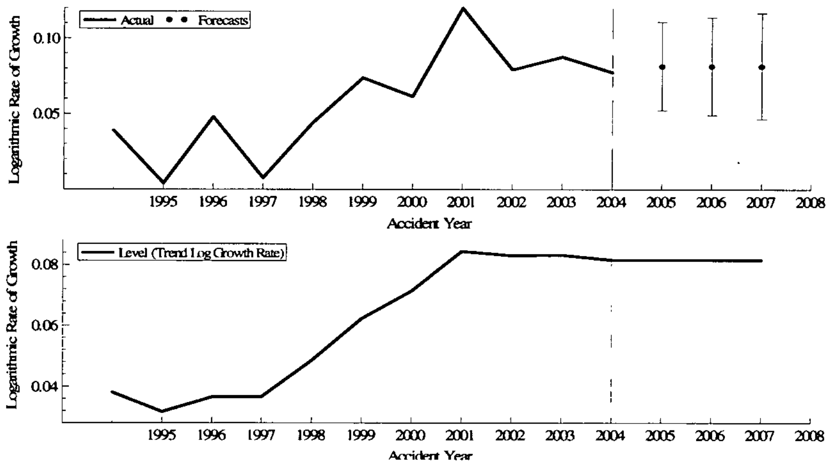
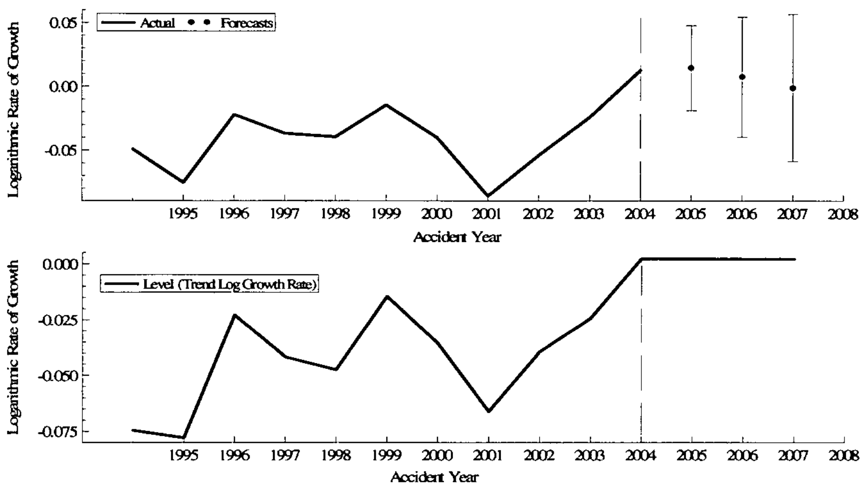


Chart 11

Forecasts (Local Level UC Model) for the Log Growth Rate of Indemnity Severity



If the growth rate of medical severity follows a random walk, as hypothesized above, then no averaging should be done. In general, if a time series follows a random walk, then the last (that is, the time T) realization serves as an unbiased forecast for any future value of this series, given the information available at time T. Hence, the last, time T observed log rate of growth for medial severity—the reading for the 2004 accident year is 7.74 percent—is an unbiased forecast for the log growth rates of periods $T+f$, $(f=1,2,3)$ (that is, accident years 2005 through 2007). Yet, as shown in Table 1, the final state vector and, hence, the Kalman-filtered forecast for the annual log rate of growth equals 8.15 percent. This 41 basis point annual difference is due to the measurement error in the observed data. Although both the last observed rate of growth (0.0774 in this realization) and the final state vector (0.0815 in this realization) reflect an unbiased forecast for any future period, these forecasts differ in precision. The Kalman-filtered forecast of 8.15 percent is likely to be much closer to the actual future outcome.

As a demonstration of the difference in typical outcome accuracy between the Kalman-filtered forecasts and forecasts that disregard a potential measurement error in the observed data, consider the holdout forecast for medical severity presented above. For the Kalman-filtered forecasts of the annual log rates of growth, the sum of the absolute forecast errors (for periods $T+1$, $T+2$, and $T+3$) equals 0.0387, and the RMSE amounts to 0.0090. When the last observed rates of growth are used, these gauges of forecast inaccuracy are as high as 0.1154 and 0.0234, respectively.

4. CONCLUSIONS

The experience of NCCI with Kalman filtered estimation of trend rates during the policy year 2006 rate filing season was encouraging. NCCI anticipates continued use of unobserved components models and structural time series models in future rate filings and the Kalman filter estimation technique. Current research at NCCI focuses on testing a Bayesian five-equation state-space forecasting model for severities, frequency, and the corresponding loss ratios—this multi-equation model, which accounts for add-up constraints and contemporaneous (cross-equation) covariances, is estimated using the Metropolis-Hastings algorithm.

Acknowledgment

We are grateful to Harry Shuford for comments. Thanks to Auntara De and Chris Doyle for research assistance. All errors are the authors'.

Equation Chapter 1 Section 1

5. REFERENCES

[1] Bryson, Arthur E.; Ho, Yu-Chi, *Applied Optimal Control: Optimization, Estimation, and Control*, **1969**, Blaisdell.

[2] De Jong, Piet, "Smoothing and Interpolation with the State Space Model," *Journal of The American Statistical Association*, **1989**, Vol. 84, 1085-1088.

[3] Durbin, James; Koopman, Siem Jan, *Time Series Analysis by State Space Models*, **2001**, Oxford University Press.

[4] Green, William H., *Econometric Analysis*, 5th ed. **2002**, Prentice Hall.

[5] Hartwig, Robert P.; Kahley, William J.; Restrepo, Tanya E.; Retterath, Ronald C., "Workers Compensation and Economic Cycles: A Longitudinal Approach," *PCAS* **1997**, Vol. LXXIV, 660-700.

[6] Harvey, Andrew C., Forecasting, *Structural Time Series Models and the Kalman Filter*, **1991**, Cambridge University Press.

[7] Kalman, Rudolf E., "A New Approach to Linear Filtering and Prediction Problems," *Transactions ASME Journal of Basic Engineering*, **1969**, Vol. D82, 35-45.

[8] Kim, Chang-Jin; Nelson, Charles R., *State-Space Models With Regime Switching*, **1999**, MIT Press.

[9] Kohn, Robert; Ansley, Craig F., "A Fast Algorithm for Signal Extraction, Influence and Cross-validation in State Space Models," *Biometrika*, **1989**, Vol. 76, 65-79.

[10] Koopman, Siem Jan; Harvey, Andrew C.; Doornik, Jurgen A.; Shephard, Neil, *Stamp Structural Time Series Analyser Modeller And Predictor*, **1995**, Timberlake Consultants Ltd.

[11] Koopman, Siem Jan; Shephard, Neil; Doornik, Jurgen A., "Statistical Algorithms for Models in State Space Using SsfPack 2.2," *Econometrics Journal*, **1999**, Vol. 2, 113-166.

[12] Lommele, Jan A.; Sturgis, Robert W., "An Economic Model of Workmen's Compensation," *PCAS* **1974**, Vol. LXI, 170-189.

[13] McGuinness, John S., "Elements of Time-Series Analysis is Liability and Property Insurance Ratemaking," *PCAS* **1968**, Vol. LV, 202-254.

[14] Plato, *The Republic*, Book VII, **390 BCE**.

[15] Van Slyke, Oakley E., "Is Econometric Modeling Obsolete?" *Casualty Actuarial Society Discussion Paper Program* **1980**, 650-680.

[16] Venter, Gary G., "Classical Partial Credibility with Application to Trend," *PCAS* **1986**, Vol. LXXIII, 27-51.

[17] Zivot, Eric; Wang, Jeffrey; Koopman, Siem Jan, "State Space Modelling in Macroeconomics and Finance Using SsfPack in S+Finance," State Space and Unobserved Component Models, **2004**, Cambridge University Press.

Abbreviations and notations

ARIMA, auto-regressive integrated moving average
ARMA, auto-regressive moving average
ML, maximum likelihood
MLE, maximum likelihood estimation
NCCI, National Council on Compensation Insurance, Inc.
RMSE, root mean squared error
STS model, structural time series model
UC model, unobserved components model

Biographies of the Authors

Jonathan Evans, FCAS, MAAA is an Actuary at the National Council on Compensation Insurance, Inc.
Frank Schmid, Dr. habil., is a Director and Senior economist at the National Council on Compensation Insurance, Inc. He is also a 2006 Hicks-Tinbergen co-medalist for the best paper published in the Journal of the European Economic Association in the period corresponding to the years 2004 and 2005.

Designing a New Automobile Insurance Pricing System in China: Actuarial and Social Considerations

Daqing Huang and J. Tim Query

"To make simple things complicated is simple; to make complicated things simple is complex."
-- Schpagin

I. Introduction

There are numerous articles extolling the future potential of the insurance industry in The People's Republic of China (hereafter referred to as China) since the implementation of economic market reforms. While the current size of the Chinese insurance market is much smaller than many industrialized nations, the rate of premium growth is among the highest in the world. The national premium income had reached nearly US$60 billion in 2005, up 13.95 percent over 2004, and 3.09 times of the amount in 2000. Although the insurance industry has garnered increasing interest from researchers, little has been written about actuarial practices in the world's most populous nation.

Supreme Court Justice Stephen Breyer recently commented that international opinion can be relevant in determining fundamental freedoms in a more global society. According to Breyer, "U.S. law is not handed down from on high even at the U.S. Supreme Court," he said. "The law emerges from a conversation with judges, lawyers, professors and law students. ... It's what I call opening your eyes as to what's going on elsewhere." While the role foreign countries should play in dictating American law is debatable (and beyond the scope of this study), social researchers, including those in economics and business, have unearthed valuable insights when researching corollary issues beyond our borders. Examining the inimitable challenges facing casualty actuaries in China, recognizing the unique history of their social and legal framework, and studying the creative solutions used to overcome these obstacles should improve the understanding of our global diversity and is warranted.

The automobile industry in China has a relatively short but impressive history. Backed by technical assistance from the Soviet Union, China started producing automobiles in the 1950s. The automobile sector was under strict central planning in terms of investment, production, and consumption until the early 1980s. Total production of automobiles was only 100,000 in 1971. Because motor vehicles were categorized as capital rather than consumer goods, private ownership of cars was not allowed. There have been a number of peaks and valleys in automobile production over the past three decades, but the overall trend since the mid-1990s has been steadily accelerating.

When China's economic reform was introduced in the 1980s, the automobile sector gained momentum and began to expand significantly. Automobile production in China exceeded one million units by 1992, and that figure was doubled eight years later. Over the past two years the number of automobiles produced in China has doubled yet again with China surpassing France to become the fourth largest producer of automobiles in the world. Around 2008, China will likely surpass Japan to become the second biggest producer of automobiles in the world (Pan, *et al.*, 2004).

See Figure 1, page 99

Originally, the People's Bank of China (PBOC) was responsible for regulating the insurance industry, but in 1998 this sector was transferred to the China Insurance Regulatory Commission (CIRC) in order to streamline intensify financial reforms, minimize financial risks, and shore up the fledgling financial services industry. The CIRC oversees insurance business operations; formulates and enforces related laws and regulations; protects the interests of policyholders; develops the insurance market, maintains order and ensures fair competition; promotes insurance industry reform and restructuring; and sets up a risk evaluation and advance warning system to minimize insurance risk. Because many firms were making up for losses in their life insurance business by borrowing from their property insurance business, in 1998 the CIRC decreed that insurance companies could no longer handle both life and non-life business (Allison, 2001).

To acquire a better understanding of the insurance industry in China, a discussion of the auto market's history and previous versions of auto insurance is constructive. The Chinese insurance industry was initiated with the first insurance company established in Shanghai in 1885. By 1949, the total branch network of the domestic insurance industry in China consisted of more than 600 firms, but its market penetration was only about 25%. In comparison, foreign insurers totaled about 60 firms, yet had a 75% market share. After the People's Republic of China was established, it created a state-run insurer, the People's Insurance Company of China in 1949. For various social, political and economic reasons, foreign insurers were increasingly motivated to discontinue their operations in China and were completely out of the country within a few years. From this point forward, the state-owned insurer, the Peoples' Insurance Company of China, was essentially the only insurer operating in China. In January of 1959, the domestic business of PICC was suspended due to the restrictions on private ownership of property and the implementation of comprehensive social entitlement programs which lasted until 1979. A major policy shift to reform and open-up China was initiated shortly after the death of Mao-Tse Tung in 1976. China re-established its insurance sector with property business-lines resuming in 1980 and life assurance in 1982. The monopoly of PICC remained in place until the creation of a second insurer, Xinjiang Corps Insurance Company in 1986, followed in 1990 by Ping An Insurance Company and CPIC.

At the beginning of the 1980s, the property casualty loss ratio was about 64%, peaking at 82% in 1986. By 1994 PICC insured 5,400,000 vehicles, 920,000 tractors, 2,190,000 motorcycles, 97,000 ships, and 12,000 fishing boats, with a total premium of 14,088 million RMB, an increase of 39.04% from 1993. The insurance contract wording during this time was relatively simple with broad coverage and few exclusions. Due to the combination of increasing exposure to thievery, higher accident frequency, and premiums lagging the rising prices for cars, accessories, and automobile parts, the loss ratio rose to nearly 79%.

In response, PICC applied to the regulator, Peoples Bank of China, for a premium increase and for permission to restrict coverage. In 1995, coverage for theft was removed, third party liability changed from unlimited to limited, and the premium nearly doubled. Insight into the ownership of an automobile in China during the late 1980's can be found in an article by Bates and Goldstein (1989), who were expatriates in Beijing at the time. Chinese cars were purchased directly from the manufacturer as China had yet to open dealerships. Cars, trucks, motorcycles, bicycles, pedi-cabs, animal-drawn carts, and pedestrians competed for limited road space. In addition, widespread public awareness of traffic regulations was not well understood by pedestrians, and traffic lights were routinely out of service during power cutoffs. Night driving posed its own challenges in China, as trucks would frequently drive with only one working headlight or none at all. Often, the Chinese authorities would hold the driver of a motor vehicle automatically liable in an accident involving a motor vehicle and a pedestrian or bicyclist. If liability was split, Chinese police would unilaterally determine the apportionment.

The decision by China to join the World Trade Organization (WTO) has resulted in numerous changes to that its insurance laws and regulations. The Standing Committee of the Ninth National People's Congress made the proclamation to modify significantly the insurance laws governing China on October 28, 2002. Guided by a combination of governmental promises, required changes contingent to joining the WTO, new objectives, and international traditions, the Committee identified 38 items for revision. The amended terms under the new insurance laws involved the areas of insurance company operations, contracts, inspection and management, insurance agent and broker duties, and general rules. In addition, substantial technical modifications on rate filing examinations and relaxation of rules regarding the item and rate approval system were modified.

The China Insurance Regulatory Commission (CIRC) asks for the same information as the U.S. for filings: actuarial memoranda, illustrations, sample policies, and details of the calculations. The CIRC does not request cash flow testing, and all reports are done on a statutory basis. Companies do not need approval before selling their products, and only unusual products receive closer attention. However, credentialed actuaries are required to sign for life filings. The CIRC also is implementing a requirement for property casualty filings to include a review by a Fellow of the Casualty Actuarial Society (Yang and Lu, 2004).

See Table 1, page 101

Another influential factor advancing these changes was the Road Traffic Safety Law of the People's Republic of China, which was adopted at the 5th Meeting of the Standing Committee of the National People's Congress on October 28, 2003, and took effect on May 1, 2004. It is China's first national law on road traffic safety and is expected to improve current road and traffic conditions. The detailed implementation methods are formulated by the State Council.

The effect on motor insurance and development in China, as expected, has been profound. For example, Article 17 of the Law requires that the State implement a third party liability insurance system on motor vehicles and set up social assistance funds for road traffic accidents. In addition, the current tort system for resolving road accidents will be replaced with a "no-fault" system.

The method of insurance supervision and management has been improved with decentralization and regulations on insurer financial strength has been toughened. The Supervision Department continues to be responsible for monitoring standards as previously, but individual insurance companies now have the latitude to design their own products and rating structures.

Currently, there are four main kinds of motor insurance coverage in China:

- Motor physical damage (without theft coverage)
- Compulsory traffic accident insurance (limited liability, with 20% no-fault)[1]
- Voluntary traffic accident insurance (fault, one of the three kinds of coverages recommended by the Insurance Association of China, or excess of the limited compulsory traffic accident insurance)
- Endorsement (designed by insurers)

As a result of these regulatory changes, the requirements for both vehicle and driver will be stricter and more standardized. Motor insurance coverage has been extended with provisions in place to mandate that insurers adjust rates for both private and commercial vehicles. Foreign insurance companies will be excluded from Part 2 coverage as compulsory insurance is not covered in China's WTO agreements with other countries.

The purpose of this paper is to provide an overview of automobile insurance regulation reforms currently taking place in China. Comparisons to and contrasts with the more mature, yet continuously challenging, automobile insurance industry in the United States and other developed nations are incorporated throughout the paper. Current actuarial principles in practice, given the low availability of credible data, are a focal point of much of the paper. Ratemaking challenges faced by casualty actuaries in China, cultural and regulatory hurdles, as well as the state of the actuarial profession in China today are also discussed.

1.a. Transformations Affecting the China Insurance Market – Automobile, Traffic, and Road Conditions

The average retail cost of an automobile in China has been dramatically reduced since 2002. The cheaper auto prices were brought about by tariff adjustments for imported cars, and increased production domestically. This trend is expected to continue. The China Auto Price Index (CAPI)

[1] Also known as compulsory third party liability insurance. All owners and/or managers of vehicles on China's roads are required to pay premiums for this compulsory insurance by October 1, three months after the new rules were enacted on July 1. The third party liability of the compulsory traffic insurance policy has a limit of 50 thousand Yuan (about US$6,000) for fatalities and injuries, 8 thousand Yuan for medical treatment, and 2 thousand Yuan for property. The average annual premium for a family car is about US$125.

shows the average prices of cars have decreased about 13% from January 1, 2004 to 2005 and continue to descend.

The rationale for tracking automobile prices is that the most important rating factor for automobile insurance in China is what's known as the "sum insured." The sum insured is basically the value of the asset being insured, which in this case is the price of a new, this year's model car (not the depreciated value of the original vehicle).[2] Therefore, insurance premiums are expected to be positively correlated with decreasing automobile prices. As can be seen in Figure 2 and Table 2 below, automobile prices are trending lower, but are still relatively more expensive than in more developed countries. Utilizing SAS to calculate the sample CPIC data from 2003 to 2005 on an accident year basis, the sum insured estimates are found to be weakly correlated with paid losses (the correlation coefficients, calculated by using Pearson's product-moment coefficient, are between 0.150 - 0.189).

See Figure 2, page 100

See Table 2, page 102

The number of vehicles registered in China has increased sharply in recent years. Changes in the national economic policy, coupled with this decrease in prices, have stimulated the demand for individuals to have their own private automobile. National car production in 2003 increased by 38.5% compared to 2002, and sales increased about 36.7%. By the end of 2003, more than 96 million motor vehicles traveled the roads of China, including 24 million private passenger automobiles. China has more than 100 million vehicle drivers. Of this total, 54 million were drivers of automobiles, according to statistics. Rising consumer wealth has been a major contributory factor to the sudden explosion in the private passenger automobile market. The purchasing power of Chinese consumers, defined as the value of a particular monetary unit in terms of the goods or services that can be purchased with it, generally measured by income, has risen to the

[2] This caused quite a number of problems, such as giving rise to the risk of fraud as the "asset" depreciated with usage. To address these problems, CPIC uses a total loss sum insured and part damage sum insured calculation.

critical level traditionally associated with car consumption in other markets (Min, 2005). Similar relationships have been found in auto insurance. Outreville (1990) developed a model specifically for property-liability insurance demand and tested it with a cross section of 55 developing countries. He found that the level of the gross domestic product combined with the level of financial development were the only factors explaining the level of development of property-liability insurance demand in developing countries.

The vehicle product structure is also undergoing changes. According to a 2004 report by the China Association of Automobile Manufacturers, the 2001 market share of annual output for trucks, buses, and cars was 63.8%, 24.8%, and 11.4%, respectively. In 2003 the percentages of vehicles produced by class were 27.7% (trucks), 26.9% (buses), and 45.4% (cars). The China Council for International Cooperation on Environment and Development has identified the following major problems associated with the automobile sector:

- First, the scale of production by individual manufacturers is small. While production efficiencies are improving, there are still geographic and supplier-related problems.
- Second, the increase in motor vehicles causes air pollution, especially in urban areas. The World Bank reported in 1997 that Beijing had one tenth the number of cars as Los Angeles, but emissions of pollutants were almost the same.
- Third, demand for oil poses a serious challenge for the energy supply in China. One response by China to the oil situation is a move toward producing liquid fuel from coal, which is more abundant in China.
- Fourth, encroachment on land for roads and parking constitutes an increasing challenge for the expansion of the automobile sector. In some cities, space is already at a premium. Rent for an underground parking space in downtown Beijing can cost more than the annual income of an unskilled worker (Pan, *et al.*, 2004).

The vehicle retention rate, i.e., number of vehicles in use, is also increasing quickly. Despite improvements in the country's infrastructure, accident numbers continue to be a problem due to the layout of many major cities. Another factor contributing to the accident rate is the lack of parking in many residential areas. According to statistical data covering the year 2005, there were 450,254 traffic accidents, resulting in 98,738 fatalities and 469,911 injuries. The direct economic impact of these accidents exceeds US$235 million. [3] Efforts to improve the safety of road conditions are not keeping up with the rapidly increasing vehicle ownership rate. As a comparison, the National Highway Traffic Safety Administration reports that the United States had 6,181,000 police-reported motor vehicle traffic crashes, 42,636 traffic crash fatalities, and 2,788,000 injuries in 2004. There were 198,889,000 registered vehicles in the United States in 2004.

1.b. Transformations Affecting the China Insurance Market – Competitive and Agency Considerations

Since China joined the WTO in 2002, insurance companies in China have faced many challenges. One major hurdle has been meeting the need for more agents in a concentrated Chinese insurance market. Statistical analysis using both the Herfindahl Hirschman Index and the x firm concentration ratio (CRx) empirically tests for market concentration and is illustrated in Table 3 below.[4]

See Table 3, page 102

Do insurers in China have to deal with hard and soft markets? In a study comparing underwriting cycles in emerging markets in Asia with developed markets, Chen, *et al.* (1999), found that second-order auto-regression results support the existence of the underwriting cycle in Asia. Their results also seem to indicate that underwriting cycles are mainly related to the pace of the economic growth in those countries, and that factors affecting changes in premiums generally differ from those found in developed nations.

The numerous small scaled non-life insurance companies are eager to increase their market share, and generally do so by pursuing a strategy of cost efficiencies and competitive pricing. As a result, many suppliers of insurance are operating below normal equilibrium levels. If this state of affairs continues, the consequences of market failure could potentially reverberate throughout the insurance industry, with a contagion effect that would negatively impact even healthy insurers. There are both

[3] For comparative purposes, according to the CIA World Factbook, China's estimated Gross Domestic Product (purchasing power parity) in 2005 is estimated at US$8.859 trillion.

[4] Two commonly used measures of market concentration are the Herfindahl Hirschman Index (HHI) and the x firm concentration ratio (CRx). The CRx is simply the sum of the market shares of the x largest firms in the market in question. The Herfindahl-Hirschman Index (HHI) is generally considered a superior measure of market concentration. The HHI is the sum of the squares of the market shares of all firms in the market.

formal and informal sharing arrangements in place in China. Effective January 1, 2005, a guaranty fund type assessment was initiated. Every property and casualty insurer must contribute one percent of their premiums (net of reinsurance) into the fund until the value equals 6% of all insurers' assets. While in the future, all insurers will be ultimately responsible for all losses, such is not the case today. Because the Chinese market is developing, most insurers are still owned directly or indirectly by the state. As a result, if a small insurer were to lose market share quickly and its solvency margin falls below regulatory minimums, shareholders can be easily "persuaded" to provide supplemental capital.

In the meantime, the number of agents is growing quickly. There were 209 entities prepared to start business as of May 15, 2003 (the most recent data available). The 209 entities consisted of 157 agencies, 18 brokerage, and 34 appraisers. There were also another 551 companies categorized as medium-sized poised to begin insurance operations in the near future.

Standardization of the overall insurance market is proving difficult, mainly due to the PICC's special position as a state-owned insurer with over half of the market share. During reformation of the motor insurance industry in 2003, PICC's near monopoly of the property insurance market resulted in an abnormally low profit balance point which posed operational difficulties for the rest of the property insurance market. This situation has been alleviated somewhat since the Chinese government privatized PICC through an Initial Public Offering on the Hong Kong and New York stock exchanges in November of 2003. Privatization is generally considered a benefit to market development - by improving the transparency of the whole industry and bringing greater security to policyholders. Increased creativity in the market stimulated by this shift to a more capitalistic environment is expected to move the insurance industry closer to the modern business enterprise system.

Another regulatory challenge is protecting the continued solvency of smaller insurance companies and their ability to compete in the area of compensation. With market conditions resulting in 10% of insurance companies owning 95% of the market and the other 5% of market share split among 90% of insurers, the competitive environment is increasing the risk of default for a number of companies.

Changes in the regulation of the insurance industry, following the transformation of China's economic policy from a centralized government system to an increasingly capitalistic system, have motivated modifications in management styles as well. For example, some insurers are focused on increasing market share as their primary objective, while others are principally concentrating on earnings as a strategic goal. The differing structures and rating systems used in China, discussed later in the paper, have resulted in divergent measures of performance and management techniques for increasingly innovative insurance companies seeking a competitive edge.

Non-life actuarial techniques are receiving increased attention in China. Refining the existing non-life actuarial methodologies has been driven by the motor vehicle rate reformation. As a result, a number of insurance companies have contracted with foreign actuarial consulting firms or individual qualified actuaries with international work experience to assist with rate design, etc. While these collaborations have accomplished much and provided important transfer of knowledge, the relationship has not been a perfect solution. Since laws, regulations, management, roads, vehicles, etc. are in a constant state of rapid change, a number of actuarial assumptions commonly found in more established insurance markets around the world are not suitable for China's unique circumstances and market conditions. The complexity of a formerly Communistic economic system gradually incorporating capitalistic principles requires practical application of actuarial theory and distinctive tools necessary for feasible solutions. Simply transferring actuarial techniques used abroad in other insurance markets is not an acceptable option at this time.

During the initial design and calculation of the automobile insurance product and rate structure, information technology is inevitably involved, albeit at a more rudimentary level than in more advanced economies. The technology techniques are still in what most countries would consider the early development stage, and China just recently started construction of a fully integrating data bank. Network quality is still in need of major improvements, and until recently, most actuarial software was limited to EXCEL™ and other worksheet based products. Fortunately, more Chinese insurers are incorporating more sophisticated actuarial software through purchases from outside vendors or by internal development. Statistical tools such as SAS™ are being used to analyze the impact of reforms to the rate and contract clauses of motor insurance policies. It should be noted that the majority of Chinese-educated actuaries are still relatively recent college graduates. As such, there is still a learning curve that the insurance industry in China will need to endure patiently until time resolves this issue.

As can be deduced from this discussion, actuarial science in the Chinese automobile insurance market is still more of an "art" than a "science," with a little luck thrown into the mix. In other words, the tools and training needed to adequately evaluate factors in ratemaking are lagging behind what is ideal under these volatile market conditions. For example, the city of Shenzhen in Guangdong province initiated a rate reformation (reduction) that was the first of its kind in China. This decision was derived from a combination of actuarial determinations and political motives. As a result, demand for the "cheap" insurance was greater than anticipated, resulting in woefully inadequate rates for most companies. The first insurance company to implement price reductions in the Guangdong province of China did not adequately anticipate the subsequent result of such action and was almost forced to withdraw from the automobile insurance market in Guangdong. However, since then it has been determined that the rates were not economically feasible and the city returned rates to their original levels. Effective October 1, 2001, the CIRC deregulated the automobile insurance market on an experimental basis in the Guangdong province by adopting the "file and use" system. Under this system, insurance firms are allowed to design more customized terms and premiums for specific vehicles and geographic areas. This arrangement was expanded countrywide on January 1, 2003, with the elimination of many generous insurance contract clauses and relaxation of many rate regulations.

The efforts of the CIRC, insurance companies, and various professional associations have resulted in four jointly determined objectives in the reformation of the automobile insurance market in China. The first objective is that the automobile insurance market becomes a steady and stable operation. The second objective is for reformation to push the property insurance industry's business model adjustment forward. This includes balancing risks versus premiums, and supplies versus needs (which includes matching the ability of insurance employees with need of insurers).

Third, that insurance company operations and management levels conform to new expectations; performance and market-oriented principles are to be gradually incorporated into the day-to-day functions of the business. Fourth, is that insurance companies eventually exert more control over their various operations, which in turn is expected to increase creativity in the product development, financing, and ratemaking aspects of the industry.

II. Motor Insurance Pricing and the Actuarial Cycle in China

The first year for actual implementation of motor insurance reforms in China was 2003 with the preparation work required of insurance companies completed in 2002. The actuarial working procedures are very similar to western countries, but there is a greater difference in the actual content versus developed countries. One of the main reasons for this dissimilarity is the weak actuarial foundation and practice in the insurance industry.

2.a. Information Collection

The largest challenge facing actuaries when attempting to apply rate-making principles in China is identifying, collecting, and standardizing the relevant data. In the United States there are services such as the Insurance Services Offices, Inc. that are the state mandated statistical agencies whose purpose is to collect, standardize, and provide aggregated data to the various insurance departments and back to the participating insurers. In Europe, there is freedom to access helpful data, but little support. There is at least one market in Europe where a reinsurer's consulting arm has developed detailed pooled data for auto pricing (Schmitt, 2000). These resources are not readily available in China. Even firm-specific data provided in-house is not considered statistically credible, including PICC. To compound the difficulty, many definitions are not uniform, even across departments within the same organization. The result is a lack of a benchmark and data quality that fails to satisfy actuarial principles of data quality.

So how does an insurance actuary in China operate under these circumstances? Very creatively! Actuaries basically operate in the mode of obtaining as much useful information as possible in any way possible. Various types of information, copyrighted or not, qualitative or quantitative -- from the automobile industry, traffic department, agents, universities, even web sites – are fair game. The validity of data obtained is sometimes verified out of necessity by professional colleagues, actuarial students, and co-workers. The general principles of data quality have been revised given this challenging environment, and are described as the following:

> ***Reliability*** – While much information is collected, a lot of it is not what would normally be considered credible. Some of it is also contradictory, and selected information is outright confusing. Often it is necessary to filter or combine pieces of data to uncover something of value. The actuary's mindset in China is that any data collected has some potential value. Obviously, in a country like China the cost-benefit analysis of data collection is even more critical. Nevertheless, enough data to satisfy ratemaking criteria is

essential. Actuaries in China like to say that "We have to try to redefine GIGO [garbage in garbage out] as garbage in leads to *gold* out" (loosely translated).

Completeness – Little active data comes from market research. Since the information collected is usually not complete, part of the responsibility of the data miner is to attempt to fill in the missing information. Often, sampling techniques or further market investigations are needed. Also, the types and coverage of motor insurance change frequently, so special skills are needed to re-classify the data, sometimes with the use of text mining techniques. For example, the Chinese name usually consists of two or three Chinese characters, and by excluding some of the characters which are not normally contained in the names of individuals the ability to identify privately owned cars is enhanced.

Timeliness - Changes are occurring in the world at a faster rate, resulting in more rapid obsolescence of historical data. Ensuring a solid infrastructure of ratemaking that will last for a long time is very expensive. At this time, insurers in China are relying on practical and inexpensive methods to reach this objective.

2b. Motor Insurance Rating Criteria

Driver Characteristics

Characteristics and demographics of insured drivers are extremely important factors in a mature automobile insurance market. This includes information about age, sex, occupation, etc. Traditionally, the characteristics of the vehicle being insured were of primary concern to Chinese insurers, with driver attributes given little or no consideration. Culturally, insurance agents and their customers are not comfortable with providing information about the potential insured to the insurance company. As one can imagine, this has been an area of dispute among foreign and domestic insurers, as well as regulators.

Objectively, the factors constituting the drivers ability to operate a motor vehicle safely are hard to describe by the simplest characteristics. In mature insurance markets, the claim frequency for young men as a risk category is usually high, but it is impractical to prove that an individual young man is necessarily dangerous. The degree of risk is normally judged by investigating the past record of a set of drivers with the same or very similar characteristics.

This is not a reasonable approach in China at this time, as most families who do own automobiles only recently became auto owners. An intuitive argument can be made that all of these drivers relatively new drivers, regardless of age, sex, or occupation and therefore, have the ability to improve their driving ability with experience. This is a major difference between China and western nations, where there is typically a longer history of numerous family members with comparatively longer driving experience. The number of new private car owners and novice drivers in China is rapidly increasing.

Another complexity is the fact that a number of pedestrians are not completely familiar with the newly instituted traffic regulations, including the vast number of Chinese that are migrating from the rural areas to large urban areas in China. The combination of these factors has resulted in unusually high frequency of accidents in some instances, amounting to 50% occurrence in some risk categories and nearly 100% occurrence for some classes of individuals. There are insureds who have submitted over thirty claims to the same insurer over a multi-year period. One proposed solution to this problem has been instigated by some insurance companies, by binding the expiring loss ratio with the No Claim Discount (NCD). The CPIC, the second largest property casualty insurer in China, has also implemented this method to revise the coefficient for reward for good drivers and a debit to premium for bad drivers to recognize their propensity toward accident frequency.

Novice and young drivers

Novice drivers are particularly crash-prone in China as elsewhere. Novice drivers are 5-7 times more likely to crash than drivers with two years of experience.[5] Like other countries, young drivers also have more claims than other age categories. However, China's novice drivers are not as correlated with younger drivers as in more developed countries. Since motorization in China is so recent, there are a large percentage of novice drivers of all ages. By the end of 2004 there were 430,000 novice drivers in Beijing, out of 3.5 million total drivers in that city, and there were 5.1 million novice drivers in all of China. Unfortunately, it is difficult to distinguish the novice drivers, as it is typical for a car to be shared by a number of drivers. While some system that assigns the

[5] Based on sampling thousands of claims from CPIC and calculating the accident frequency with drivers of different experience levels.

highest risk driver to the vehicle for purposes of insurance rating is preferable, there is currently no system available to accurately identify such drivers. Because insurers in China have very limited information on drivers, related factors are sometimes used in ratemaking, such as correlating new cars with new drivers. Obviously, this is only remotely feasible while the automobile industry is in relative infancy. Insurer databases are not sufficiently accurate, and brokers will not provide the true risk characteristics of customers for fear of losing the business. At this point in time such a system is scheduled for implementation as soon as possible after compulsory insurance requirements go into effect on July 1, 2006. In the short run this will negatively affect automobile insurer loss ratios. China is a big country, with a large rural population that has not participated in the economic growth of the nation. The gap between the rich and less affluent is growing. The automobile insurance situation should improve in the major cities with time. However, the challenges discussed in this paper are expected to persist for decades on a country-wide basis.

Risk of Vehicles

Another important risk factor is information regarding the vehicle(s) being insured, which can include the price of the vehicle, the age, depreciation, etc. At this time, the "sum insured" (value of the vehicle) is the most important risk factor in China. In contrast to other countries, Chinese insurers do not use vehicle years as the exposure unit instead of sum insured. The mechanical application of techniques and methods, for a system in which the exposure base is the value of the vehicle rather than vehicle years, will create difficulties. The answer may be to use another model, such as that of homeowners insurance, where value, not home years is the base. In a mature motor insurance market, the risk factors associated with vehicles are generally easy to measure. Car repair organizations like the Research Council for Automobile Repairs (RCAR) can classify vehicles into groups. After attaining the vehicle's model and type, finding out the risk class of that vehicle can be done by checking the relevant manual. There is no such classification method in China. At the same time, with rapidly increasing numbers of vehicles, the automobile brands offered are somewhat chaotic. Even for the same make, vehicles may have totally different characteristics with large differences in risk. Generally speaking, a relationship exists between the automobile type and sum insured. However, while a reasonable classification of the vehicle type can alleviate the risk exposure problem on sum insured differences, it does not identify properly the real risk of a loss when it uses only using the insured sum.

The automobile insurance industry's loss ratio is increasing rapidly because of the continuous decrease in vehicle price accompanied by decreases in premium, the rising price of accessories, and the relatively fixed labor rate and other monopolistic operations associated with repairs and parts suppliers.

Environment of Vehicle Use

One of the characteristics of automobile insurance is that the insured object is mobile and the differences in territorial differentials large. Territory of usage is one of the main risk characteristics. The territory factor in China not only pertains to the location where the vehicle is garaged, but also incorporates the natural environment, economic environment, customs and habits etc., that includes the safety record of the state, native road conditions, and population density (Ng and Schipper, 2005). The risk factors applicable to the vehicle and environment in which it will be used can all be represented by the characteristic values to describe the risk's category. A two-dimensional table is applied to rate construction.

Distribution

In China, the distribution system is somewhat unique. Usually the car dealer and salesman exert control over all auto-related transactions. As such, they are also often the conduit for new automobile owners to find an appropriate auto insurer. If an insurance company wants to connect with customers, it has to either make arrangements with automobile dealers or spend major resources through media marketing. Insurance companies are also reliant on automobile repair workshops for controlling repair cost expenses and indirectly for the service quality of claims.

Other factors - In addition to rating factors like personal characteristics, vehicle, and territory, there are other dynamics that need to be managed. As with insurers in other regions of the world, China's insurance industry struggles with insurance fraud. Gaps in the risk classification system can invite dishonest behavior if the system is not reputable and rigorous. Insurance companies attempt to reduce this occurrence with close analyses of various methodologies with the goal of more classification groups and smaller ranges. Sometimes, this will cause interaction between the factors.

According to the CIRC regulations, the total price change is limited to a 30% range up and down now, and the premium will not allowed to be readjusted again within six months.

2c. Lack of Data Sharing

A fact that cannot be over-emphasized in this analysis of automobile insurance in China is the competitive environment. This competitive market structure makes it very difficult to share information. As mentioned earlier, the market share of insurance is still very concentrated, with a handful of companies owning a significant market share in China.

These large companies have extensive branch and sub-branch networks that are vertically integrated in every aspect of the insurance market. There is an understandable reluctance by these companies to share resources and information with smaller companies. This leads to heterogeneity in underwriting practices for each company, and information that produces bigger deviations than the actual market. For example, one vehicle may not be a bad risk in the market and yet have less than necessary and sufficient data backing it than another type of vehicle. Loss ratio calculation results may show that its loss ratio is unacceptable. This conclusion will restrict the underwriting of this type of vehicles producing a "vicious circle" with less data available in the future. In fact, this kind of system deviation exists extensively, be it within an insurance company, a whole profession, or perhaps in a country. In China only a few insurers dominate the market, and individual insurers are left to collect relevant information on their own instead of relying on an industry or governmental agency. During the actual automobile insurance rate-making process, horizontal correction should be made for this kind of error and deviation by the use of other data or resources such as manual rating information.

2d. Data Cleaning

Accurate data and complete information are basic conditions for ratemaking. For a variable collected from insurers with seemingly ever-changing markets, data cleaning is the most important work among the basic operations during automobile insurance ratemaking. Assuring that data are accurate, complete, timely, and adequate is a critical problem in China. Although the automobile insurance database is more complete than for some other insurance products in China, the quality of automobile insurance data used would still be considered poor by global standards. For larger insurers such as CPIC, the second-largest property insurance company in China, the data quantity is sufficient. However, even these data are insufficient for satisfying actuarial requests due to technical problems (business information switching, data connection variation and Chinese support functions), and general business management problems (changing products, variety in clauses, and management methods).

An investigation into these phenomena has resulted in the following conclusions. First, data standardization has been low in the past, there is an abundance of garbage data, the information value of data is insufficient, and the support function of the information technology system is relatively weaker in such areas as historical data management, which usually accompanies a system switch. Second, since there were few private passenger automobiles in the past, in many respects all customers, whether commercial or non-commercial, are combined into one general customer base. A key to efficient and equitable personal car insurance ratemaking in China is to use suitable methods to distinguish the customer. In addition, some information needs to be filtered to reduce problems such as insurance fraud and the artificial interference factor, which refers to a reasonable adjustment related to the clash between aims of regulatory rules with the operations of the insurer.

During the process of data cleaning, choosing suitable data as the benchmark needs careful consideration. More localized data, such as daily reports provided by branch managers, is preferred. On occasion, some indices can use the headquarters' average values as the reference. Moreover, samples-drawing methods for certifications can be adapted to confirm some information if circumstances and conditions are appropriate. Data mining is a potentially useful tool in this stage.

2e. Changed Rates

Despite national unified clauses practiced by domestic insurers prior to the sweeping automobile insurance reforms, there were extensive differences in actual operational processes. The past unified clause constituted by the CIRC and enforced by all general insurance companies in China is very extensive, which allows for differences in its supervisory application at the local level. Discount phenomena and other market factors should be filtered and solved for this portion of errors, and is a challenging problem for rate-making professionals, unless one doesn't use loss ratio methods when making rates. Many branch companies' rates fluctuated multifariously in the year 2005's motor insurance data and the range was not very small. This factor needs be considered when the rate is being calculated.

2f. The Unsound Reputation and Insurance Fraud Problem

The market economy is still developing in China. Deception by both businesses and customers is still rampant. For example, potential insureds often intentionally give false information in an attempt to obtain insurance coverage at a favorable rate. Agents have been caught providing a dummy address to an insurer in a scheme to get renewal commissions. At this time, the ability to verify much of this information is cost prohibitive. The Chinese culture includes a high standard of morality, the ability to tell right from wrong. Therefore, many Chinese feel it is shameful that some individuals have violated the cultural standards of utmost good faith are violated.

Customarily, customer's information is usually controlled by the salesman and not by the company; so customer's information in the record-keeping systems can incompletely reflect the customer's true circumstances. Insurance companies run the risk of their salespeople changing jobs and taking part of their customers' information with them. There have also been problems with insurance salespeople, who can represent more than one insurer, essentially selling their services to the highest bidder. Motor insurance companies have taken advantage of this lapse in moral judgment by offering bribes for salespeople to promote their own product offerings over their rivals' products.

III. The Pricing Model

Issues normally associated with pricing models in developed countries may include such areas as class and territorial relativity analysis, reinsurance pricing, and assessment of the value of reinsurance structures to reinsured. Other required services might consist of reserve reviews, retention studies, and Dynamic Financial Analysis. The pricing models used in China are still relatively rudimentary, and an examination of the current environment provides valuable insight into the condition of the profession.

3a. Organization of Data

Organizing data is critical in a changing market. The most recent year's claims experience was chosen for analysis because this data was available on the new database. Because all claims are not yet reported for the more recent incurred months of this data, an allowance for IBNR claims will need to be added to the claim history collected. China is a rapidly growing market, and the quantity of underwriting and claims data will increase quickly with the passage of time. Since the environment is also changing, recent horizontal data is also segmented. For example, although there are some differences in males and females when estimating driving risks, other conditions are similar, enabling the entire information data supply to be utilized.

3b. The Treatment of Large Claims

Large claims or catastrophes are random events, with very high costs. Two methods are usually utilized to decide the critical value of large claims: the claim value and percentage of claim frequency. For example, when considering physical damage to the insured vehicle, a small car will not cause a large value of claim. Because China uses sum insured as the measure of risk exposure instead of vehicle year, it will produce different results from that obtained using traditional treatments. These events, if included in a particular rating classification, may make that rating classification look abnormally bad. Large claims, which will distort the data analysis, should be excluded or truncated, with the excess cost reallocated over all policies at a later stage. Since the insured sum are taken as exposure unit and there exists high relativity between the high insured sum and large claims, a great deal of the high insured sum's indemnity will be shared by the low insured sum policies. This will also lower the indemnity ratio of the high insured sum policies if the actuary uses of the usual and customary method to distribute large claims. Obviously, this is distorting results. For this reason, large claims need to be included in ratemaking in a functional relationship to the insured sum so as to correct the bias.

3c. The Usage of Loss Distribution

Zero claims usually arise if there is a dispute as to who was at fault in the accident, or if the amount of damage in the accident was small. If the amount of damage is small, the driver might not pursue the claim, because the loss of the no claims discount on next year's premium would exceed the total cost of damage. Some actuaries have suggested excluding these zero finalized claims from the analysis, as including them would lower the average claim cost to unrealistic levels. This might be a reasonable course of action if the percentage of zero finalized claims is changing over time.

Generally, the research and investigation of those policies on which the indemnity is greater than zero is more sufficient than that on which the claims are equal to zero. In fact, high-quality customers are often included in the part of claims that is zero so that the research is positively skewed for those customers and policies.

3d. Rate Smoothing

Smoothing can be achieved by first looking for linear relationships between resultant risk premiums within a risk category. If the fit resembles a linear relationship, then the actuary can smooth the rates to fit to this linear relationship, and this will make rates easier to quote. At this point in time, some insurers fit the rate curve with a changeable or fixed inclined ratio respectively.

3e. Bonus-Malus Systems

The traditional no claims discount system is designed with the assumption that the percentage of novice drivers is small relative to the total number of drivers, and that the drivers who hold the licenses driving experience is proportional to their age. This is not necessarily the case in China. Most drivers of private passenger vehicles are new drivers getting their driver's license for the first time, regardless of their age. Given that novice drivers have more accidents than experienced drivers, the claim frequency for the insureds in China is much higher than that in developed countries. If traditional Bonus-Malus Systems are used in China, there will be extensive losses suffered due to inexperienced drivers (those newly licensed and those holding a license but having never driven before).

3e. Generalized Linear Modeling

The main focus here is to seek the best-fitting model after making certain of the risk classifications. The Generalized Linear Modeling (GLM) has been one of the most popular tools used to rate motor insurance for decades. It is a highly valid methodology for vehicle insurance ratemaking, and can easily handle a large quantity of risk combinations being examined and establish complex relationships related to claim experiences. GLM is used extensively by actuaries world-wide and is the core technique for most rate-calculating software.

GLM primarily includes the additive and multiplicative models. The additive model, despite imperfect theoretical deductions, is created under some assumptions which cannot necessarily be satisfied in actual applications.

Chin Pacific Insurance Company, for example has encountered some problems during the process of additive model application. The sum of all increments totaled more than 100%, which exceeds the regulated limitation value (50%) mandated by the CIRC. Actuaries have had to construct a method of selecting one factor from among the three factors initial factors. It has been very challenging to implement in practice due to difficulty in distinguishing the loss factors. As a result, the applicability of the multiplicative model is considered superior to the additive model, and is being applied most extensively now in China.

3f. MAX model

There are some problems in appropriately identifying which factors are highly correlative with the risk dynamic. Different factors can have similar results; repetitive usage of single-variable techniques may cause repetition of calculations. Therefore, one of the obvious advantages of multi-dimensional analysis is to deal with the interaction and connection of many risk factors in order to seek the optimal combinations. Results can therefore be achieved more accurately than under a one-way analysis of claims experience. On the other hand, as the number of rating factors being examined increases, the volume in each data cell decreases. Multi-way analysis can create unwarranted fluctuations in the data simply due to random variation rather than any inherent differences in the data.

The key to the GLM application is in handling correlations of pricing factors effectively. Generally, actuarial theory suggests selecting the most important factor and discarding the others. As mentioned above, this approach is not feasible for ratemaking in China. Additional factors, whether correlated or not, are used to narrow the gap among relativities. Policyholders providing incorrect information will not receive a substantial discount from the rate system, and there is enough redundancy to correct it. For example, actuaries in China can use the price of the car, size of cylinder, manufacturer or type, age of vehicle, etc. to estimate the risk of the vehicle, although they are all correlated. It does not satisfy the assumptions of GLM. Other factors also may make it difficult to distinguish risks. For example, consider the case of an experienced driver with a defective car or vice versa. Often major problems can be traced to one specific factor, while the other factors in the model are considered accurate and reliable.

The Max model is recommended to solve these problems,. For those correlated risk factors K_i (i=1, 2, ...N), a variable MAX(K_i) is utilized in the calculation in order to improve the accuracy. This method can be still further extended to that non-correlated risk factor. It can also be combined with the traditional model as a mixed model.

For example, driving record, age and no claim discount are correlated or interaction effects. If for policyholder "A" is age 25 with a decent driving record and the relativities of the risk classifications are 1.50, 1.20 and 1.30 (no claim discount is compared with the average) separately; therefore, the relativity of these risk factors is 1.5 if the Max model is used, compared with 1.50*1.20*1.30 = 2.34 using the multiplicative model and 1.0+0.5+0.2+0.3 = 2.0 using the additive model.

For policyholder "B" age 45 whose driving record is bad and the relativities of the risk classifications are 0.85, 1.50 and 1.20, therefore the relativity of these risk factors is 1.50 if using the Max model, compared with 0.85*1.50*1.20 = 1.53 using the multiple model and 1.00-0.150+0.50+0.20 = 1.55 using the additive model. It should be noted that the Max model can be generalized with a special function G(k1,k2,...,kn), and mixed with other models to fit the actual risk classifications. This approach has limited practical applications for a variety of reasons that go beyond the scope of this paper.

3g. Forecasting the Target Market's Price

The purpose of the various analyses is to forecast of the price of the target market in the future. It is the core competency of the for-profit corporation. Loss development analysis estimates the number and value of claims not yet reported as well as adverse (or favorable) development on known claims, with each different type of risk, such as third party versus property claims requiring a separate development pattern. It generally takes more time for insurers to receive and settle third-party bodily injury claims than third-party property damage claims. Claims for damage to the insured's vehicle take the least time to be reported and settled. Furthermore, claim settlement takes much longer in countries with a common law legal system than in those with a civil law legal system.

To estimate expected claim severity and frequency projected into the future, historical estimates need to be trended to consider future changes (e.g., inflation).

Determining accurate loss reserves is one of the most challenging tasks facing the actuary. In a rapidly changing environment, an approach that previously provided accurate results may no longer be appropriate (see Lester and Fisher, 1975). During the process of calculating the automobile insurance rate, the cash flows and funds receivable accounts should be considered. The receivable account ratio, an important supervising and management index for some insurers, has increased up to 20% for certain special customers and channels in China. This perpetuates the existence of fraud premium and indemnity, by artificially boosting management incentive measures.

Financial checks as a support function of rate-making remains weak in China. The ABC method (Activity-Based Costing) is still considered to be in the early stages. The method for expense-sharing is somewhat simplistic at this time, so there is no reasonable and valid method for calculating the shared expenses for automobile insurance and non-auto insurance, such as the fixed and variable expenses, commission and brokerage expenses, and administrative expenses. This will directly affect the accuracy of rate-making and also result in difficulty analyzing operation and profits capacity and improving the management level.

Distribution channels are becoming increasingly important in China, and are influencing the indemnity ratio (total indemnity divided by total premium). Since the market is far from standardized, there are obvious differences in the service charges among the channels and even among the customers for the same channels, and this has a direct influence on the property insurance company.

The loss ratio has proven to be a very confusing concept in China. In the past a simple paid loss ratio was used. It was defined as the value of paid losses over gross premiums during a period of time. A similar definition of loss ratio is used today, but the value of the case estimates and IBNR are usually not calculated correctly, the raw data varies quite significantly, and it is extremely difficult to adjust the data and satisfy the assumptions. A significant percentage of shareholders at most insurers in China are still associated with the state, so there is not stockholder pressure to increase profits. These shareholders are more responsive to such measures as scale and market share. With ongoing efforts to privatize the financial services industry in China, i.e., the initial public offering of PICC, there is a greater interest in areas directly related to profits and earnings, although there remains a contingent of investors still focused on bigger scale. The diversity of assessments used to evaluate a company's progress by various stakeholders has created an interesting environment for corporate strategic planners and investor relations.

3.h. The Pricing Enviroment in Today's China

Pricing is not only a science; it is also an art. When results are finally obtained, it is after almost all the data have been adjusted. Most of the information collected has been filtered out, and added dummy variables to the model. In spite of this extensive fine-tuning and tweaking of results, occasionally absurd conclusions result. Under those circumstances, experience and intuition may come into play to judge the usefulness of such results.

As with other regions of the world, the most difficult and important course of action in China involves communication with top management. Since the history of non-life actuarial science in China is short, actuaries usually have middle and low level positions in the company. This results in certain barriers for communicating with administrative officers. There is also a balancing act involved when sharing important information with company executives. On the one hand, if the topic is too technical, the officers do not understand what it means and therefore are not willing to provide their support; on the other hand, if the topic is too elementary there is a tendency to believe that it is nothing new and has limited value at best. There is also the importance of balancing conflicts of interest within different units, which often requires an actuary to quickly develop diplomatic skills. Interestingly, company presidents must sign off on all claims over $50,000 to ensure that no fraudulent claims are paid. This is in contrast of the practice in North America where the claim VP signs off on all declined claims to ensure all legitimate claims are paid (Yang and Lu, 2004).

Presentation of the pricing structure is also very important. It must be clearly explained and straightforward for the underwriter or agent to use, in a manner that facilitates easy understandability for the insured as well. Otherwise, the insured will go to other insurers where the structure is easier to follow or because they are used to the old structure. In China, PICC dominated the market for a long time, so people are used to the pricing configuration in use there. Because of well entrenched habits, it will take time for consumers to become used to an innovative pricing composition.

IV. Other Considerations

A rate is an estimate of the expected value of future costs, and automobile insurance is a critical consumer service. So it follows that regulation in this market has an important direct impact on consumer welfare. There are many factors to consider and adjust in such a changing environment, yet as alluded to throughout this paper, most of the factors are full of uncertainty making it hard to model with accuracy. In reality, only rough estimates can be provided, making actuarial activities more intuitive than discrete in many respects. Unfortunately, a number of people, including company management, believe the results should be exact. Some of the designs under consideration are discussed during the ratemaking application in this section.

4a. Social and Cultural Environment

On May 1, 2004 the benchmark for compensating personal injury claims was revised in China. Additional items associated with the claim were included, and some of the existing items were valued higher than before. These changes are based on writings contained in a document with the lengthy title of "Interpretation of the Supreme People's Court of Some Issues Concerning the Application of Law for the Trial of Cases on Compensation for Personal Injury." The financial consequences of these changes are estimated to result in an increase in the cost of Bodily Injury of about 260%.

In recent years, Chinese has managed to grow its economy and maintain very low inflation or even deflation. Since 2003, however, signs of inflation have begun to emerge. The Consumer Price Index came in at about 4.4% in May 2004. Experience shows that the rate of inflation applicable to automobile insurance is usually higher than the inflation rate for all items.

The use of deductibles is very popular in other countries, but is not acceptable in China by most consumers. As one can imagine, this causes a high frequency rate, with the cost of claims settlement sometimes exceeding the actual loss. PICC set up a deductible within many of its policies for RMB500, which caused many of their policyholders to become angry and switch to another carrier. [6] It is a considered a major reason for the loss of market share by the company.

Despite the migration of young people from the rural areas to the urban areas, there are currently 30 million motor vehicles, more than 60 million motorcycles, and 10 million agricultural vehicles - including 8 million tractors – in the rural parts of China.[7] This is important because vehicles associated with the countryside and agriculture usually have experienced a very high loss ratio. If an actuarially fair premium was derived and charged to these rural insureds, most vehicle owners will be unable to afford it. With large numbers of such vehicles, it has been difficult to come up with solutions to resolve these problems.

To illustrate the vast gulf between insurance products designed in China compared to the United States, consider the following. A controversial new insurance policy that provides coverage for drunken driving activities in China has been approved by the China Insurance Regulatory Commission. Offered by Tian'an Insurance Company, the policy stipulates that the insurer will compensate a third party for injuries or property losses caused by a policyholder as a result of drunken driving. It is a common practice for Chinese businessmen to have dinner and drink alcohol with their colleagues, which they claim improves their relationships and business opportunities. (www.starinfo.net.cn)

4b. Compulsory Third Party Liability

The primary objective of compulsory automobile insurance in China is "to provide affordable, fair, and accessible treatment, rehabilitation, and compensation for bodily injury to, or the death of, third party road accident victims." Secondary objectives are to provide education and information to the community on scheme entitlements/procedures, and to promote road safety awareness with the aim of reducing road accident rates and resulting injuries and disabilities.

Although the Road and Traffic Safety Law already took effect on May 1, 2004, compulsory automobile insurance laws have not yet been enacted. There are a number of problems that need to

[6] The exchange rate was approximately 8.3 Renminbi (RMB) Yuan = 1 U.S. Dollar at the time of this article.

[7] Source: http://auto.news.hexun.com (a Chinese-language website)

be resolved before it becomes effective. Scenarios such as compulsory Third Party Liability Bodily Injury (TPL) with no-fault, and Physical Damage with fault or no-fault tort system must be carefully examined. Some estimates show that premiums will increase sharply if Physical Damage is within a no-fault tort system. Other issues currently being debated are: Should voluntary third party liability still be maintained under a fault basis? What should the limits of compensation be?

Social assistance funds for road traffic accidents will be created and maintained by legal mandate. This is expected to raise the premiums of automobile insurance as well. It is estimated that less than 30% of vehicles in China have any insurance coverage at all. With continued economic development and accompanying wage increases, the percentage of insured vehicles is projected to eventually rise to about 80%. The loss experience of vehicles not insured will continue to be much higher than vehicles that already have insurance.

There is concern within China that if it selects the no-fault basis of tort system, issues of fraud, abuse, and overuse will result in a very challenging environment for the insurance industry. Hopefully China will closely examine the no-fault tort system in place throughout the United States to assist in making an informed decision. The insurance industry in China is also working to build a platform for sharing information among insurers in the market.

At the present time, foreign providers are not allowed to enter the third-party motor insurance market. However, there is not a great desire to enter this line of insurance, as it requires more capital and is much riskier. If restrictions are lifted in the future, it will make the market more attractive for overseas companies.

4c. Additional Regulatory Issues

Although the history of the CIRC is very short, the criterion and standards of administration are developing. CIRC is still in the midst of a paradigm change. Areas such as the basis of insurance accounting and solvency margin techniques as they relate to regulation are under review, and will hopefully be improved. One impediment to smoothly and rapidly implementing these improvements is the shortage of skilled staff and board members with knowledge and experience in the insurance industry. There is also a shortage of experienced actuaries qualified to write financial condition reports.

4.d. China's Actuarial Profession

Actuarial services are developing in China, but still have a long way to go before it can be considered comparable to the standards experienced in other countries. The actuarial profession in China is relatively young, but the greatest immediate concern is the shortage of experience in the operations of well-run companies. As recently as November 2003, the Asia Insurance Review e-weekly noted that, "with the explosion of the insurance industry in China, there is now a shortage of actuaries as out of the 400 plus people working as actuaries today in the industry, only 30% are qualified, according to industry sources."

The first actuarial exam center in China was established in 1993 as an extension of a comprehensive training program for the People's Insurance Company of China by Manulife (Shen, 2000). However, formalized non-life actuarial education only began in earnest during the fall of 2004. Actuarial students have the option of taking the Fellow of the Society Actuaries, FIA, or China's internal exams. Most take the FSA exams as the FIA exams require better English skills and the Chinese internal exams have exceptionally challenging calculation problems. Lack of qualified and experienced actuaries limits the ability of companies to provide financial condition reports to an adequate standard. As a result, many companies employ well qualified and experienced actuaries from Hong Kong or overseas to set appropriate standards for the financial condition reports, and non-life actuarial work usually is a very closed practice of business. Ideally, when professionals and consultants from outside China come in to perform this work, they need to spend sufficient time to clarify the meaning of data. Sometimes when faced with the choice of explaining Chinese data to foreign actuaries, and spending the time and resources to train local staff on specific facets of actuarial principles, it is more efficient to do the latter. In many respects, actuarial principles are similar around the world, but practical applications can be quite different.

The most difficult problem when applying actuarial science in China is not the techniques, but the theory and communication. Actuaries in administrative positions are considered to be in relatively lower standing than other disciplines. Subsequently, in situations where theory is quite complicated, the administration usually gravitates to the simple choices. Educating the industry on what exactly actuarial science is, and its importance to the insurance concept, takes patience.

Conclusion

The motor insurance market in China is developing at a rapid pace, and business models are continuously changing to avoid being left behind. It was only a short time ago that non-life actuarial techniques were introduced in the rate-making and motor insurance product design segments of China. Cultural and governmental hurdles need to be overcome before traditional techniques will be widely applicable in China. As a result, many projects to integrate generally accepted actuarial principles into the unique Chinese model are being studied and gradually implemented. The Chinese saying "May you live in interesting times" certainly applies to the casualty actuarial profession in China both today and into the future.

References

Allison, Tony, 2001, "Risks and Rewards in China's Insurance Market." *Asia Times – Special Reports, February 16.*

Anonymous, 2004, "The Development of China's Auto Industry." *Presentation at the China Association of Automobile Manufacturers (CAAM).*

Bates, Laurence W. and Andre T. Goldstein, 1989, "Sunday Drivers – All Week Long." *The China Business Review*, December: 36-39.

Chen, Renbao, Kie Ann Wong, and Hong Chew Lee, 1999, "Underwriting Cycles in Asia." *Journal of Risk & Insurance*, 66: 29-47.

Hall, Eric, 2005, "Foreign Insurers in China: Opportunity and Risk." KPMG-Reuters joint briefing paper.

"Insurance Offer Triggers Debate." http://www.starinfo.net.cn/english/news/enews/65.htm.

Lester, Edward P. and Wayne H. Fisher, 1975, "Loss Reserve Testing in a Changing Environment." Proceedings of the Casualty Actuarial Society, Vol. LXII: 154-171.

Min, Zhao, 2005, "Five Competitive Forces in China's Automobile Industry." *Journal of American Academy of Business, Cambridge*, 7: 99-105.

National Highway Traffic Safety Administration, "Traffic Safety Facts 2004." National Center for Statistics and Analysis (NHTSA).

Ng, Wei-shiuen and Lee Schipper, 2005, "China Motorization Trends: Policy Options in a World of Transport Challenges." Chapter in Growing in the Greenhouse: Protecting the Climate by Putting Development First, World Resources Institute, December: 48-67.

Outreville, J. Francois, 1990, "The Economic Significance of Insurance Markets in Developing Countries." Journal of Risk & Insurance v57: p.487-498.

Pan, Jiahua, Huaiguo Hu, Fei Yu, and Limin Cheng, "Automobiles." An Environmental Impact Assessment of China's WTO Accession, 2004, A Report by the Task Force on WTO and Environment, China Council for International Cooperation on Environment and Development

"Research & Markets Issues Study of China's Insurance Industry." Insurance Journal, September 1, 2006.

Satyaprakash, PGPIB, 2005, "China's Automobile Industry Post-WTO." White Paper, K.J. Somaiya Institute of Management Studies & Research.

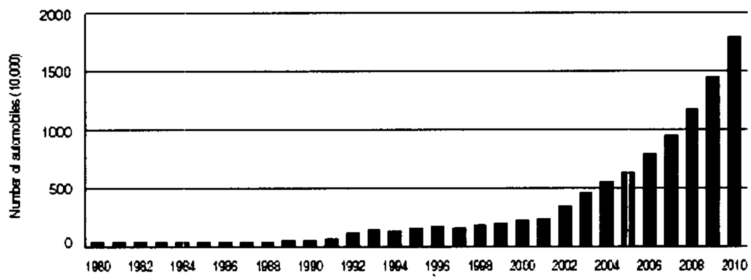
Schmitt, Karen E., 2000, "European Auto Insurance Pricing Considerations." Casualty Actuarial Society Forum, Winter: 141-158.

Shen, Yiming, 2000, "China's Insurance Market: Opportunity, Competition and Market Trends." *The Geneva Papers on Risk and Insurance*, 25: 335-355.

World Factbook, 2006, Central Intelligence Agency.

Yang, Nian-Chih and Wilfred Lu, 2005, "China Insurance Industry." Chinese Actuarial Club Article, January.

Figure 1
China Automobile Production, 1980-2010



Source: Data for 1980-1999, Yearbook of the China Automobile Industry 2000; for 2000-2003, actual production figures from various sources; for 2004-2010, estimates from various Web page sources.

Figure 2

New Automobile Price Trends in China

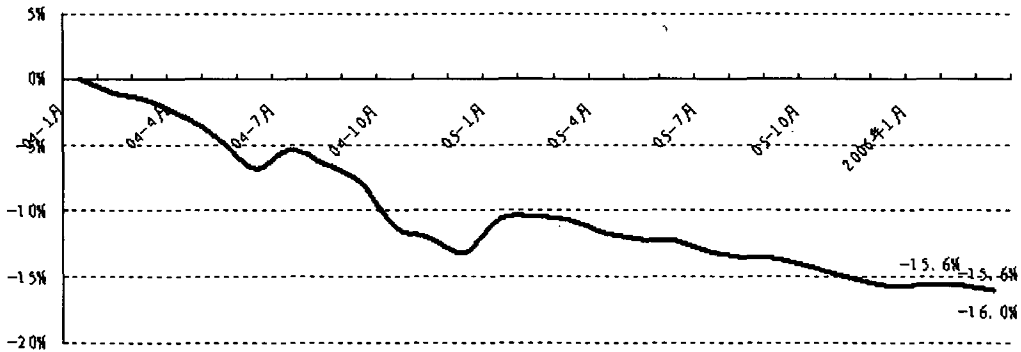


Table 1

Changes in Automobile Industry Post-WTO

	Before Entry into WTO	After entry into WTO
Tariffs	200% in 1980s:80-100% in 1990s	25% by 2006
Import quotas	30.000 vehicles a year allowed from foreign car makers	Quota increases 20%a year, phased out by 2006
Local content	40% in final year of production increasing to 60%.80% in second and third years, respectively	No local-content ratio required
Foreign participation in sales, distribution	Limited to wholesaling through JVs; prohibited from consolidating sales organizations of imports, JVs	Will be allowed to own vehicle wholesale, retail organisations;integrated sales organisations permitted by 2006
Auto financing for Chinese customers	Foreign, non-banking financial institutions prohibited from providing financing	Foreign, non-banking financing permitted in selected cities prior to gradual national roll out.

Source: Satyaprakash, PGPIB, 2005, "China's Automobile Industry Post-WTO." White Paper, K.J. Somaiya Institute of Management Studies & Research.

Table 2
Comparison of Auto Prices in China and World Markets (2003)

	Type	World market price (10,000 RMB)	China (10,000 RMB)	China/ world
Higher price spectrum	Audi A6	17.8	37.9–42.7	> 2.0
Middle spectrum	Honda/Accord	19.6	25.9	1.32
	Mazda	15.3	26.3	1.72
Lower spectrum	VW/Jetta	7.5–8.4	10.00	
	Aoto/Xiali	3.0	4.0	
Imported	Toyota/Camry	18.5	> 41.6	2.25
	Benz S600	100	> 200	> 2.0

Survey made by Automobile Digest, September 23, 2003.

Table 3
Market Concentration Indices for China Auto Insurance

	No. of domestic Insurers	CR1	CR2	CR3	CR4	HHI
2002	10	70.67	84.21	95.3	96.41	5304
2003	10	70.75	82.51	91.64	94.68	5242
2004	11	68.69	76.94	85.42	89.44	4904
2005	17	53.34	67.84	76.64	82.70	3218

Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Models

Noriszura Ismail and Abdul Aziz Jemain

Abstract

In actuarial literature, researchers suggested various statistical procedures to estimate the parameters in claim count or frequency model. In particular, the Poisson regression model, which is also known as the Generalized Linear Model (GLM) with Poisson error structure, has been widely used in the recent years. However, it is also recognized that the count or frequency data in insurance practice often display overdispersion, i.e., a situation where the variance of the response variable exceeds the mean. Inappropriate imposition of the Poisson may underestimate the standard errors and overstate the significance of the regression parameters, and consequently, giving misleading inference about the regression parameters. This paper suggests the Negative Binomial and Generalized Poisson regression models as alternatives for handling overdispersion. If the Negative Binomial and Generalized Poisson regression models are fitted by the maximum likelihood method, the models are considered to be convenient and practical; they handle overdispersion, they allow the likelihood ratio and other standard maximum likelihood tests to be implemented, they have good properties, and they permit the fitting procedure to be carried out by using the Iterative Weighted Least Squares (IWLS) regression similar to those of the Poisson. In this paper, two types of regression model will be discussed and applied; multiplicative and additive. The multiplicative and additive regression models for Poisson, Negative Binomial and Generalized Poisson will be fitted, tested and compared on three different sets of claim frequency data; Malaysian private motor third party property damage data, ship damage incident data from McCullagh and Nelder, and data from Bailey and Simon on Canadian private automobile liability.

Keywords: Overdispersion; Negative Binomial; Generalized Poisson; Multiplicative; Additive; Maximum likelihood.

1. INTRODUCTION

In property and liability insurance, the determination of premium rates must fulfill four basic principles generally agreed among the actuaries; to calculate "fair" premium rates whereby high risk insureds should pay higher premium and vice versa, to provide sufficient funds for paying expected losses and expenses, to maintain adequate margin for adverse deviation, and to produce a reasonable return to the insurer. The process of establishing "fair" premium rates for insuring uncertain events requires estimates which were made of two important elements; the probabilities associated with the occurrence of such event, i.e., the frequency, and the magnitude of such event, i.e., the severity. The frequency and severity estimates were usually calculated through the use of past experience for groups of similar risk characteristics. The process of grouping risks with similar risk characteristics to establish "fair" premium rates in an insurance system is also known as risk classification. In this paper, risk classification will be applied to estimate claim frequency rate which is equivalent to the claim count per exposure unit.

In the last forty years, researchers suggested various statistical procedures to estimate the parameters in risk classification model. For example, Bailey and Simon [1] suggested the minimum chi-squares, Bailey [2] devised the zero bias, Jung [3] produced a heuristic method for minimum modified chi-squares, Ajne [4] proposed the method of moments also for minimum modified chi-squares, Chamberlain [5] used the weighted least squares, Coutts [6] produced the method of orthogonal weighted least squares with logit transformation, Harrington [7] suggested the maximum likelihood procedure for models with functional form, and Brown [8] proposed the bias and likelihood functions for minimum bias and maximum likelihood models.

In the recent actuarial literature, research on the estimation methods for risk classification model is still continuing and developing. For example, Mildenhall [9] merged the models which were introduced by Bailey and Simon, i.e., the minimum bias models, with the Generalized Linear Models (GLMs), i.e., the maximum likelihood models. Besides providing strong statistical justifications for the minimum bias models which were originally based on a non-parametric approach, his effort also allowed a variety of parametric models to be chosen from. Later, Feldblum and Brosius [10] summarized the minimum bias procedure and provided intuition for several bias functions, which include zero bias, least squares, minimum chi-squares and maximum likelihood, for practicing actuary. Anderson *et al.* [11] provided foundation for GLMs statistical theory also for practicing actuary. Their study provided practical insights and realistic output for the analysis of GLMs. Fu and Wu [12] developed the models of Bailey and Simon by following the same approach which was created by Bailey and Simon, i.e., the non-parametric approach. As a result, their research offers a wide range of non-parametric models to be created and applied. Ismail and Jemain [13] found a match point that merged the available parametric and non-parametric models, i.e., minimum bias and maximum likelihood models, by rewriting the models in a more generalized form. They solved the parameters by using weighted equation, regression approach and Taylor series approximation.

Besides statistical procedures, research on multiplicative and additive models has also been carried out. Among the pioneer studies, Bailey and Simon [1] compared the systematic bias of multiplicative and additive models and found that the multiplicative model overestimates the high risk classes. Their result was later agreed by Jung [3] and Ajne [4] who also found that the estimates for multiplicative model are positively biased. Bailey [2] compared the multiplicative and additive models by producing two statistical criteria, namely, the minimum chi-squares and average absolute difference. In addition, he also suggested the multiplicative model for percents classes and additive model for cents classes. Freifelder [14] predicted the pattern of over and under estimation for multiplicative and additive models if true models were misspecified, Jee [15] compared the predictive accuracy of multiplicative and additive models, Brown [8] discussed and summarized the additive and multiplicative models which were derived from the maximum likelihood and minimum bias approaches, Holler *et al.* [16] compared the initial values sensitivity of multiplicative and additive models, Mildenhall [9] identified the Generalized Linear Models for identity and log link functions with the additive and multiplicative models which were discussed in Brown [8], and Ismail and Jemain [13] discussed and compared the parameter estimates and goodness-of-fit of the additive and multiplicative regression models.

In insurance practice, the Poisson regression model, which is also known as the Generalized Linear Model with Poisson error structure, has been widely used for modeling claim count or frequency data in the recent years. For example, Aitkin *et al.* [17] and Renshaw [18] each respectively fit the Poisson model to two different sets of U.K. motor claim count data. For insurance practitioners, the Poisson regression model has been considered as practical and convenient; besides allowing the statistical inference and hypothesis tests to be determined by statistical theories, the model also permits the fitting procedure to be carried out easily by using any statistical package containing a routine for the Iterative Weighted Least Squares (IWLS) regression.

However, at the same time it is also recognized that the count or frequency data in insurance practice often display overdispersion or extra-Poisson variation, a situation where the variance of the response variable exceeds the mean. Inappropriate imposition of the Poisson may underestimate the standard errors and overstate the significance of the regression parameters, and consequently, giving misleading inference about the regression parameters.

Based on the actuarial literature, the Poisson quasi likelihood model has been suggested to accommodate overdispersion in claim count or frequency data. For example, McCullagh and Nelder [19], using the data provided by Lloyd's Register of Shipping, applied the quasi likelihood model for damage incidents caused to the forward section of cargo-carrying vessels, to allow for possible inter-ship variability in accident proneness. The same quasi likelihood model was also fitted to the count data of U.K. own damage motor claims by Brockman and Wright [20], to take into account the possibility of within-cell heterogeneity.

For insurance practitioners, the most likely reason for using Poisson quasi likelihood is that the model can still be fitted without knowing the exact probability function of the response variable, as long as the mean is specified to be equivalent to the mean of Poisson, and the variance can be written as a multiplicative constant of the mean. To account for overdispersion, the Poisson quasi likelihood produces parameter estimates equivalent to the Poisson, and standard errors larger than those of the Poisson.

On the contrary, the maximum likelihood approach suggested in this paper differs from the quasi likelihood approach such that it requires the complete probability of the response variable, thus, allowing the likelihood ratio and other standard maximum likelihood tests to be implemented. With this objective in mind, this paper suggests the Negative Binomial and Generalized Poisson regression models for handling overdispersion. If the Negative Binomial and Generalized Poisson were fitted by the maximum likelihood method, the models may also be considered as convenient and practical; they allow the likelihood ratio and other standard maximum likelihood tests to be implemented, they have good properties, they permit the fitting procedure to be carried out by using Iterative Weighted Least Squares (IWLS) regression similar to those of the Poisson, and last but not least, they handle overdispersion. In this paper, two types of regression models will be discussed and applied; multiplicative and additive models. Specifically, the multiplicative and additive regression models for Poisson, Negative Binomial and Generalized Poisson will be fitted, tested and compared on three different sets of claim frequency data; Malaysian private motor third party property damage data, ship damage incident data from McCullagh and Nelder [19], and data from Bailey and Simon [1] on Canadian private automobile liability.

2. MULTIPLICATIVE REGRESSION MODELS

2.1 Poisson

Let Y_i be the random variable for claim count in the *i*th class, $i = 1,2,\ldots,n$, where n denotes the number of rating classes. If Y_i follows a Poisson distribution, the probability density function is,

$$\Pr(Y_i = y_i) = \frac{\exp(-\lambda_i)\lambda_i^{y_i}}{y_i!}, \quad y_i = 0,1,\ldots \tag{2.1}$$

with mean and variance, $E(Y_i) = Var(Y_i) = \lambda_i$.

To incorporate covariates and to ensure non-negativity, the mean or the fitted value is assumed to be multiplicative, i.e., $E(Y_i \mid \mathbf{x}_i) = \lambda_i = e_i \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, where e_i denotes a measure of exposure, $\mathbf{x}_i$ a $p \times 1$ vector of explanatory variables, and $\boldsymbol{\beta}$ a $p \times 1$ vector of regression parameters.

If $\boldsymbol{\beta}$ is estimated by the maximum likelihood method, the likelihood equations are,

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_j} = \sum_i (y_i - \lambda_i) x_{ij} = 0, \quad j = 1,2,\ldots,p. \tag{2.2}$$

Since Eq.(2.2) is also equal to the weighted least squares, the maximum likelihood estimates, $\hat{\boldsymbol{\beta}}$, may be solved by using the Iterative Weighted Least Squares (IWLS) regression.

2.2 Negative Binomial I

Under the Poisson, the mean, λ_i, is assumed to be constant or homogeneous within the classes. However, by defining a specific distribution for λ_i, heterogeneity within the classes is now allowed. For example, by assuming λ_i to be a Gamma with mean $E(\lambda_i) = \mu_i$ and variance $Var(\lambda_i) = \mu_i^2 v_i^{-1}$, and $Y_i \mid \lambda_i$ to be a Poisson with conditional mean $E(Y_i \mid \lambda_i) = \lambda_i$, it can be shown that the marginal distribution of Y_i follows a Negative Binomial distribution with probability density function,

$$\Pr(Y_i = y_i) = \int \Pr(Y_i = y_i \mid \lambda_i) f(\lambda_i) d\lambda_i = \frac{\Gamma(y_i + v_i)}{\Gamma(y_i + 1)\Gamma(v_i)} \left(\frac{v_i}{v_i + \mu_i} \right)^{v_i} \left(\frac{\mu_i}{v_i + \mu_i} \right)^{y_i}, \tag{2.3}$$

where the mean is $E(Y_i) = \mu_i$ and the variance is $Var(Y_i) = \mu_i + \mu_i^2 v_i^{-1}$.

Different parameterization can generate different types of Negative Binomial distributions. For example, by letting $v_i = a^{-1}$, Y_i follows a Negative Binomial distribution with mean $E(Y_i) = \mu_i$ and variance $Var(Y_i) = \mu_i(1 + a\mu_i)$, where a denotes the dispersion parameter (see Lawless [21]; Cameron and Trivedi [22]).

If a equals zero, the mean and variance will be equal, $E(Y_i) = Var(Y_i)$, resulting the distribution to be a Poisson. If $a > 0$, the variance will exceed the mean, $Var(Y_i) > E(Y_i)$, and the distribution allows for overdispersion as well. In this paper, the distribution will be called as Negative Binomial I.

If it is assumed that the mean or the fitted value is multiplicative, i.e., $E(Y_i \mid \mathbf{x}_i) = \mu_i = e_i \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, the likelihood for Negative Binomial I regression model may be written as,

$$\ell(\boldsymbol{\beta}, a) = \sum_i \left\{ \sum_{r=1}^{y_i - 1} \log(1 + ar) \right\} - y_i \log(a) - \log(y_i!) + \\ y_i \log(a\mu_i) - (y_i + a^{-1}) \log(1 + a\mu_i). \tag{2.4}$$

Therefore, the maximum likelihood estimates, $(\hat{\boldsymbol{\beta}}, \hat{a})$, may be obtained by maximizing $\ell(\boldsymbol{\beta}, a)$ with respect to $\boldsymbol{\beta}$ and a. The related equations are,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial \beta_j} = \sum_i \frac{(y_i - \mu_i) x_{ij}}{1 + a\mu_i} = 0, \quad j = 1, 2, \ldots, p, \tag{2.5}$$

and,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial a} = \sum_i \left\{ \sum_{r=1}^{y_i - 1} \left(\frac{r}{1 + ar} \right) \right\} + a^{-2} \log(1 + a\mu_i) - \frac{(y_i + a^{-1})\mu_i}{(1 + a\mu_i)} = 0. \tag{2.6}$$

The maximum likelihood estimates, $(\hat{\boldsymbol{\beta}}, \hat{a})$, may be solved simultaneously, and the procedure involves sequential iterations. In the first sequence, by using an initial value of a, $a_{(0)}$, $\ell(\boldsymbol{\beta}, a)$ is maximized with respect to $\boldsymbol{\beta}$, producing $\boldsymbol{\beta}_{(1)}$. The related equation is Eq.(2.5) which is also equivalent to the weighted least squares. Therefore, with a slight modification, this task can be performed by using the IWLS regression similar to those of the Poisson. In the second sequence, by holding $\boldsymbol{\beta}$ fixed at $\boldsymbol{\beta}_{(1)}$, $\ell(\boldsymbol{\beta}, a)$ is maximized with respect to a, producing $a_{(1)}$. The related equation is Eq.(2.6), and the task can be carried out by using the Newton-Raphson iteration. By iterating and cycling between holding a fixed and holding $\boldsymbol{\beta}$ fixed, the maximum likelihood estimates, $(\hat{\boldsymbol{\beta}}, \hat{a})$, will be obtained. Further explanation on the fitting procedure will be discussed in Section 4.

An easier approach to estimate a is by using the moment estimation suggested by Breslow [23], i.e., by equating the Pearson chi-squares statistic with the degrees of freedom,

$$\sum_i \frac{(y_i - \mu_i)^2}{\mu_i (1 + a\mu_i)} = n - p, \tag{2.7}$$

where n denotes the number of rating classes and p the number of regression parameters. The sequential iteration procedure similar to the one mentioned above can also be used, this time producing maximum likelihood estimates of $\boldsymbol{\beta}$ and moment estimate of a, $(\hat{\boldsymbol{\beta}}, \tilde{a})$.

In this paper, when a is estimated by the maximum likelihood, the model will be called as Negative Binomial I (MLE). Likewise, when a is estimated by the method of moment, the model will be called as Negative Binomial I (moment).

2.3 Negative Binomial II

By letting $v_i = \mu_i a^{-1}$, another type of Negative Binomial distribution is produced, this time with mean $E(Y_i) = \mu_i$ and variance $Var(Y_i) = \mu_i(1+a)$ (see Nelder and Lee [24]; Cameron and Trivedi [22]). If a equals zero, the mean and variance will be equal, resulting the distribution to be a Poisson. If $a > 0$, the variance will exceed the mean and the distribution allows for overdispersion as well. In this paper, the distribution will be called as Negative Binomial II.

If it is assumed that the mean or the fitted value is multiplicative, i.e., $E(Y_i \mid \mathbf{x}_i) = \mu_i = e_i \exp(\mathbf{x}_i^{\mathrm{T}} \boldsymbol{\beta})$, the likelihood for Negative Binomial II regression model may be written as,

$$\begin{aligned}\ell(\boldsymbol{\beta}, a) = \sum_i \log(\Gamma(y_i + \mu_i a^{-1})) - \log(\Gamma(\mu_i a^{-1})) - \log(y_i!) - \\ \mu_i a^{-1} \log(a) - (y_i + \mu_i a^{-1}) \log(1 + a^{-1}).\end{aligned} \tag{2.8}$$

Therefore, the maximum likelihood estimates, $(\hat{\boldsymbol{\beta}}, \hat{a})$, may be obtained by maximizing $\ell(\boldsymbol{\beta}, a)$ with respect to $\boldsymbol{\beta}$ and a. The related equations are,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial \beta_j} = \sum_i \mu_i x_{ij} a^{-1} \left\{ \sum_{r=0}^{y_i - 1} (\mu_i a^{-1} + r)^{-1} - \log(1 + a) \right\} = 0, \quad j = 1, 2, \ldots, p, \tag{2.9}$$

and,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial a} = -\sum_i \mu_i a^{-2} \left\{ \sum_{r=0}^{y_i - 1} (\mu_i a^{-1} + r)^{-1} - \log(1 + a) \right\} + \sum_i \frac{y_i - \mu_i}{(1 + a) a} = 0. \tag{2.10}$$

However, the maximum likelihood estimates, $\hat{\boldsymbol{\beta}}$, are numerically difficult to be solved because the related equation, Eq.(2.9), is not equal to the weighted least squares. As an alternative, since the Negative Binomial II has a constant variance-mean ratio, the method of weighted least squares is suggested, i.e., by equating,

$$\sum_i \frac{y_i - \mu_i}{Var(Y_i)} \frac{\partial \mu_i}{\partial \beta_j} = \sum_i \frac{(y_i - \mu_i) x_{ij}}{1+a} = 0, \quad j = 1,2,\dots, p, \tag{2.11}$$

to produce the least squares estimates, $\tilde{\boldsymbol{\beta}}$.

It is shown that in the presence of a modest amount of overdispersion, the least squares estimates were highly efficient for the estimation of a moment parameter of an exponential family distribution (Cox [25]). Since Eq.(2.11) is also equivalent to the likelihood equation of the Poisson, i.e., Eq.(2.2), the same IWLS regression which is used for the Poisson can be applied to estimate the least squares estimates, $\tilde{\boldsymbol{\beta}}$. As a result, the least squares estimates are also equal to the maximum likelihood estimates of Poisson, but the standard errors are equal or larger than the Poisson because they are multiplied by $\sqrt{1+a}$ where $a \geq 0$.

For simplicity, a is suggested to be estimated by the method of moment, i.e., by equating the Pearson chi-squares statistic with the degrees of freedom,

$$\sum_i \frac{(y_i - \mu_i)^2}{(1+a)\mu_i} = n - p, \tag{2.12}$$

which involves a straightforward calculation and produces a moment estimate, $\tilde{a}$.

In this paper, the estimates which were produced by the multiplicative regression models of Negative Binomial I (MLE), Negative Binomial I (moment) and Negative Binomial II will be denoted respectively by $(\hat{\boldsymbol{\beta}}, \hat{a})$, $(\hat{\boldsymbol{\beta}}, \tilde{a})$ and $(\tilde{\boldsymbol{\beta}}, \tilde{a})$.

2.4 Generalized Poisson I

The advantage of using the Generalized Poisson distribution is that it can be fitted for both overdispersion, $Var(Y_i) > E(Y_i)$, as well as underdispersion, $Var(Y_i) < E(Y_i)$. In this paper, two different types of Generalized Poisson will be discussed; each will be referred to as Generalized Poisson I and Generalized Poisson II. For Generalized Poisson I distribution, the probability density function is (Wang and Famoye [26]),

$$\Pr(Y_i = y_i) = \left(\frac{\mu_i}{1+a\mu_i}\right)^{y_i} \frac{(1+ay_i)^{y_i - 1}}{y_i!} \exp\left(-\frac{\mu_i(1+ay_i)}{1+a\mu_i}\right), \quad y_i = 0,1,\dots, \tag{2.13}$$

with mean $E(Y_i) = \mu_i$ and variance $Var(Y_i) = \mu_i(1 + a\mu_i)^2$:

The Generalized Poisson I is a natural extension of the Poisson. If a equals zero, the Generalized Poisson I reduces to the Poisson, resulting $E(Y_i) = Var(Y_i)$. If $a > 0$, the variance is larger than the mean, $Var(Y_i) > E(Y_i)$, and the distribution represents count data with overdispersion. If $a < 0$, the variance is smaller than the mean, $Var(Y_i) < E(Y_i)$, so that now the distribution represents count data with underdispersion.

If it is assumed that the mean or the fitted value is multiplicative, i.e., $E(Y_i \mid \mathbf{x}_i) = \mu_i = e_i \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, the likelihood for Generalized Poisson I regression model may be written as,

$$\ell(\boldsymbol{\beta}, a) = \sum_i y_i \log\left(\frac{\mu_i}{1 + a\mu_i}\right) + (y_i - 1)\log(1 + ay_i) - \frac{\mu_i(1 + ay_i)}{1 + a\mu_i} - \log(y_i!). \quad (2.14)$$

Therefore, the maximum likelihood estimates, $(\hat{\boldsymbol{\beta}}, \hat{a})$, may be obtained by maximizing $\ell(\boldsymbol{\beta}, a)$ with respect to $\boldsymbol{\beta}$ and a. The related equations are,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial \beta_j} = \sum_i \frac{(y_i - \mu_i)x_{ij}}{(1 + a\mu_i)^2} = 0, \quad j = 1,2,\dots, p, \quad (2.15)$$

and,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial a} = \sum_i -\frac{y_i\mu_i}{1 + a\mu_i} + \frac{y_i(y_i - 1)}{1 + ay_i} - \frac{\mu_i(y_i - \mu_i)}{(1 + a\mu_i)^2} = 0. \quad (2.16)$$

The sequential iteration procedure similar to the Negative Binomial I regression model may also be implemented to obtain the maximum likelihood estimates, $(\hat{\boldsymbol{\beta}}, \hat{a})$. For the sequential iteration, the IWLS regression can be applied because Eq.(2.15) is also equal to the weighted least squares.

An easier approach to estimate a is by using the moment estimation, i.e., by equating the Pearson chi-squares statistic with the degrees of freedom,

$$\sum_i \frac{(y_i - \mu_i)^2}{\mu_i(1 + a\mu_i)^2} = n - p, \quad (2.17)$$

producing $(\hat{\boldsymbol{\beta}}, \tilde{a})$.

In this paper, when a is estimated by the maximum likelihood, the model will be called as Generalized Poisson I (MLE). Likewise, when a is estimated by the method of moment, the model will be called as Generalized Poisson I (moment).

2.5 Generalized Poisson II

For Generalized Poisson II, the probability density function may be written in the form of (Consul and Famoye [27]),

$$\Pr(Y_i = y_i) = \begin{cases} \mu_i(\mu_i + (a-1)y_i)^{y_i-1} a^{-y_i} \dfrac{\exp(-a^{-1}(\mu_i + (a-1)y_i))}{y_i!}, & y_i = 0,1,\ldots \\ 0, & y_i > m, a < 1 \end{cases}, \quad (2.18)$$

where $a \geq \max(\frac{1}{2}, 1 - \frac{\mu_i}{4})$, and m the largest positive integer for which $\mu_i + m(a-1) > 0$ when $a < 1$. For this distribution, the mean is equal to $E(Y_i) = \mu_i$, whereas the variance is equivalent to $Var(Y_i) = a^2 \mu_i$.

The Generalized Poisson II is also a natural extension of the Poisson. If a equals one, the Generalized Poisson II reduces to the Poisson. If $a > 1$, the variance is larger than the mean and the distribution represents count data with overdispersion. If $\frac{1}{2} \leq a < 1$ and $\mu_i > 2$, the variance is smaller than the mean so that now the distribution represents count data with underdispersion.

If it is assumed that the mean or the fitted value is multiplicative, i.e., $E(Y_i \mid \mathbf{x}_i) = \mu_i = e_i \exp(\mathbf{x}_i^T \boldsymbol{\beta})$, the likelihood for Generalized Poisson II regression model may be written as,

$$\begin{aligned} \ell(\boldsymbol{\beta}, a) = \sum_i \log(\mu_i) + (y_i - 1)\log(\mu_i + (a-1)y_i) - \\ y_i \log(a) - a^{-1}(\mu_i + (a-1)y_i) - \log(y_i!). \end{aligned} \quad (2.19)$$

Therefore, the maximum likelihood estimates, $(\hat{\boldsymbol{\beta}}, \hat{a})$, may be obtained by maximizing $\ell(\boldsymbol{\beta}, a)$ with respect to $\boldsymbol{\beta}$ and a. The related equations are,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial \beta_j} = \sum_i \left\{ \mu_i^{-1} - a^{-1} + \frac{y_i - 1}{\mu_i + (a-1)y_i} \right\} \mu_i x_{ij}, \quad j = 1,2,\ldots,p, \quad (2.20)$$

and,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial a} = \sum_i \frac{y_i(y_i - 1)}{\mu_i + (a-1)y_i} - y_i a^{-1} + (\mu_i - y_i)a^{-2} = 0. \tag{2.21}$$

However, the maximum likelihood estimates, $\hat{\boldsymbol{\beta}}$, are numerically difficult to be solved because the related equation, Eq.(2.20), is not equal to the weighted least squares. Since the Generalized Poisson II has a constant variance-mean ratio, the method of weighted least squares is suggested as an alternative, i.e., by equating,

$$\sum_i \frac{(y_i - \mu_i)x_{ij}}{a^2} = 0, \quad j = 1,2,\ldots,p, \tag{2.22}$$

to produce the least squares estimates, $\tilde{\boldsymbol{\beta}}$. The same Poisson IWLS regression may be used to estimate $\tilde{\boldsymbol{\beta}}$ because Eq.(2.22) is also equivalent to the Poisson likelihood equation, i.e., Eq.(2.2). As a result, the least squares estimates are also equal to the maximum likelihood estimates of Poisson. However, the standard errors could be equal, larger or smaller than the Poisson because they are multiplied by a where $a \geq 1$ or $\frac{1}{2} \leq a < 1$.

For simplicity, a is suggested to be estimated by the method of moment, i.e., by equating the Pearson chi-squares statistic with the degrees of freedom,

$$\sum_i \frac{(y_i - \mu_i)^2}{a^2 \mu_i} = n - p, \tag{2.23}$$

involving a straightforward calculation and producing a moment estimate, $\tilde{a}$.

In this paper, the estimates which were produced by the regression models of Generalized Poisson I (MLE), Generalized Poisson I (moment) and Generalized Poisson II will be denoted respectively by $(\hat{\boldsymbol{\beta}}, \hat{a})$, $(\hat{\boldsymbol{\beta}}, \tilde{a})$ and $(\tilde{\boldsymbol{\beta}}, \tilde{a})$.

To summarize the multiplicative regression models which were discussed in this section, Table 1 shows the methods and equations for solving the estimates of $\boldsymbol{\beta}$ and a.

Table 1. Methods and equations for solving $\boldsymbol{\beta}$ and a in multiplicative regression models

Models	Estimation of β		Estimation of a	
	Method	Equation	Method	Equation
Poisson	Maximum Likelihood	$\sum_i (y_i - \mu_i)x_{ij} = 0$	-	-
NBI(MLE)	Maximum Likelihood	$\sum_i \frac{(y_i - \mu_i)x_{ij}}{1 + a\mu_i} = 0$	Maximum Likelihood	$\sum_i \left\{ \sum_{r=1}^{y_i - 1} \left(\frac{r}{1 + ar} \right) + a^{-2} \log(1 + a\mu_i) - \frac{(y_i + a^{-1})\mu_i}{(1 + a\mu_i)} \right\} = 0$
NBI(moment)	Maximum Likelihood	$\sum_i \frac{(y_i - \mu_i)x_{ij}}{1 + a\mu_i} = 0$	Moment	$\sum_i \left\{ \frac{(y_i - \mu_i)^2}{\mu_i(1 + a\mu_i)} \right\} - (n - p) = 0$
NBII	Weighted Least Squares	$\sum_i \frac{(y_i - \mu_i)x_{ij}}{1 + a} = 0$	Moment	$\sum_i \left\{ \frac{(y_i - \mu_i)^2}{\mu_i(1 + a)} \right\} - (n - p) = 0$
GPI(MLE)	Maximum Likelihood	$\sum_i \frac{(y_i - \mu_i)x_{ij}}{(1 + a\mu_i)^2} = 0$	Maximum Likelihood	$\sum_i \left\{ -\frac{y_i\mu_i}{1 + a\mu_i} + \frac{y_i(y_i - 1)}{1 + ay_i} - \frac{\mu_i(y_i - \mu_i)}{(1 + a\mu_i)^2} \right\} = 0$
GPI(moment)	Maximum Likelihood	$\sum_i \frac{(y_i - \mu_i)x_{ij}}{(1 + a\mu_i)^2} = 0$	Moment	$\sum_i \left\{ \frac{(y_i - \mu_i)^2}{\mu_i(1 + a\mu_i)^2} \right\} - (n - p) = 0$
GPII	Weighted Least Squares	$\sum_i \frac{(y_i - \mu_i)x_{ij}}{a^2} = 0$	Moment	$\sum_i \left\{ \frac{(y_i - \mu_i)^2}{\mu_i a^2} \right\} - (n - p) = 0$

3. GOODNESS-OF-FIT TESTS

In this section, several goodness-of-fit measures will be briefly discussed, including the Pearson chi-squares, deviance, likelihood ratio test, Akaike Information Criteria (AIC) and Bayesian Schwartz Criteria (BSC). Since these measures are already familiar to those who used the Generalized Linear Model with Poisson error structure for claim count or frequency analysis, the same measures may also be implemented to the regression models of Negative Binomial and Generalized Poisson as well.

3.1 Pearson chi-squares

Two of the most frequently used measures for goodness-of-fit in the Generalized Linear Models are the Pearson chi-squares and the deviance. The Pearson chi-squares statistic is equivalent to,

$$\sum_i \frac{(y_i - \mu_i)^2}{Var(Y_i)}. \tag{3.1}$$

For an adequate model, the statistic has an asymptotic chi-squares distribution with $n-p$ degrees of freedom, where n denotes the number of rating classes and p the number of parameters.

3.2 Deviance

The deviance is equal to,

$$D = 2(\ell(\mathbf{y};\mathbf{y}) - \ell(\boldsymbol{\mu};\mathbf{y})), \tag{3.2}$$

where $\ell(\boldsymbol{\mu};\mathbf{y})$ and $\ell(\mathbf{y};\mathbf{y})$ are the model's log likelihood evaluated respectively under $\boldsymbol{\mu}$ and $\mathbf{y}$. For an adequate model, D also has an asymptotic chi-squares distribution with $n-p$ degrees of freedom. Therefore, if the values for both Pearson chi-squares and D are close to the degrees of freedom, the model may be considered as adequate.

The deviance could also be used to compare between two nested models, one of which is a simplified version of the other. Let D_1 and df_1 be the deviance and degrees of freedom for such model, and D_2 and df_2 be the same values by fitting a simplified version of the model. The chi-squares statistic is equal to $(D_2 - D_1)/(df_2 - df_1)$ and it should be compared to a chi-squares distribution with $df_2 - df_1$ degrees of freedom.

3.3 Likelihood ratio

The advantage of using the maximum likelihood method is that the likelihood ratio test may be employed to assess the adequacy of the Negative Binomial I (MLE) or the Generalized Poisson I (MLE) over the Poisson because both Negative Binomial I (MLE) and Generalized Poisson I (MLE) will reduce to the Poisson when the dispersion parameter, a, equals zero.

For testing Poisson against Negative Binomial I (MLE), the hypothesis may be stated as $H_0 : a = 0$ against $H_1 : a > 0$. The likelihood ratio statistic is,

$$T = 2(\ell_1 - \ell_0), \tag{3.3}$$

where ℓ_1 and ℓ_0 are the model's log likelihood under the respective hypothesis. T has an asymptotic distribution of probability mass of one-half at zero and one-half-chi-squares distribution with one degrees of freedom (see Lawless [21]; Cameron and Trivedi [22]). Therefore, to test the null hypothesis at the significance level of α, the critical value of chi-squares distribution with significance level 2α is used, i.e., reject H_0 if $T > \chi^2_{(1-2\alpha,1)}$.

For testing Poisson against Generalized Poisson I (MLE), the hypothesis may be stated as $H_0 : a = 0$ against $H_1 : a \neq 0$. The likelihood ratio is also equal to Eq.(3.3) and under null hypothesis, T has an asymptotic chi-squares distribution with one degrees of freedom (see Wang and Famoye [26]).

3.4 AIC and BIC

When several maximum likelihood models are available, one can compare the performance of alternative models based on several likelihood measures which have been proposed in the statistical literature. Two of the most regularly used measures are the Akaike Information Criteria (AIC) and the Bayesian Schwartz Information Criteria (BIC). The AIC is defined as (Akaike [28]),

$$\text{AIC} = -2\ell + 2p, \tag{3.4}$$

where ℓ denotes the log likelihood evaluated under $\boldsymbol{\mu}$ and p the number of parameters. For this measure, the smaller the AIC, the better the model is.

The BIC is defined as (Schwartz [29]),

$$\text{BIC} = -2\ell + p\log(n), \tag{3.5}$$

where ℓ denotes the log likelihood evaluated under $\boldsymbol{\mu}$, p the number of parameters and n the number of rating classes. For this measure, the smaller the BIC, the better the model is.

4. FITTING PROCEDURE

As mentioned previously, the estimates of $\boldsymbol{\beta}$ and a for Negative Binomial I (MLE), Negative Binomial I (moment), Generalized Poisson I (MLE) and Generalized Poisson I (moment), may be solved simultaneously and the fitting procedure involves sequential iterations. The sequential iterations involve two steps of maximization in each sequence; maximizing $\ell(\boldsymbol{\beta}, a)$ with respect to $\boldsymbol{\beta}$ by holding a fixed, and maximizing $\ell(\boldsymbol{\beta}, a)$ with respect to a by holding $\boldsymbol{\beta}$ fixed.

4.1 Maximizing $\ell(\boldsymbol{\beta}, a)$ with respect to $\boldsymbol{\beta}$

By using the Newton-Rahpson iteration and the method of Scoring, the iterative equation in the standard form of IWLS regression may be written as,

$$\boldsymbol{\beta}_{(r)} = \boldsymbol{\beta}_{(r-1)} + \mathbf{I}_{(r-1)}^{-1}\mathbf{z}_{(r-1)}, \tag{4.1}$$

where $\boldsymbol{\beta}_{(r)}$ and $\boldsymbol{\beta}_{(r-1)}$ denote the vectors for $\boldsymbol{\beta}$ in the *r*th and *r*-1th iteration, $\mathbf{I}_{(r-1)}$ the information matrix containing negative expectation of the second derivatives of log likelihood evaluated at $\boldsymbol{\beta}_{(r-1)}$, and $\mathbf{z}_{(r-1)}$ the vector containing first derivatives of log likelihood evaluated at $\boldsymbol{\beta}_{(r-1)}$.

For an easier demonstration, an example for Poisson's IWLS regression will be shown and the notation for Poisson mean, λ_i, will be replaced by μ_i. The first derivatives of Poisson log likelihood, which is shown by Eq.(2.2), can also be written as,

$$\mathbf{z} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{k}, \tag{4.2}$$

where $\mathbf{X}$ denotes the matrix of explanatory variables, $\mathbf{W}$ the diagonal weight matrix whose *i*th diagonal element is,

$$w_i^P = \mu_i, \tag{4.3}$$

and $\mathbf{k}$ the vector whose ith row is equal to,

$$k_i = \frac{y_i - \mu_i}{\mu_i}. \tag{4.4}$$

The negative expectation of the second derivatives of Poisson log likelihood may be derived and it is equivalent to,

$$-E\left(\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial\beta_j \partial\beta_s}\right) = \sum_i \mu_i x_{ij} x_{is}, \quad j,s = 1,2,\ldots,p. \tag{4.5}$$

Therefore, the information matrix, $\mathbf{I}$, which contains negative expectation of the second derivatives of log likelihood, may be written as,

$$\mathbf{I} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}, \tag{4.6}$$

where the ith diagonal element of the weight matrix is also equal to Eq.(4.3).

Finally, the iterative equation shown by Eq.(4.1) may be rewritten as,

$$\boldsymbol{\beta}_{(r)} = \boldsymbol{\beta}_{(r-1)} + (\mathbf{X}^{\mathrm{T}}\mathbf{W}_{(r-1)}\mathbf{X})^{-1}(\mathbf{X}^{\mathrm{T}}\mathbf{W}_{(r-1)}\mathbf{k}_{(r-1)}). \tag{4.7}$$

It can be shown that with a slight modification in the weight matrix, the same iterative equation, i.e., Eq.(4.7), can also be used to obtain the maximum likelihood estimates, $\hat{\boldsymbol{\beta}}$, for Negative Binomial I and Generalized Poisson I as well.

The related equations for the first derivatives of log likelihood for Negative Binomial I and Generalized Poisson I are shown by Eq.(2.5) and Eq.(2.15). Both equations may also be written as Eq.(4.2), where the ith row of vector $\mathbf{k}$ is also the same as Eq.(4.4). However, the ith diagonal element of the weight matrix is,

$$w_i^{NBI} = \frac{\mu_i}{1 + a\mu_i}, \tag{4.8}$$

for Negative Binomial I and,

$$w_i^{GPI} = \frac{\mu_i}{(1+a\mu_i)^2}, \tag{4.9}$$

for Generalized Poisson I.

The negative expectation of the second derivatives of log likelihood for Negative Binomial I and Generalized Poisson I may be derived, and the respective equations are,

$$-E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, a)}{\partial\beta_j \partial\beta_s}\right) = \sum_i \frac{\mu_i x_{ij} x_{is}}{1+a\mu_i}, \quad j,s = 1,2,\ldots,p, \tag{4.10}$$

and,

$$-E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, a)}{\partial\beta_j \partial\beta_s}\right) = \sum_i \frac{\mu_i x_{ij} x_{is}}{(1+a\mu_i)^2}, \quad j,s = 1,2,\ldots,p. \tag{4.11}$$

Therefore, the information matrix may also be written as Eq.(4.6), where the *i*th diagonal element of the weight matrix for Negative Binomial I and Generalized Poisson I are respectively equal to Eq.(4.8) and Eq.(4.9).

The same iterative equation for the Poisson may also be used for Negative Binomial II and Generalized Poisson II because the weighted least squares equations, i.e., Eq.(2.11) and Eq.(2.22), are equivalent to the likelihood equations of the Poisson, i.e., Eq.(2.2).

The matrices and vectors for solving $\boldsymbol{\beta}$ in multiplicative regression models are summarized in Table 2.

Table 2. Matrices and vectors for solving $\boldsymbol{\beta}$ in multiplicative regression models

Models		
	Matrices and vectors for $\boldsymbol{\beta}_{(r)} = \boldsymbol{\beta}_{(r-1)} + \mathbf{I}_{(r-1)}^{-1}\mathbf{z}_{(r-1)}$, where $\mathbf{I}_{(r-1)} = \mathbf{X}^{\mathrm{T}}\mathbf{W}_{(r-1)}\mathbf{X}$, $\mathbf{z}_{(r-1)} = \mathbf{X}^{\mathrm{T}}\mathbf{W}_{(r-1)}\mathbf{k}_{(r-1)}$, js-th element of matrix $\mathbf{I} = i_{js} = -E\left(\frac{\partial^2 \ell}{\partial\beta_j \partial\beta_s}\right)$, j-th row of vector $\mathbf{z} = z_j = \frac{\partial \ell}{\partial \beta_j}$.	
Poisson/ NBII/ GPII	matrix $\mathbf{I}$	$i_{js} = \sum_i \mu_i x_{ij} x_{is} \rightarrow \mathbf{I} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}$
	weight matrix $\mathbf{W}$	$w_i^P = \mu_i$
	vector $\mathbf{z}$	$z_j = \sum_i \mu_i \frac{(y_i - \mu_i)}{\mu_i} x_{ij} \rightarrow \mathbf{z} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{k}$
	vector $\mathbf{k}$	$k_i = \frac{y_i - \mu_i}{\mu_i}$
NBI(MLE)/ NBI(moment)	matrix $\mathbf{I}$	$i_{js} = \sum_i \frac{\mu_i}{1 + a\mu_i} x_{ij} x_{is} \rightarrow \mathbf{I} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}$
	weight matrix $\mathbf{W}$	$w_i^{NBI} = \frac{\mu_i}{1 + a\mu_i}$
	vector $\mathbf{z}$	$z_j = \sum_i \frac{\mu_i}{1 + a\mu_i} \frac{(y_i - \mu_i)}{\mu_i} x_{ij} \rightarrow \mathbf{z} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{k}$
	vector $\mathbf{k}$	$k_i = \frac{y_i - \mu_i}{\mu_i}$
GPI(MLE)/ GPI(moment)	matrix $\mathbf{I}$	$i_{js} = \sum_i \frac{\mu_i}{(1 + a\mu_i)^2} x_{ij} x_{is} \rightarrow \mathbf{I} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}$
	weight matrix $\mathbf{W}$	$w_i^{GPI} = \frac{\mu_i}{(1 + a\mu_i)^2}$
	vector $\mathbf{z}$	$z_j = \sum_i \frac{\mu_i}{(1 + a\mu_i)^2} \frac{(y_i - \mu_i)}{\mu_i} x_{ij} \rightarrow \mathbf{z} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{k}$
	vector $\mathbf{k}$	$k_i = \frac{y_i - \mu_i}{\mu_i}$

4.2 Maximizing $\ell(\boldsymbol{\beta}, a)$ with respect to a

The maximization of $\ell(\boldsymbol{\beta}, a)$ with respect to a can be carried out by applying one-dimensional Newton-Raphson iteration,

$$a_{(r)} = a_{(r-1)} - \frac{f'(a_{(r-1)})}{f''(a_{(r-1)})}, \tag{4.12}$$

where f' denotes the first derivatives of function f and f'' the second derivatives of function f. The respective f' for Negative Binomial I (MLE), Negative Binomial I (moment), Generalized Poisson 1 (MLE) and Generalized Poisson I (moment) are Eq.(2.6), Eq.(2.7), Eq.(2.16) and Eq.(2.17).

The f'' equations for Negative Binomial I (MLE), Negative Binomial I (moment), Generalized Poisson I (MLE) and Generalized Poisson (moment) may be derived, and the respective equations are,

$$\sum_i \left\{ -\sum_{r=0}^{y_i-1} \left(\frac{r}{1+ar} \right)^2 - 2a^{-3} \log(1 + a\mu_i) + \frac{2a^{-2}\mu_i}{1 + a\mu_i} + \frac{(y_i + a^{-1})\mu_i^2}{(1 + a\mu_i)} \right\}, \tag{4.13}$$

$$-\sum_i \frac{(y_i - \mu_i)^2}{(1 + a\mu_i)^2}, \tag{4.14}$$

$$\sum_i \frac{y_i \mu_i^2}{(1 + a\mu_i)^2} - \frac{y_i^2 (y_i - 1)}{(1 + ay_i)^2} + \frac{2\mu_i^2 (y_i - \mu_i)}{(1 + a\mu_i)^3}, \tag{4.15}$$

and,

$$-2\sum_i \frac{(y_i - \mu_i)^2}{(1 + a\mu_i)^3}. \tag{4.16}$$

The process of finding the moment estimate, $\tilde{a}$, for Negative Binomial II and Generalized Poisson II does not involve any iteration. The moment estimate can be obtained directly from Eq.(2.12) and Eq.(2.23).

The equations for solving a in multiplicative regression models are summarized in Table 3.

Table 3. Equations for solving a.

Models	Equations for $a_{(r)} = a_{(r-1)} - \frac{f'(a_{(r-1)})}{f''(a_{(r-1)})}$	
NBI(MLE)	$f'(a)$	$\sum_i \left\{ \sum_{r=0}^{y_i-1}\left(\frac{r}{1+ar}\right) + a^{-2}\log(1+a\mu_i) - \frac{(y_i+a^{-1})\mu_i}{(1+a\mu_i)} \right\}$
	$f''(a)$	$\sum_i \left\{ -\sum_{r=0}^{y_i-1}\left(\frac{r}{1+ar}\right)^2 - 2a^{-3}\log(1+a\mu_i) + \frac{2a^{-2}\mu_i}{1+a\mu_i} + \frac{(y_i+a^{-1})\mu_i^2}{(1+a\mu_i)} \right\}$
NBI(moment)	$f'(a)$	$\sum_i \frac{(y_i-\mu_i)^2}{\mu_i(1+a\mu_i)} - (n-p)$
	$f''(a)$	$-\sum_i \frac{(y_i-\mu_i)^2}{(1+a\mu_i)^2}$
NBII	Straightforward calculation	$a = \sum_i \frac{(y_i-\mu_i)^2}{\mu_i(n-p)} - 1$
GPI(MLE)	$f'(a)$	$\sum_i \left\{ -\frac{y_i\mu_i}{1+a\mu_i} + \frac{y_i(y_i-1)}{1+ay_i} - \frac{\mu_i(y_i-\mu_i)}{(1+a\mu_i)^2} \right\}$
	$f''(a)$	$\sum_i \frac{y_i\mu_i^2}{(1+a\mu_i)^2} - \frac{y_i^2(y_i-1)}{(1+ay_i)^2} + \frac{2\mu_i^2(y_i-\mu_i)}{(1+a\mu_i)^3}$
GPI(moment)	$f'(a)$	$\sum_i \frac{(y_i-\mu_i)^2}{\mu_i(1+a\mu_i)^2} - (n-p)$
	$f''(a)$	$-2\sum_i \frac{(y_i-\mu_i)^2}{(1+a\mu_i)^3}$
GPII	Straightforward calculation	$a = \sqrt{\sum_i \frac{(y_i-\mu_i)^2}{\mu_i(n-p)}}$

4.3 Restrictions on Generalized Poisson I

The iterative programming for Generalized Poisson I distribution should also allows for restrictions on a because the probability density function, Eq.(2.13), indicates that the value of a must satisfy both $1+a\mu_i > 0$ and $1+ay_i > 0$. Therefore, after obtaining estimate of a in each iteration, the program should check that when $a < 0$ (underdispersion), a must also fulfilled the condition for both $1+a\mu_i > 0$ and $1+ay_i > 0$.

In other words, for condition $1+a\mu_i > 0$, the program should checks if $a > -\frac{1}{\max(\mu_i)}$ is true. If this condition is not true, a new estimate for a is set as $-\frac{1}{\max(\mu_i)+1}$. A similar check is then implemented for $1+ay_i > 0$. Finally, if both conditions of $a > -\frac{1}{\max(\mu_i)}$ and $a > -\frac{1}{\max(y_i)}$ are not true, a new estimate for a is set as $\min(-\frac{1}{\max(\mu_i)+1}, -\frac{1}{\max(y_i)+1})$.

4.4 Variance-covariance matrix for $\hat{\boldsymbol{\beta}}$

The variance-covariance matrix, $Var(\hat{\boldsymbol{\beta}})$, for Negative Binomial I and Generalized Poisson I regression models is also equal to the variance-covariance matrix of Poisson regression model, i.e.,

$$Var(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}. \tag{4.17}$$

However, the *i*th diagonal element of the weight matrix differs for each model, i.e., it is equal to Eq.(4.3) for Poisson, Eq.(4.8) for Negative Binomial I and Eq.(4.9) for Generalized Poisson I.

The variance-covariance matrix for Negative Binomial II and Generalized Poisson II is multiplied by a constant and they are equal to,

$$Var(\hat{\boldsymbol{\beta}}) = (1+a)(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}, \tag{4.18}$$

for Negative Binomial II and,

$$Var(\hat{\boldsymbol{\beta}}) = a^2(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}, \tag{4.19}$$

for Generalized Poisson II, where the *i*th diagonal element of the weight matrix is equal to the Poisson weight matrix, i.e., Eq.(4.3).

Examples of S-PLUS programming for Negative Binomial I (moment) and Generalized Poisson I (moment) are given in Appendix A and Appendix B. Similar programming can also be used for all of the multiplicative regression models which were discussed in this paper. Each programming is differentiated only by four distinguishable elements:

- Types of iteration.

  The sequential iterations are required for Negative Binomial I and Generalized Poisson I. For Poisson, Negative Binomial II and Generalized Poisson II, the standard iterations are adequate.
- Weight matrix.

  The weight matrix for Negative Binomial II and Generalized Poisson II is equal to the Poisson. Each of Negative Binomial I and Generalized Poisson I has its own weight matrix.
- Equation for estimating a.

  Each of Negative Binomial I (MLE), Negative Binomial I (moment), Generalized Poisson I (MLE) and Generalized Poisson I (moment) has its own equation for estimating a.
- Restriction on a.

  The restriction on a is required only in Generalized Poisson I.

5. EXAMPLES

5.1 Malaysian data

In this paper, the data for private car Third Party Property Damage (TPPD) claim frequencies from an insurance company in Malaysia will be considered. Specifically, the TPPD claim covers the legal liability for third party property loss or damage caused by or arising out of the use of an insured motor vehicle. The data, which was based on 170,000 private car policies for a three-year period of 1998-2000, was supplied by the General Insurance Association of Malaysia (PIAM). The exposure was expressed in terms of a car-year unit and the incurred claims consist of claims which were already paid as well as outstanding. Table 4 shows the rating factors and rating classes for the exposures and incurred claims, and altogether, there were $2\times2\times3\times4\times5=240$ cross-classified rating classes of claim frequencies to be estimated. The complete data, which contains the exposures, claim counts, rating factors and rating classes, is shown in Appendix C.

Table 4. Rating factors and rating classes for Malaysian data

Rating factors	Rating classes
Coverage type	Comprehensive Non-comprehensive
Vehicle make	Local Foreign
Vehicle use and driver's gender	Private-male Private-female Business
Vehicle year	0-1 year 2-3 year 4-5 year 6+ year
Location	Central North East South East Malaysia

The claim counts were first fitted to the Poisson multiplicative regression model. The fitting involves only 233 data points because seven of the rating classes have zero exposures. Several models were fitted by including different rating factors; first the main effects only, then the main effects plus each of the paired interaction factors. By using the deviance and degrees of freedom, the chi-squares statistics were calculated and compared to choose the best model. Table 5 gives the results of fitting several Poisson regression models to the count data.

Table 5. Analysis of deviance for Poisson

Model	deviance	df	Δdeviance	Δdf	χ^2	p-value
Null	2202	232	-	-	-	-
+ Coverage type	1924	231	278	1	278	0.00
+ Use-gender	997	229	927	2	464	0.00
+ Vehicle year	522	226	475	3	158	0.00
+ Vehicle location	369	222	153	4	38	0.00
+ Vehicle make	358	221	11	1	11	0.00
+ Vehicle make*vehicle year	255	218	103	3	34	0.00

Based on the deviance analysis, the best model indicates that all of the main effects are significant and one of the paired interaction factors, i.e., vehicle make and vehicle year, is also significant. Therefore, it is suggested that the rating factors for both vehicle make and vehicle year are combined to take into account the interaction between these two rating factors. The number of rating factors is now reduced from five to four. Table 6 shows the parameter estimates for the four-factor model.

Table 6. Parameter estimates for Poisson four-factor model

Parameter		estimate	std.error	p-value
β_1	Intercept	-2.37	0.04	0.00
β_2	Non-comprehensive	-0.68	0.07	0.00
β_3	Female	-0.51	0.03	0.00
β_4	Business	-6.04	1.00	0.00
β_5	Local, 2-3 year	-0.48	0.04	0.00
β_6	Local, 4-5 year	-0.82	0.05	0.00
β_7	Local, 6+ year	-1.06	0.05	0.00
β_8	Foreign, 0-1 year	-0.59	0.07	0.00
β_9	Foreign, 2-3 year	-0.68	0.05	0.00
β_{10}	Foreign, 4-5 year	-0.77	0.06	0.00
β_{11}	Foreign, 6+ year	-0.84	0.05	0.00
β_{12}	North	-0.22	0.03	0.00
β_{13}	East	-0.43	0.06	0.00
β_{14}	South	-0.01	0.04	0.78
β_{15}	East Malaysia	-0.50	0.06	0.00
Df		218.00		
Pearson χ^2		404.67		
Deviance		254.60		
Log L		-387.98		

The p-value for β_{14} (South) is equivalent to 0.78, and this value indicates that the parameter estimate is not significant. Therefore, the location for South is suggested to be combined with Central (Intercept) because both locations have almost similar risks. Table 7 shows the parameter estimates for the four-factor-combined-location model.

Table 7. Parameter estimates for Poisson four-factor-combined-location model

Parameter		estimate	std.error	*p*-value
β_1	Intercept	-2.37	0.03	0.00
β_2	Non-comprehensive	-0.68	0.07	0.00
β_3	Female	-0.51	0.03	0.00
β_4	Business	-6.04	1.00	0.00
β_5	Local, 2-3 year	-0.48	0.04	0.00
β_6	Local, 4-5 year	-0.82	0.05	0.00
β_7	Local, 6+ year	-1.06	0.05	0.00
β_8	Foreign, 0-1 year	-0.59	0.07	0.00
β_9	Foreign, 2-3 year	-0.68	0.05	0.00
β_{10}	Foreign, 4-5 year	-0.77	0.06	0.00
β_{11}	Foreign, 6+ year	-0.84	0.05	0.00
β_{12}	North	-0.22	0.03	0.00
β_{13}	East	-0.42	0.06	0.00
β_{14}	East Malaysia	-0.50	0.06	0.00
Df		219.00		
Pearson χ^2		404.47		
Deviance		254.67		
Log L		-388.02		

The result shows that all of the parameter estimates are significant. As a conclusion, based on the deviance analysis and parameter estimates, the best model is provided by the four-factor-combined-location model if the claim counts were fitted to the Poisson.

If the same four-factor-combined-location model was fitted to the multiplicative regression models of Negative Binomial and Generalized Poisson, the parameter estimates and standard errors may be compared. The comparisons are shown in Table 8 and Table 10.

Table 8. Poisson vs. Negative Binomial

Parameters		Poisson			Negative Binomial I (MLE)			Negative Binomial I (moment)			Negative Binomial II		
		est.	std. error	p-value	est.	std. error	p-value	est.	std. error	p-value	est.	std. error	p-value
a		-	-	-	0.02	-	-	0.15	-	-	0.85	-	-
β_1	Intercept	-2.37	0.03	0.00	-2.36	0.07	0.00	-2.37	0.15	0.00	-2.37	0.05	0.00
β_2	Non-comp	-0.68	0.07	0.00	-0.73	0.09	0.00	-0.79	0.13	0.00	-0.68	0.09	0.00
β_3	Female	-0.51	0.03	0.00	-0.54	0.05	0.00	-0.57	0.09	0.00	-0.51	0.04	0.00
β_4	Business	-6.04	1.00	0.00	-6.05	1.00	0.00	-6.06	1.00	0.00	-6.04	1.36	0.00
β_5	Local,2-3	-0.48	0.04	0.00	-0.51	0.09	0.00	-0.49	0.18	0.01	-0.48	0.06	0.00
β_6	Local,4-5	-0.82	0.05	0.00	-0.87	0.09	0.00	-0.87	0.19	0.00	-0.82	0.07	0.00
β_7	Local,6+	-1.06	0.05	0.00	-1.04	0.09	0.00	-0.98	0.18	0.00	-1.06	0.07	0.00
β_8	Foreign,0-1	-0.59	0.07	0.00	-0.62	0.10	0.00	-0.63	0.20	0.00	-0.59	0.09	0.00
β_9	Foreign,2-3	-0.68	0.05	0.00	-0.69	0.09	0.00	-0.65	0.19	0.00	-0.68	0.07	0.00
β_{10}	Foreign,4-5	-0.77	0.06	0.00	-0.76	0.10	0.00	-0.76	0.19	0.00	-0.77	0.08	0.00
β_{11}	Foreign,6+	-0.84	0.05	0.00	-0.81	0.09	0.00	-0.76	0.18	0.00	-0.84	0.07	0.00
β_{12}	North	-0.22	0.03	0.00	-0.16	0.06	0.00	-0.12	0.11	0.28	-0.22	0.04	0 00
β_{13}	East	-0.42	0.06	0.00	-0.43	0.08	0.00	-0.46	0.13	0.00	-0.42	0.08	0.00
β_{14}	EastM'sia	-0.50	0.06	0.00	-0.51	0.08	0.00	-0.49	0.13	0.00	-0.50	0.08	0.00
Df			219.00			218.00			218.00			218.00	
Pearson χ^2			404.47			293.71			219.00			-	
Deviance			254.67			155.99			90.72			-	
Log L			-388.02			-368.72			-391.64			-	

Table 8 shows the comparison between Poisson and Negative Binomial multiplicative regression models. The regression parameters for all models give similar estimates. The Negative Binomial I (MLE) and Negative Binomial II give similar inferences about the regression parameters, i.e., their standard errors are slightly larger than the Poisson's. However, the Negative Binomial I (moment) gives a relatively large values for the standard errors and hence, resulted in an insignificant regression parameter for β_{12}.

The deviance for Poisson regression model is relatively larger than the degrees of freedom, i.e., 1.16 times larger, and thus, indicating possible existence of overdispersion. To test for overdispersion, the likelihood ratio test of Poisson against Negative Binomial I (MLE) is implemented. The likelihood ratio statistic of $T = 38.6$ is significant, implying that the Negative Binomial I (MLE) is a better model. Further comparison can be made by using the results of likelihood ratio, AIC and BIC as shown in Table 9. Based on the likelihood ratio, AIC and BIC, the Negative Binomial I (MLE) is better than the Poisson.

Table 9. Likelihood ratio, AIC and BIC

Test/Criteria	Poisson	Negative Binomial I (MLE)
Likelihood ratio	-	38.6
AIC	804.0	767.4
BIC	809.2	773.0

Table 10 shows the comparison between Poisson and Generalized Poisson multiplicative regression models. Both Negative Binomial II and Generalized Poisson II give equal values for parameter estimates and standard errors. However, this result is to be expected because both regression models were fitted by using the same procedure.

The comparison between Poisson and Generalized Poisson also shows that the regression parameters for all models give similar estimates. The Generalized Poisson I (MLE) and Generalized Poisson II give similar inferences about the regression parameters. The Generalized Poisson I (moment) gives a relatively large values for the standard errors and this resulted in an insignificant regression parameter for β_{12}.

Based on the likelihood ratio test of Poisson against Generalized Poisson I (MLE), the likelihood ratio statistic of $T = 37.7$ is significant. Therefore, the Generalized Poisson (MLE) is also a better model compared to the Poisson.

Table 11 gives further comparison between Poisson and Generalized Poisson I (MLE). The comparison, which was based on the likelihood ratio, AIC and BIC, indicates that the Generalized Poisson I (MLE) is also a better model compared to the Poisson.

Table 10. Poisson vs. Generalized Poisson

Parameters		Poisson			Generalized Poisson I (MLE)			Generalized Poisson I (moment)			Generalized Poisson II		
		est.	std. error	p-value	est	std. error	p-value	est.	std. error	p-value	est.	std. error	p-value
a		-	-	-	0.007	-	-	0.035	-	-	1.359	-	-
β_1	Intercept	-2.37	0.03	0.00	-2.35	0.07	0.00	-2.37	0.16	0.00	-2.37	0.05	0.00
β_2	Non-comp	-0.68	0.07	0.00	-0.74	0.09	0.00	-0.80	0.13	0.00	-0.68	0.09	0 00
β_3	Female	-0.51	0.03	0.00	-0 55	0.05	0.00	-0.59	0.09	0.00	-0.51	0.04	0.00
β_4	Business	-6.04	1.00	0.00	-6.06	1.00	0.00	-6.08	1.00	0.00	-6.04	1.36	0.00
β_5	Local,2-3	-0.48	0.04	0.00	-0.52	0.09	0.00	-0.49	0.20	0.01	-0.48	0.06	0.00
β_6	Local,4-5	-0.82	0.05	0.00	-0.89	0.09	0.00	-0.88	0.19	0.00	-0.82	0.07	0.00
β_7	Local,6+	-1.06	0.05	0.00	-1.05	0.09	0.00	-0.94	0.19	0.00	-1.06	0.07	0.00
β_8	Foreign,0-1	-0.59	0.07	0.00	-0.63	0.10	0.00	-0.63	0.19	0.00	-0.59	0.09	0 00
β_9	Foreign,2-3	-0.68	0.05	0.00	-0.71	0.10	0.00	-0.64	0.19	0.00	-0.68	0.07	0.00
β_{10}	Foreign,4-5	-0.77	0.06	0.00	-0.77	0.10	0.00	-0.75	0.19	0.00	-0.77	0.08	0.00
β_{11}	Foreign,6+	-0.84	0.05	0.00	-0.81	0.09	0.00	-0.74	0.18	0.00	-0.84	0.07	0.00
β_{12}	North	-0.22	0.03	0.00	-0.14	0.06	0.00	-0.09	0.12	0.46	-0.22	0.04	0.00
β_{13}	East	-0.42	0.06	0.00	-0.43	0.08	0.00	-0.45	0.13	0.00	-0.42	0.08	0.00
β_{14}	EastM'sia	-0.50	0.06	0.00	-0.51	0.08	0.00	-0.51	0.12	0.00	-0.50	0.08	0.00
Df			219.00			218.00			218.00			218.00	
Pearson χ^2			404.47			294.72			219.00			-	
Deviance			254.67			159.21			98.52			-	
Log L			-388.02			-369.19			-392.92			-	

Table 11. Likelihood ratio, AIC and BIC

Test/Criteria	Poisson	Generalized Poisson I (MLE)
Likelihood ratio	-	37.7
AIC	804.0	766.4
BIC	809.2	773.9

The deviance analysis should also be implemented to both Negative Binomial I (MLE) and Generalized Poisson I (MLE) multiplicative regression models because the aim of our analysis is to obtain the simplest model that reasonably explains the variation in the data.

Following the same procedure as the Poisson, several Negative Binomial I (MLE) and Generalized Poisson I (MLE) regression models were fitted by including different rating factors; first the main effects only, then the main effects plus each of the paired interaction factors. By using the deviance and degrees of freedom, the chi-squares statistics were calculated and compared to choose the best model. Table 12 and Table 13 give the results of fitting several Negative Binomial I (MLE) and Generalized Poisson I (MLE) multiplicative regression models to the count data.

Table 12. Analysis of deviance for Negative Binomial I (MLE)

Model	deviance	df	Δdeviance	Δdf	χ^2	p-value
Null	207	231	-	-	-	-
+ Use-gender	166	229	41.63	2	20.82	0.00
+ Covarage type	149	228	16.54	1	16.54	0.00

Table 13. Analysis of deviance for Generalized Poisson I (MLE)

Model	deviance	df	Δdeviance	Δdf	χ^2	p-value
Null	262	231	-	-	-	-
+ Use-gender	180	229	81.90	2	40.95	0.00
+ Covarage type	159	228	20.73	1	20.73	0.00

Based on the deviance analysis, the best model indicates that only two of the rating factors, i.e., coverage type and use-gender, are significant and none of the paired interaction factor is significant. The parameter estimates for the two-factor models are shown in Table 14.

The two-factor models give significant parameter estimates. As a conclusion, based on the deviance analysis and parameter estimates, the best model for Negative Binomial I (MLE) and Generalized Poisson I (MLE) regression models is provided by the two-factor model.

Table 14. Parameter estimates for Negative Binomial I (MLE) and Generalized Poisson I (MLE)

Parameter		Negative Binomial I (MLE)			Generalized Poisson I (MLE)		
		estimate	std.error	p-value	estimate	std.error	p-value
a		0.16	-	-	0.04	-	-
β_1	Intercept	-3.15	0.06	0.00	-3.17	0.07	0.00
β_2	Non-comprehensive	-0.94	0.12	0.00	-0.92	0.12	0.00
β_3	Female	-0.55	0.09	0.00	-0.55	0.09	0.00
β_4	Business	-6.02	1.00	0.00	-6.01	1.00	0.00
Df			228.00			228.00	
Pearson χ^2			259.53			275.51	
Deviance			149.12			158.91	
Log L			-423.69			-425.97	

Based on the comparison between Poisson, Negative Binomial and Generalized Poisson multiplicative regression models, several remarks can be made:

- The Poisson, Negative Binomial and Generalized Poisson regression models give similar parameter estimates.
- The Negative Binomial and Generalized Poisson regression models give larger values for standard errors. Therefore, it is shown that in the presence of overdispersion, the Poisson overstates the significance of the regression parameters.
- The best regression model for Poisson indicates that all rating factors and one paired interaction factor are significant. However, the best regression model for Negative Binomial I (MLE) and Generalized Poisson I (MLE) indicates that only two rating factors are significant. Therefore, it is shown that in the presence of overdispersion, the Poisson overstates the significance of the rating factors.

5.2 Ship damage data

The ship damage incidents data of McCullagh and Nelder [19] was based on the damage incidents caused to the forward section of cargo-carrying vessels. The data provides information on the number and exposure for ship damage incidents, where the exposure was expressed in terms of aggregate number of month service. The risk of ship damage incidents was associated with three rating factors; ship type, year of construction and period of operation. The fitting procedure only involves thirty-four data points because six of the rating classes have zero exposures. The data, which was provided by Lloyd's Register of Shipping, can also be accessed from the Internet by using the following website address, http://sunsite.univie.ac.at/statlib/datasets/ships.

Since the same data was analyzed in some detail by both McCullagh and Nelder [19] and Lawless [21], the related remarks and discussions from their studies will be reported here. McCullagh and Nelder detected that there was some inter-ship variability in accident-proneness which could lead to overdispersion. For these reasons, McCullagh and Nelder assumed that,

$$Var(Y_i) = a\mu_i,$$

where,

$$a = \frac{\sum_i \frac{(y_i - \mu_i)^2}{\mu_i}}{n - p},$$

i.e., a is equal to the Pearson chi-squares divided by the degrees of freedom.

By using the fitting procedure which is similar to the Poisson IWLS regression, the McCullagh and Nelder's model was fitted to the main effects data. The parameter estimates for the model are equal to the Poisson, but the standard errors are equal or larger than the Poisson because they are multiplied by $\sqrt{a}$ where $a \geq 0$.

The same main effects data was also fitted to the multiplicative regression models of Negative Binomial I (MLE) and Negative Binomial I (moment) by Lawless [21]. However, the Negative Binomial I (MLE) produced $a = 0$, and this result is equivalent to fitting the data to the Poisson multiplicative regression model.

To confirm Lawless's result, we also run the S-PLUS programming for Negative Binomial I (MLE) to the ship data. We found that the parameter estimates for the ship data did not converge and therefore concluded that the data is better to be fitted by the Poisson. Table 15 shows the comparison between Poisson, Negative Binomial and McCullagh and Nelder multiplicative regression models.

Table 15. Poisson, Negative Binomial and McCullagh and Nelder regression models

Parameters		Poisson/Negative Binomial I (MLE)			Negative Binomial I (moment)			Negative Binomial II/ McCullagh and Nelder		
		est.	std. error	p-value	est.	std. error	p-value	est.	std. error	p-value
a			-	-	0.15	-	-	0.69/ 1.69	-	-
β_1	Intercept	-6.41	0.22	0.00	-6.45	0.41	0.00	-6.41	0.28	0.00
β_2	Ship type B	-0.54	0.18	0.00	-0.50	0.30	0.10	-0.54	0.23	0.02
β_3	Ship type C	-0.69	0.33	0.04	-0.56	0.41	0.18	-0.69	0.43	0.11
β_4	Ship type D	-0.08	0.29	0.79	-0.11	0.41	0.79	-0.08	0.38	0.84
β_5	Ship type E	0.33	0.24	0.17	0.46	0.35	0.19	0.33	0.31	0.29
β_6	Const'n 65-69	0.70	0.15	0.00	0.72	0.35	0.04	0.70	0.19	0.00
β_7	Const'n 70-74	0.82	0.17	0.00	0.91	0.34	0.01	0.82	0.22	0.00
β_8	Const'n 75-79	0.45	0.23	0.05	0.46	0.42	0.27	0.45	0.30	0.13
β_9	Oper'n 75-79	0.38	0.12	0.00	0.34	0.23	0.14	0.38	0.15	0.01
Df			25.00			24.00			24.00	
Pearson χ^2			42.28			25.00			-	
Deviance			38.70			25.01			-	
Log L			-68.28			-72.83			-	

The parameter estimates and standard errors for both Negative Binomial II and McCullagh and Nelder are equal because the models were fitted by using the same procedure.

McCullagh and Nelder [19] found that the main effects model fits the data well, i.e., all of the main effects are significant and none of the paired interaction factor is significant. According to McCullagh and Nelder, if the Poisson regression model was fitted, there was an inconclusive evidence of an interaction between ship type and year of construction. However, this evidence vanished completely if the data is fitted by the overdispersion model.

Lawless [21] reported that the regression models for both McCullagh and Nelder and Negative Binomial I (moment) fit the data well. According to Lawless, both models gave the same estimates for the regression parameters and similar inferences about the regression effects. Lawless also remarked that even though there was no strong evidence of overdispersion under the Negative Binomal I (moment) or McCullagh and Nelder regression models, the method for fitting the models has a strong influence on the standard errors. In particular, the Poisson and Negative Binomial I (moment) respectively produced the smallest and largest standard errors, whereas the McCullagh and Nelder's were somewhere in between. In addition, the effects of ship type are not significant under the Negative Binomial I (moment), whereas they are under the Poisson and to a lesser extent under the McCullagh and Nelder.

If the same main effects data was fitted to the multiplicative regression models of Generalized Poisson, the parameter estimates and standard errors may also be compared. The comparisons are shown in Table 16.

Table 16. Poisson vs. Generalized Poisson

Parameters		Poisson/Generalized Poisson I (MLE)			Generalized Poisson I (moment)			Generalized Poisson II		
		est.	std. error	p-value	est.	std. error	p-value	est.	std. error	p-value
a		0.00	-	-	0.06	-	-	1.30	-	-
β_1	Intercept	-6.41	0.22	0.00	-6.46	0.45	0.00	-6.41	0.28	0.00
β_2	Ship type B	-0.54	0.18	0.00	-0 49	0.33	0.14	-0.54	0.23	0.02
β_3	Ship type C	-0.69	0.33	0.04	-0.56	0.41	0.17	-0.69	0.43	0.11
β_4	Ship type D	-0.08	0.29	0.79	-0.11	0.41	0.80	-0.08	0.38	0.84
β_5	Ship type E	0.33	0.24	0.17	0.49	0.36	0.17	0.33	0.31	0.29
β_6	Const'n 65-69	0.70	0.15	0.00	0.73	0.41	0.07	0.70	0.19	0.00
β_7	Const'n 70-74	0.82	0.17	0.00	0.94	0.39	0.02	0.82	0.22	0.00
β_8	Const'n 75-79	0.45	0.23	0.05	0.46	0.46	0.31	0.45	0.30	0.13
β_9	Oper'n 75-79	0.38	0.12	0.00	0.34	0 26	0.19	0.38	0.15	0.01
Df			25.00			24.00			24.00	
Pearson χ^2			42.28			25.00			-	
Deviance			38.70			25.29			-	
Log L			-68.28			-74.22			-	

The parameter estimates and standard errors for Generalized Poisson II, Negative Binomial II and McCullagh and Nelder are equal because the regression models were fitted by using the same procedure.

Similar to the Negative Binomial I (MLE), the Generalized Poisson I (MLE) also does not give converged values for its parameter estimates. Therefore, it will be assumed that the

Generalized Poisson I (MLE) produces $a = 0$ and this is also equivalent to fitting the data to the Poisson.

All models give similar estimates for the regression parameters. The Poisson and Generalized Poisson I (moment) respectively produced the smallest and largest standard errors, whereas the Generalized Poisson II's were somewhere in between. The effects of ship type are also not significant under the Generalized Poisson I (moment), whereas they are under the Poisson and to a lesser extent under the Generalized Poisson II.

5.3 Canadian data

The Canadian private automobile liability insurance data from Bailey and Simon [1] provides information on the number of claims incurred and exposures, where the exposure was expressed in terms of number of earned car years. The data was classified into two rating factors, merit rating and class rating. Altogether, there were twenty cross-classified rating classes of claim frequencies to be estimated. The data can also be accessed from the Internet by using the following website address, http://www.casact.org/library/astin/vol1no4/ 192.pdf.

Table 17 and Table 18 show the comparison between Poisson, Negative Binomial and Generalized Poisson multiplicative regression models for the main effects data.

The Negative Binomial II and Generalized Poisson II give equal values for parameter estimates and standard errors. The regression parameters for all models give similar estimates. The smallest standard errors are given by the Poisson, the largest are by the Negative Binomial II and Generalized Poisson II, whereas the standard errors for Negative Binomial I (MLE), Negative Binomial I (moment), Generalized Poisson I (MLE) and Generalized Poisson I (moment) lie somewhere in between.

The likelihood ratio test for Poisson against Negative Binomial I (MLE) produces likelihood ratio statistic of $T = 514.94$. The likelihood ratio is very significant, indicating that the Negative Binomial I (MLE) is a better model compared to the Poisson.

The likelihood ratio test for Poisson against Generalized Poisson I (MLE) also produces a very significant likelihood ratio statistic, $T = 525.44$. Therefore, the Generalized Poisson I (MLE) is also a better model compared to the Poisson.

Table 17. Poisson vs. Negative Binomial

Parameters		Poisson			Negative Binomial I (MLE)			Negative Binomial I (moment)			Negative Binomial II		
		est.	std. error	*p*-value	est.	std. error	*p*-value	est.	std. error	*p*-value	est.	std. error	*p*-value
a		-	-	-	0.001	-	-	0.002	-	-	47.15	-	-
β_1	Intercept	-2.53	0.00	0.00	-2.45	0.02	0.00	-2.45	0.03	0.00	-2.53	0.01	0.00
β_2	Class 2	0.30	0.01	0.00	0.24	0.03	0.00	0.24	0.03	0.00	0.30	0.05	0.00
β_3	Class 3	0.47	0.01	0.00	0.43	0.03	0.00	0.43	0.03	0.00	0.47	0.03	0.00
β_4	Class 4	0.53	0.01	0.00	0.46	0.03	0.00	0.46	0.03	0.00	0.53	0.04	0.00
β_5	Class 5	0.22	0.01	0.00	0.14	0.03	0.00	0.14	0.04	0.00	0.22	0.07	0.00
β_6	Merit X	0.27	0.01	0.00	0.22	0.03	0.00	0.22	0.03	0.00	0.27	0.05	0.00
β_7	Merit Y	0.36	0.01	0.00	0.27	0.03	0.00	0.27	0.03	0.00	0.36	0.04	0.00
β_8	Merit B	0.49	0.00	0.00	0.41	0.02	0 00	0.41	0.03	0.00	0.49	0.03	0.00
Df		12 00			11.00			11.00			11.00		
Pearson χ^2		577.83			17.56			12.00			-		
Deviance		579.52			17.67			12.08			-		
Log L		-394.96			-137.49			-138.11			-		

Table 18. Poisson vs. Generalized Poisson

Parameters		Poisson			Generalized Poisson I (MLE)			Generalized Poisson I (moment)			Generalized Poisson II		
		est.	std. error	*p*-value	est.	std. error	*p*-value	est.	std. error	*p*-value	est.	std. error	*p*-value
a		-	-	-	0.0002	-	-	0.0002	-	-	6.94	-	-
β_1	Intercept	-2.53	0.00	0.00	-2.41	0.03	0.00	-2.41	0.03	0.00	-2.53	0.01	0.00
β_2	Class 2	0.30	0.01	0.00	0.22	0.03	0.00	0.22	0.03	0.00	0.30	0.05	0.00
β_3	Class 3	0.47	0.01	0.00	0.42	0.02	0.00	0.42	0.03	0.00	0.47	0.03	0.00
β_4	Class 4	0.53	0.01	0.00	0.43	0.02	0.00	0.43	0.03	0.00	0.53	0.04	0.00
β_5	Class 5	0.22	0.01	0.00	0.12	0.03	0.00	0.12	0.03	0.00	0.22	0 07	0.00
β_6	Merit X	0.27	0 01	0.00	0.20	0.02	0.00	0.20	0.02	0.00	0.27	0.05	0.00
β_7	Merit Y	0.36	0.01	0.00	0.24	0.02	0.00	0.24	0.02	0.00	0.36	0.04	0.00
β_8	Merit B	0.49	0.00	0.00	0.38	0.02	0.00	0.38	0.38	0.00	0.49	0.03	0 00
Df		12.00			11.00			11.00			11.00		
Pearson χ^2		577.83			15.04			12.00			-		
Deviance		579.52			15.31			12.20			-		
Log L		-394.96			-132.24			-132.46			-		

Table 19 gives further comparison between Poisson, Negative Binomial I (MLE) and Generalized Poisson I (MLE). The comparison, which was based on the likelihood ratio, AIC and BIC, indicates that the Negative Binomial I (MLE) and Generalized Poisson I (MLE) are better models compared to the Poisson.

Table 19. Likelihood ratio, AIC and BIC

Test/Criteria	Poisson	Negative Binomial I (MLE)	Generalized Poisson I (MLE)
Likelihood ratio	-	514.94	525.44
AIC	805.92	292.98	282.48
BIC	800.33	286.69	276.19

6. ADDITIVE REGRESSION MODELS

In this section, the estimation procedure for the additive regression models will be briefly discussed. However, a slightly different approach is taken to compute the regression parameters.

6.1 Poisson

Let r_i, y_i and e_i denote the claim frequency rate, claim count and exposure for the *i*th class so that the observed frequency rate is equal to,

$$r_i = \frac{y_i}{e_i}. \tag{6.1}$$

If the random variable for claim count, Y_i, follows a Poisson distribution, the probability density function can be written as,

$$f(y_i) = g(r_i) = \frac{\exp(-e_i f_i)(e_i f_i)^{e_i r_i}}{(e_i r_i)!}, \tag{6.2}$$

where the mean and variance for the claim count is equal to $E(Y_i) = Var(Y_i) = e_i E(R_i) = e_i f_i$.

For Poisson regression model, the likelihood equations are equal to,

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_j} = \sum_i \frac{e_i(r_i - f_i)}{f_i}\frac{\partial f_i}{\partial \beta_j} = 0, \qquad j = 1,2,\ldots, p. \tag{6.3}$$

If the Poisson follows an additive model, the mean or the fitted value for frequency rate can be written as,

$$E(R_i) = f_i = \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta}, \tag{6.4}$$

so that,

$$\frac{\partial f_i}{\partial \beta_j} = x_{ij}. \tag{6.5}$$

Therefore, the first derivatives of log likelihood for Poisson are,

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_j} = \sum_i \frac{(r_i - f_i)e_i x_{ij}}{f_i} = 0, \qquad j = 1,2,\ldots p, \tag{6.6}$$

and the negative expectation of the second derivatives of log likelihood are,

$$-E\left(\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \beta_j \partial \beta_s}\right) = \frac{e_i}{f_i} x_{ij} x_{is}, \qquad , j,s = 1,2,\ldots, p. \tag{6.7}$$

The information matrix, $\mathbf{I}$, which contains negative expectation of the second derivatives of log likelihood, may be written as,

$$\mathbf{I} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}, \tag{6.8}$$

where $\mathbf{X}$ denotes the matrix of explanatory variables, and $\mathbf{W}$ the diagonal weight matrix whose *i*th diagonal element is equal to,

$$w_i^P = \frac{e_i}{f_i}. \tag{6.9}$$

The first derivatives of log likelihood, i.e., Eq.(6.6), can be written as,

$$\mathbf{z} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{k}, \tag{6.10}$$

where $\mathbf{W}$ is the diagonal weight matrix whose *i*th diagonal element is also equivalent to Eq.(6.9), and $\mathbf{k}$ the vector whose *i*th row is equal to,

$$k_i = r_i - f_i. \tag{6.11}$$

6.2 Negative Binomial I

If the mean or the fitted value for frequency rate is assumed to follow an additive regression model, the first derivatives of log likelihood for Negative Binomial I are,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial \beta_j} = \sum_i \frac{(r_i - f_i) e_i x_{ij}}{f_i (1 + a e_i f_i)} = 0, \qquad j = 1, 2, \ldots p, \tag{6.12}$$

and the negative expectation of the second derivatives of log likelihood are,

$$-E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, a)}{\partial \beta_j \partial \beta_s}\right) = \frac{e_i}{f_i (1 + a e_i f_i)} x_{ij} x_{is}, \qquad j, s = 1, 2, \ldots, p. \tag{6.13}$$

Therefore, the information matrix, $\mathbf{I}$, may also be written as Eq.(6.8). However, the *i*th diagonal element of the weight matrix, $\mathbf{W}$, is equal to,

$$w_i^{NBI} = \frac{e_i}{f_i (1 + a e_i f_i)}. \tag{6.14}$$

The first derivatives of log likelihood, i.e., Eq.(6.12), can also be written as Eq.(6.10) where $\mathbf{k}$ is the vector whose *i*th row is equal to Eq.(6.11). However, the *i*th diagonal element of the weight matrix, $\mathbf{W}$, is equal to Eq.(6.14).

6.3 Negative Binomial II

The maximum likelihood estimates, $\hat{\boldsymbol{\beta}}$, for Negative Binomial II additive regression model are numerically difficult to be solved from the likelihood equations. However, the regression parameters are easier to be approximated by using the least squares equations,

$$\sum_i \frac{e_i (r_i - f_i)}{f_i (1 + a)} \frac{\partial f_i}{\partial \beta_j} = \sum_i \frac{(r_i - f_i) e_i x_{ij}}{f_i (1 + a)} = 0, \qquad j = 1, 2, \ldots, p, \tag{6.15}$$

because the distribution of Negative Binomial II has a constant variance-mean ratio. Since Eq.(6.15) is also equal to the likelihood equations of the Poisson, i.e., Eq.(6.6), the least squares estimates, $\tilde{\boldsymbol{\beta}}$, are also equivalent to the Poisson maximum likelihood estimates. However, the standard errors are equal or larger than the Poisson because they are multiplied by $\sqrt{1+a}$ where $a \geq 0$.

6.4 Generalized Poisson I

If the mean or the fitted value for frequency rate is assumed to follow an additive regression model, the first derivatives of log likelihood for Generalized Poisson I are,

$$\frac{\partial \ell(\boldsymbol{\beta}, a)}{\partial \beta_j} = \sum_i \frac{(r_i - f_i)e_i x_{ij}}{f_i(1 + ae_i f_i)^2}, \qquad j = 1,2,\ldots p, \tag{6.16}$$

and the negative expectation of the second derivatives of log likelihood are,

$$-E\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, a)}{\partial \beta_j \partial \beta_s}\right) = \frac{e_i}{f_i(1 + ae_i f_i)^2} x_{ij} x_{is}, \qquad , j,s = 1,2,\ldots, p. \tag{6.17}$$

Therefore, the information matrix, $\mathbf{I}$, may also be written as Eq.(6.8). However, the *i*th diagonal element of the weight matrix, $\mathbf{W}$, is equal to,

$$w_i^{GPI} = \frac{e_i}{f_i(1 + ae_i f_i)^2}. \tag{6.18}$$

The first derivatives of log likelihood, i.e., Eq.(6.16), can be written as Eq.(6.10), where $\mathbf{k}$ is the vector whose *i*th row is equal to Eq.(6.11). However, the *i*th diagonal element of the weight matrix, $\mathbf{W}$, is equivalent to Eq.(6.18).

6.5 Generalized Poisson II

The maximum likelihood estimates, $\hat{\boldsymbol{\beta}}$, for Generalized Poisson II additive regression model are also numerically difficult to be solved from the likelihood equations. However, by using the least squares equations,

$$\sum_i \frac{(r_i - f_i)e_i x_{ij}}{a^2 f_i} = 0, \qquad j = 1,2,\ldots, p. \tag{6.19}$$

the regression parameters are easier to be calculated because the distribution of Generalized Poisson II also has a constant variance-mean ratio. Since Eq.(6.19) is equal to the likelihood equations of the Poisson, i.e., Eq.(6.6), the the least squares estimates, $\tilde{\boldsymbol{\beta}}$, are also equivalent to the Poisson maximum likelihood estimates. However, the standard errors are equal, larger or smaller than the Poisson because they are multiplied by a where $a \geq 1$ or $\frac{1}{2} \leq a < 1$.

The methods and equations for solving $\boldsymbol{\beta}$ in additive regression models are summarized in Table 20. The matrices and vectors for solving $\boldsymbol{\beta}$ in additive regression models are summarized in Table 21. An example of S-PLUS programming for the additive regression model of Negative Binomial I (moment) is given in Appendix D.

Table 20. Methods and equations for solving $\boldsymbol{\beta}$ in additive regression models

Models	Estimation of β	
	Method	Equation
Poisson	Maximum Likelihood	$\sum_i \frac{(r_i - f_i)e_i x_{ij}}{f_i} = 0$
Negative Binomial I	Maximum Likelihood	$\sum_i \frac{(r_i - f_i)e_i x_{ij}}{f_i(1 + ae_i f_i)} = 0$
Negative Binomial II	Weighted Least Squares	$\sum_i \frac{(r_i - f_i)e_i x_{ij}}{(1 + a)f_i} = 0$
Generalized Poisson I	Maximum Likelihood	$\sum_i \frac{(r_i - f_i)e_i x_{ij}}{f_i(1 + ae_i f_i)^2} = 0$
Generalized Poisson II	Weighted Least Squares	$\sum_i \frac{(r_i - f_i)e_i x_{ij}}{a^2 f_i} = 0$

Table 21. Matrices and vectors for solving $\boldsymbol{\beta}$ in additive regression models

Models	Matrices and vectors for $\boldsymbol{\beta}_{(r)} = \boldsymbol{\beta}_{(r-1)} + \mathbf{I}_{(r-1)}^{-1}\mathbf{z}_{(r-1)}$, where $\mathbf{I}_{(r-1)} = \mathbf{X}^{\mathrm{T}}\mathbf{W}_{(r-1)}\mathbf{X}$, $\mathbf{z}_{(r-1)} = \mathbf{X}^{\mathrm{T}}\mathbf{W}_{(r-1)}\mathbf{k}_{(r-1)}$, js-th element of matrix $\mathbf{I} = i_{js} = -E\left(\dfrac{\partial^2 \ell}{\partial\beta_j \partial\beta_s}\right)$, j-th row of vector $\mathbf{z} = z_j = \dfrac{\partial \ell}{\partial \beta_j}$.	
Poisson/ NBII/ GPII	matrix $\mathbf{I}$	$i_{js} = \sum_i \dfrac{e_i}{f_i} x_{ij} x_{is} \rightarrow \mathbf{I} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}$
	weight matrix $\mathbf{W}$	$w_i^P = \dfrac{e_i}{f_i}$
	vector $\mathbf{z}$	$z_j = \sum_i \dfrac{e_i}{f_i}(r_i - f_i)x_{ij} \rightarrow \mathbf{z} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{k}$
	vector $\mathbf{k}$	$k_i = r_i - f_i$
NBI	matrix $\mathbf{I}$	$i_{js} = \sum_i \dfrac{e_i}{f_i(1+ae_i f_i)} x_{ij} x_{is} \rightarrow \mathbf{I} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}$
	weight matrix $\mathbf{W}$	$w_i^{NBI} = \dfrac{e_i}{f_i(1+ae_i f_i)}$
	vector $\mathbf{z}$	$z_j = \sum_i \dfrac{e_i}{f_i(1+ae_i f_i)}(r_i - f_i)x_{ij} \rightarrow \mathbf{z} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{k}$
	vector $\mathbf{k}$	$k_i = r_i - f_i$
GPI	matrix $\mathbf{I}$	$i_{js} = \sum_i \dfrac{e_i}{f_i(1+ae_i f_i)^2} x_{ij} x_{is} \rightarrow \mathbf{I} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}$
	weight matrix $\mathbf{W}$	$w_i^{GPI} = \dfrac{e_i}{f_i(1+ae_i f_i)^2}$
	vector $\mathbf{z}$	$z_j = \sum_i \dfrac{e_i}{f_i(1+ae_i f_i)^2}(r_i - f_i)x_{ij} \rightarrow \mathbf{z} = \mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{k}$
	vector $\mathbf{k}$	$k_i = r_i - f_i$

6.6 Examples

Following the same examples as the multiplicative regression models, the additive regression models were also fitted on three different sets of claim frequency data; Malaysian data, ship damage incident data, and Canadian data. Unfortunately, the Malaysian data did not give converged parameter solutions for any of the Poisson, Negative Binomial and Generalized Poisson regression models. However, the parameter solutions are obtainable for both ship damage incident data and Canadian data. Table 22 shows the comparison between Poisson, Negative Binomial and Generalized Poisson additive regression models for the ship damage incident data.

Table 22. Poisson, Negative Binomial and Generalized Poisson for ship data

Parameters		Poisson/NBI(MLE)/ GPI(MLE)			NBI(moment)			GPI(moment)			NBII/GPII		
		est. ($\times 10^3$)	std. error ($\times 10^3$)	p-value	est. ($\times 10^3$)	std. error ($\times 10^3$)	p-value	est. ($\times 10^3$)	std. error ($\times 10^3$)	p-value	est. ($\times 10^3$)	std. error ($\times 10^3$)	p-value
a		0.00	-	-	133.73	-	-	52.94	-	-	599.25/ 1264.61	-	-
β_1	Intercept	2.60	0.72	0.00	2.19	1.04	0.03	2.16	1.05	0.04	2.60	0.91	0.00
β_2	Ship type B	-1.73	0.71	0.01	-1.33	1.01	0.19	-1.30	1.02	0.20	-1.73	0.90	0.05
β_3	Ship type C	-1.89	0.86	0.03	-1.52	1.12	0.17	-1.52	1.11	0.17	-1.89	1.09	0.08
β_4	Ship type D	-0.79	1.10	0.47	-1.05	1.32	0.43	-1.13	1.28	0.38	-0.79	1.39	0.57
β_5	Ship type E	1.87	1.30	0.15	2.72	1.91	0.15	2.87	1.92	0.13	1.87	1.64	0.25
β_6	Cons. 65-69	1.05	0.24	0.00	0 87	0.56	0.12	0.78	0.63	0.22	1.05	0.31	0.00
β_7	Cons. 70-74	1.58	0.38	0.00	2.15	0.76	0.00	2.33	0.84	0.01	1.58	0.47	0.00
β_8	Cons. 75-79	0.69	0.55	0.22	0.77	0.94	0.42	0.76	0.98	0.44	0.69	0 70	0.33
β_9	Oper. 75-79	0.79	0.24	0.00	0.79	0.52	0.13	0.81	0.58	0.16	0.79	0.31	0.01
Df		25.00			24.00			24.00			24.00		
Pearson χ^2		39.98			25.00			25.00			-		
Deviance		38.44			25.65			26.12			-		
Log L		-68.15			-72.44			-73.48			-		

After running the S-PLUS programming for Negative Binomial I (MLE) and Generalized Poisson I (MLE) to the ship data, we found that the models did not give converged parameter solutions and concluded that the data is better to be fitted by the Poisson. Since the Poisson is a special case of the Negative Binomial I (MLE) and Generalized Poisson I (MLE), the result of fitting the Poisson is also equivalent to the result of fitting the Negative Binomial I (MLE) or Generalized Poisson I (MLE) which produces $a = 0$.

The parameter estimates and standard errors for Negative Binomial II and Generalized Poisson II are equal because both models were fitted by using the same procedure.

The smallest standard errors are given by the Poisson, the largest are by the Negative Binomial I (moment) and Generalized Poisson I (moment), whereas the standard errors for Negative Binomial II and Generalized Poisson II are somewhere in between.

Table 23 shows the comparison between Poisson, Negative Binomial and Generalized Poisson additive regression models for the Canadian data.

The parameter estimates and standard errors for Negative Binomial II and Generalized Poisson II are equal because both models were fitted by using the same procedure.

The smallest standard errors are given by the Poisson, the largest are by the Negative Binomial I (moment) and Generalized Poisson I (moment), whereas the standard errors for Negative Binomial I (MLE), Generalized Poisson I (MLE), Negative Binomial II and Generalized Poisson II are somewhere in between.

Table 23. Poisson, Negative Binomial and Generalized Poisson for Canadian data

Parameters		Poisson			NBI(MLE)			NBI(moment)		
		est. (×10²)	std. error (×10²)	p-value	est. (×10²)	std. error (×10²)	p-value	est. (×10²)	std. error (×10²)	p-value
a		-	-	-	0.06	-	-	0.12	-	-
β_1	Intercept	7.88	0.02	0.00	7.98	0.17	0.00	8.00	0.22	0.00
β_2	Class 2	3.13	0.09	0.00	2.99	0.25	0.00	2.99	0.32	0.00
β_3	Class 3	5.24	0.07	0.00	5.66	0.27	0.00	5.70	0.35	0.00
β_4	Class 4	6.53	0.08	0.00	6.36	0.28	0.00	6.34	0.36	0.00
β_5	Class 5	2.17	0.12	0.00	1.88	0.26	0.00	1.81	0.32	0.00
β_6	Merit X	2.76	0.08	0.00	2.74	0.24	0.00	2.72	0.30	0.00
β_7	Merit Y	3.86	0.08	0.00	3.55	0.24	0.00	3.50	0.31	0.00
β_8	Merit B	5.88	0.06	0.00	5.63	0.25	0.00	5.59	0.32	0.00
Df			12.00			11.00			11.00	
Pearson χ^2			95.93			19.19			12.00	
Deviance			96.07			19.36			12.10	
Log L			-153.24			-132.31			-133.24	

Parameters		GPI(MLE)			GPI(moment)			NBII/GPII		
		est. (×10²)	std. error (×10²)	p-value	est. (×10²)	std. error (×10²)	p-value	est. (×10²)	std. error (×10²)	p-value
a		0.02	-	-	0.02	-	-	699.38/ 282.73	-	-
β_1	Intercept	8.24	0.28	0.00	8.29	0.35	0.00	7.88	0.05	0.00
β_2	Class 2	2.84	0.30	0.00	2.84	0.35	0.00	3.13	0 24	0.00
β_3	Class 3	5.84	0.31	0.00	5.90	0.38	0.00	5.24	0.19	0.00
β_4	Class 4	6.21	0.31	0.00	6.19	0.38	0.00	6.53	0.23	0.00
β_5	Class 5	1.72	0.31	0.00	1.64	0.36	0.00	2.17	0.33	0.00
β_6	Merit X	2.55	0.27	0.00	2 51	0.32	0.00	2.76	0.23	0.00
β_7	Merit Y	3.28	0.27	0.00	3.19	0.32	0.00	3.86	0.22	0.00
β_8	Merit B	5.37	0.28	0.00	5.28	0.33	0.00	5.88	0.18	0.00
Df			11.00			11.00			11.00	
Pearson χ^2			17.08			12.00			-	
Deviance			17.51			12 34			-	
Log L			-132.12			-132.61			-	

7. CONCLUSIONS

This paper proposed the Negative Binomial and Generalized Poisson regression models as alternatives for handling overdispersion. Specifically, four types of distributions, i.e., Negative Binomial I, Negative Binomial II, Generalized Poisson I and Generalized Poisson II, and two types of regression models, i.e., multiplicative and additive, were discussed. Since the likelihood equations for the multiplicative and additive regression models of the Negative Binomial I and Generalized Poisson I are equal to the weighted least squares, the fitting procedure can be carried out easily by using the Iterative Weighted Least Squares (IWLS) regression.

The estimation of the dispersion parameter, a, can be implemented by using either the maximum likelihood method or the method of moment. In this paper, the models where a is estimated by the maximum likelihood method are denoted by Negative Binomial I (MLE) and Generalized Poisson I (MLE). Similarly, the Negative Binomial I (moment) and Generalized Poisson I (moment) represent the models where a is estimated by the method of moment.

The maximum likelihood estimates for Negative Binomial II and Generalized Poisson II are numerically difficult to be solved because their likelihood equations are not equal to the weighted least squares. As an alternative, the method of least squares is suggested because both Negative Binomial II and Generalized Poisson II have constant variance-mean ratios.

Table 1 and Table 20 summarize the methods and equations for solving $\beta_j, j = 1,2,\ldots, p$, in multiplicative and additive regression models. The matrices and vectors for solving $\boldsymbol{\beta}$ in multiplicative and additive regression models are summarized in Table 2 and Table 21. Finally, Table 3 summarizes the equations for solving a.

This paper also briefly discussed several goodness-of-fit measures which were already familiar to those who used Generalized Linear Model with Poisson error structure for claim count or frequency analysis. The measures, which are also applicable to the Negative Binomial as well as the Generalized Poisson regression models, are the Pearson chi-squares, deviance, likelihood ratio test, Akaike Information Criteria (AIC) and Bayesian Schwartz Information Criteria (BIC).

In this paper, the multiplicative and additive regression models of the Poisson, Negative Binomial and Generalized Poisson were fitted, tested and compared on three different sets of claim frequency data; Malaysian private motor third party property damage data, ship damage incident data from McCullagh and Nelder [19], and data from Bailey and Simon [1] on Canadian private automobile liability. Unfortunately, none of the additive regression models give converged parameter solutions for the Malaysian data.

This paper shows that even though the Poisson, the Negative Binomial and the Generalized Poisson produce similar estimate for the regression parameters, the standard errors for the Negative Binomial and the Generalized Poisson are larger than the Poisson. Therefore, the Poisson overstates the significance of the regression parameters in the presence of overdispersion. An example can be seen from the results of fitting the Poisson, the Negative Binomial and the Generalized Poisson to the ship damage data. The effects of ship type are not significant under the NBI-moment or the GPI-moment, whereas they are under the Poisson, and to a lesser extent under the McCullagh and Nelder or the NBII or the GPII.

This paper also shows that in the presence of overdispersion, the Poisson overstates the significance of the rating factors. An example can be seen from the results of implementing the deviance analysis to the Malaysian data. The best regression model for the Poisson indicates that all rating factors and one paired interaction factor are significant. However, the best regression model for NBI-MLE and GPI-MLE indicates that only two rating factors are significant. Another example can be seen from the ship damage data. According to McCullagh and Nelder [19], there was an evidence of interaction between ship type and year of construction if the Poisson regression was fitted. However, the evidence vanished completely if the data is fitted by the overdispersion model.

In addition, this paper shows that the maximum likelihood approach has several advantages compared to the quasi likelihood approach, which was suggested in the actuarial literature, to accommodate overdispersion in claim count or frequency data. Besides having good properties, the maximum likelihood approach allows the likelihood ratio and other standard maximum likelihood tests to be implemented.

The Negative Binomial and the Generalized Poisson models are not that difficult to be understood. Even though the probability density function for both Negative Binomial and Generalized Poisson involve mathematically complex formulas, the mean and variance for both models are conceptually simpler to be interpreted. The mean for both Negative Binomial and Generalized Poisson models are equal to the Poisson. The variance of the Negative Binomial is equal or larger than the Poisson, and this allows the Negative Binomial model to handle overdispersion. The variance of the Generalized Poisson is equal, larger or smaller than the Poisson, and this allows the Generalized Poisson to handle either overdispersion or underdispersion.

The Negative Binomial and Generalized Poisson are also not that difficult to be fitted. The fitting procedure can be carried out by using the Iterative Weighted Least Squares regression which was used in the Poisson fitting procedure. The only difference is that the Negative Binomial and the Generalized Poisson has their own weight matrix, and the iteration procedure for calculating the dispersion parameter, a, has to be added in the fitting procedure.

Acknowledgment

The authors gratefully acknowledge the financial support received in the form of a research grant (IRPA RMK8: 09-02-02-0112-EA274) from the Ministry of Science, Technology and Innovation (MOSTI), Malaysia. The authors are also pleasured to thank the General Insurance Association of Malaysia (PIAM), in particular Mr. Carl Rajendram and Mrs. Addiwiyah, for supplying the data.

Appendix A: S-PLUS programming for Negative Binomial I (moment) multiplicative regression model

```
NB.moment <- function(data)
{
# To identify matrix X, vector count and vector exposure from the data
    X          <- as.matrix(data[, -(1:2)])
    count      <- as.vector(data[, 1])
    exposure   <- as.vector(data[, 2])
# To set initial values for a and beta
    new.a      <- c(0.001)
    new.beta   <- rep(c(0.001), dim(X)[2])
# To start iterations
    for (i in 1:50)
    {
# To start the first sequence
            a           <- new.a
            beta        <- new.beta
            miul        <- exposure*exp(as.vector(X%*%beta))
            W           <- diag(miul/(1+a*miul))
            I.inverse   <- solve(t(X)%*%W%*%X)
            k           <- (count-miul)/miul
```

```
            z           <- t(X)%*%W%*%k
            new.beta    <- as.vector(beta+I.inverse%*%z)
            new.miul    <- exposure*exp(as.vector(X%*%new.beta))
# To start the second sequence
            G           <- sum((count-new.miul)^2/(new.miul*(1+a*new.miul)))-
                              (dim(X)[1]-dim(X)[2])
            G.prime     <- -(sum((count-new.miul)^2/(1+a*new.miul)^2))
            new.a       <- a-G/G.prime
    }
# To calculate the variance and standard error
    varians     <- as.vector(diag(I.inverse))
    std.error   <- sqrt(varians)
# To list the programming output
    list (a=new.a, beta=new.beta, std.error=std.error, df=dim(X)[1]-dim(X)[2]-1)
}
```

Appendix B: S-PLUS programming for Generalized Poisson I (moment) multiplicative regression model

```
GP.moment <- function(data)
{
# To identify matrix X, vector count and vector exposure from the data
    X           <- as.matrix(data[, -(1:2)])
    count       <- as.vector(data[, 1])
    exposure    <- as.vector(data[, 2])
# To set initial values for a and beta
    new.a       <- c(0.001)
    new.beta    <- rep(c(0.001), dim(X)[2])
# To start iterations
    for (i in 1:50)
    {
# To start the first sequence
            a           <- new.a
            beta        <- new.beta
            miul        <- exposure*exp(as.vector(X%*%beta))
```

```
        W           <- diag(miul/(1+a*miul)^2)
        I.inverse   <- solve(t(X)%*%W%*%X)
        k           <- (count-miul)/miul
        z           <- t(X)%*%W%*%k
        new.beta    <- as.vector(beta+I.inverse%*%z)
        new.miul    <- exposure*exp(as.vector(X%*%new.beta))
# To start the second sequence
        G           <- sum((count-new.miul)^2/(new.miul*(1+a*new.miul)^2))-
                                                    (dim(X)[1]-dim(X)[2])
        G.prime     <- -(sum(2*(count-new.miul)^2/(1+a*new.miul)^3))
        new.a       <- a-G/G.prime
# To set restrictions for a
          if ((new.a<0)*(new.a<=-1/max(count)))
            new.a <- -1/(max(count)+1)
            else
              if ((new.a<0)*(new.a<=-1/max(new.miul)))
                new.a <- -1/(max(new.miul)+1)
                else
                  if ((new.a<0)*(new.a<=-1/max(count))*
                                        (new.a<=-1/max(new.miul)))
                    new.a <- min(-1/(max(count)+1),-1/(max(new.miul)+1))
                    else
                      new.a <- new.a
    }
# To calculate the variance and standard error
    varians <- as.vector(diag(I.inverse))
    std.error <- sqrt(varians)
# To list the programming output
    list(a=new.a, beta=new.beta, std.error=std.error, df=dim(X)[1]-dim(X)[2]-1)
}
```

Appendix C: Malaysian data

Rating factors					Exposures	Claim counts
Coverage type	Vehicle make	Use-gender	Vehicle year	Location		
Comprehensive	Local	Private-male	0-1 year	Central	4243	381
				North	2567	146
				East	598	44
				South	1281	161
				East Malaysia	219	8
			2-3 year	Central	6926	422
				North	4896	203
				East	1123	41
				South	2865	164
				East Malaysia	679	19
			4-5 year	Central	6286	276
				North	4125	145
				East	1152	29
				South	2675	115
				East Malaysia	700	17
			6+ year	Central	6905	223
				North	5784	150
				East	2156	39
				South	3310	89
				East Malaysia	1406	33
		Private-female	0-1 year	Central	2025	165
				North	1635	55
				East	301	12
				South	608	23
				East Malaysia	126	6
			2-3 year	Central	3661	147
				North	2619	72
				East	527	12
				South	1192	39
				East Malaysia	359	8
			4-5 year	Central	2939	56
				North	1927	36
				East	439	7
				South	959	23
				East Malaysia	376	2
			6+ year	Central	2215	51
				North	1989	38
				East	581	5
				South	937	23
				East Malaysia	589	9
		Business	0-1 year	Central	290	0
				North	66	0
				East	24	0
				South	52	0
				East Malaysia	6	0
			2-3 year	Central	572	0
				North	148	0
				East	40	0
				South	91	0
				East Malaysia	17	0
			4-5 year	Central	487	0
				North	100	0
				East	40	0
				South	59	0

			East Malaysia	22	0
		6+ year	Central	468	0
			North	93	1
			East	33	0
			South	77	0
			East Malaysia	25	0
Foreign	Private-male	0-1 year	Central	1674	94
			North	847	47
			East	377	21
			South	740	38
			East Malaysia	518	6
		2-3 year	Central	3913	202
			North	1930	85
			East	618	21
			South	1768	65
			East Malaysia	833	23
		4-5 year	Central	4002	157
			North	1777	85
			East	534	15
			South	1653	73
			East Malaysia	840	24
		6+ year	Central	6891	245
			North	4409	151
			East	1345	44
			South	2735	113
			East Malaysia	2108	64
	Private-female	0-1 year	Central	1222	29
			North	632	11
			East	209	2
			South	452	17
			East Malaysia	345	6
		2-3 year	Central	2111	46
			North	1068	41
			East	283	5
			South	857	13
			East Malaysia	493	10
		4-5 year	Central	1699	39
			North	793	15
			East	188	0
			South	637	16
			East Malaysia	367	11
		6+ year	Central	1922	47
			North	1376	35
			East	336	6
			South	710	9
			East Malaysia	792	10
	Business	0-1 year	Central	457	0
			North	135	0
			East	70	0
			South	86	0
			East Malaysia	101	0
		2-3 year	Central	1134	0
			North	315	0
			East	113	0
			South	284	0
			East Malaysia	205	0
		4-5 year	Central	1030	0
			North	252	0
			East	70	0

				South	208	0
				East Malaysia	221	0
			6+ year	Central	1075	0
				North	297	0
				East	78	0
				South	231	0
				East Malaysia	282	0
Non-comprehensive	Local	Private-male	0-1 year	Central	8	0
				North	14	0
				East	5	0
				South	8	0
				East Malaysia	3	0
			2-3 year	Central	34	3
				North	65	0
				East	26	0
				South	51	1
				East Malaysia	21	0
			4-5 year	Central	71	1
				North	180	5
				East	47	0
				South	48	1
				East Malaysia	39	0
			6+ year	Central	349	9
				North	496	5
				East	143	2
				South	233	4
				East Malaysia	141	2
		Private-female	0-1 year	Central	2	0
				North	6	0
				East	6	0
				South	3	0
				East Malaysia	3	0
			2-3 year	Central	12	0
				North	23	1
				East	22	0
				South	14	0
				East Malaysia	21	0
			4-5 year	Central	36	0
				North	66	1
				East	19	0
				South	13	0
				East Malaysia	29	0
			6+ year	Central	133	1
				North	213	0
				East	50	0
				South	55	0
				East Malaysia	85	1
		Business	0-1 year	Central	1	0
				North	2	0
				East	0	0
				South	0	0
				East Malaysia	0	0
			2-3 year	Central	1	0
				North	5	0
				East	1	0
				South	1	0
				East Malaysia	1	0
			4-5 year	Central	18	0
				North	8	0

			East	1	0
			South	1	0
			East Malaysia	0	0
		6+ year	Central	57	0
			North	27	0
			East	1	0
			South	133	0
			East Malaysia	3	0
Foreign	Private-male	0-1 year	Central	4	0
			North	11	0
			East	2	0
			South	5	0
			East Malaysia	8	0
		2-3 year	Central	41	0
			North	54	3
			East	7	0
			South	30	2
			East Malaysia	25	0
		4-5 year	Central	68	0
			North	132	3
			East	20	0
			South	55	0
			East Malaysia	48	3
		6+ year	Central	3164	49
			North	3674	71
			East	920	6
			South	2067	56
			East Malaysia	1985	22
	Private-female	0-1 year	Central	2	0
			North	8	0
			East	1	0
			South	3	0
			East Malaysia	6	0
		2-3 year	Central	10	0
			North	47	0
			East	0	0
			South	12	0
			East Malaysia	26	0
		4-5 year	Central	29	0
			North	66	0
			East	2	0
			South	14	0
			East Malaysia	25	0
		6+ year	Central	875	14
			North	1177	15
			East	190	2
			South	411	6
			East Malaysia	555	3
	Business	0-1 year	Central	1	0
			North	1	0
			East	0	0
			South	2	0
			East Malaysia	2	0
		2-3 year	Central	4	0
			North	6	0
			East	0	0
			South	5	0
			East Malaysia	14	0
		4-5 year	Central	17	0

		North	14	0
		East	4	0
		South	7	0
		East Malaysia	20	0
	6+ year	Central	157	0
		North	141	0
		East	22	0
		South	89	0
		East Malaysia	152	0
Total			170,749	5,728

Appendix D: S-PLUS programming for Negative Binomial I (moment) additive regression model

```
NBmoment.add <- function(data)
{
# To identify matrix X, vector count, vector exposure and vector frequency from the data
    X <- as.matrix(data[,-(1:2)])
    count <- as.vector(data[,1])
    exposure <- as.vector(data[,2])
    rate <- count/exposure
# To set initial values for a and beta
    new.beta <- rep(c(0.001), dim(X)[2])
    new.a <- c(0.001)
# To start iterations
    for (i in 1:50)
    {
# To start the first sequence
        beta <- new.beta
        a <- new.a
        fitted <- as.vector(X%*%beta)
        W <- diag(exposure/(fitted*(1+a*exposure*fitted)))
        I.inverse <- solve(t(X)%*%W%*%X)
        k <- rate-fitted
        z <- t(X)%*%W%*%k
        new.beta <- as.vector(beta+I.inverse%*%z)
        new.fitted <- as.vector(X%*%new.beta)
```

```
# To start the second sequence
        G <-   sum((exposure*(rate-new.fitted)^2)/(new.fitted*(1+a*exposure*new.fitted)))-
                    (dim(X)[1]-dim(X)[2])
        G.prime <- -sum((exposure^2*(rate-new.fitted)^2)/(1+a*exposure*new.fitted)^2)
        new.a <- a-G/G.prime
    }
# To calculate the variance and standard error
    varians <- as.vector(diag(I.inverse))
    std.error <- sqrt(varians)
    }
# To list the programming output
    list(a=new.a, beta=new.beta, std.error=std.error, df=dim(X)[1]-dim(X)[2]-1)
}
```

8. REFERENCES

[1] R.A. Bailey, L.J. Simon, "Two Studies in Automobile Insurance Ratemaking", *ASTIN Bulletin*, **1960**, Vol. 4, No. 1, 192-217.

[2] R.A. Bailey, "Insurance Rates with Minimum Bias", *Proceedings of the Casualty Actuarial Society*, **1963**, Vol. 50, No. 93, 4-14.

[3] J. Jung, "On Automobile Insurance Ratemaking", *ASTIN Bulletin*, **1968**, Vol. 5, No. 1, 41-48.

[4] B. Ajne, "A Note on the Multiplicative Ratemaking Model", *ASTIN Bulletin*, **1975**, Vol. 8, No. 2, 144-153.

[5] C. Chamberlain, "Relativity Pricing through Analysis of Variance", *Casualty Actuarial Society Discussion Paper Program*, **1980**, 4-24.

[6] S.M. Coutts, "Motor Insurance Rating, an Actuarial Approach", *Journal of the Institute of Actuaries*, **1984**, Vol. 111, 87-148.

[7] S.E. Harrington, "Estimation and Testing for Functional Form in Pure Premium Regression Models", *ASTIN Bulletin*, **1986**, Vol. 16, 31-43.

[8] R.L., Brown, "Minimum Bias with Generalized Linear Models", *Proceedings of the Casualty Actuarial Society*, **1988**, Vol. 75, No. 143, 187-217.

[9] S.J. Mildenhall, "A Systematic Relationship Between Minimum Bias and Generalized Linear Models", *Proceedings of the Casualty Actuarial Society*, **1999**, Vol.86, No. 164, 93-487.

[10] S. Feldblum, J.E. Brosius, "The Minimum Bias Procedure: A Practitioner's Guide", *Proceedings of the Casualty Actuarial Society*, **2003**, Vol. 90, No. 172, 196-273.

[11] D. Anderson, S. Feldblum, C. Modlin, D. Schirmacher, E. Schirmacher, N. Thandi, "A Practitioner's Guide to Generalized Linear Models", *Casualty Actuarial Society Discussion Paper Program*, **2004**, 1-115.

[12] L. Fu, C.P. Wu, "Generalized Minimum Bias Models", *Casualty Actuarial Society Forum*, **2005**, Winter, 72-121.

[13] N. Ismail, A.A. Jemain, "Bridging Minimum Bias and Maximum Likelihood Methods through Weighted Equation", *Casualty Actuarial Society Forum*, **2005**, Spring, 367-394.

[14] L. Freifelder, "Estimation of Classification Factor Relativities: A Modeling Approach", *Journal of Risk and Insurance*, **1986**, Vol. 53, 135-143.

[15] B. Jee, "A Comparative Analysis of Alternative Pure Premium Models in the Automobile Risk Classification System", *Journal of Risk and Insurance*, **1989**, Vol. 56, 434-459.

[16] K.D. Holler, D. Sommer, G. Trahair, "Something Old, Something New in Classification Ratemaking with a Novel Use of GLMs for Credit Insurance", *Casualty Actuarial Society Forum*, **1999**, Winter, 31-84.
[17] M. Aitkin, D. Anderson, B. Francis, J. Hinde, *Statistical Modelling in GLIM*, **1990**, Oxford University Press, New York.
[18] A.E. Renshaw, "Modelling the Claims Process in the Presence of Covariates", *ASTIN Bulletin*, **1994**, Vol. 24, No. 2, 265-285.
[19] P. McCullagh, J.A. Nelder, *Generalized Linear Models (2nd Edition)*, **1989**, Chapman and Hall, London.
[20] M.J. Brockmann, T.S. Wright, "Statistical Motor Rating: Making Effective Use of Your Data", *Journal of the Institute of Actuaries*, **1992**, Vol. 119, No. 3, 457-543.
[21] J.F. Lawless, "Negative Binomial and Mixed Poisson Regression", *The Canadian Journal of Statistics*, **1987**, Vol. 15, No. 3, 209-225.
[22] A.C. Cameron, P.K. Trivedi, "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests", *Journal of Applied Econometrics*, **1986**; Vol. 1, 29-53.
[23] N.E. Breslow, "Extra-Poisson Variation in Log-linear Models", *Journal of the Royal Statistical Society (Applied Statistics)*, **1984**, Vol. 33, No. 1, 38-44.
[24] J.A. Nelder, Y. Lee, "Likelihood, Quasi-likelihood and Pseudolikelihood: Some Comparisons", *Journal of the Royal Statistical Society (B)*, **1992**, Vol. 54, No. 1, 273-284.
[25] D.R. Cox, "Some Remarks on Overdispersion", *Biometrika*, **1983**, Vol. 70, No. 1, 269-274.
[26] W. Wang, F. Famoye, "Modeling Household Fertility Decisions with Generalized Poisson Regression", *Journal of Population Economics*, **1997**, Vol. 10, 273-283.
[27] P.C. Consul, F. Famoye, "Generalized Poisson Regression Model", *Communication Statistics (Theory & Methodology)*, **1992**, Vol. 2, No. 1, 89-109.
[28] H. Akaike, "Information Theory and An Extension of the Maximum Likelihood Principle", *Second International Symposium on Inference Theory*, **1973**, Akademiai Kiado, Budapest, 267-281.
[29] G. Schwartz, "Estimating the Dimension of a Model", *Annals of Statistics*, **1978**, Vol. 6, 461-464.

Biographies of the Authors

Noriszura Ismail is a lecturer in National University of Malaysia since July 1993, teaching Actuarial Science courses. She received her MSc. (1993) and BSc. (1991) in Actuarial Science from University of Iowa, and is now currently pursuing her PhD in Statistics. She has presented several papers in Seminars, including those locally and in South East Asia, and published several papers in local and Asian Journals.

Abdul Aziz Jemain is an Associate Professor in National University of Malaysia, teaching in Statistics Department since 1982. He received his MSc. (1982) in Medical Statistics from London School of Hygiene and Tropical Medicine and PhD (1989) in Statistics from University of Reading. He has written several articles, including local and international Proceedings and Journals, and co-authors several local books.

"An Exposure Based Approach to Automobile Warranty Ratemaking and Reserving"

John Kerper, FSA, MAAA

Lee Bowron, ACAS, MAAA

ABSTRACT: Existing actuarial techniques for automobile warranty ratemaking and reserving rely heavily on emerging experience (loss development) for the pricing and unearned premium reserving of these products. Since terms for automobile warranties can extend up to 10 years, such data is typically not available or not credible to the degree that the actuary can take great reliance on it. In addition, changing coverage terms in the auto warranty products can often make past development even less meaningful. Exposure techniques that have been developed (Cheng, 1993) rely on overall averages for some critical assumptions instead of distributions or individual policy characteristics.

We propose a "miles-driven" approach in which claims are assumed to arise from auto warranties in proportion to the miles driven times a weight assigned to the overall mileage of the vehicle. The method we employ is much more complex than traditional methods, but relies on data that is typically available at warranty writers. Important data elements would include the mileage of the vehicle at the time of a claim and if the contract cancels. In addition, the underlying manufacturer's warranty is also critical.

In order to provide an accurate model of pricing, a distributional approach is utilized for each policy to model the different driving habits of the policyholders. For example, claim costs can be developed using 5 different driving habits for each policy.

Such a method is very useful for the pricing and premium reserving of new coverages or at start-up companies.

The method proposed utilizes "policy-event based loss estimation methodology" in which a predicted claim cost is derived from each warranty individually.

1. The Continuing Problem of Extended Warranty Coverages

Pricing issues continue to plague the extended warranty industry for vehicles, often known as "vehicle service contracts." Some of these issues are due to the structure of the industry which has historically had a low barrier to entry and a significant number of players with capital constraints. As such, the market can attract inexperienced players that are unaware of the complexities of this insurance product.

Warranties may be written as traditional insurance products, or may be in risk retention groups or captives. In some cases, warranties may not be classified as insurance for regulatory purposes. Regulation of warranty products varies widely and is constantly changing. Due to the fragmented nature of the industry and the variety of forms that warranties may take, it is difficult to compile industry level statistics.

The long warranty period gives rise to a long payout pattern that can mask optimistic pricing and reserving assumptions for several years. Terms for automobile warranties can range up to 10 years. For new car coverages, the effective coverage provided by the warranty over this time period is not uniform. For the first several years, relatively few claims are paid as manufacturer's warranty will cover most claims. As the manufacturer's warranties begin to expire, claims will begin to rise dramatically. Claims also should moderate at the end of the contract as many contract holders will "mile out" of their coverage – that is they will drive the allowed miles before the time has expired. In addition, the policyholder may sell or otherwise dispose of the vehicle without transferring the warranty to the new owner.

In general, this paper will use the term "warranty" which is common in the actuarial industry. However, the term "service contract" is increasingly being used in the industry. For the purposes of this paper, these terms are interchangeable.

2. The Structure of Automobile Extended Warranty Industry

Extended warranty or service contract underwriting is structurally different from other property/casualty products and an understanding of the structure and terminology may be helpful for the actuary who is unfamiliar with the business.

Although there are many different models, a common practice is that the extended warranty is sold at the dealership at the time of purchase of a new or used vehicle. Typically, the consumer may encounter several ancillary products which are sold at the time the vehicle is purchased. These would not only include extended warranties, but also pre-paid maintenance, GAP insurance (which covers the difference between the actual cash value and the loan balance at the time of an insurable event if the vehicle is a total loss), VIN etch, etc. These products are almost always financed with the vehicle. Once an extended warranty has been sold, the amount charged for the warranty will be divided into several components. These include:

- Retail markup (for the auto dealer)
- Agent's commission
- Administrator Fee
- Warranty Reserve

An administrator typically will perform all the processing and servicing of the warranty. An agent will represent the administrator to the dealer clients. The warranty reserve is remitted to an insurance company, which may or may not be owned by the administrator. For the actuary, there are two items of note:

1. The terminology of reserve is misleading because "reserve" in extended warranty typically refers to all funds used to pay claims, not just the outstanding portion, and is more analogous to written premium. For our purposes, we will use the term premium.
2. Since the vast majority of expenses are paid prior to the remittance of funds to the insurance company, the expected loss ratio is higher than other property/casualty products. Often, a book will be priced at an expected loss ratio of 95 to 100 percent. Because these contracts are generally single premium and long term, there is a significant amount of investment income associated with extended warranties.

While this paper only concerns the calculation of expected loss costs for extended warranties, these techniques could also be used by administrators to recognize their fees in proportion to the expected claims from service contracts.

3. Warranty Exposure Bases

In general, exposure bases are measurements for insurers that tell of the relationship that exists between insurable objects and critical conditions where a claim can occur, that note the proportional size of hazard as measured by the losses (magnitude), and that are preferably practical and already in use. This means that exposure bases should have certain qualities, namely, accurate in measure of exposure to loss, easy to determine, and difficult to manipulate.[1]

The purpose of exposure bases is to determine the exposure to loss for an insurer based on the expected loss determined by a series of accepted calculations in order to use the simple and reliable data to develop correct premiums for the insurer and equitably distribute the premiums among the insureds.

For vehicle service contracts, exposure bases are somewhat unique in that the exposure base used to price and rate the coverage (Miles/Time) is not the exposure base that has been commonly used to evaluate the experience (Projected Claim Reporting Pattern).

Deriving an appropriate exposure base for vehicle warranty coverage is a fundamental question when analyzing this line. Fortunately, changing the exposure base in the analysis of the product does not imply changing the exposure base used to market the product.

- Time (Earned Warranty Year) is a poor choice. Warranty claims are not uniform during the policy period. For an extended warranty sold for a new car, the claims pattern will be especially non-uniform, with few claims arising during the initial period that is covered by the manufacturer's warranty. The majority of claims will occur after the manufacturer's warranty expires. In addition, there will be a drop in claims at the end of the warranty as many vehicles exceed the maximum mileage allowed under the warranty or are sold without the transfer of the warranty coverage.
- Indicated Claims Reporting Pattern – This is the most common exposure base used today. This is formulated by developing incremental pure premiums (Cheng, 1993) or simply developing losses by reporting period. This is typically done by loss triangulation. However, instead of aging the claims since the time of the accident, the age of claims are measured from the inception of the policy. This method is appropriate, however, only if:

    1. There is enough data to make these assumptions. While extended warranty achieves credibility at low volumes due to the high frequency/low severity nature of this coverage, there may be limited or no data at the latter points of the coverage being analyzed. If there is no data, common practice is to

[1] See Bouska, 1989

revert back to a benchmark pattern which may not be appropriate for the book being analyzed.

2. The data is homogenous in each cell. This assumption is difficult in that the underlying warranties analyzed may change over time. For example, if the average new car warranty on cars sold five years ago was 36 months but it has now increased to 48 months, the historical pure premium at 60 months will not be predictive of the projected pure premium. In addition, the mix of business may change (European makes typically have higher costs than Asian makes, for example). Another problem is that the coverage offered typically changes due to market conditions.

- Mileage Driven – This is the exposure base proposed in this paper. If mileage is hypothesized as an exposure base, then there is an assumption that claims are basically a function of the number of miles driven by the vehicle. This method is helpful for a number of reasons:

  1. Underlying warranty information is typically available at the individual contract level. Therefore, one could explicitly model the miles driven inside and outside the manufacturer's warranty.
  2. Historical claims information at the end of the contract is not necessary to make an estimate of future claims. Future claims can be modeled as a function of miles driven and the underlying cost per mile. While the claims cost per mile will increase with age, this assumption can also be modeled and tested.

3. A Different Approach

A better approach than loss development for estimating ultimate costs for either pricing or reserving is an exposure based modeling basis, where future losses are modeled for all contracts. This approach has been suggested for modeling other insurance liabilities, such as environmental and asbestos claims (Bouska, 1996). There are several advantages to modeling at the exposure level.

Unlike many insurance products, extended warranty is a high frequency/low severity coverage. It is common for most extended warranties to experience several claims during the life of the warranty. Because of the nature of extended warranty claims, loss data at specific evaluations is credible at relatively low levels, if credibility is defined by the number of claims reported.

The difficulty is estimating the exposure base. This paper proposes an exposure base consisting of the miles driven for the vehicle, so that each mile driven under the warranty is considered an exposure unit.

A miles based exposure base over the term of the contract is closely matched to the actual exposure of the vehicle, as claims can be considered a function of the miles driven during the contract.

One problem with using miles as an exposure base is that there will be some increase in claims per mile during the latter periods of the contract when the frequency of claims will rise due to the age and mileage of the vehicle. This problem can be alleviated by a trend factor, though for newer sets of contracts it will remain a source of uncertainty.

4. A Warranty Pure Premium For a New Vehicle Using a Mileage Function

T_{BASIC} = Maximum term of manufacturer basic (full) warranty in months
T_{PT} = Maximum term of manufacturer power train warranty in months

T_{START} = Age of Vehicle in months (since in service date) at start date of extended warranty
T_{BA_REM} = Remaining term of manufacturer basic (full) warranty in months at start date of extended warranty
= Max $(0, T_{BASIC} - T_{START})$
T_{PT_REM} = Remaining term of manufacturer power train warranty in months at start date of extended warranty
= Max $(0, T_{PT} - T_{START})$
T_{EXT} = Maximum term in Months of extended warranty at start date of extended warranty

M_{BASIC} = Maximum term of manufacturer basic (full) warranty in miles (actual odometer reading)
M_{PT} = Maximum term of manufacturer power train warranty in miles (actual odometer reading)

M_{START} = Actual odometer reading in miles at start date of extended warranty
M_{BA_REM} = Remaining miles of manufacturer basic (full) warranty at start date of extended warranty
= Max $(0, M_{BASIC} - M_{START})$
M_{PT_REM} = Remaining miles of manufacturer power train warranty at start date of extended warranty
= Max $(0, M_{PT} - M_{START})$
M_{EXT} = Maximum term of extended warranty in miles at start date of extended warranty

The following two formulas are based on the assumption that the miles driven for any particular vehicle is proportionate to time and that the number of miles driven per time period for each vehicle, A, is randomly distributed as a lognormal function.

A lognormal may be a reasonable approximation for the distribution of driving habits since it is positively skewed and one can model the "high mileage" drivers in the tail of the function.

$m(t)$	=	miles driven at time t in months
	=	At
$t(m)$	=	time in months at which m miles have been driven
	=	m/A
t_0	=	start of extended warranty
	=	0
t_1	=	time of true expiration of manufacturer basic (full) warranty, measured in months from start of extended warranty
	=	$\text{Min}\ (T_{BA\_REM}, t[M_{BA\_REM}])$
t_2	=	time of true expiration of manufacturer power train warranty, measured in months from start of extended warranty
	=	$\text{Min}\ (T_{PT\_REM}, t[M_{PT\_REM}])$
t_3	=	time of true expiration of extended warranty, measured in months from start of extended warranty
	=	$\text{Min}\ (T_{EXT}, t[M_{EXT}])$
$Cost_{BASIC}$	=	Extended Warranty cost per mile while manufacturer basic (full) warranty is in effect
$Cost_{PT}$	=	Extended Warranty cost per mile after manufacturer basic (full) warranty expires and while manufacturer power train warranty is in effect
$Cost_{EXT}$	=	Extended Warranty cost per mile after both manufacturer basic (full) and power train warranties have expired
$m(t)$	=	mileage driven during extended warranty
$k(t)$	=	trend of repair costs
$p(t)$	=	trend rate of probability of claims and size of the claims as the vehicle ages
Prem	=	Extended warranty pure premium (4.1)

$$= \int_0^{t_1} Cost_{BASIC} * m(t) * k(t) * p(t + T_{START}) * dt$$

$$+ \int_{t_1}^{t_2} Cost_{PT} * m(t) * k(t) * p(t + T_{START}) * dt$$

$$+ \int_{t_2}^{t_3} Cost_{EXT} * m(t) * k(t) * p(t + T_{START}) * dt$$

5. A Simple Example

In the example, we will use a new vehicle for an extended warranty. For a used vehicle, there is typically not an underlying warranty, so a similar analysis can be performed. A "Wrap Coverage" is often sold for vehicles with a long manufacturer's warranty and provides coverage in areas that the manufacturer's warranty excludes. This product can also be modeled using a similar technique.

Assume a contract is sold for a new vehicle for 6 years/72,000 miles for a vehicle with a 3 year/36,000 mile manufacturer's warranty. Assume that the inflation rate is 3% and claims will increase in proportion to the miles driven another 4%. In this example, the driver is assumed to drive 15,000 miles per year.

Warranty Example

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Year	Cumulative Miles	Miles in Manufacturer's Warranty	Cumulative Exposed Miles	Incremental Miles	Trend	Mileage Factor	Adjusted Exposed Miles	Percent Exposure
1	15,000	15,000	0	0	1.000	1.000	-	0%
2	30,000	30,000	0	0	1.030	1.040	-	0%
3	45,000	36,000	9,000	9,000	1.061	1.082	10,332	23%
4	60,000	36,000	24,000	15,000	1.093	1.125	18,444	41%
5	75,000	36,000	36,000	12,000	1.126	1.170	15,809	35%
6	90,000	36,000	36,000	0	1.160	1.217	-	0%
							44,585	100%

Assumptions:

15,000	Miles per Year
72,000	Contract Miles
36,000	Miles for the Manufacturer's Warranty
3.0%	Trend Rate for Repair Costs
4.0%	Mileage Trend

Column 1 represents the cumulative miles driven during the contract.

Column 2 is the cumulative miles covered by the manufacturer.

Column 3 is Column 1 – Column 2, subject to the limitations of the contract. In this example the warranty covers the 36,000 miles between the odometer readings of 36,000 and 72,000.

Column 4 is the incremental miles in Column 3 for each year

Column 5 is an estimate of the increase in repair costs.

Column 6 is an estimate of the rate of increase in claims due to the increased wear-and-tear on the vehicle.

Column 7 is Column 4 x Column 5 x Column 6. This is the adjusted miles.

Column 8 is the percentage of Column 7.

So in this example, we could assume that the earnings pattern should be 23% in Year 3, 41% in Year 4, and 35% in Year 5. Nothing would earn in Year 6 due to contract expiring due to miles. Years 1 and 2 would also earn nothing due to the manufacturer's warranty.

Issues with the Simplified Example

The example above is too simplified to utilize for a couple of reasons.

1. The assumption that no claims occur during the manufacturer's warranty is probably erroneous. Most contracts contain minimal coverage during the warranty period. This can be modeled by assuming the percentage of ultimate claims paid during the manufacturer's warranty.
2. Knowledge of the specific driving habits of a contract holder is unknown. In this example, we have assumed that the driver's mileage exceeds the maximum covered by the warranty in Year 5. That may be true for average driver on the book, but one could expect some earnings in the 6^{th} y ear for drivers who are driving fewer miles than the average for the book.

The next section will more closely examine estimating the average miles driven under Vehicle Service Contracts.

6. Estimating Miles Driven from the Contracts [m(t)]

A mileage function can be estimated from the average miles driven and therefore the percentage of the premium that ought to be earned in each period. One can examine all contracts that had a claim or cancellation (or both) and look at the average miles driven per month as of the last recorded event. This data will typically be available since coverage must be confirmed at the time of a claim and cancellations are typically "pro-rata" as to the greater of miles or time.

Instead of estimating a probability distribution for the mileage driven as shown above, it may be more practical to use a discrete approximation.

For our purposes, we will split the insured vehicles into five equal groups based on average miles driven per year at the time of the claim or cancellation with the arithmetic average calculated for each group. Then factors are calculated for each contract group assuming that claims are proportional to covered miles driven (miles under the contract but not under manufacturer warranty) and that the vehicle for each contract was driven at the respective average yearly rates. The final factor applied is the average of these five factors.

The factors thus derived for a new book of business may overstate earnings because the average miles generally decline as the warranty runs to expiration. This declining pattern is due to two factors - early claims are much more prevalent on cars with the most miles driven per month and as the higher mileage cars use up coverage, the average naturally declines.

Therefore, one can triangulate the data and project to ultimate the average miles driven per year.

For a new book of business, there may not be data available. In this case, the actuary may simply assume a distribution of miles or obtain driving mileage data from an external source.

For this example, the averages for the book have been estimated at the mileage rates below:

Estimated Mileage of Warranties Divided Into 5 Equal Groupings

	Base Average	Minimum Yearly	Maximum Yearly
Group 1	8,400	-	10,200
Group 2	12,000	10,201	13,200
Group 3	14,400	13,201	16,200
Group 4	18,000	16,201	20,400
Group 5	22,800	20,401	

7. A Better Example

Now we will redo the initial example with two changes. First, we will assume that 3% of claims occur during the manufacturer's warranty. Second, we will utilize the "5 bucket" approximation noted above and calculate the exposures for each scenario.

The results are shown below:

(1) (2) (3) (4) (5) (6) (7)

Year	Adjusted Exposed Miles 8,400 per year	Adjusted Exposed Miles 12,000 per year	Adjusted Exposed Miles 14,400 per year	Adjusted Exposed Miles 18,000 per year	Adjusted Exposed Miles 22,800 per year	Exposure Average	Percent Exposure
1	252	360	432	540	684	454	1%
2	252	360	432	540	9,708	2,258	7%
3	252	360	7,200	17,460	22,116	9,478	28%
4	8,148	11,640	13,968	17,460	3,492	10,942	32%
5	8,148	11,640	13,968	-	-	6,751	20%
6	8,148	11,640	-	-	-	3,958	12%
Total	25,200	36,000	36,000	36,000	36,000	33,840	100.00%

8. Developing a Coverage Factor

The use of a "coverage factor" when calculating mileage can be a simplifying assumption. For example, one can calculate the mileage driven inside the manufacturer's warranty, inside the Power Train warranty, and outside the warranty. Claims can be aggregated by examining the mileage on the claim in relation to the underlying warranty.

Calculation of Coverage Factors
(Miles 000)

Warranty	Initial Covered Miles	Reported Losses	Cost per Mile	Coverage Adjustment	Adjusted Covered Miles
Manufacturers	174,831	349,662	0.002	0.071	12,413
Power Train	33,082	496,230	0.015	0.536	17,732
None	324,504	9,086,112	0.028	1.000	324,504

In this example, the cost per mile for each type of warranty is placed in ratio to the cost per mile for claims outside the manufacturer's warranty. Miles inside the warranty are then adjusted downward to reflect the substantially lower claims during this period. In this case the cost per mile during no manufacturer's warranty is 2.8 cents per mile (9,086,122/324,504,000).

9. Estimating the Trend [k(t), p(t)]

As noted above, there are two types of trend that impact the vehicle as the warranty ages:

The first type of trend [k(t)] is the general increase in repair costs. Information concerning repair costs can be estimated from industry repair information or by using the Consumer Price Index (CPI). While repair costs increase due to general inflation, it is important to realize that this trend has been tempered in the past by the increasing reliability of automobiles.

The second type of trend [p(t)] is the increase in costs due to the age of the vehicle. Theoretically, this would be offset by decreasing claims consciousness as the vehicle ages, i.e. a vehicle owner may be more accepting of minor issues as the car ages. In addition, the owner of the vehicle may not know the warranty is in effect. While the warranty can typically be transferred or cancelled for refund by a vehicle owner when the vehicle is sold, there may be some cases where this does not occur.

One could also estimate the two trends simultaneously, since the observed data will have trends due to both the inflationary [k(t)] and aging [p(t)] impact

Using this methodology, there is an assumption that all differences in loss costs between development periods are due to changing costs due to inflation and the aging of the vehicle. Therefore, one should be aware of any changes outside of these factors that would have a significant impact on the loss ratios. These would include:

- Changes in coverages. Administrators may change the coverages offered from time-to-time which can result in different expected loss costs.
- Changes in claims settlement practices. There appears to be significant leeway in how claims are settled. It is common that administrators place more resources in denying or reducing marginal claims when results are above the expected level.

Losses should now be segregated by the time since policy inception, and mileage calculated by the methodology above, also dividing the mileage into periods since policy inception and adjusting the mileage by the coverage levels above.

At this point, one can compare the cost per mile for various ages to calculate the underlying trend for both the aging of the vehicle and the underlying inflation rate. In the example below, used car experience will be used since it is easier to display and more credible at lower mileage levels.

Trend Estimation (Calculation of P(t), K(t))
(Miles 000)

Covered *Miles During Policy Age Months* Undriven

Make	Term	Coverage	Miles	0-12	12-24	24-36	36-48	Miles
European	36	Used	47,520	24,948	7,128	3,564		11,880
American	36	Used	69,863	32,696	12,575	5,030		19,562
Asian	36	Used	74,199	38,346	15,789	2,256		17,808
European	48	Used	38,475	17,006	5,233	2,616	1,308	12,312
American	48	Used	69,925	27,271	9,999	5,454	2,727	24,474
Asian	48	Used	54,667	16,531	6,298	2,624	787	28,427

Make	Term	Coverage	Overall Average Cost per Mile	Cost per Mile in Successive Time Periods 0-12	12-24	24-36	36-48
European	36	Used	0.0414	0.0403	0.0429	0.0458	
American	36	Used	0.0258	0.0250	0.0269	0.0285	
Asian	36	Used	0.0152	0.0149	0.0156	0.0175	
European	48	Used	0.0465	0.0446	0.0473	0.0519	0.0568
American	48	Used	0.0316	0.0304	0.0317	0.0347	0.0374
Asian	48	Used	0.0209	0.0202	0.0212	0.0234	0.0256

Change in Cost per Mile over Time

Make	Term	Coverage	12-24	24-36	36-48
European	36	Used	6.5%	6.8%	
American	36	Used	7.6%	5.9%	
Asian	36	Used	4.7%	12.2%	
European	48	Used	6.1%	9.7%	9.4%
American	48	Used	4.3%	9.5%	7.8%
Asian	48	Used	5.0%	10.4%	9.4%
	Weighted Avg*		5.7%	8.6%	8.5%
	Selected Trend		5.7%	8.6%	8.5%

* Weighted by covered miles

Note in this example the trends for each year range from 5.7% to 8.6%. One must be careful to anticipate that the trend may increase in the outlying years. It might be advisable to simulate different trend levels, especially on the later years, to check the sensitivity of the loss estimate to the trend assumption.

The trend can either be modeled directly into the mileage function (by increasing the estimated miles in proportion to the selected trend) or by directly trending the results. The first method may be more practical when the selected trend varies significantly by product, term, or other variable.

10. Calculating the Future Claims Rate ($Cost_{BASIC}$, $Cost_{PT}$, $Cost_{EXT}$)

As noted above, future claims costs is a function of the expected mileage driven times the cost per mile. The historical cost per mile can be easily calculated by taking the reported claims divided by the historical estimated miles. For future claims, a claims rate should be calculated for each contract based on the characteristics for this contract. Important characteristics one should consider are:

- The type and term of the coverage
- The deductible of the coverage.
- The mileage of the vehicle when the contract was purchased. It is important to segregate contracts from "new" vehicles from "nearly new" vehicles (vehicles with perhaps 1,000 miles on them) because they are typically significant claims differences at this level.
- A general grouping of the vehicle type. Typical groupings are by vehicle national origin (American, European, and Asian) with a couple of sub groupings for each type to differentiate between high cost makes and low cost makes. Certain make groups exhibit different claims characteristics. For example, Asian makes tend to exhibit lower claims costs than North American makes, which in turn exhibit lower claims costs than European makes.
- Other differences that you can model with the available data. For example, some books may have different distribution sources. A common structure is a "Producer Owned Reinsurance Company" where the ultimate liability for covering the claim will be at the servicing dealer. Not surprisingly, these books can exhibit significantly lower clams costs than books with claims paid by a third-party.

In general the actuary should model all available variables and discard those with little relation to claims costs.

In modeling the claims costs, an iterative minimum bias approach is recommended since many variables with have significant correlations. Generalized linear modeling may also provide good results.

Once again, the high frequency/low severity nature of this line will tend to provide more credible relativities at lower loss levels than other property/casualty lines.

11. Cancellations

Future cancellations should also be considered when evaluating a book of business. In general, cancellations will result in a refund of premium equal to the lesser of the proportion of the miles remaining to total miles or time remaining to the total term of the warranty. No consideration of underlying manufacturer's warranty is usually given. For example, if the warranty holder with a 6 year/72,000 mile contract cancels after three years and 50,000 miles, the warranty holder will receive approximately 31% of the premium as a refund ((72000-50000)/72000). This is true even though the majority of the exposure of the warranty remains. In effect, the refund is stated as pro-rata to miles driven or time, but the impact is that of a short-rate cancellation. Therefore, it is generally advantageous for the underwriter of new vehicles for the warranty to be cancelled.

12. Case Reserves and IBNR

Case reserves may or may not be held by an administrator, and are generally not a significant liability compared to the unearned premium reserve. Amounts held for pure incurred but not reported claims are rare since most claims must be pre-approved by the administrator before work can commence. Since the date of loss is typically the date of approval from the administrator, this should eliminate unreported claims except for supplemental payments beyond the initial estimate to repair the vehicle.

If reported losses are used to analyze a book, it should not be necessary to include additional reserves in your estimate. If paid losses are used, the actuary can do a paid loss analysis for a development pattern and add this to observed cost per mile or extend the terms of the contracts by the average delay between claim report and claim payment date.

13. Building the Indicated Rates

Indicated rates should be trended by the inflationary measure [p(t)] from the average accident date on the book until the average accident date of the proposed rates. Assuming terms offered are similar, it is simpler to trend from the effective date of the contract until the effective date of the new rate change. The final indicated loss cost is defined by:

(Reported Losses + Future Claims) x Cost Trend/Number of Warranties

where Future Claims is the Adjusted Mileage (adjusted for trend and coverage factors x Future Claims Rate.

Depending on the situation, other expenses such as taxes, underwriting expense, profit and contingencies, administrator fees, dealer commission and retail markup must be considered. However, some of these items may be either a flat dollar amount or percentage.

14. Conclusion

The methodology proposed in this article is certainly more complex, but should estimate costs better than traditional methodology. Fortunately, the data required to do this type of analysis is typically available from a vehicle service contract database. The unique characteristics of the book (such as term, coverages, and underlying warranties) are explicitly modeled using such an approach.

Because of the high credibility of extended warranty losses, detailed analysis can be done with small and immature books. Indeed, this type of analysis is even more appropriate for such books since a traditional "triangle analysis" will not have enough data for a good estimate.

By explicitly modeling the exposures, the actuary is forced to consider the specific elements such as the trend rate which will have the most impact on the estimate.

15. References

1) Bouska, Amy S. "Exposure Bases Revisited," *Proceedings of the Casualty Actuarial Society*, **1989**, 1-23.

2) Cheng, Joseph S.; Bruce, Stephen J., "A Pricing Model for New Vehicle Extended Warranties," *Casualty Actuarial Society Forum* Year, **1993** Vol: Special Edition Page(s): 1-24

3) Hayne, Roger M. "Extended Service Contracts," *Casualty Actuarial Society*, **1994**, LXXXI, 243-302.

4) Noonan, Simon J. "The Use of Simulation in Addressing Auto Warranty Pricing and Reserving Issues," *Casualty Actuarial Society Forum*, **1993**, Special Ed. 25-52.

5) Weltmann, Jr., L. Nicholas; Mulhonen, David. "Extended Warranty Ratemaking," *Casualty Actuarial Society*, **2001**, Winter, 187-216.

6) Bouska, Amy S. "From Disability to Mega-Risks; Policy Based Loss Estimation," *Casualty Actuarial Society*, **1996**, Summer, 291-320.

Pricing the Hybrid

R. Stephen Pulis, ACAS, MAAA.

Abstract

The current literature describes pricing and reserving of medical malpractice insurance as written on either an occurrence or a claims-made basis. In current practice, many policies allow the reporting of incidents before a claim is submitted, to attach the claim to the current claims-made policy. This creates experience with characteristics of both types of experience. This paper addresses the blend of the two types of experience based on the acceleration of the attachment of claims from their true assertion date back into the claims-made period. The goal is to assign exposure in proportion to expected claims, and to determine the number of claims and the related reserves to expect to be assigned to the current claims-made policy and to the residual tail exposure, and to reflect the change in the final pricing of the policy.

Keywords. Medical malpractice, claims-made, pricing, reserving, Monte Carlo modeling

1. Introduction

Insurance contracts have been bound to provide coverage for events that occurred during the contract period since the inception of insurance. In the ninety years since its inception, the Casualty Actuarial Society has published papers outlining problems and methods to address these concerns in analyzing property/casualty experience for reserving and pricing.

The problems of estimating professional liability costs in the late 1970's led to the emphasis of providing insurance on a "claims-made" coverage basis. The claims-made coverage facilitated the analysis by concentrating on reserving and pricing the events that would be newly reported and deferred the more difficult effort to evaluate future reported claims. The claims-made policy continues to be used extensively for professional liability, and has been adopted for use on other difficult lines such as Directors and Officers Liability.

The occurrence policy attaches responsibility for the claim to the policy in effect when the event giving rise to the claim took place. While this definition seems precise, there has been substantial controversy and litigation over identifying a precise moment of occurrence, especially when a continuous event is taking place. It is not the purpose of this paper to investigate making this assignment, but to recognize that once this definition is accepted, the claim is attached to the occurrence policy in effect on the occurrence date even though it may be reported a substantial amount of time after the occurrence date. Once the claim is reported, a determination is made and the count of the claim and the costs for the claim are assigned back to the "occurrence" period. Tracking the history of the reporting and change in cost estimates provides historical development patterns. For simplicity, assume that these periods are 12-month continuous periods that will be called years.

The assignment of a claim and its associated costs back to the occurrence year means that there will be future changes to be anticipated in the number of claims reported, the costs of the new claims and any revisions in the estimate of the costs on claims previously reported. The estimate of the costs on future reported claims is the "pure" portion of what is normally referred to as IBNR (Incurred But Not Reported).

Under the claims-made policy, the assignment of a claim to the insuring policy is simplified. When the insurer receives notice of a claim, either directly or through its agent (either the insurance agent or the insured acting as a conduit to the insurer), the claim attaches to the policy in effect on that date. While there may be a short delay from the acceptance of the notice until the matter is recorded by the insurer, the "pure" IBNR is zero as all claims are known by the end of the policy term. There will not be any increases in claim counts except for the occasional clerical lag or mishandling. Any development of the case incurred losses will be from adjustments made on known claims, and the general "IBNR" fund for these changes is only for this more limited need. The claims-made insured will have some lingering exposure that will attach subsequent to the expiration of the current policy, and this is referred to as the "tail" of the experience.

Some claims-made policies provide for a claim to be attached to a current policy if the insured gives the insurer notice that an incident has occurred that may result in a claim being asserted in the future. The "assertion" of the claim is the official submission of a request for damages from the claimant to the insured/insurer. The traditional "report date" corresponds to the "assertion date" referred to in this paper. To distinguish from the pure IBNR claims, these reported but not asserted claims will be called "RBNA", and the remaining unknown claims will be the incurred but not known, "IBNK".

Marker and Mohl initially state as Principle #4[1], "Claims-made policies incur no liability for IBNR claims ..." and later state[2] that at the introduction of the claims-made policy, it was "assumed that, on average, claims would be reported sooner" and "that there would be

[1] Joseph Marker and James Mohl, "Rating Claims-Made Insurance Policies", CAS 1980 Discussion Paper Program, page 278.

[2] Joseph O. Marker and F. James Mohl, ibid, page 293.

some additional reporting of incidents that would never have come in under the occurrence policy". This acceleration was viewed as a one-time occurrence at the transition. Their on-going approach did not identify the RBNA component within the claims-made year, and treated its emergence in their backward-recursive development factors. However, an on-going acceleration of claim reporting may adversely affect the adequacy of the renewal premium.

There are pros and cons as to why an insured may give the insurer notice of a potential claim beyond simply providing the insurer additional time to prepare the insurer's defense of the potential claim. If the insured believes that this claim and the aggregate of the other expected claims for this period are within the limits currently purchased, then it is to the insured's benefit to submit an incident report to the insurer during the current policy period. This will maximize the benefit of the coverage already purchased and reduce the future liability under either another claims-made policy, or a "tail" coverage policy. If the insured is switching to self-insurance without purchasing a tail policy, then the reporting of incidences can only reduce potential self-insurance costs. A tort reform change may also simulate a change in the reporting and assertion pattern.

If the frequency of claims or magnitude of a particular claim would exceed current coverage, then there is a disincentive to report the incident until an actual assertion of a claim is received. There is also an incentive for the insured to purchase increased coverage in future policies when there is an increased likelihood of a need for such expanded coverage. When the renewal policy is for limits greater than the expiring limits, an endorsement could be attached that applies the expiring limits to claims reported subsequent to the occurrence year. If the underwriter is really concerned about this possibility, the new policy will not be permitted to have the limit changed.

2. Analysis of the Hybrid

It is the reporting of the incident prior to the claim's assertion that creates a hybrid between the claims-made and occurrence policy. The maximum number of potential claims will be known at the end of the policy, but the number of asserted claims will emerge over time and, therefore, have some characteristics similar to an occurrence policy. Not all potential claims that occur during the policy period will be recognized and reported as an incident within the policy period. The future asserted claims that were not reported as an incident in the first policy period (incurred but not known or IBNK) will attach to a future claims-made policy. The claim count on a hybrid policy will be a blend of:

- Claims that occurred during this period and are asserted during this period.

- Claims asserted during this period that were IBNK at the end of the prior period

- Claims that occurred during this period that are reported but not asserted (RBNA).

- Claims reported during this period that were IBNK at the end of the prior period, but that are not asserted yet (RBNA).

The reserve needed for this period consists of a provision for adjustments on case reserves on asserted claims (the first two types) plus a provision for RBNA claims as of the evaluation date (the last two types). The residual IBNK reserve is a separate issue to be handled as "tail" coverage or in a subsequent claims-made policy

For analysis purposes, some companies may set a subjective reserve and probability of assertion on individual incident reports if there is a substantial likelihood of a future claim with a payment. The hybrid therefore has reserves for development on known case reserves plus reserves on claims reported as incidents but not asserted (RBNA). The subjective reserves are part of the RBNA. They generally are not carried on the books as official reserves, but are used only in the reserve analysis for estimating case incurred development, claim frequency, and claim severity distributions.

In a perfect world, all risks would have experience available and be sufficiently large to be given full credence. If complete information were available, development triangles unique for this business could be calculated and applied. Lacking this, an estimate of the impact using a broad based distribution, and information and assumptions about the particular segment of business are used. This paper assumes complete information is not available and presents an approach to estimate the RBNA reserve. This approach is particularly useful when the pure IBNR is to be modeled using a Monte Carlo simulation.

The method requires knowledge of the claim reporting distribution between the occurrence date and the assertion date. If this distribution is defined in terms of the number of days between these dates, then an assumption should be made, such as "claims occur uniformly throughout the year", and the distribution converted into the portion reported by the end of 12 months, 24 months, etc. Edward Weissner's paper[3], "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood", describes a procedure for estimating the distribution when the final claim reporting is still unknown. For this paper's purpose, Exhibit 1 creates a claims-made reporting pattern using an estimated[4] probability distribution of the number of months between the occurrence and the assertion of a claim. Column (a) is the age, in months, from an occurrence that will produce a claim until the time of its assertion. Column (b) is the probability that the claim will be asserted in that month. Column (c) is the sum of the probabilities that the claim has been asserted by the end of the indicated month. Claims are assumed to occur uniformly throughout the year. An occurrence year would have an equal expected number of claims from each month but with varying ages of maturity. Column (d) calculates the 12-month rolling average of the monthly data by summing Column (c) for the 12 months ending at this age, and divides by 12. Note that if the occurrence year is a partial year (less than 12 months old), the rolling average needs to be adjusted for the period incurred.

Knowledge of the acceleration due to the incident reporting needs to be quantified when analyzing the hybrid. A development triangle of asserted claims by claims-made year can be compared to the distribution described in the previous paragraph. The claims-made distribution by report date defines how claims are assigned to current and future report years, and once assigned there is no development of claim counts. The measured development from the claims-made triangle is all emergence on claims reported as incidents by 12 months. At one extreme, if there is no incident reporting until a claim is asserted, the acceleration is 0% and the reporting distribution is a standard claims-made reporting pattern. At the other extreme, if incident reports are made on every situation inclusive of all claims ultimately asserted, then the acceleration is 100% and the resulting distribution is the same as the reporting pattern for an occurrence policy. Exhibit 2 is a table of the cumulative number of claims asserted for each evaluation of the hybrid year where the claim attaches.

---

[3] Edward W. Weissner, "Estimation of Distribution of Report Lags by the Method of Maximum Likelihood", PCAS LXV, page 1.

[4] The distributions used in this paper have been created to produce realistic results similar to observed data.

Acceleration could also be measured based on the additional change measured on developing losses. Quantifying the additional development resulting from late asserted claims over the case development on claims asserted within the first year, may be difficult to identify and measure. The probability of a severe claim having a higher likelihood of being reported early or late is debatable. Operating on the wrong part of the body or excessive anesthetics can be severe and immediately known damage. Missing a diagnosis, or leaving a foreign object in the body, may take years to recognize and cause irreparable harm or extended pain and suffering. Several large insurers now reflect different reporting patterns by specialty. It has been assumed in this paper that the severity of the claim is independent of the length of time for the claim to be asserted. Therefore, measuring the change in reporting patterns tracks with the associated costs. An adjustment for payment patterns is addresses later in the paper.

If the claim development shows that the number of claims reported at 12 months will ultimately increase by 48%, as in the example on Exhibit 2, then the quantity of 48% times the percentage of claims asserted at 12 months, divided by the percentage unreported at 12 months, gives a measure of the accelerated claim reporting. This calculation can be made at each successive12-month evaluation to determine the accelerated portion reported by that date. It is not obvious that an insured will be better at identifying and reporting an incident that will be asserted in the third year verses being asserted in the fifth year. It may be possible to report incidents occurring near the end of the policy period that are more likely to be asserted in the next 12 months. A uniform acceleration has been used in this paper, and is reflected in Exhibit 3.

A second possible measure of the acceleration can be estimated based on the frequency of the incident reports compared to a standard reporting frequency. If the underlying claim frequency is expected to be the same, then the ratio of incident frequency to asserted frequency is a measure of the acceleration. An adjustment may need to be applied to reflect a probability of less than 100% that all the incidents reported will result in an asserted claim. The initial incident reports should have a much higher probability of predicting an assertion. As the number of incident reports increase, the probability of identifying a future assertion should stay the same or decrease as marginal incidents are added. It is unlikely that an insured will be able to report all incidents that will result in an asserted claim without reporting an excessive number of incidents that will not result in an asserted claim. The example of neglecting to remove a foreign object from the body after an operation, will either be immediately known and treated, or remain unknown until such time that it is discovered and an immediate claim assertion is made. There is little expectation that an occurrence has taken place between those times that would warrant an incident report.

The hybrid year claim distribution, resulting from applying the acceleration, is separated into attachment years (hybrid year) on Exhibit 3. The change in the cumulative assertions (Exhibit 3, line (c)) is the amount of assertions during the calendar year as represented by each column. The probability that an RBNA will be asserted during the current calendar year is the ratio of asserted claims to the RBNA at the end of the prior year.

The cumulative development factor from Exhibit 2 provides a measure of the acceleration as a ratio of the projected the future assertions 0.09093 [=(0.18917)(0.481)] to the unasserted claims at the end of the first year 0.81083. The ratio indicates 11.214% of what would be claims in future claims-made years will now be attached to the current hybrid year. Assuming that the acceleration is uniform, the cumulative portion of occurrence claims attached is the sum of the claims asserted to date plus the acceleration ratio times the portion of claims not asserted as of the evaluation. The calendar year change in the cumulative total is the hybrid year's ultimate portion.

On Exhibit 4, the portion of the occurrence year accelerated and attached within the hybrid year is split into the amount asserted at each subsequent evaluation date, and the portion remaining as RBNA. These are expressed as proportions of the original occurrence-based incurred. Exhibit 4 assigns the ultimate hybrid year total [Exhibit 3, row (i)] to the initial subtotal for the hybrid year on Exhibit 4. The assertions during the calendar year [Exhibit 3 row (c)] correspond to the Total New Assertions at the bottom of Exhibit 4. The assertions during the calendar year are distributed between active hybrid years in proportion to the RBNA existing at the beginning of each the calendar year. Subtracting the asserted claims from the beginning RBNA produces the RBNA at the end of the current calendar year that will also be the RBNA at the beginning of the next calendar year. The probability that a RBNA will be asserted is the ratio of the assertions during the year to the RBNA at the beginning of the year.

Exhibits 4a and 4b provide the same information as Exhibit 4 but Exhibit 4a has 0% acceleration and, therefore, resembles a pure claims-made policy, and Exhibit 4b assumes 100% acceleration and, therefore, resembles an occurrence policy. As the acceleration increases, the tail diminishes as the exposure is shifted back into the prior years.

A multi-year analysis is modeled on Exhibit 5a and 5b. If the insureds are large self-insured hospitals or physician groups written on a claims-made policy. They want to know three things:

1. What is the reserve need at the end of the policy period?
2. What funding is needed for the next year?
3. What residual liability exists beyond next year?

The development on asserted claims can be measured using the standard actuarial techniques; however, care must be used not to include the pure IBNR emergence that is calculated separately. The cost of the unasserted and future claims is essentially a frequency time severity projection: multiplying frequency estimates times the underlying exposure, and multiplying the resulting expected number of asserted claims times an average claim cost amount.

A full-time equivalent exposure (FTE) is calculated as the sum of the product of the unit exposure and the rating relativities: such items as classification, territory, step factor[5], and fractional year exposed. These relativities recognize the variation in costs by medical specialty (classification), tendency for more or larger settlements depending on the location within the state (territory), number of years written under a claims-made policy (step factor), and portion of a year insured (fractional year). The historical claims are adjusted to a closed with payment basis, and developed to an ultimate occurrence basis for use in determining the underlying claim frequency. The historic claim frequency is used to project the ultimate frequency for each period under review. The product of the ultimate frequency and FTE produces the expected number of ultimate claims for each period.

On Exhibit 5a, the hybrid year proportions [Exhibit 3, row (i)] are multiplied times the calendar year exposures to distribute the exposures over the years in proportion to the expected claim assertions. The column can be summed to obtain the hybrid year total. A simplifying assumption could be made that either no exposure growth exists or that a fixed percentage of growth applies over all years. With these assumptions a modified distribution can be derived and applied to only the current calendar year exposure. This has not been done here. The proposed procedure has the benefits of: being sensitive to uneven growth that may arise from such things as general expansion of business or acquisitions; provides

[5] The step factor represents the cumulative percentage of an occurrence year that has been insured.

details of where the expected asserted claims were incurred; and facilitates applying trends and/or discounts related to the time lags. Exhibit 5a displays the allocation of the total exposures in proportion to the expected claim attachment distribution. Exhibit 5b multiplies the exposures times a frequency to project the expected ultimate claims for the occurrence year, and then uses the hybrid year proportions to distribute claims to the hybrid year.

The ultimate claims underlying the three desired quantities are found on Exhibit 5b. The ultimate claims for Hybrid Year 0 and prior are the asserted and RBNAs as of the experience evaluation date (claims in columns (r) thru (w)). The claims enclosed in the box produce the tail exposure at the Year0 year end evaluation. The claims under Year+1 (135 claims) will produce the loss experience to be funded for the next year, and the new tail subsequent to next year (160 claims) will be the losses produced by the claims in Columns (y) through (ad).

The separation of the asserted and the RBNA claims for Year 0 and prior is calculated on Exhibit 6. The ultimate claims on the upper portion of Exhibit 6 were calculated on Exhibit 5b. For each occurrence year, a line is shown with its contribution to the hybrid years in each column. The RBNA is the product of the ultimate occurrence year claims times the RBNA ratio for that assertion year and evaluation lag. The 12 RBNAs for Year0 is the product of 136 ultimate claims [column (b)] time 0.09093 on Exhibit 4 for Year0.

If a change in the acceleration has or is expected to take place, a probability of assertion can be calculated for each hybrid year and evaluation lag. The probability of assertion would be multiplied times the RBNA to determine the number of new asserted claims, and the remaining RBNA count. The probability of assertion may also be adjusted to reflect impacts of tort reform legislation. The cumulative emerged claims equals the ultimate minus the ending RBNAs. The hybrid year count is the total of the column.

The case incurred on known claims can be projected to ultimate using loss development factors if sufficient historical experience is available. However, including the open counts with the RBNA counts provides a mechanism to determine a range around the ultimate losses. Only the claims where a high likelihood that the case incurred is correct are treated as equivalent to a closed claim. The projected RBNA reserve is added to the "closed incurred" to determine the ultimate incurred.

The approached used to project the RBNA reserve is a Monte Carlo model similar to that described by Bickerstaff[6]. The loss dollars on closed claims and the subjective estimates for RBNAs with a high likelihood of payment, are trended to a common date, and fit with distribution curve(s), usually a single or compound log-normal curve(s) to project unlimited losses. A set of simulations (usually 1,000) are run to project first the number of claims based on a claim count distribution (a Poisson distribution is often used) with the expected number of claims as the mean. And second, for each random claim drawn by the Poisson, a random claim size is generated using the mean and variance of the severity-modeled lognormals. The lag between the time the incident is reported to the closing date can be accounted for by trending the (unlimited) severity mean used to generate the claim size.

A loss expense adjustment cost is also generated for each claim. On average, the loss expense increases as the size of the loss increases. Bickerstaff[7] demonstrated the development of a conditional Defense and Cost Containment (DCC) distribution. Its parameters and the generated loss size are used to generate a random DCC for the unlimited loss-size claim. After generating the DCC, the claim-size is limited to the policy provisions. If the policy terms include DCC within the coverage limit, then the combination of loss and DCC is limited and prorated.

The losses and DCC are summed for each sample, and the samples used to calculate the expected value, and the funding needed to meet the desired probability levels of confidence of adequate reserves. An additional loading is added for the reported incidents that are expected to produce loss adjustment expenses, but no indemnity payments.

One factor to consider for the hybrid is whether the paid development will be the same for claims reported and asserted in the first year, compared to claims asserted in future years. One large insurer has developed statistics that show the payout on claims asserted after the occurrence year is longer from occurrence than for claims asserted in the occurrence year, but when comparing the development from the year asserted, the payout is faster on the claims asserted after the year in which the event occurred giving rise to the claim. The speed up is faster during the first year after the assertion, and the differences diminish with age. This introduces a new dimension into determining the discounted value of the reserves.

[6] Dave Bickerstaff, "Hospital Self-Insurance Funding: A Monte Carlo Approach", CAS Forum, Spring 1989 Edition, page 89.

[7] Dave Bickerstaff, "Hospital Self-Insurance Funding: A Monte Carlo Approach", CAS Forum, Spring, 1989 Edition, page 105.

The statewide rate level change is based on comparing the indicated average premium to the current on-level average premium. Medical Malpractice policies generally carry high limits. It is a common practice to limit the analysis (premiums and losses) to a selected lower limit, such as 200/600 or 500/1000, to reduce the parameter variability. The fixed expenses are included in the premium as an Expense Constant added to the variable portion of the premium. The variable portion is the product of a "base rate" multiplied by relativity factors to adjust for territory, classification, time insured by claims-made coverage, and the other credit and debit adjustments. For the remainder of the paper I will use the more common term "claims-made" as inclusive of the "hybrid" coverage unless stated.

The current base rate is a know quantity. The average current relativity is calculated by sequentially applying the current relativity and measuring the average factor resulting from the application of a rating element. Exhibit 7 shows the determination of the average relativity as each rating element is added. The sequential calculation also facilitates measuring changes in relativities; however, none are taking place in this review. The product of the exposure, based on head-count, times the sum of the expense constant plus the base rate times the average factors (Exhibit 8) develops the premium at current rates.

The incurred losses and DCC expenses need to be increased for the Adjusting and Other expenses (AO, formerly known as unallocated loss adjustment expenses (ULAE)). Countrywide experience from the Annual Statement's Schedule P provides incurred Loss, DCC and AO experience. Ratios of the AO to loss plus DCC are calculated (see Exhibit 9) for the last 5 years.. A loading is selected and applied to the state loss plus DCC to determine the ultimate incurred for all loss and loss adjustment expense.

The incurred loss and loss adjustment expense needs to be adjusted to the level expected under the new rates. A pure premium per base class equivalent exposure is calculated on Exhibit 10. Curves are fit by least squares to the average pure premiums for several lengths of time, and the best fit for each time span is shown. An annual trend amount is selected and used to project the historic loss and loss adjustment expense to the mid-point proposed under the new rates.

The expense loadings are separated between variable costs and fixed costs. The General Expenses and Other Acquisition are allocated on a per exposure insured basis to recognize that the costs to write and issue a policy do not materially vary with the location or classification of the risk. For this allocation the actual exposure are divided into the dollars of fixed expenses. The variable expenses are typically dependent on the state where the premium will be charged. The taxes, licenses and fees are dependent on the state laws. The brokerage and commissions are dependent on the contracts that will apply under the new rates. The adjustment for investment income recognizes the investment income on the available funds generated by the cash-flow and prevailing rates of return and taxes.

There are many papers on investment income calculations. This paper will not delve into a particular method, but it should be noted that with the shortened life of a claim under a claims-made policy, the investment income is significantly less than that realized under an occurrence policy. The hybrid policy will realize a return between the occurrence and pure claims-made amounts based on its payout pattern.

The premium from the expense constant will be subject to taxes, commissions, etc. The fixed expenses are loaded for these elements by dividing by the variable expense factor. The premium for fixed expenses is divided by the number of exposures that will be assessed the expense constant. One expense constant will be charged for every exposure, and will only be modified for a shortened policy term.

The statewide rate level indication uses premiums and losses limited to $500,000 per claim/$1,000,000 aggregate basic limit. These losses and loss adjustment expenses are trended to the average loss date under the proposed rates, and divided by the base class equivalent exposures to determine the indicated base pure premium at the future rate level. A base pure premium is selected, and a percentage, say 5%, is added for Death, Disability and Retirement[8]. The result is divided by the variable expense factor to determine the indicated base rate. The indicated average premium is the product of the base rate, the average proposed base class factor (which includes all factors other than the increased limit factor), and the average increased limit factor, and, as the final step, the expense constant is added. Dividing the indicated average premium by the current level average premium produces the indicated change.

This paper does not include revisions being made to the rate relativities, but the off-balance from each is used to adjust the base rate, and maintain the selected overall average premium.

[8] The Death, Disability and Retirement provision is a loading in an on-going business to provide for the average cost of tail coverage on individuals who have ceased to practice through death, disability or retirement.

3. Conclusion

A critical factor in evaluating medical malpractice insurance is to determine the period where claims will attach, and to align the losses and exposures. The claims-made policy provision allowing an insured to report an incident of a potential claim, and thereby attach that claim to a particular policy, creates experience that is a hybrid between a claims-made policy and an occurrence policy. The more aggressively the insured reports incidents in advance of the actual assertion of the claim, the greater the experience will resemble the experience expected under an occurrence policy. The procedure described in this paper facilitates measuring the shift and the calculation of the pure IBNR created for the claims-made policy by the acceleration of the attachment of the claims.

The shift in claims covered from a pure claims-made coverage, increases the pure premium needed, increases the step factors that apply, and increases the investment income. The amount of acceleration allowed determines the degree that the change moves from a pure claims-made basis to an occurrence basis.

4. References

Bickerstaff, Dave, "Hospital Self-Insurance Funding: A Monte Carlo Approach", CAS Forum. Spring 1989 Edition, 89-138.

Marker, Joseph, and James Mohl, "Rating Claims-Made Insurance Policies", CAS 1980 Discussion Paper Program, 265-304.

Weisnner, Edward W., "Estimation of Distribution of Report Lags by the Method of Maximum Likelihood", PCAS LXV, 1-9.

Abbreviations and notations

AO, all other loss adjustment expense
DCC, Defense & Cost Containment expense
FTE, full-time equivalent exposure
IBNK, incurred but not known
IBNR, incurred but not reported
RBNA, reported but not asserted
ULAE, unallocated loss adjustment expense

Biography of the Author

R. Stephen Pulis, ACAS, MAAA, is a consulting actuary at Actuarial Services and Programs in Houston, Texas. He has a Bachelor of Science Degree in Mathematics from Michigan State University. He is a past president of SWAF, and has participated on industry committees, and CAS research.

IRR, ROE, and PVI/PVE

Ira Robbin, PhD, Senior Pricing Actuary, PartnerRe

Abstract:

This paper presents three related measures of the return on a Property-Casualty insurance policy. These measures are based on a hypothetical Single Policy Company model. Accounting rules are applied to project the Income and Equity of the company and the flows of money between the company and its equity investors. These are called Equity Flows. The three measures are: i) the Internal Rate of Return (IRR) on Equity Flows, ii) the Return on Equity (ROE), and iii) the Present Value of Income over the Present Value of Equity (PVI/PVE). The IRR is the yield achieved by an equity investor in the Single Policy Company. The ROE is the Growth Model Calendar Year ROE computed on a book of steadily growing Single Policy business. The PVI/PVE is computed by taking present values of the projected Income and Equity of the Single Policy Company. The paper includes new results relating the PVI/PVE and ROE to the IRR. Beyond developing the foundation and theory of these return measures, the other main goal of the paper is to demonstrate how to use the measures to obtain risk-sensitive prices. To do this, Surplus during each calendar period is set to a theoretically required amount based on the risk of the venture. The main source of risk arises from uncertainty about the amount and timing of subsequent loss payments. With the IRR and PVI/PVE, the indicated prices are those needed to achieve a fixed target return. The indicated price using the Growth Model is that needed to hit the target return at a specified growth rate. With the Growth Model, one can also compute the premium-to-surplus leverage ratio for the Book of Business when it achieves equilibrium. The ability to relate indicated pricing to a leverage ratio, growth rate, and return is an advantage of Growth Model and could lead to greater acceptance of its results. The paper includes sensitivity analysis on the returns and on the indicated profit provisions. In the presentation, the analysis of return is initially done for a single loss scenario. Later, there is discussion on how to model the return when losses are a random variable instead of a single point estimate. Finally, there is a comparison of the approach in this paper versus that of the Discounted Cash Flow model.

Keywords: ROE, IRR, PVI/PVE

1. INTRODUCTION

In this paper, we will present three related ways to measure the return on an insurance policy. The three measures are:

- The Internal Rate of Return on Equity Flows (IRR)

- The Growth Model Calendar Year Return on Equity (ROE)
- The Present Value of Income Over Present Value of Equity (PVI/PVE)

Then we will demonstrate how to use these measures to price Property-Casualty insurance products. We will do this from the perspective of a pricing actuary conducting analysis for a stock insurance company. Whether any of these methods is appropriate in another context is a subject outside the scope of our discussion.

There is nothing novel about using measures of return to price products. The idea is simple enough: any venture with return above a given target hurdle rate is presumably profitable enough to be undertaken. The indicated price for a product can then be defined as the one at which its expected return hits the target. Within the context of internal corporate pricing analysis, corporate management usually sets the target return and a common target is generally used for all insurance ventures.

A significant problem in Property and Casualty insurance pricing applications is that there is no one universally accepted measure of return. The sale of an insurance policy leads to cash flows, underwriting income, investment income, income taxes, and equity commitments that may span several years. How do we distill all this into one number, the return on the policy?

Our three measures are based on two related, but distinct, notions of return on a policy. The first idea is to define return from the perspective of an equity investor who supplies all the capital required to support the policy and who in return receives all the profits it generates. The other idea is to generalize the return achieved by a corporation so that it can be applied to a policy. GAAP ROE (Return on Equity) is a commonly accepted measure of corporate calendar year return. We have two ways to adapt this to a single policy. One is to extend GAAP ROE beyond a single calendar year so that it can handle multi-year ventures. The other is to generate a hypothetical book of business and then measure its ROE. Thus we will end up with three measures of return.

To ensure necessary precision in our analysis, we will define our measures of return by modeling a hypothetical company, the Single Policy Company, which writes a particular policy, the Single Policy. The Single Policy Company writes no other business and is liquidated when the last loss and expense payment is made. Suppose we consider a particular loss scenario and have a model for its anticipated premium, loss, and expense cash flows. We can then apply accounting rules to derive the underwriting income for the Single Policy Company. With other assumptions about investment returns, Statutory Surplus requirements, and taxes, we can derive the company's Investment Income, Income Tax, GAAP After-Tax-Income and GAAP Equity for each accounting period. We will also model a related hypothetical company, the Book of Business Company. This company has a portfolio consisting entirely of Single Policy business. Each period it writes a policy that is a scaled version of the Single Policy. The Book of Business Company begins operations when it writes its first policy and is liquidated after the last loss and expense payment is made on the last policy. We can project the Income Statement and Balance Sheet for the Book of Business Company. Our three profitability measures are defined from the Single Policy and Book of Business Company constructs.

The IRR on Equity Flows is the return that would be achieved by an equity investor in the Single Policy Company. It is a total return measure that reflects the equity requirements, underwriting income, investment income, and taxes associated with the policy by accounting period over time.

PVI/PVE is another measure of profitability based on the Single Policy Company model. It is a generalization of GAAP ROE defined as the ratio of the present value of income valued as of the end of year 1 over the present value of equity. We will show that PVI/PVE will also equal IRR if the present values are computed using a rate equal to the IRR.

Growth Model Return on Equity (ROE) is defined as the Calendar Year ROE that will eventually be achieved by the Book of Business Company if it grows at a constant rate. Under the constant growth assumption, the company will attain an equilibrium in which its Calendar

Year ROE stays constant. We will show that Growth Model ROE equals IRR if the growth rate is also the IRR.

We will derive indicated prices from our return measures. We want these indicated prices to be consistent and sensitive to risk. We also want them to reasonably reflect management's risk-return preferences. To achieve this, we will set Surplus in our model based on a theoretical requirement, and not on an allocation of actual Surplus. Since each of our return measures is sensitive to the effects of leverage, the resulting prices will vary with risk. There are several ways to derive theoretical Surplus requirements and we will not advocate any particular method. We will assume that one has been chosen and that it incorporates any necessary portfolio correlation and order adjustments.

We have said Surplus in our model is a theoretically required amount based on the risk of the venture. But what risk are we talking about? While there is some risk related to the investment of assets, the principal risk in Property and Casualty insurance ventures stems from uncertainty about the timing and amount of loss payments[1]. That is the sole risk we will consider in setting Surplus for our model.

Our initial Surplus is based on the distribution of the present value of ultimate losses. This seemingly innocuous statement has major implications in pricing analysis. For if we vary the premium, we do not change the losses and therefore do not change the amount of surplus. The conclusion is that variations in pricing should lead to variations in the premium-to-surplus

---

[1] Robbin and DeCouto[15] argue that the risk measure should act on the present value of underwriting cash outflow, where underwriting cash outflow is loss plus expense less premium. This allows consistent treatment of swing rating plans and contingent commissions, where the premium or expense may be functions of the loss. We will simplify matters in this discussion and assume premium and expense are not adjusted retrospectively.

ratio. In order to see this, consider an example in which the required surplus is derived from the loss distribution and is equal to $50. Suppose the initial premium is $100, so that the initial premium-to-surplus ratio is 2.00. Now consider the situation when the premium is changed to $110. Since the loss distribution is unchanged while the premium has been increased, the required initial surplus should still suffice[2]. Let us suppose it stays at $50. Even though the required surplus has not changed, the leverage ratio is now 2.20 (2.20= 110/50).

The situation is even more complicated when we consider the duration of surplus commitments. Following our logic one step further, we should set surplus at each point in time based on the risk associated with unpaid losses. Since it may take many years for all loss to be paid on a policy, the surplus will evolve over several years. This underscores the conclusion that when pricing analysis is being conducted the proper way to set surplus is not with a fixed premium-to-surplus ratio. This does not mean that in a different context, such as in solvency regulation or rating agency analysis, that comparisons against fixed premium-to-surplus ratios would not be appropriate.

As a caution we should note that our discussion has not addressed the question of comparability between insurance ventures and alternative non-insurance ventures. Since delving into this larger question would take us too far afield from our main topic, we will not consider it further. Also, we should note that in the modeling examples in this paper, Surplus is set simply as a fixed percentage of the expectation of the present value of unpaid losses. This is done in order to clarify the presentation. In any actual application, this loading percentage should vary with the risk by policy and development age.

---

[2] Robbin and DeCouto [15] discuss two sorts of capital requirements. One is called Level Sensitive and it declines as the premium rate is increased. The other is called Deviation Sensitive and it stays invariant when the premium rate changes. The approach in this paper is equivalent to the Deviation Sensitive approach.

An equivalent, but different, approach to pricing can likely be obtained by using a fixed and common Surplus requirement for all insurance ventures in conjunction with target returns that vary with risk. In order to avoid debate on which approach is better, we will allow that our preference for using a fixed target return on risk-sensitive capital may be largely aesthetic.

The IRR on Equity Flows has already been presented in the Robbin [13] and Feldblum [8] Study Notes. It has also been used in NCCI rate filings. Appel and Butler [1] have previously addressed some criticisms of the IRR approach. The PVI/PVE has also been presented by Robbin [13] and it appears to be equivalent to the NVP Return developed by Bingham [2].

The Growth Model ROE has some connection to previous work done by Roth [16]. In it, he showed how to convert calendar year figures into a true measure of current year return. He also advocated a target return that includes provision for growth as well as the current return needed for shareholders. The Growth Model ROE provides a way to implement these ideas in a pricing context. With it, the actuary can relate indicated pricing with a calendar year ROE, growth rate, and leverage ratio. These are metrics of interest to insurance company executives and could lead to greater acceptance of the results.

Our analysis will also touch on some of the differences between alternative approaches. First it is important to clarify differences between different IRR models. Some authors have discussed an IRR that is an IRR on underwriting cash flows (paid premium less paid loss and paid expense). There has rightly been criticism that this IRR may not even be defined when the flows switch sign more than once. This may not happen frequently in such models, but the counterexamples given by critics are not unduly atypical.[3] However, as we shall later see, it would be very unusual for the Equity Flows we define to change sign more than once. So this criticism generally does not apply to our IRR on Equity Flows.

---

[3] See D'Arcy [5] p525.

Discounted Cash Flow models have many features in common with our three models, but there are important differences. Perhaps most notable is the tautological point that they are focused on underwriting cash flows. As a consequence, they either omit or need to graft on factors such as the accounting treatment of expenses and Surplus requirements. Consider that these methods have no direct way to reflect the conservative treatment of expenses under Statutory Accounting or, equivalently, no direct way to reflect the Deferred Acquisition Balance under GAAP. While some DCF methods do account for taxes on investment income related to Surplus, their results are relatively insensitive to the leverage effects of Surplus. As well, there is no way to study the impact on return from holding discounted loss reserves.

In Section 2, we will present the Single Policy Model. We will use it to define the IRR on Equity Flows in Section 3 and the PVI/PVE Measure in Section 4. In Section 5 we will construct the Book of Business Growth model and define the Growth Model Equilibrium Calendar Year ROE. In Section 6, we will consider modeling returns when the loss can be a random variable instead of a single point estimate. In Section 7, we will study the sensitivity of our return measures to the premium, Surplus level, the interest rate, and the loss payout pattern. We will do this with reserves held at full value or discounted. Then, in Section 8, we will show how to use these measures to derive profit provisions. We will examine the sensitivity of these profit provisions to the Surplus level, the interest rate, and the loss payout pattern. In Section 9 we will compare our approach against the Risk-Adjusted Discounted Cash Flow procedure.

2. THE SINGLE POLICY COMPANY MODEL

Our objective here is to show how to model the accounts of the Single Policy Company based on assumptions about the underwriting results and cash flows of the Single Policy. Our specific goal is to derive the Income and Equity of the Single Policy Company. We will often make simplifying assumptions as this will make it easier to understand the procedure[4]. When modeling actual policies for business analysis, sufficient detail should be incorporated.

An initial assumption we will make is that results are exactly as anticipated. Thus, we will derive a return that is really a return "if all goes just as planned". Later, we will discuss modeling when there is a distribution of possible outcomes.

Before modeling the various income statement, cash flow, and balance sheet accounts, we need to carefully state our indexing conventions. We will use a subscript, j, to denote the value of an income item or cash flow occurring at the end of the j^{th} accounting period. Similarly, a balance sheet account with a subscript, j, denotes its value as of the end of the j^{th} accounting period. We use the subscript, $j=0$, for a cash flow to indicate the flow takes place at policy inception. As well we use the $j=0$ subscript for a balance sheet account to denote its initial value. However, we will assume that income can only be declared at the end of an accounting period so that any income item with a $j=0$ subscript is automatically zero. This is an important assumption. If we were working with an accounting system with some income or loss declared at inception, we would adopt a modified accounting system that would defer that income to the end of the first period and post the appropriate deferred balance as a debit or credit to surplus. To simplify the analysis, we will also assume that no cash flows take place at intermediate times and that the value of a balance sheet account stays constant during a period. This implies the average value of a balance sheet account <u>during</u> the $(j+1)^{st}$ period is equal to its value as of the end of the $(j)^{th}$ period. We will use annual accounting in presenting our model. We will later add a few comments on refining the accounting to a quarterly or monthly basis. Finally, we will assume that the last loss payment is made exactly "n" periods after policy inception and that the Single Policy Company is then liquidated.

---

[4] See Feldblum [8] for a more extensive discussion of modeling details.

As regards accounting conventions, our general approach will be to use Statutory Accounting and make some of the adjustments needed to derive GAAP Income and GAAP Equity. Our Income and Equity will also reflect some simplifications. Nonetheless, unless there is a need to make a distinction, we will refer to our Income and Equity as "GAAP".

With these conventions we define booked underwriting income for the j^{th} accounting period:

(2.1

$$U_j = EP_j - IL_j - IX_j$$
$$\text{for } j = 1, 2, \ldots, n$$

Here U is underwriting gain, EP is earned premium, IL is incurred loss, and IX is incurred underwriting and general expense. The loss includes loss adjustment expense. The incurred loss is calculated on a calendar period accounting basis so that it reflects posted IBNR adjustments as well as case incurred losses. However, the loss reserve is not necessarily held at full value, but could be discounted. In the examples in the Exhibits, we compute expense as the sum of a fixed amount plus a component that varies with premium. We assume the Statutory Incurred Expenses are incurred according to a fixed pattern, while the GAAP Expenses are incurred as premium is earned. The difference between Statutory and GAAP Incurred Expense to date is called the Deferred Acquisition Cost Balance (DAC). To keep matters simple, we ignore policyholder dividends.

Next we turn to the very critical question of how Equity is handled in our model. Our assumption is that Equity will be derived from Statutory Surplus and that the Statutory Surplus will adhere to pre-set requirements. We define S_j as the Required Surplus as of the end of the j^{th} period. In later examples, we will always set Required Surplus as a fixed percentage of the expected discounted unpaid loss. However, for our initial purposes, it is not so important how it is set, as the fact that it is set in advance. We can then derive Q_j, the required GAAP Equity.

We make the simplifying assumption that the only difference between GAAP and Statutory Accounting is in the treatment of initial expenses. Thus, we only need to adjust Q_0 for Deferred Acquisition Costs (DAC). Under this hypothesis we have:

(2.2

$$Q_0 = S_0 + DAC$$
$$Q_j = S_j \text{ for } j = 1,2,\ldots,n$$

Note that $Q_n = 0$ since that is the time when the last loss is paid.

Next we define assets as the sum of Statutory Reserves and Statutory Surplus:

(2.3

$$A_j = UEPR_j + XRSV_j + LRSV_j + S_j$$
$$\text{for } j = 0, 1, 2, \ldots n$$

This equation embodies the fundamental accounting principle that the balance sheet must balance. Here UEPR is the Unearned Premium, XRSV is the Statutory Expense Reserve, LRSV is the Loss Reserve, and S is the Surplus. The Loss Reserve is the calendar period loss reserve, inclusive of IBNR as well as case reserves. We could write a similar equation under GAAP. While the Equity would differ from Statutory Surplus and the expense reserves would be different, the resulting assets would be the same under the simplifying assumptions we have made[5]. Note the basic balance sheet formula is used here to define the assets. In contrast, when evaluating real companies, the assets are given and it is the surplus that is then derived by subtracting the liabilities.

---

[5] As long as there are no GAAP assets such as Goodwill that do not exist in Statutory Accounting, we will have equality between GAAP and Statutory Assets even though the liabilities may differ.

Next we derive invested assets:

(2.4

$$IA_j = A_j - RECV_j$$

$$\text{for } j = 0, 1, 2, \ldots, n$$

In this formula, we use RECV to denote receivables and amounts recoverable.

With invested assets we can compute investment income for each accounting period. Letting "i" denote the risk-free return on invested assets, we have:

(2.5

$$II_j = i \cdot IA_{j-1}$$

$$\text{for } j = 1, 2, \ldots, n$$

We define pre-tax income as the sum of investment income and underwriting income:

(2.6

$$INCPTX_j = U_j + II_j$$

$$\text{for } j = 1, 2, \ldots, n$$

To handle taxes, we define taxable underwriting income, UITX, and taxable investment income IITX by period. We let t_U denote the tax rate on underwriting income and t_I the tax rate on the taxable investment income. We then compute the tax each period via:

(2.7

$$TAX_j = t_U UITX_j + t_I IITX_j$$
$$\text{for } j = 1, 2, \ldots, n$$

Note we are allowing income taxes to be negative. Also note that taxes in our simplified model are paid when the income is declared. A more realistic approach might utilize carry-forwards and carry-backs in the tax calculation. We would also apply the reserve discounting, unearned premium disallowance, and other provisions of the current US tax code. As well, we would model GAAP in more detail by setting up a deferred tax balance to reflect differences between tax basis and accounting basis income. While the model could be made more elaborate and realistic along these lines, we will avoid complications by using our simplified approach in this paper. In any real-world application, the actual tax code should be modeled in detail. A final note on taxes is that in our examples we will simplify matters by using a common tax rate for underwriting and investment income.

Finally, we define after-tax income:

(2.8

$$I_j = INCPTX_j - ITAX_j$$
$$\text{for } j = 1, 2, \ldots, n$$

Now that we have the Income and Equity accounts of the Single Policy Company, we are ready to define the return on the Single Policy.

3. THE IRR ON EQUITY FLOWS

We now define equity flows as the flows of money between an equity investor and a company. The flows of money could be due to the purchase of stock, the payment of dividends, or the repurchase of stock. We suppose the equity flows are given by the reconciliation formula: equity flow equals income less the change in the equity balance[6]. This presumes any capital shortfall will by corrected by using equity capital[7]. Under this definition, flows of investor capital into the company carry a negative sign, while payments from the company to the investors carry a positive sign.

To compute the Equity Flow, F, we add the Income and subtract the increase in the Single Policy Company's Equity:

For j= 0, we set

(3.1

$$F_0 = I_0 - Q_0 = -Q_0$$

For j= 1,2, ..., n, we set:

(3.2

$$F_j = I_j - (Q_j - Q_{j-1}) = I_j - \Delta Q_{j-1}$$

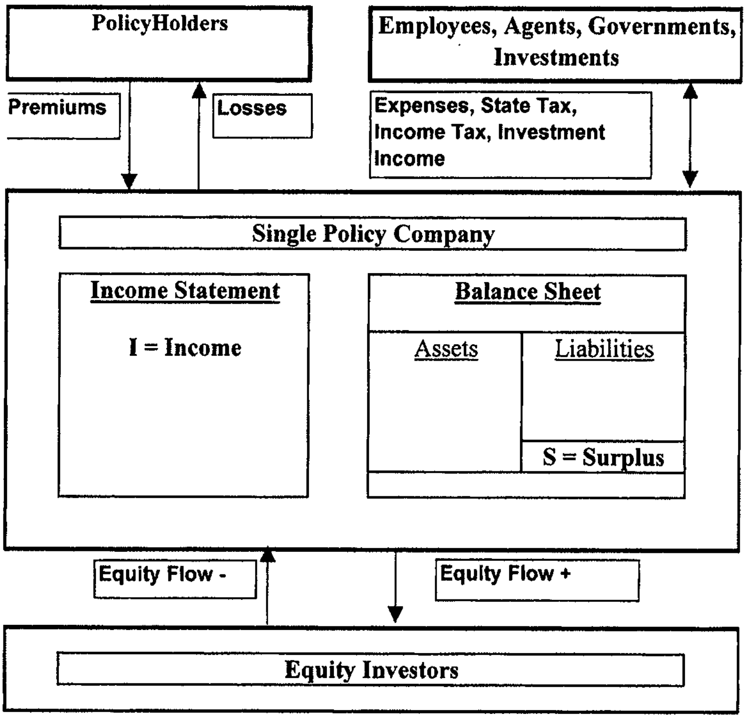
Figure 1 depicts this general construction.

---

[6] This is a simplified version of the formula in Roth[15].

[7] In other words, we will not consider the use of debt and other non-equity capital in meeting the Surplus requirements.

Figure 1



For the insurance applications we are considering, the initial equity flow, F_0, will always be negative. There are two reasons for this. First, the initial commitment of equity needed to fund the Surplus, S_0, contributes the amount $-S_0$, to the initial equity flow. Second, there is a commitment of equity associated with the Deferred Acquisition Cost balance. This is also called the "Equity in the Unearned Premium Reserve". It arises from the conservative treatment of expenses in Statutory Accounting under which acquisition expenses are incurred up-front rather than as the premium is earned.

The IRR on Equity Flows, y, solves the IRR equation:

(3.3

$$\sum_{j=0} F_j \cdot (1+y)^{-j} = \sum_{j=0} F_j \cdot w^j = 0 \quad \text{where } w = (1+y)^{-1}$$

The IRR, if it exists and is unique, is comparable to the interest rate on a loan or the yield rate on a bond. However, since IRR is in general the solution to a nth degree polynomial, there might be multiple real roots. In that case, for each real root, the equity flows can be decomposed into a series of lending and borrowing transactions at the rate of interest equal to that root. For example, if the flows are (–200, +420, -220), the roots are 0% and 10%. With 0%, a loan of 200 is made from A to B and paid back after one year, and then a loan of 220 is made from B to A and it is paid back a year later. The decomposition is: (-200, 420, -220) = (-200, 200, 0) + (0, 220, -220). For the 10% interest rate, the decomposition is (-200,420, -220) = (-200, +220, 0) +(0, 200, -220). This is shown in the following chart.

Figure 2

	Combined Flow from A to B	**Loan From A to B**		**Loan From B to A**	
Time	**FV Flows**	**FV Flows**	**PV @ 10%**	**FV Flows**	**PV @ 10%**
0	-200	-200	-200	0	0
1	420	220	200	200	181.82
2	-220	0	0	-220	-181.82

While multiple roots are a general problem for IRR analysis[8], they do not arise, except in pathological cases, when computing the IRR on the anticipated Equity Flows for a Single Policy. This is because the Equity Flows in our model only switch signs once. As previously noted, the initial Equity Flow is negative due to the up-front commitment of Surplus and the posting of the Deferred Acquisition Cost balance. After that, during the period the premium is earned, the Equity Flows could be negative or positive depending on the amount of underwriting loss and expense in relation to premium and on whether reserves are held at full value or are discounted. Thereafter, the Equity Flows are all positive. This is due to the earning of investment income and the takedown of Surplus[9]. Also, note that anticipated deferred premium payments or salvage and subrogation loss recoveries and other factors that could lead to reversals in the sign of the net underwriting cash flows do not lead to reversals in the sign of the Equity Flows we have defined. This is true because such payments do not impact the booked underwriting gain. With only one sign change in the Equity Flows, there will be only one root to the polynomial equation and the IRR on the Equity Flows will thus be unique.

On the first sheets of Exhibits 2 and 3 are examples showing the accounts of the Single Policy Company for a hypothetical policy. In each case, the resulting equity flows switch signs once and as a result the IRR is unique. Exhibit 2 is the base case. In Exhibit 3 we show results when loss reserves are discounted. Our ability to do this stems from having an underlying corporate structure with balance sheets and income statements. With Discounted Cash Flow models, there is no natural way to model the distinction.

---

[8] Sign reversals are a problem for single policy cash flow analysis as shown in D'Arcy [5].

[9] There is also an implicit assumption that reserves, if discounted, will be discounted at a consistent rate that is less than the anticipated risk-free immunized investment rate. Pathological examples can be constructed by abruptly altering the reserve discount rate from one period to the next. This could lead to reversals in the sign of the Equity Flows.

Two objections that have been raised against IRR are, first, it may not exist due to multiple roots to the IRR equation, and, second, it has an implicit reinvestment assumption at a rate different from the market rate. Appel and Butler [1] have already answered these on general grounds. To eliminate the sign changes that lead to multiple roots, they introduced preferential borrowing and lending rules between a firm and a project under the assumption that "...a transfer of a loan to a future date must be accomplished at the market rate of interest". While we agree with Appel and Butler on general grounds, we do not need such a sweeping argument. We may grant there are general problems with IRR analysis when the flows change sign more than once, but the Equity Flows we are analyzing only experience one sign change. So, for our particular application, that is not an issue.

4. THE PVI/PVE MEASURE

While the IRR on the Equity Flows is an intuitive measure comparable to the interest rate on a loan, we would also like to define a single policy ROE, a measure expressed as the ratio of income over equity. In calendar year accounting it makes perfect sense to take the ratio of income for the year over the initial (or average) equity for the year. However, the Single Policy generates Income over many years and it has Equity requirements that may span more than one year. To summarize the multi-year Income and Equity associated with the Single Policy, we will take present values. The result is a measure of return, PVI/PVE, the ratio of the present value of income over the present value of equity.

Let r_I be the interest rate we will use to discount Income and let r_Q be the interest rate we will use to discount Equity. We set $w_I = (1+r_I)^{-1}$ and $w_Q = (1+r_Q)^{-1}$. Assume the last loss payment for the Single Policy is made at the end of "n" years. Then PVI/PVE is given as:

(4.1

$$\mathrm{PVI/PVE} = \frac{(1+r_I)\cdot\sum_{j=1}^{n} I_j \cdot w_I^{-j}}{\sum_{j=0}^{n-1} Q_j \cdot w_Q^{-j}}$$

Note the formula is effectively discounting income to the end of the first year. This is done to make the definition of return consistent with the usual definitions of ROE and interest rate. In those definitions, income is taken at the end of the year and is not discounted. Note that under our definition a one-year venture has PVI/PVE equal to the interest rate and is independent of the rates used for discounting[10].

We have allowed for possibly different rates to be used for discounting numerator and denominator. However, our favored approach is to discount both at the same rate and we will henceforth assume a common rate is used unless otherwise stated. Also, we believe that in the PVI/PVE context, the appropriate rate for discounting is the cost of capital. We favor the cost of capital over the risk-free rate because the Single Policy Company can borrow at the cost of capital. The thought is that the Single Policy Company could use borrowed money to give its equity investors the PVI/PVE return each year. The income generated by the Single Policy Company in subsequent years would be used to repay the loans. We have previously mentioned a criticism against IRR: that it uses implicit rates of reinvestment at non-market rates of interest. It is hard to raise a similar criticism against PVI/PVE when the discounting is done using the cost of capital. The rate is explicit and it is the market rate for the company.

For a numerical example, suppose the Single Policy has a two-year payout pattern and assume the Single Policy Company will have Equity of 40.0 for year one, and 22.0 for year two.

---

[10] If \$100 is put in a bank account at the start of the year and earns \$10 of interest paid at the end of the year, the return is 10%. The \$10 is not discounted.

Using our indexing notation, we would have $Q_0 = 40.0$, $Q_1 = 22.0$, and $Q_3 = 0$. Now assume income of 5.0 for year one and 4.4 for year two. With our notation, this would translate to $I_0 = 0.0$, $I_1 = 5.0$ and $I_2 = 4.4$. Using a 10.0% rate for discounting, the present value of the income at the end of year one would be 9.0 (5.0 + 4.4/1.1). The present value of the equity would be 60.0 (40.0 + 22.0/1.1). Thus the resulting PVI/PVE would be 15.0% (9.0/60.0).

Next we will show that PVI/PVE is equal to the IRR if the rates for discounting are set equal to the IRR.

(4.2

Result Relating PVI/PVE and IRR: If $r_I = r_Q$ = IRR, then PVI/PVE = IRR.

Proof: Let y = IRR and $w = (1+y)^{-1}$. Then from the IRR Equation we have

(4.3

$$0 = \sum_{j=1}^{n} I_j \cdot w^{-j} - Q_0 - \sum_{j=1}^{n-1} (Q_j - Q_{j-1}) \cdot w^{-j} + Q_{n-1} w^n$$

It follows that:

(4.4

$$\sum_{j=1}^{n} I_j \cdot w^{-j} = Q_0 + (Q_1 - Q_0)w + (Q_2 - Q_1)w^2 + \ldots + (Q_{n-1} - Q_{n-2})w^{n-1} - Q_{n-1}w^n$$

$$= (1-w) \cdot \sum_{j=0}^{n-1} Q_j \cdot w^{-j}$$

Dividing both sides by the present value of the equity, we obtain:

(4.5

$$1 - w = \frac{\sum_{j=1}^{n} I_j \cdot w^{-j}}{\sum_{j=1}^{n} Q_j \cdot w^{-j}}$$

and multiplying by $(1+y)$ leads to the desired result.

This result can be viewed as a way to interpret IRR. Under this interpretation, IRR is a PVI/PVE measure in which the rates for discounting change with the profitability of the policy. Note the idea that these rates should change is antithetical to the PVI/PVE approach. Under the PVI/PVE approach, these rates are, in principle, fixed before modeling the particular result for a policy. In Exhibits 2 and 3 we show the two PVI/PVE that result from use of two different discount rates. The first is based on a common rate of 12.0% and the second is based on a rate equal to the IRR.

Now, suppose we set the target IRR, target PVI/PVE, and the PVI/PVE discounting rates equal to the cost of capital and derive the resulting profit provisions. According to our theory, the two measures will generate identical profit provisions. So in the end, as far as indicated profit provisions are concerned, we arrive at the same answer whether we use IRR or PVI/PVE. In that situation, PVI/PVE does not provide an alternative to IRR, but rather another justification for the validity of an indicated IRR-derived profit provision.

5. BOOK OF BUSINESS GROWTH MODEL

We will construct a book of Single Policy business by writing a policy at the start of each accounting period. Each policy is a scaled version of the Single Policy. By summing contributions from all prior policies we can derive the income statement items, cash flows, and balances for the Book of Business Growth Company. If the scaling factors are generated from a uniform growth rate, we can express the accounts for the Book of Business Company as polynomial functions of the growth rate. We will see that the company goes through a start-up phase during which its reserves, assets, surplus and investment income all increase at a rate higher than the generating growth rate. Eventually, the company reaches an equilibrium growth phase at which point all accounts increase at the generating growth rate. We will measure the calendar period return for the Book of Business Growth Company.

Before we can properly analyze the Book of Business Company, we need to convert our indexing notation from one that refers to timing to one that refers to accounting period. We do this by introducing beginning of period (BOP) and end of period (EOP) suffix notation. The conversion is straightforward. Balance sheet accounts having a subscript, "0", get converted to accounts with a suffix BOP and a subscript "1". In other words, the balance at time t=0 is viewed as the balance for the beginning of period 1. For a balance sheet account, B_t, with time value index, t, strictly larger than zero, we define the ending balance at the end of period "t", $BEOP_t$, to be equal to B_t. Under our assumptions, this is the starting value for the next period, so that we have: $BBOP_{t+1} = B_t$. Also, since we have assumed that income is only declared at the end of time periods, the translation is very easy for income accounts: an account with a timing subscript t becomes an end of period account for period t. Figure 3 provides a simple numerical example of the conversion to accounting period notation.

Figure 3

Single Policy - Timing Notation		
t	Equity Q	Income I
0	40.0	0.0
1	22.0	5.0
2	0.0	4.4

Single Policy - Accounting Notation			
	Equity		Income
Period	QBOP	QEOP	IEOP
1	40.0	22.0	5.0
2	22.0	0.0	4.4

Next, we will extend this notation to the Book of Business Growth Company, by adding a prefix G in front of a Single Policy Company variable. We assume the business premium volume is growing at a fixed rate of growth, g, and that a new scaled version of the Single Policy is added to the Growth Company at the start of each period. We let "n" denote the number of periods till all loss is paid for the Single Policy. We can then translate a Single Policy Balance Sheet account, B, to the corresponding beginning of period and end of period balances for the Book of Business Growth Company using the following formulas:

(5.1

$$GBBOP_k = \sum_{i=0}^{k-1} B_i \cdot (1+g)^{k-1-i}$$

(5.2

$$GBEOP_k = \sum_{i=1}^{k} B_i \cdot (1+g)^{k-i}$$

For example, the Equity at the beginning of year two would be $GQBOY_2 = Q_0(1+g)+Q_1$ and the Equity at the end of year two would be $GQEOY_2 = Q_1(1+g)+Q_2$.

The summations in formulas 5.1 and 5.2 can be readily understood with a policy contribution diagram:

Figure 4

Book of Business with n=2
Balance Sheet Account Growth - Policy Contribution Diagram

	Year 1		Year 2		Year 3		Year 4	
Policy	**BOY**	**EOY**	**BOY**	**EOY**	**BOY**	**EOY**	**BOY**	**EOY**
1	B_0	B_1	B_1					
2			$(1+g)*B_0$	$(1+g)*B_1$	$(1+g)*B_1$			
3					$(1+g)^2*B_0$	$(1+g)^2*B_1$	$(1+g)^2*B_1$	
4							$(1+g)^3*B_0$	$(1+g)^3*B_1$

To provide a numerical example, suppose the Single Policy had Equity balances: $Q_0 = 40.0$ and $Q_1 = 22.0$, and $Q_2 = 0.$. Assume the Growth Company writes the Single Policy at the beginning of year one and writes a 10% larger version of the Single Policy at the start of year two. Using 5.1 and 5.2, the total Equity for the two policies at the beginning of year two would be 66.0 (40.0*1.1+ 22.0). The total Equity would then drop to 24.2 (22.0*1.1) at the end of year two. Using our growth model notation, we would write $GQBOY_1 = 40.0$, $GQEOY_1 = 22.0$, $GQBOY_2 = 66.0$, and $GQEOY_2 = 24.2$.

It is important to note that, even though we have assumed end of period balances for one period are identical to the starting balances for the next period for the Single Policy, the same is not true for the Growth Company. This is true because a new policy is added to the Growth Company portfolio at the start of the next period. The balances from the new policy show up in beginning of period balances for that next period.[11].

---

[11] For example, since a new policy is written on 1/1/(y+1), the unearned premium balance on 12/31/y is different from the unearned premium balance on 1/1/(y+1).

We will next write a formula for Growth Company income statement accounts. However, under our assumptions, the beginning of period income will always be zero. So we only need supply a formula for "end of period" income items:

(5.3

$$\mathrm{GIEOP}_k = \sum_{j=1}^{k} I_j \cdot (1+g)^{k-j}$$

Again, a policy contribution diagram can be useful in understanding the summation:

Figure 5

Book of Business with n=2
Income Account Growth - Policy Contribution Diagram

	Year 1		Year 2		Year 3		Year 4	
Policy	**BOY**	**EOY**	**BOY**	**EOY**	**BOY**	**EOY**	**BOY**	**EOY**
1		I_1		I_2				
2				$(1+g)*I_1$		$(1+g)*I_2$		
3						$(1+g)^2*I_1$		$(1+g)^2*I_2$
4								$(1+g)^3*I_1$

To continue with our numerical example, suppose the Single Policy had income of 5.0 at time t=1 and income of 4.4 at time t=2. Under the Growth Model, this would translate to income of 5.0 at the end of year one and 4.4 at the end of year two. Again supposing a 10% larger version of the policy was written at the start of year two, the total income for the Book of Business Company would be 5.0 at the end of year one and 9.9 (9.9 = 5.0*1.1 + 4.4) at the end of year two. The ROE for year two would be 15.0% (.15 = 9.9/66.0).

Now we consider what happens when the Growth Company has been growing for "n" periods. After that, all income statement and balance sheet accounts will be increasing at the growth rate and we say the business is in the Equilibrium Growth Phase. When this equilibrium has been reached, the formulas can be written as:

(5.4

$$GBBOP_{n+k} = (1+g)^k \sum_{j=0}^{n-1} B_j \cdot (1+g)^{n-1-j} = (1+g)^{k+n-1} \sum_{j=0}^{n-1} B_j \cdot (1+g)^{-j}$$

(5.5

$$GBEOP_{n+k} = (1+g)^k \sum_{j=1}^{n-1} B_j \cdot (1+g)^{n-j} = (1+g)^{k+n-1} \sum_{j=1}^{n-1} B_j \cdot (1+g)^{-(j-1)}$$

So, for example, if n=2, the Equity at the beginning of the fourth year would be given as:

(5.6

$$GQBOY_4 = (1+g)^3 (Q_0 + Q_1 (1+g)^{-1})$$

The Equity at the end of the fourth year would be:

(5.7

$$GQEOY_4 = (1+g)^3 (Q_1)$$

The general formula for income in the k^th^ year of equilibrium is:

(5.8

$$GIEOP_{n+k} = (1+g)^k \sum_{j=1}^{n} I_j \cdot (1+g)^{n-j} = (1+g)^{k+n-1} \sum_{j=1}^{n} I_j \cdot (1+g)^{-(j-1)}$$

We can now compute ROE when the Book of Business Growth Company is in the Equilibrium Growth Phase. Our ROE will be defined as the ratio of end of period Income over beginning of period Equity. For any year in the Equilibrium Growth Phase, the ratio will be:

(5.9

$$\mathrm{ROE} = \frac{\sum_{j=1}^{n} I_j \cdot (1+g)^{-(j-1)}}{\sum_{j=0}^{n-1} Q_j \cdot (1+g)^{-j}}$$

A key observation is that Equilibrium Growth ROE is a function of the growth rate. We are now ready to show that if the growth rate is equal to the IRR on Equity Flows, then the ROE will also equal that IRR.

(5.10

Result Relating IRR and CY Growth ROE: Calendar Year ROE in the Equilibrium Growth phase will equal IRR if the Book of Business is growing at a uniform growth rate equal to the IRR..

Proof. Let g= IRR and set $w = (1+g)^{-1}$. We rewrite the IRR defining equation 2.11 as follows

(5.11

$$\sum_{j=1}^{n} I_j \cdot w^j = Q_0 + \sum_{j=1}^{n-1} (Q_j - Q_{j-1}) \cdot w^j - Q_{n-1} w^n$$

Expanding the right hand side and regrouping, we have

(5.12

$$\sum_{j=1}^{n} I_j \cdot w^j = (1-w)\cdot Q_0 + (1-w)Q_1 w^1 + \ldots$$
$$= \frac{g}{1+g}\sum_{j=0}^{n-1} Q_j \cdot w^j$$

Therefore it follows that:

(5.13

$$\frac{\sum_{j=1}^{n} I_j \cdot w^j}{\sum_{j=0}^{n-1} Q_j \cdot w^j} = \frac{g}{1+g}$$

From that we derive:

(5.14

$$g = \frac{\sum_{j=1}^{n} I_j \cdot (1+g)^{-(j-1)}}{\sum_{j=0}^{n-1} Q_j \cdot (1+g)^{-j}} = \text{ROE}$$

Thus we have proved our desired result.

The reader may note that this proof is essentially the same as the proof for the PVI/PVE result, with the growth rate playing the role of the rate used for discounting. The Growth Model ROE also provides another interpretation of IRR. Consider that once in the Equilibrium Growth Phase the Equity increases from one year to the next by the factor, (1+g). When g equals the IRR, our result says that ROE is equal to the growth rate g. The conclusion is that all the Income is being used to support growth and that the Income generated is all that is needed to support growth at that rate. In other words the end of period Income from one period equals the increase in beginning of period Equity for the next period. So, when we find IRR we are finding the maximal self-sustaining growth rate. It is self-sustaining in the sense that equity investors need supply no more capital once the Equilibrium Growth Phase is reached.

In Exhibits 2 and 3 we show Growth Model accounts for our example. We do this in two stages. First in Sheet 2 of these exhibits, we restate the Single Policy Model accounts using our Beginning of Year (BOY) and End of Year (EOY) accounting conventions. Then, we show growth results in Sheet 3, all at a common growth rate of 5.0%. We compute ROE for each year in the Growth Model. A summary table displays IRR and ROE results. The ROE summary results are for the Equilibrium Growth Phase. In Exhibits 2 and 3, we also have a Sheet 4 that displays accounts where the calculations have been done using a growth rate equal to the IRR. For those scenarios, the ROE equals the IRR, thus demonstrating our theoretical result. For the Sheet 3 scenarios, the two measures are not equal.

If we compare Sheet 3 ROE results by year in Exhibit 2, which is based on full value reserves, versus the comparable ones in Exhibit 3, which is based on discounted reserves, we find that they are nearly identical in equilibrium. However, during the start-up years, the ROE based on discounted reserves is quite a bit higher. This is true even though leverage ratios are unrealistically high in the initial years in both models. Were the leverage ratios reduced in those initial years, the ROEs would decline in both cases. So, in the case when reserves are held full value, the pattern of low ROEs in the initial years rising up to the equilibrium value would be even more pronounced. This leads us by example to a general observation: rapid growth tends to depress ROE, but this can be countered by discounting reserves. Thus, our theory tends to make us more apt to scrutinize the adequacy of reserves and capital in a rapidly growing firm that posts a high ROE and has a heavy concentration in long-tailed lines of business.

We have presented models constructed on an annual basis. It is straightforward to build comparable models on a quarterly or monthly basis, because the accounting rules allow us to do so. Quarterly equity flows can thus be computed and a quarterly effective IRR can be derived from them. PVI/PVE presents a little bit of a problem. Because we have four equity values each year instead of one, our PVE denominator will be roughly four times as large as the PVE from the annual model. On the other hand, the PVI numerator does not necessarily increase or decrease in moving from an annual model to a quarterly one. Two alternatives that have been proposed to deal with this are: i) view the return as a quarterly effective return or ii) annualize the return by dividing the Equity roughly by 4.[12] For ROE we have comparable choices. We could take income for a quarter and divide it by the equity for that quarter. The result would be a quarterly return. The alternative is to take a full year's income and divide it by the average equity for the four quarters. We will not do that in our demonstration. Our point is simply that it is not terribly difficult to extend our models to a quarterly basis. That would allow us to achieve greater accuracy.

6. RETURNS WHEN LOSS IS A RANDOM VARIABLE

We have derived our return measures by modeling results of hypothetical corporations under the assumption all goes as planned. In particular, we have modeled loss as a single point estimate. We now explore how to compute the returns when loss is a random variable. Assume we have a loss distribution consisting of a finite number of loss scenarios and associated probabilities. To be complete, we could also have a more complicated set of scenarios, each consisting of a loss amount and a loss payout pattern. But, for our current work, we will assume it is only the loss amount that varies.

---

[12] See Robbin [13] for a more in-depth discussion of annualization.

Our plan is to model the Income, Surplus, and Equity Flows of each scenario. At first this would seem to be easy. We could just plug the loss amount for each scenario into our model and let it run. However, the problem is a bit harder than that. We can identify at least three major related issues that need to be resolved. The first is whether to let our Single Policy Company go bankrupt in adverse scenarios. The second is the related issue of how to set Surplus. The third is how to model the timing of when the actual ultimate loss is recognized.

We could let our Single Policy go bankrupt in very unprofitable scenarios. The opposite approach is to keep it afloat by implicitly assuming the equity investors will pump in as much money as is needed. This is over and above the initial or planned commitment of Capital. A compromise position is to assume the equity investors post some fixed amount of extra money that could be tapped if needed. The rental of this extra capital should carry a charge. In a setup suggestive of the shared assets paradigm for insurance developed by Mango [10], we could model a Holding Company that would back a portfolio of different Single Policy Company subsidiaries. The Holding Company would assess a "use of extra equity" charge against each Single Policy Company and would be an intermediary between the equity investors and these subsidiary companies. The required segregated Holding Company capital would then depend on the amount of capital in each Single Policy Company subsidiary, the odds each subsidiary would need to draw on Holding Company funds, and the covariance between results of the subsidiaries. While this is conceptually attractive as well as more realistic, it is complicated. We will leave implementation of this approach as a topic for future research. Instead, we will model a company that does not go bankrupt. While this approach has some conceptually debatable underpinnings, it is the easiest to implement. Further, as we will later argue, it provides a conservative estimate of what would result from a more complete model.

In regard to what Surplus requirement should be used, we believe, on theoretical grounds, that all scenarios should start with the same initial Surplus. The reason is simple: at the outset there is no way to know what scenario will ensue. Under our procedure, the initial Surplus would thus be set as a percentage of the expected present value of unpaid losses. The expectation would be taken with respect to all scenarios. After that the situation gets more complicated. As results are posted for the first accounting period, company management may have a better idea than at the start which scenarios are more likely than others. In theory it would then set the Surplus based on its revised estimate of present value of unpaid losses. While this is in some sense realistic as well as conceptually appealing, it is complicated. For our current purposes, we will opt again for the simplest approach and assume a common amount of Surplus at each point in time for all scenarios. The common amount of Surplus would be set at a given point in time as a percentage of the expected present value of unpaid losses. In concept, the percentage would be based on a risk measure operating on the distribution of the present value of unpaid losses. In the examples we use the same percentage for all evaluations.

Now we turn to the question of when to recognize the ultimate loss in a given scenario. Initially, we know only the expected loss over all scenarios. Within any particular scenario, the discrepancy must eventually be recognized on the books of the Single Policy Company. The timing of this recognition will impact underwriting income, loss reserves, investment income, income taxes, and equity flows. Our approach is to recognize the difference at the end of the first accounting period.[13] An alternative is to set reserves equal to the expected ultimate loss times the percent of loss unpaid. The expectation is over all scenarios. Under this approach, the difference between the ultimate loss in the particular scenario and the expected ultimate over all scenarios would be recognized piecemeal as the losses are paid. Various intermediate recognition algorithms could also be used and all the methods could be adjusted to handle reserve discounting. While it is somewhat unrealistic to assume complete recognition of the ultimate loss at the first evaluation, this leads to the simplest algorithm. As well, we will argue that it is the most conservative approach.

Use of our simplest solutions to each of these problems leads to a very convenient modeling result: the average income, average equity, and average equity flow over all scenarios

[13] In a quarterly model, we would recognize one fourth of the difference at the end of each of the first four quarters.

are the same as those resulting when the model is run on the average scenario. In Exhibit 5, we illustrate this with a three-point loss distribution. What this means is that we do not need to separately model all the scenarios to find the returns. Our results for the average scenario will suffice.

An important caveat is that this observation only applies when the premium and expenses are fixed and do not vary with the loss. With Retrospective Rating plans, for example, the premium varies with the loss, and is further subject to Maximum and Minimum Retro Premium restrictions. The average underwriting loss for such a plan does not in general equal the underwriting loss that results from the average loss scenario. So we would need to model the full distribution when dealing with a Retro Plan. However, when complications of that sort are not present, we have found that our simplifying assumptions will allow us to legitimately reduce the distribution of losses to a single scenario.

What have we lost by adopting these simplifications? The answer is that the major factor we are missing is consideration of the default scenarios in which the Single Policy Company fails to meet its obligations to policyholders. We have incorrectly assumed the equity investors would keep the company afloat rather than letting it become insolvent. In effect, we have neglected to put a cap on the downside risk to the equity investors. Because we have not done so, the amounts lost by the investors in adverse scenarios are greater in our model than those that would be indicated in a more sophisticated model. The conclusion is that our model leads to a more conservative average result. In other words, our returns are lower than what they would be if we had modeled the default option. Though our simplified approach would thus be inappropriate for some applications, such as modeling Guarantee Fund assessments, its conservative answers are arguably the answers that are most useful in internal corporate pricing analysis. In that context, the more complete models can exhibit inadequate sensitivity to the tail of the loss distribution. While increasing the relative weight of the tail does increase the risk measure and thus the required Surplus, this is partly offset by the assumption that the equity investors can walk away from the big events. With our simplified approach, there is no walking away and, therefore, no offset. Thus the returns we derive are sensitive to tail events. We feel this is more appropriate in the pricing context of our discussion.

7. SENSITIVITY OF RETURNS

Before going further, it is useful to study how our three measures of return respond to changes in premium, Surplus, interest rate, and payout pattern. We will do this with a simple example. Base case assumptions are shown in Exhibit 1.

The sensitivity of return with respect to premium is of interest when pricing a particular product or policy. Perhaps the return on a product is initially below target at the premium suggested by an agent or broker. Knowing the sensitivity to premium will provide us an intuition about much more premium it will it take to get to the target. Summary premium sensitivity results for our example are shown in Exhibit 4 on Sheets 1 and 2. Reserves are held at full value for Sheet 1 and are discounted in Sheet 2. All Growth Model results assume a 5.0% growth rate and all PVI/PVE results assume discounting at 12.0%. These selections would be appropriate if we suppose that corporate management has targeted a 5.0% growth rate and a 12.0% calendar year ROE. As might be expected, due to the fact that all three models share a common foundation, there is not much difference in the results. Only when returns are negative in the low premium scenario do we see any real difference and even that is fairly modest. In that scenario, the IRR is not quite as negative as the PVI/PVE.

As premiums increase by a constant increment, the returns increase, but in a slightly nonlinear fashion. The IRR goes up at a slightly increasing rate, while the PVI/PVE and ROE rise at a slightly decreasing rate. While a full explanation of the nonlinearities would require detailed analysis, we can at least indicate that our assumptions regarding Deferred Acquisition are part of the explanation as regards PVI/PVE and Growth Model ROE. According to these assumptions, an increase in premium leads to an increase in DAC and thus to an increase in PVE and GAAP Equity in the respective models. The increase in the DAC component of Equity slightly moderates the increase in returns caused by the premium increase. Another consequence of our modeling assumptions is that, counterintuitively, an increase in premium can lead to a reduction in investment income in the second year of the policy. This happens since we have supposed some premium is not paid till the second year. The assets in that year are equal to Reserves plus Surplus and do not change when premium is increased. However the rise in premium boosts the Receivables and thus decreases the investible assets.

Note that the different premium scenarios have different premium-to-surplus leverage ratios. This is in accord with our assumption that the Surplus requirement is driven by the loss distribution. Since all the premium sensitivity scenarios thus have the same amount of Surplus and differing amounts of premium, they end up with different leverage ratios. Another observation is that the change in Equilibrium Growth Model ROE as the result of a change in premium is the same whether reserves are held at full value or are discounted. This makes intuitive sense since the amount of Equity in our model is independent of whether actual reserves are held at full value or are discounted.

Now we examine the sensitivity of our returns to changes in the level of Surplus. This might be of interest when comparing products with different levels of risk. The different levels of risk would translate into different Surplus loading factors for the products. The results for our example are shown in Exhibit 4, Sheet 3. There is nothing surprising: more Surplus produces returns closer to the after-tax yield on investment, no matter which of our return measures is used. However, the sensitivity is perhaps lower than might be guessed in advance. As we increase our loading factor for Surplus so that the Growth Model premium-to-surplus ratio drops from around 3.0 to around 2.0, the returns drop by a bit less than 2 points. The major reason for this is that the after-tax return on investment of the Surplus is fixed and immune to the effects of leverage. So, of the roughly 11.7% returns we get in our low Surplus scenario, nearly 4.0% is achieved on the Surplus itself and only the increment of 7.7% is due to the insurance venture. To get a rough estimate of the Surplus sensitivity in moving from leverage of 3.0 to leverage of 2.0, we would multiply the 7.7% by 2/3 to get 5.1%. The difference of 2.6% is higher than our observed difference of nearly 2.0%, but it suggests that the observed sensitivity is plausible.

We next look at the sensitivity of our returns to changes in the interest rate. As is to be expected, the higher interest rates yield higher returns. They are even a bit higher than one might initially have guessed. This is due to our method of setting Surplus values as a percentage of the present value of unpaid loss. As the interest rate increases, these present values decline. This reduces the amount of Surplus, and so the Growth Model leverage ratios increase.

Finally we turn to examine sensitivity due to changes in the payout pattern. To make the analysis cleaner, we changed our Surplus-loading factor between scenarios so that all scenarios would have the same Growth Model leverage ratio. Implicitly we are assuming that the longer tailed scenarios have lower risk that just offsets the larger commitment of Surplus due to their longer duration. The results are just as expected: longer payout patterns lead to higher returns. The effects are significant. We see that a change in duration of half a year can change the return by over 2 points. This result is sensitive to the interest rate assumption of 6.0% used in our analysis. With a higher rate, we would see even greater sensitivity.

To summarize, the returns exhibit appropriate sensitivities that we can intuitively explain after the fact, even if we did not entirely foresee them beforehand. We should caution that the particular results we have presented are critically dependent on our modeling assumptions. The results would differ if the required Surplus or the Deferred Acquisition balance were computed differently.

8. INDICATED PROFIT PROVISIONS

We define the Indicated Profit Provisions and Indicated Premiums for each of our measures by solving for the profit provision and resulting premium that yields a return equal to the selected target return. Results are shown in Exhibit 5 assuming a target of 12.0%. All results assume reserves are held at full value. Recall that for PVI/PVE we also need to choose a rate for discounting income and equity. We again chose 12.0% under our logic that the cost of capital is a natural target and the natural rate to use for such discounting. However, according to our result relating IRR and PVI/PVE, when the same rate is used for the target and for discounting, we will end up with a PVI/PVE equal to the IRR. Thus our indicated profit provisions for IRR and PVI/PVE are identical. With the Growth Model ROE, we used a growth rate target of 5.0%. If we had used a growth target of 12.0%, results for ROE would have also been the same as for IRR. However, we have no logic that compels such a choice. Rather, we have assumed that management has specified a long-term growth target of 5.0% and a target calendar year return of 12.0%.

In Sheet 1, we examine sensitivity of the Indicated Profit Provisions to changes in the level of Surplus. We change the level of Surplus by varying the Surplus-loading factor. As we would anticipate, higher Surplus loading factors give rise to higher profit provisions. However, the leverage ratios do not follow a direct inverse relation with the loading factors. The divergence arises because the premium is also changing between scenarios. As shown in Exhibit 5, the ROE profit provision moves from –1.97% to –0.13% in response to a change in Surplus loading factors that reduces the Growth Model leverage ratio from 3.09 to 2.15.

Next we examine sensitivity of indicated premiums to a changes interest rates while keeping the target return fixed. Results are shown in Sheet 2. Raising the interest rate leads to a reduction in the profit provision. This is in accord with our intuition. With more investment income we need less underwriting income to achieve the target. The IRR and ROE results are similar, but not identical. With our loss payout pattern duration of only 2.0 years, moving the interest rate up one point reduces the indicated profit provision by a bit less than 2.0 points. The result also depends on our Surplus-loading factor. With a higher loading factor, we could drive sensitivity down. The results can also be explained by noting that interest rates impact the leverage ratio in our model. On the one hand, increasing the interest rate reduces the present value of unpaid loss. That reduces the Surplus. On the other hand, higher interest rates reduce the indicated premium, assuming the target return stays fixed. This happens because they reduce the difference between that target return and the after-tax investment return as well as increase the investment income on our full value reserves. The net tradeoff between the reduction in Surplus and the reduction in Premium as seen in our results is that the leverage ratios decrease modestly with an increase in the interest rate.

Finally, we turn to sensitivity analysis of the indicated profit provisions with respect to changes in the loss payout pattern. Results are shown in Sheet 3. To facilitate comparisons, we adjust our loading factors for Surplus in order to achieve a constant Equilibrium Growth Model leverage ratio in all scenarios. We see, as expected, that the results show significant response to the duration of the payout pattern. Increasing the duration by half a year moves the profit provision down by just over 2.0 points when the interest rate is 6.0%.

To summarize, despite a few subtleties, the models produce Indicated Premiums that are appropriately responsive to changes in key inputs. Next, we will compare our corporate structure approach with the Risk-Adjusted Cash Flow Model.

9. COMPARISON TO THE RISK-ADJUSTED DISCOUNTED CASH FLOW MODEL

The Risk-Adjusted Discounted Cash Flow Model (RA DCF) has often been used in pricing. However, it takes a different approach to pricing than the one we have taken. Instead of finding the Indicated Premium needed to hit a fixed target return on Risk-sensitive Surplus, the RA DCF approach is to find the Fair Premium directly. The Fair Premium is defined as the sum of loss, expense, and income tax cost components. Each component is discounted. However, since losses are a risky cash flow, they are discounted at a risk-adjusted rate.

In words, the formula is

(9.1

Fair Premium =
PV of Loss at the Risk - Adjusted Rate + PV of Expense
+ PV of Tax on Investment Income on Surplus and Premium net of Expense
+ PV of Tax on Underwriting Income from Premium less Expense
- PV of Tax Reduction for Losses at the Risk - Adjusted Rate

For a single period example, we can write the formula in mathematical symbols as follows:

(9.2

$$P = \frac{L}{(1+r_A)} + \frac{X}{(1+r_f)} + \frac{T_I \cdot r_f \cdot (P - X + S)}{(1+r_f)} + \frac{T_U \cdot (P - X)}{(1+r_f)} - \frac{T_U \cdot L}{(1+r_A)}$$

Here P stands for premium, L is loss, and X is expense. The losses are discounted at a risk adjusted rate, r_A, which is less than or equal to the risk –free rate, r_f. The tax rate on investment income is T_I and the tax rate on underwriting income is T_U. Here S stands for Surplus. Note that the Fair Premium includes a provision for the tax on the investment income from both the Surplus and the balance of underwriting cash flows.

The risk-adjusted rate is a key parameter in the RA DCF model. As D'Arcy and Dyer [6] note, determination of this rate is a "thorny issue"[14]. They describe two approaches. One is to view the adjustment "as a form of compensation to the insurer for placing its capital at risk in the insurance contract"[15]. The second is to derive the risk-adjustment from principles of the Capital Asset Pricing Model (CAPM). This is the approach used by Myers and Cohn [12] in

[14] D'Arcy and Dyer [6], p.342.

[15] D'Arcy and Dyer [6], p.342.

their original paper introducing the model. Under CAPM, there should be no charge for process risk, only for systematic risk related to the covariance of insurance losses with returns on the stock market. This covariance is known as "beta. The determination of beta has been the subject of some disagreement. Some believe beta is close to zero. For example, Vaughn [17] notes:: " For many P/L lines, indemnity losses possess very little systematic risk. As such, the risk-free rate is often used as an acceptable approximation ..."[16]. However, Derrig [7] and others have used a non-zero, CAPM-based beta in rate filings.

This short introduction to the RA DCF model is necessarily incomplete, but it will suffice to allow us to reasonably compare that model against the procedure we have presented. The most obvious distinction is that the RA DCF is a method to determine premium without need to assume a target return. In our models, the Indicated Premium is that needed to achieve a given target return (or target return at a given target growth rate for the Growth Model).

The next major distinction is that the RA DCF model has no underlying corporate or accounting structure, while such a framework is the basis for defining our returns. Because of this, the RA DCF has no natural way to reflect the conservative treatment of expenses under Statutory Accounting. In our corporate model, this was handled by making an adjustment to GAAP Equity for Deferred Acquisition Costs. As well, there is no natural way in the RA DCF framework to reflect reserve discounting. While reserve discounting does not impact underwriting cash flows, it does impact the flow of funds to equity investors. Our corporate model of Equity Flows takes this into account.[17]

---

[16] Vaughn [17], p. 406

[17] Another anomaly caused by lack of an accounting substructure is that the balance of investible assets does not automatically decay to zero. However, since it usually decays to a positive or negative balance close to zero and the RADCF provision is for the present value of taxes on the investment income on the balance, the practical impact of the non-disappearing balance is usually negligible.

The next point of distinction concerns the role of Surplus. In the RA DCF, it plays no direct major role. There is a provision in the Fair Premium for the present value of the tax on the investment of the Surplus, but this is usually small. Consider a one-year example assuming a 3.0 leverage ratio, 6.0% interest rate, and a 35% tax rate. The full value tax in that case would come to around 0.69%. Not only is the effect small, the sensitivity to changes in Surplus is even smaller. Reducing the leverage ratio to 2.0 in our example produces a full value tax of 1.05%. The difference of 0.36% is significantly smaller than the 1.84% difference (-0.13% -(-1.97%)) seen in our Growth Model ROE results. Further, if the tax rate were zero, the Fair Premium would be independent of Surplus. In contrast, in our models the leverage effect of Surplus has a critical impact on the results. It is revealing that in some RA DCF models[18], Surplus is assumed to be larger than the amount needed to ensure that there is essentially no chance of insolvency. This view of Surplus is effectively tantamount to regarding it as a "free" good; there is more than enough of it to go around. However, in the corporate context of our models, Surplus is in scarce supply.

Another major difference between the models concerns their sensitivity to risk. As we previously noted, risk sensitivity in the RA DCF model depends on how beta is selected. Yet, that selection is problematic. If we follow Vaughn and use no risk-adjustment, RA DCF pricing would have no sensitivity to risk. Since we believe pricing ought to be risk-sensitive, we would disagree with this implementation of the RADCF: it is an RADCF without the "RA". If we follow others who use CAPM to derive a non-zero beta, we would have some risk sensitivity. However, those methods have typically been applied at a line of business level for the industry. It is not obvious how to extend them to pricing different products within a line for a single company.

Finally, we could follow those who set the beta so as to provide an adequate return on risk-sensitive capital. In that case, we would look to our approach to arrive at the Indicated

---

[18] See Vaughn [17].

Premium and solve for the beta that leads to the same answer. While the presentation of that result as a RA DCF calculation might be useful in some situations, it forces us to think about risk sensitivity in terms of changes in beta. Within our framework, risk sensitivity depends on the Surplus requirement formula and the spread between the target return and the after-tax yield on investment. We believe actuaries and insurance company management find it more intuitive to think in those terms. Further, though there are disagreements about how to set theoretical Surplus, they are not as severe as the disagreements over beta.

Ultimately we feel the methods arise in different contexts and reflect different perspectives in pricing. Others have noted these differences[19]. Management, we believe, will be far less interested in knowing the Fair Premium for a product than it will be in knowing the Indicated Premium needed to attain its risk-return objectives. One the other hand, as the title of the Myers and Cohn paper [12], " A Discounted Cash Flow Approach to Property-Liability Insurance Rate Regulation" makes clear, that model was originally developed to handle pricing in a regulatory arena. From a policyholder or regulatory perspective, there may be much greater concern with finding the Fair Premium than knowing whether the premium is adequate for shareholders to achieve the expected return they desire. While the Fair Premium may contain some compensation for the equity investors of the insurance company, those investors may or may not find that compensation acceptable.

One other issue that must be clarified is that there are discounting methods, such as the one developed by Butsic [4], in which the losses are discounted at a risk-adjusted rate, yet which are closer to our method than to the RA DCF approach. In Butsic's model, the rate adjustment depends explicitly on the equity requirement and a given target return. Butsic sets the equity requirement as a percentage of the discounted loss reserve. He also computes an IRR that is conceptually the same as our IRR on Equity Flows. He finds the premium needed to hit a given target return. What Butsic shows is that if reserves are discounted at just the right rate,

---

[19] See Bingham [2].

then the ROE for each year is equal to the IRR and the target return. His rate for discounting losses is given as:

(9.3

$$r_A = i - e(R - i)$$

Here i is the risk-free rate, R is the target return, and e is the equity loading factor relative to the discounted reserve.

What Butsic has done is to show how to modify the accounting system to bring it into accord with economic reality so the anticipated calendar year returns each year would be the same as the IRR. If we were to discount reserves in our model according to Butsic's formula, we would obtain the same results.

10. CONCLUSION

We have covered many topics and now it is time to summarize what has been accomplished. The first step in our journey was to define our three measures based on a hypothetical corporate structure. Looking back we can see that this structure enforced a certain discipline in our analysis. We had to be precise about the amount of Surplus being held and about the flows of money to and from equity investors. The structure allowed us to reflect the impact of the DAC adjustment in GAAP and the effect of reserve discounting. Having a corporate structure that incorporates accounting rules is a critical aspect of our approach. Further we can conclude that models without sufficient corporate structure cannot fully capture key aspects of the return on an insurance venture, at least not the return to an equity investor or to the insurance company.

We proved results relating PVI/PVE and ROE to IRR and used these to provide new interpretations of IRR. We found that, with some simplifying assumptions, we could conveniently use a single average loss scenario to obtain the average return when the loss is a random variable. We then argued that these simplifying assumptions led to a conservative answer that was appropriate in the internal corporate context of our pricing analysis. With examples, we explored the sensitivity of our returns to changes in premium adequacy, Surplus level, interest rate, and payout pattern.

Our examination of the sensitivity of indicated profit provisions showed that these models should lead to reasonably responsive risk-sensitive prices for insurance products. The risk-sensitive pricing was obtained by using risk-sensitive Surplus requirements in conjunction with a fixed target return.

We have seen the Growth Model ROE emerge as a very strong contender to the IRR on Equity Flows. While there was not much of a difference in the results, the Growth Model allows us to directly relate product pricing to long-term calendar year ROE and growth rate targets. It also produces a calendar year premium-to-surplus leverage ratio for the Book of Business in equilibrium. This could be compared against industry benchmarks.

We have discussed why results from our models would differ from those of others such as the Risk-Adjusted Discounted Cash Flow model. This was done in an attempt to increase understanding. While some of our comments could be taken as critical, we have not gone so far as to say there is anything inappropriate about using other approaches in other contexts. In some regulatory situations, it may well be better to use the RA DCF model than any of the three we have presented.

There already is a significant body of literature on other ways of pricing in general[20] and on other ways of pricing insurance products in particular[21]. However, we feel we have demonstrated a methodology for deriving indicated prices that should be appropriate for internal corporate pricing analysis. We believe each of our three measures of return could reasonably be used in that context. Methods similar to ours are in common use and we hope our work furthers their acceptance. In conclusion, while we have left some theoretical questions unresolved and frequently adopted simplifying assumptions, we believe we have nonetheless demonstrated three variants of an approach to pricing that is both sound and practical.

---

[20] For example, the Black-Scholes formula for pricing options does not use a target return.

[21] See D'Arcy and Dyer [6], Derrig [7], and Robbin [13] for various alternative approaches to pricing property and casualty insurance products.

References

[1] David Appel and Richard Butler, "Internal Rate of Return Criteria in Ratemaking", NCCI Digest, Vol. IV Issue III, September 1990.

[2] Russel E. Bingham, "Rate of Return – Policyholder, Company, and Shareholder Perspectives", Proceedings of the CAS, 1993, Vol. LXXX, pp 110-147.

[3] Russell E. Bingham, "Cash Flow Models in Ratemaking: A Reformulation of Myers-Cohn NPV and IRR Models for Equivalency", *Actuarial Considerations Regarding Risk and Return in Property-Casualty Insurance Pricing*, Edited by Oakley E. Van Slyke, 1999, pp27-60.

[4] Robert Butsic, "Determining the Proper Interest Rate for Loss Reserve Discounting: An Economic Approach", *Evaluating Insurance Company Liabilities*, CAS Discussion Paper Program, 1988, pp 147-186.

[5] Stephen P. D'Arcy, "Investment Issues in Property-Liability Insurance", Foundations of Casualty Actuarial Science, 2nd Edition, Chapter 8, CAS, 1990, pp 485-534.

[6] Stephen P. D'Arcy and Michael A. Dyer, "Ratemaking: A Financial Economics Approach", Proceedings of the CAS, 1997, Vol. LXXXIV, pp 301-389.

[7] Richard A. Derrig, " Investment Income, Underwriting Profit and Contingencies: Financial Models", CAS Forum, Fall 1988, pp 133-176.

[8] Sholom Feldblum, "Pricing Insurance Policies: The Internal Rate of Return Model", CAS Study Note

[9] J. Robert Ferrari, "The Relationship of Underwriting, Investment, Leverage, and Exposure to Total Return on Owners' Equity", Proceeding of the CAS, 1968, Vol. LV, pp 295-302.

[10] Don Mango, "The Shared Asset Approach to Capital Allocation", ASTIN Bulletin, Vol. 35, Issue 2, November, 2005, pp 471-486.

[11] Charles L. McClenahan, "Insurance Profitability", *Actuarial Considerations Regarding Risk and Return in Property-Casualty Insurance Pricing*, CAS, 1999, Chapter 8.

[12] Stewart Myers and Richard A. Cohn, " A Discounted Cash Flow Approach to Property-Liability Insurance Rate Regulation", *Fair Rate of Return in Property-Liability Insurance*, 1987, pp 55-78.

[13] Ira Robbin, "The Underwriting Profit Provision", CAS Study Note as updated in 1992.

[14] Ira Robbin, "Theoretical Premiums for Property and Casualty Insurance Coverage – A Risk-Sensitive, Total Return Approach", *Actuarial Considerations Regarding Risk and Return in Property –Casualty Insurance Pricing*, Edited by Oakley E. Van Slyke, 1999, pp 105-112.

[15] Ira Robbin and Jesse DeCouto, "Coherent Capital for ROE Pricing Calculations", CAS Forum, Spring 2005.

[16] Richard Roth, "Analysis of Surplus and Rate of Return Without Using Leverage Ratios", *Insurer Financial Solvency*, CAS Discussion Paper Program, 1992, Volume I, pp 439-464.

[17] Trent Vaughn, "Misapplications of Internal Rate of Return Models in Property/Liability Insurance Ratemaking", CAS Forum, Winter 2002, pp 341-416.

[18] Michael A. Wacek, "Discussion of 'Ratemaking: A Financial Economics Approach' by D'Arcy and Dyer", Proceedings of the CAS, 2004, Vol. XCI, pp 42-59.

Acknowledgement

The author gratefully acknowledges the contributions of C. W. (Walt) Stewart. It was Walt who first introduced the author to the key ideas that are the foundation of this paper.

Glossary of Exhibits

Exhibit 1		Assumptions
Exhibit 2	Sheet 1	Single Policy Model -Base Case Premium and FV Reserve
Exhibit 2	Sheet 2	Single Policy Model -Base Case Premium and FV Reserve-BOY/EOY Accounting
Exhibit 2	Sheet 3	Book of Business Company-Base Case Premium and FV Reserve-Target Growth
Exhibit 2	Sheet 4	Book of Business Company-Base Case Premium and FV Reserve-Growth at IRR
Exhibit 3	Sheet 1	Single Policy Model -Base Case Premium and Discounted Reserve
Exhibit 3	Sheet 2	Single Policy Model -Base Case Premium and Discounted Reserve-BOY/EOY Accounting
Exhibit 3	Sheet 3	Book of Business Company-Base Case Premium and Discounted Reserve-Target Growth
Exhibit 3	Sheet 4	Book of Business Company-Base Case Premium and Discounted Reserve-Growth at IRR
Exhibit 4	Sheet 1	Sensitivity of Return Measures to the Premium Amount When Reserves are Full Value
Exhibit 5	Sheet 2	Sensitivity of Return Measures to the Premium Amount When Reserves are Discounted
Exhibit 4	Sheet 3	Sensitivity of Return Measures to the Surplus Level
Exhibit 4	Sheet 4	Sensitivity of Return Measures to the Interest Rate
Exhibit 4	Sheet 5	Sensitivity of Return Measures to the Payout Pattern
Exhibit 5	Sheet 1	Sensitivity of Indicated Premiums to the Surplus Level
Exhibit 5	Sheet 2	Sensitivity of Indicated Premiums to the Interest Rate
Exhibit 5	Sheet 3	Sensitivity of Indicated Premiums to the Payout Pattern
Exhibit 6	Sheet 1	Summary of Results by Loss Scenario
Exhibit 6	Sheet 2	Single Policy Company Accounts Under Loss Scenario 1
Exhibit 6	Sheet 3	Single Policy Company Accounts Under Loss Scenario 2
Exhibit 6	Sheet 4	Single Policy Company Accounts Under Loss Scenario 3

Exhibit 1

Profit Measure Examples
Assumptions

Rates	
Investment Return	6.00%
Tax Rate	35.00%
PVI/PVE Discount Rate Selection	12.00%
Growth Rate Target	5.00%

Surplus Requirements	
Ratio to PV Unpaid Loss	31.5%
Rate for Discounting Unpaid Loss	6.00%

Patterns

Earning and Incurral

Year	Earned Premium	Full Value Incurred Loss	Stat Incurred Expense	GAAP Incurred Expense
0	0.0%	0.0%	60.0%	0.0%
1	100.0%	100.0%	40.0%	100.0%
2	0.0%	0.0%	0.0%	0.0%
3	0.0%	0.0%	0.0%	0.0%
4	0.0%	0.0%	0.0%	0.0%

Payment Patterns

Year	Paid Premium	Paid Loss	Paid Expense
0	75.0%	0.0%	30.0%
1	20.0%	25.0%	45.0%
2	5.0%	50.0%	20.0%
3	0.0%	25.0%	5.0%
4	0.0%	0.0%	0.0%
Total	100.0%	100.0%	100.0%
PV Factor (t=0)	0.9832	0.8908	0.9445

Underwriting	Loss	Expense
Fixed	72.00	10.00
Variable	0.0%	20.0%

Exhibit 2
Sheet 1

Single Policy Company

UW Assumptions			Financial Assumptions		IRR and PVI/PVE Results		
	Amount	Ratio	Interest Rate	6.00%	IRR	10.74%	10.74%
Premium	100.0	100.0%	Tax Rate	35.00%	PVI/PVE Discount Rate	12.00%	10.74%
Loss	72.0	72.0%	Rsv Discount Rate	0.00%	PVI	6.05	6.10
Expense	30.0	30.0%	S as % of PV Unpaid Loss	31.50%	PVE	56.52	56.78
Combined	102.0	102.0%	PV Loss Discount for S Calc	6.00%	PVI/PVE	10.71%	10.74%

Year	Earned Premium	Incurred Loss	Stat Incurred Expense	Stat UW Income	Paid Premium	Paid Loss	Paid Expense	UW Cash Flow
0	0.0	0.0	18.0	-18.0	75.0	0.0	9.0	66.0
1	100.0	72.0	12.0	16.0	20.0	18.0	13.5	-11.5
2	0.0	0.0	0.0	0.0	5.0	36.0	6.0	-37.0
3	0.0	0.0	0.0	0.0	0.0	18.0	1.5	-19.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	100.0	72.0	30.0	-2.0	100.0	72.0	30.0	-2.0

Year	Unearned Premium	Loss Reserve	PV Unpaid Loss	Stat Expense Reserve	Total Stat Reserves	Surplus	Assets	Receivables	Invested Assets	Inv Income
0	100.0	0.0	64.1	9.0	109.0	20.2	129.2	25.0	104.2	
1	0.0	54.0	50.0	7.5	61.5	15.7	77.2	5.0	72.2	6.3
2	0.0	18.0	17.0	1.5	19.5	5.3	24.8	0.0	24.8	4.3
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	DAC	GAAP Equity	GAAP Incurred Expense	GAAP UW Income	GAAP Pre-tax Income	Income Tax	Income	Change in Equity	Equity Flow
0	18.0	38.2	0.0	0.0	0.0	0.0	0.0	38.2	-38.2
1	0.0	15.7	30.0	-2.0	4.3	1.5	2.8	-22.5	25.2
2	0.0	5.3	0.0	0.0	4.3	1.5	2.8	-10.4	13.2
3	0.0	0.0	0.0	0.0	1.5	0.5	1.0	-5.3	6.3
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total			30.0	-2.0	10.1	3.5	6.6	0.0	6.6

Exhibit 2
Sheet 2

Single Policy Company- BOY and EOY Accounting

UW Assumptions			Financial Assumptions		IRR and PVI/PVE Results		
	Amount	Ratio	Interest Rate	6.00%	IRR	10.74%	10.74%
Premium	100.0	100.0%	Tax Rate	35.00%	PVI/PVE Discount Rate	12.00%	10.74%
Loss	72.0	72.0%	Rsv Discount Rate	0.00%	PVI	6.05	6.10
Expense	30.0	30.0%	S as % of PV Unpaid Loss	31.50%	PVE	56.52	56.78
Combined	102.0	102.0%	PV Loss Discount for S Calc	6.00%	PVI/PVE	10.71%	10.74%

Year	Earned Premium EOY	Incurred Loss EOY	GAAP Incurred Expense EOY	GAAP UW Income EOY	Paid Premium BOY	Paid Premium EOY	Paid Loss EOY	UW Cash Flow BOY	UW Cash Flow EOY
1	100.0	72.0	30.0	-2.0	75.0	20.0	18.0	66.0	-11.5
2	0.0	0.0	0.0	0.0	0.0	5.0	36.0	0.0	-37.0
3	0.0	0.0	0.0	0.0	0.0	0.0	18.0	0.0	-19.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	Unearned Premium BOY	Unearned Premium EOY	Loss Reserve BOY	Loss Reserve EOY	Stat Expense Reserve BOY	Stat Expense Reserve EOY	Total Stat Reserves BOY	Total Stat Reserves EOY
1	100.0	0.0	0.0	54.0	9.0	7.5	109.0	61.5
2	0.0	0.0	54.0	18.0	7.5	1.5	61.5	19.5
3	0.0	0.0	18.0	0.0	1.5	0.0	19.5	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	Surplus BOY	Surplus EOY	Assets BOY	Assets EOY	Receivables BOY	Receivables EOY	Invested Assets BOY	Investment Income EOY
1	20.2	15.7	129.2	77.2	25.0	5.0	104.2	6.3
2	15.7	5.3	77.2	24.8	5.0	0.0	72.2	4.3
3	5.3	0.0	24.8	0.0	0.0	0.0	24.8	1.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	DAC BOY	DAC EOY	GAAP Equity BOY	GAAP Equity EOY	GAAP Pre-tax Income EOY	Income Tax EOY	GAAP Income EOY
1	18.0	0.0	38.2	15.7	4.3	1.5	2.8
2	0.0	0.0	15.7	5.3	4.3	1.5	2.8
3	0.0	0.0	5.3	0.0	1.5	0.5	1.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Exhibit 2
Sheet 3

Book of Business Growth Company

UW Assumptions	Amount	Ratio	Financial Assumptions		IRR and ROE Results	
Premium	100.0	100.0%	Interest Rate	6.00%	IRR	10.74%
Loss	72.0	72.0%	Tax Rate	35.00%	EQ Growth ROE	10.90%
Expense	30.0	30.0%	Rsv Discount Rate	0.00%	EQ Growth P/S	2.50
Combined	102.0	102.0%	S as % of PV Unpaid Loss	31.50%	Growth Rate	5.00%
			PV Loss Discount for S Calc	6.00%		

Year	Earned Premium EOY	Incurred Loss EOY	GAAP Incurred Expense EOY	GAAP UW Income EOY	Paid Premium BOY	Paid Premium EOY	Paid Loss EOY	UW Cash Flow BOY	UW Cash Flow EOY
1	100.0	72.0	30.0	-2.0	75.0	20.0	18.0	66.0	-11.5
2	105.0	75.6	31.5	-2.1	78.8	26.0	54.9	69.3	-49.1
3	110.3	79.4	33.1	-2.2	82.7	27.3	75.6	72.8	-71.0
4	115.8	83.3	34.7	-2.3	86.8	28.7	79.4	76.4	-74.6

Year	Unearned Premium BOY	Unearned Premium EOY	Loss Reserve BOY	Loss Reserve EOY	Stat Expense Reserve BOY	Stat Expense Reserve EOY	Total Stat Reserves BOY	Total Stat Reserves EOY
1	100.0	0.0	0.0	54.0	9.0	7.5	109.0	61.5
2	105.0	0.0	54.0	74.7	17.0	9.4	176.0	84.1
3	110.3	0.0	74.7	78.4	19.3	9.8	204.2	88.3
4	115.8	0.0	78.4	82.4	20.3	10.3	214.5	92.7

Year	Surplus BOY	Surplus EOY	Assets BOY	Assets EOY	Receivables BOY	Receivables EOY	Invested Assets BOY	Investment Income EOY	P/S
1	20.2	15.7	129.2	77.2	25.0	5.0	104.2	6.3	4.95
2	37.0	21.9	212.9	106.0	31.3	5.3	181.7	10.9	2.84
3	44.2	23.0	248.4	111.3	32.8	5.5	215.6	12.9	2.50
4	46.4	24.1	260.8	116.8	34.5	5.8	226.4	13.6	2.50

Year	DAC BOY	DAC EOY	GAAP Equity BOY	GAAP Equity EOY	GAAP Pre-tax Income EOY	Income Tax EOY	GAAP Income EOY	GAAP ROE
1	18.0	0.0	38.2	15.7	4.3	1.5	2.8	7.23%
2	18.9	0.0	55.9	21.9	8.8	3.1	5.7	10.24%
3	19.8	0.0	64.0	23.0	10.7	3.8	7.0	10.90%
4	20.8	0.0	67.2	24.1	11.3	3.9	7.3	10.90%

Exhibit 2
Sheet 4

Book of Business Growth Company

UW Assumptions			Financial Assumptions		IRR and ROE Results	
	Amount	Ratio	Interest Rate	6.00%	IRR	10.74%
Premium	100.0	100.0%	Tax Rate	35.00%	EQ Growth ROE	10.74%
Loss	72.0	72.0%	Rsv Discount Rate	0.00%	EQ Growth P/S	2.58
Expense	30.0	30.0%	S as % of PV Unpaid Loss	31.50%	Growth Rate	10.74%
Combined	102.0	102.0%	PV Loss Discount for S Calc	6.00%		

Year	Earned Premium EOY	Incurred Loss EOY	GAAP Incurred Expense EOY	GAAP UW Income EOY	Paid Premium BOY	Paid Premium EOY	Paid Loss EOY	UW Cash Flow BOY	UW Cash Flow EOY
1	100.0	72.0	30.0	-2.0	75.0	20.0	18.0	66.0	-11.5
2	110.7	79.7	33.2	-2.2	83.1	27.1	55.9	73.1	-49.7
3	122.6	88.3	36.8	-2.5	92.0	30.1	79.9	80.9	-74.6
4	135.8	97.8	40.7	-2.7	101.9	33.3	88.5	89.6	-82.6

Year	Unearned Premium BOY	Unearned Premium EOY	Loss Reserve BOY	Loss Reserve EOY	Stat Expense Reserve BOY	Stat Expense Reserve EOY	Total Stat Reserves BOY	Total Stat Reserves EOY
1	100.0	0.0	0.0	54.0	9.0	7.5	109.0	61.5
2	110.7	0.0	54.0	77.8	17.5	9.8	182.2	87.6
3	122.6	0.0	77.8	86.2	20.8	10.9	221.3	97.0
4	135.8	0.0	86.2	95.4	23.1	12.0	245.0	107.4

Year	Surplus BOY	Surplus EOY	Assets BOY	Assets EOY	Receivables BOY	Receivables EOY	Invested Assets BOY	Investment Income EOY	P/S
1	20.2	15.7	129.2	77.2	25.0	5.0	104.2	6.3	4.95
2	38.1	22.8	220.3	110.4	32.7	5.5	187.6	11.3	2.91
3	47.6	25.2	268.8	122.2	36.2	6.1	232.6	14.0	2.58
4	52.7	27.9	297.7	135.4	40.1	6.8	257.6	15.5	2.58

Year	DAC BOY	DAC EOY	GAAP Equity BOY	GAAP Equity EOY	GAAP Pre-tax Income EOY	Income Tax EOY	GAAP Income EOY	GAAP ROE
1	18.0	0.0	38.2	15.7	4.3	1.5	2.8	7.23%
2	19.9	0.0	58.0	22.8	9.0	3.2	5.9	10.13%
3	22.1	0.0	69.6	25.2	11.5	4.0	7.5	10.74%
4	24.4	0.0	77.1	27.9	12.7	4.5	8.3	10.74%

Exhibit 3
Sheet 1

Single Policy Company

UW Assumptions	Amount	Ratio	Financial Assumptions		IRR and PVI/PVE Results		
Premium	100.0	100.0%	Interest Rate	6.00%	IRR	10.99%	10.99%
Loss	72.0	72.0%	Tax Rate	35.00%	PVI/PVE Discount Rate	12.00%	10.99%
Expense	30.0	30.0%	Rsv Discount Rate	6.00%	PVI	6.22	6.23
Combined	102.0	102.0%	S as % of PV Unpaid Loss	31.50%	PVE	56.52	56.73
			PV Loss Discount for S Calc	6.00%	PVI/PVE	11.01%	10.99%

Year	Earned Premium	Incurred Loss	Stat Incurred Expense	Stat UW Income	Paid Premium	Paid Loss	Paid Expense	UW Cash Flow	
0	0.0	0.0	18.0	-18.0	75.0	0.0	9.0	66.0	
1	100.0	68.0	12.0	20.0	20.0	18.0	13.5	-11.5	
2	0.0	3.0	0.0	-3.0	5.0	36.0	6.0	-37.0	
3	0.0	1.0	0.0	-1.0	0.0	18.0	1.5	-19.5	
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
Total	100.0	72.0	30.0	-2.0	100.0	72.0	30.0	-2.0	

Year	Unearned Premium	Loss Reserve	PV Unpaid Loss	Stat Expense Reserve	Total Stat Reserves	Surplus	Assets	Receivables	Invested Assets	Inv Income
0	100.0	0.0	64.1	9.0	109.0	20.2	129.2	25.0	104.2	
1	0.0	50.0	50.0	7.5	57.5	15.7	73.2	5.0	68.2	6.3
2	0.0	17.0	17.0	1.5	18.5	5.3	23.8	0.0	23.8	4.1
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	DAC	GAAP Equity	GAAP Incurred Expense	GAAP UW Income	GAAP Pre-tax Income	Income Tax	Income	Change in Equity	Equity Flow	
0	18.0	38.2	0.0	0.0	0.0	0.0	0.0	38.2	-38.2	
1	0.0	15.7	30.0	-2.0	8.3	2.9	5.4	-22.5	27.8	
2	0.0	5.3	0.0	-3.0	1.1	0.4	0.7	-10.4	11.1	
3	0.0	0.0	0.0	-1.0	0.4	0.1	0.3	-5.3	5.6	
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
Total			30.0	-2.0	9.8	3.4	6.4	0.0	6.4	

Exhibit 3
Sheet 2

Single Policy Company- BOY and EOY Accounting

UW Assumptions	Amount	Ratio	Financial Assumptions		IRR and PVI/PVE Results		
Premium	100.0	100.0%	Interest Rate	6.00%	IRR	10.99%	10.99%
Loss	72.0	72.0%	Tax Rate	35.00%	PVI/PVE Discount Rate	12.00%	10.99%
Expense	30.0	30.0%	Rsv Discount Rate	6.00%	PVI	6.22	6.23
Combined	102.0	102.0%	S as % of PV Unpaid Loss	31.50%	PVE	56.52	56.73
			PV Loss Discount for S Calc	6.00%	PVI/PVE	11.01%	10.99%

Year	Earned Premium EOY	Incurred Loss EOY	GAAP Incurred Expense EOY	GAAP UW Income EOY	Paid Premium BOY	Paid Premium EOY	Paid Loss EOY	UW Cash Flow BOY	UW Cash Flow EOY
1	100.0	68.0	30.0	2.0	75.0	20.0	18.0	66.0	-11.5
2	0.0	3.0	0.0	-3.0	0.0	5.0	36.0	0.0	-37.0
3	0.0	1.0	0.0	-1.0	0.0	0.0	18.0	0.0	-19.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	Unearned Premium BOY	Unearned Premium EOY	Loss Reserve BOY	Loss Reserve EOY	Stat Expense Reserve BOY	Stat Expense Reserve EOY	Total Stat Reserves BOY	Total Stat Reserves EOY
1	100.0	0.0	0.0	50.0	9.0	7.5	109.0	57.5
2	0.0	0.0	50.0	17.0	7.5	1.5	57.5	18.5
3	0.0	0.0	17.0	0.0	1.5	0.0	18.5	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	Surplus BOY	Surplus EOY	Assets BOY	Assets EOY	Receivables BOY	Receivables EOY	Invested Assets BOY	Investment Income EOY
1	20.2	15.7	129.2	73.2	25.0	5.0	104.2	6.3
2	15.7	5.3	73.2	23.8	5.0	0.0	68.2	4.1
3	5.3	0.0	23.8	0.0	0.0	0.0	23.8	1.4
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	DAC BOY	DAC EOY	GAAP Equity BOY	GAAP Equity EOY	GAAP Pre-tax Income EOY	Income Tax EOY	GAAP Income EOY
1	18.0	0.0	38.2	15.7	8.3	2.9	5.4
2	0.0	0.0	15.7	5.3	1.1	0.4	0.7
3	0.0	0.0	5.3	0.0	0.4	0.1	0.3
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Exhibit 3
Sheet 3

Book of Business Growth Company

UW Assumptions	Amount	Ratio	Financial Assumptions		IRR and ROE Results	
Premium	100.0	100.0%	Interest Rate	6.00%	IRR	10.99%
Loss	72.0	72.0%	Tax Rate	35.00%	EQ Growth ROE	10.85%
Expense	30.0	30.0%	Rsv Discount Rate	6.00%	EQ Growth P/S	2.50
Combined	102.0	102.0%	S as % of PV Unpaid Loss	31.50%	Growth Rate	5.00%
			PV Loss Discount for S Calc	6.00%		

Year	Earned Premium EOY	Incurred Loss EOY	GAAP Incurred Expense EOY	GAAP UW Income EOY	Paid Premium BOY	Paid Premium EOY	Paid Loss EOY	UW Cash Flow BOY	UW Cash Flow EOY
1	100.0	68.0	30.0	2.0	75.0	20.0	18.0	66.0	-11.5
2	105.0	74.4	31.5	-0.9	78.8	26.0	54.9	69.3	-49.1
3	110.3	79.1	33.1	-1.9	82.7	27.3	75.6	72.8	-71.0
4	115.8	83.1	34.7	-2.0	86.8	28.7	79.4	76.4	-74.6

Year	Unearned Premium BOY	Unearned Premium EOY	Loss Reserve BOY	Loss Reserve EOY	Stat Expense Reserve BOY	Stat Expense Reserve EOY	Total Stat Reserves BOY	Total Stat Reserves EOY
1	100.0	0.0	0.0	50.0	9.0	7.5	109.0	57.5
2	105.0	0.0	50.0	69.5	17.0	9.4	171.9	78.8
3	110.3	0.0	69.5	72.9	19.3	9.8	199.0	82.8
4	115.8	0.0	72.9	76.6	20.3	10.3	209.0	86.9

Year	Surplus BOY	Surplus EOY	Assets BOY	Assets EOY	Receivables BOY	Receivables EOY	Invested Assets BOY	Investment Income EOY	P/S
1	20.2	15.7	129.2	73.2	25.0	5.0	104.2	6.3	4.95
2	37.0	21.9	208.9	100.7	31.3	5.3	177.6	10.7	2.84
3	44.2	23.0	243.2	105.8	32.8	5.5	210.4	12.6	2.50
4	46.4	24.1	255.3	111.0	34.5	5.8	220.9	13.3	2.50

Year	DAC BOY	DAC EOY	GAAP Equity BOY	GAAP Equity EOY	GAAP Pre-tax Income EOY	Income Tax EOY	GAAP Income EOY	GAAP ROE
1	18.0	0.0	38.2	15.7	8.3	2.9	5.4	14.07%
2	18.9	0.0	55.9	21.9	9.8	3.4	6.4	11.38%
3	19.8	0.0	64.0	23.0	10.7	3.7	6.9	10.85%
4	20.8	0.0	67.2	24.1	11.2	3.9	7.3	10.85%

Exhibit 3
Sheet 4

Book of Business Growth Company

UW Assumptions			Financial Assumptions		IRR and ROE Results	
	Amount	Ratio	Interest Rate	6.00%	IRR	10.99%
Premium	100.0	100.0%	Tax Rate	35.00%	EQ Growth ROE	10.99%
Loss	72.0	72.0%	Rsv Discount Rate	6.00%	EQ Growth P/S	2.58
Expense	30.0	30.0%	S as % of PV Unpaid Loss	31.50%	Growth Rate	10.99%
Combined	102.0	102.0%	PV Loss Discount for S Calc	6.00%		

Year	Earned Premium EOY	Incurred Loss EOY	GAAP Incurred Expense EOY	GAAP UW Income EOY	Paid Premium BOY	Paid Premium EOY	Paid Loss EOY	UW Cash Flow BOY	UW Cash Flow EOY
1	100.0	68.0	30.0	2.0	75.0	20.0	18.0	66.0	-11.5
2	111.0	78.5	33.3	-0.8	83.2	27.2	56.0	73.3	-49.8
3	123.2	88.1	37.0	-1.9	92.4	30.2	80.1	81.3	-74.7
4	136.7	97.8	41.0	-2.1	102.5	33.5	88.9	90.2	-82.9

Year	Unearned Premium BOY	Unearned Premium EOY	Loss Reserve BOY	Loss Reserve EOY	Stat Expense Reserve BOY	Stat Expense Reserve EOY	Total Stat Reserves BOY	Total Stat Reserves EOY
1	100.0	0.0	0.0	50.0	9.0	7.5	109.0	57.5
2	111.0	0.0	50.0	72.5	17.5	9.8	178.5	82.3
3	123.2	0.0	72.5	80.4	20.9	10.9	216.5	91.3
4	136.7	0.0	80.4	89.3	23.2	12.1	240.3	101.4

Year	Surplus BOY	Surplus EOY	Assets BOY	Assets EOY	Receivables BOY	Receivables EOY	Invested Assets BOY	Investment Income EOY	P/S
1	20.2	15.7	129.2	73.2	25.0	5.0	104.2	6.3	4.95
2	38.2	22.8	216.6	105.1	32.7	5.5	183.9	11.0	2.91
3	47.7	25.3	264.3	116.7	36.3	6.2	227.9	13.7	2.58
4	53.0	28.1	293.3	129.5	40.3	6.8	253.0	15.2	2.58

Year	DAC BOY	DAC EOY	GAAP Equity BOY	GAAP Equity EOY	GAAP Pre-tax Income EOY	Income Tax EOY	GAAP Income EOY	GAAP ROE
1	18.0	0.0	38.2	15.7	8.3	2.9	5.4	14.07%
2	20.0	0.0	58.1	22.8	10.3	3.6	6.7	11.48%
3	22.2	0.0	69.9	25.3	11.8	4.1	7.7	10.99%
4	24.6	0.0	77.6	28.1	13.1	4.6	8.5	10.99%

Exhibit 4
Sheet 1

Return Measures
Sensitivity to Premium **Full Value Reserve**

Scenario	1	2	3	4	5	6	7
Premium	80.00	85.00	90.00	95.00	100.00	105.00	110.00
Combined Ratio	122.50%	116.47%	111.11%	106.32%	102.00%	98.10%	94.55%
Resulting Growth Model P/S	2.00	2.12	2.25	2.37	2.50	2.62	2.75
Returns							
IRR	-7.00%	-2.74%	1.65%	6.15%	10.74%	15.40%	20.10%
PVI/PVE	-9.21%	-4.07%	0.96%	5.89%	10.71%	15.43%	20.05%
ROE	-8.47%	-3.47%	1.42%	6.21%	10.90%	15.49%	19.99%
Change in Returns							
IRR		4.27%	4.39%	4.50%	4.59%	4.66%	4.70%
PVI/PVE		5.14%	5.03%	4.92%	4.82%	4.72%	4.62%
ROE		5.00%	4.89%	4.79%	4.69%	4.59%	4.50%

Assumptions for All Scenarios

Financial

Interest Rate	6.00%
Tax Rate	35.00%
Reserve Discount Rate	0.00%
Surplus as % of PV Unpaid Loss	31.50%
Rate for PV Calculation	6.00%
Rate for PVI/PVE Discounting	12.00%
ROE Growth Rate	5.00%

Assumptions for All Scenarios

Underwriting

Premium		Loss		Expense	
	varies	Fixed	72.00	Fixed	10.00
				Variable	20.0%
Premium Payment		Loss Payout		Expense Payout	
Year	%	Year	%	Year	%
0	75.0%	0	0.0%	0	30.0%
1	20.0%	1	25.0%	1	50.0%
2	5.0%	2	50.0%	2	25.0%
3	0.0%	3	25.0%	3	0.0%
4	0.0%	4	0.0%	4	0.0%

Exhibit 4
Sheet 2

Return Measures
Sensitivity to Premium **Discounted Reserve**

Scenario	1	2	3	4	5	6	7
Premium	80.00	85.00	90.00	95.00	100.00	105.00	110.00
Combined Ratio	122.50%	116.47%	111.11%	106.32%	102.00%	98.10%	94.55%
Resulting Growth Model P/S	2.00	2.12	2.25	2.37	2.50	2.62	2.75
Returns							
IRR	-7.74%	-3.23%	1.42%	6.16%	10.99%	15.87%	20.79%
PVI/PVE	-8.89%	-3.75%	1.27%	6.19%	11.01%	15.73%	20.34%
ROE	-8.52%	-3.53%	1.36%	6.15%	10.85%	15.44%	19.94%
Change in Returns							
IRR		4.52%	4.64%	4.75%	4.83%	4.88%	4.92%
PVI/PVE		5.14%	5.03%	4.92%	4.82%	4.72%	4.62%
ROE		5.00%	4.89%	4.79%	4.69%	4.59%	4.50%

Assumptions for All Scenarios

Financial

Interest Rate	6.00%
Tax Rate	35.00%
Reserve Discount Rate	6.00%
Surplus as % of PV Unpaid Loss	31.50%
Rate for PV Calculation	6.00%
Rate for PVI/PVE Discounting	12.00%
ROE Growth Rate	5.00%

Assumptions for All Scenarios

Underwriting

Premium		Loss		Expense	
	varies	Fixed	72.00	Fixed	10.00
				Variable	20.0%
Premium Payment		Loss Payout		Expense Payout	
Year	%	Year	%	Year	%
0	75.0%	0	0.0%	0	30.0%
1	20.0%	1	25.0%	1	50.0%
2	5.0%	2	50.0%	2	25.0%
3	0.0%	3	25.0%	3	0.0%
4	0.0%	4	0.0%	4	0.0%

Exhibit 4
Sheet 3

Return Measures
Sensitivity to Surplus

Scenario	1	2	3	4	5	6	7
Surplus as % of PV Unpaid Loss	25.50%	27.50%	29.50%	31.50%	33.50%	35.50%	37.50%
Resulting Growth Model P/S	3.08	2.86	2.67	2.50	2.35	2.22	2.10
Returns							
IRR	11.73%	11.37%	11.04%	10.74%	10.46%	10.21%	9.97%
PVI/PVE	11.72%	11.35%	11.02%	10.71%	10.42%	10.16%	9.92%
ROE	11.96%	11.57%	11.22%	10.90%	10.60%	10.33%	10.09%

Assumptions for All Scenarios

Financial

Interest Rate	6.00%
Tax Rate	35.00%
Reserve Discount Rate	0.00%
Surplus as % of PV Unpaid Loss	varies
Rate for PV Calculation	6.00%
Rate for PVI/PVE Discounting	12.00%
ROE Growth Rate	5.00%

Assumptions for All Scenarios

Underwriting

Premium		Loss		Expense	
Fixed	100.00	Fixed	72.00	Fixed	10.00
				Variable	20.0%
Premium Payment		Loss Payout		Expense Payout	
Year	%	Year	%	Year	%
0	75.0%	0	0.0%	0	30.0%
1	20.0%	1	25.0%	1	50.0%
2	5.0%	2	50.0%	2	25.0%
3	0.0%	3	25.0%	3	0.0%
4	0.0%	4	0.0%	4	0.0%

Exhibit 4
Sheet 4

Return Measures
Sensitivity to Interest Rate

Scenario	1	2	3	4	5	6	7
Interest Rate	4.50%	5.00%	5.50%	6.00%	6.50%	7.00%	7.50%
Resulting Growth Model P/S	2.44	2.46	2.48	2.50	2.52	2.53	2.55
Returns							
IRR	7.48%	8.56%	9.65%	10.74%	11.84%	12.93%	14.04%
PVI/PVE	7.38%	8.48%	9.59%	10.71%	11.83%	12.96%	14.10%
ROE	7.54%	8.65%	9.77%	10.90%	12.03%	13.18%	14.33%

Assumptions for All Scenarios

Financial

Interest Rate	varies
Tax Rate	35.00%
Reserve Discount Rate	0.00%
Surplus as % of PV Unpaid Loss	31.50%
Rate for PV Calculation	varies
Rate for PVI/PVE Discounting	12.00%
ROE Growth Rate	5.00%

Assumptions for All Scenarios

Underwriting

Premium		Loss		Expense	
Fixed	100.00	Fixed	72.00	Fixed	10.00
				Variable	20.0%
Premium Payment		Loss Payout		Expense Payout	
Year	%	Year	%	Year	%
0	75.0%	0	0.0%	0	30.0%
1	20.0%	1	25.0%	1	50.0%
2	5.0%	2	50.0%	2	25.0%
3	0.0%	3	25.0%	3	0.0%
4	0.0%	4	0.0%	4	0.0%

Exhibit 4
Sheet 5

Return Measures
Sensitivity to Payout Pattern

Scenario		1=Base	2	3	4	5	6	7
Loss Pattern	Year							
	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	1	25.00%	100.00%	50.00%	0.00%	0.00%	0.00%	0.00%
	2	50.00%	0.00%	50.00%	100.00%	50.00%	0.00%	0.00%
	3	25.00%	0.00%	0.00%	0.00%	50.00%	100.00%	50.00%
	4	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	50.00%
Surplus % of PV Unpaid Loss		31.50%	58.96%	40.72%	31.10%	25.68%	21.87%	19.32%
Resulting Growth Model P/S		2.50	2.50	2.50	2.50	2.50	2.50	2.50
Indicated Profit Margins								
IRR Method		10.74%	6.34%	8.60%	10.82%	12.85%	14.83%	16.61%
PVI/PVE Method		10.71%	6.33%	8.55%	10.79%	12.88%	14.97%	16.92%
ROE Method		10.90%	6.35%	8.65%	10.95%	13.15%	15.34%	17.43%

Assumptions for All Scenarios

Financial

Interest Rate	6.00%
Tax Rate	35.00%
Reserve Discount Rate	0.00%
Surplus as % of PV Unpaid Loss	varies
Rate for PV Calculation	12.00%
Rate for PVI/PVE Discounting	12.00%
Growth Rate	5.00%

Assumptions for All Scenarios

Underwriting

Premium		Loss		Expense	
		Fixed	72.00	Fixed	10.00
				Variable	20.0%
Premium Payment		Loss Payout		Expense Payout	
Year	%	Year	%	Year	%
0	75.0%	0	varies	0	30.0%
1	20.0%	1	varies	1	45.0%
2	5.0%	2	varies	2	20.0%
3	0.0%	3	varies	3	5.0%
4	0.0%	4	varies	4	0.0%

Exhibit 5
Sheet 1

Indicated Profit
Sensitivity to Surplus

Scenario	1	2	3	4	5	6	7
Surplus as % of PV Unpaid Loss	25.50%	27.50%	29.50%	31.50%	33.50%	35.50%	37.50%
Resulting Growth Model P/S	3.09	2.87	2.69	2.53	2.38	2.26	2.15
Indicated Profit Margins							
IRR Method	-1.79%	-1.49%	-1.20%	-0.90%	-0.61%	-0.32%	-0.03%
PVI/PVE Method	-1.79%	-1.49%	-1.20%	-0.90%	-0.61%	-0.32%	-0.03%
ROE Method	-1.97%	-1.65%	-1.34%	-1.04%	-0.73%	-0.43%	-0.13%

Assumptions for All Scenarios

Financial

Interest Rate	6.00%
Tax Rate	35.00%
Reserve Discount Rate	0.00%
Surplus as % of PV Unpaid Loss	varies
Rate for PV Calculation	6.00%
IRR Target Return	12.00%
PVI/PVE Target Return	12.00%
Rate for PVI/PVE Discounting	12.00%
ROE Target Return	12.00%
ROE Target Growth Rate	5.00%

Assumptions for All Scenarios

Underwriting

Premium		Loss		Expense	
		Fixed	72.00	Fixed	10.00
				Variable	20.0%
Premium Payment		Loss Payout		Expense Payout	
Year	%	Year	%	Year	%
0	75.0%	0	0.0%	0	30.0%
1	20.0%	1	25.0%	1	25.0%
2	5.0%	2	50.0%	2	50.0%
3	0.0%	3	25.0%	3	25.0%
4	0.0%	4	0.0%	4	0.0%

Exhibit 5
Sheet 2

Indicated Profit
Sensitivity to Interest Rate

Scenario	1	2	3	4	5	6	7
Interest Rate	4.50%	5.00%	5.50%	6.00%	6.50%	7.00%	7.50%
Resulting Growth Model P/S	2.56	2.55	2.54	2.53	2.52	2.50	2.49
Indicated Profit Margins							
IRR Method	1.91%	0.98%	0.05%	-0.90%	-1.86%	-2.82%	-3.80%
PVI/PVE Method	1.91%	0.98%	0.05%	-0.90%	-1.86%	-2.82%	-3.80%
ROE Method	1.88%	0.92%	-0.05%	-1.04%	-2.03%	-3.03%	-4.05%

Assumptions for All Scenarios

Financial

Interest Rate	varies
Tax Rate	35.00%
Reserve Discount Rate	0.00%
Surplus as % of PV Unpaid Loss	31.50%
Rate for PV Calculation	varies
IRR Target Return	12.00%
PVI/PVE Target Return	12.00%
Rate for PVI/PVE Discounting	12.00%
ROE Target Return	12.00%
ROE Target Growth Rate	5.00%

Assumptions for All Scenarios

Underwriting

Premium		Loss		Expense	
		Fixed	72.00	Fixed	10.00
				Variable	20.0%
Premium Payment		Loss Payout		Expense Payout	
Year	%	Year	%	Year	%
0	75.0%	0	0.0%	0	30.0%
1	20.0%	1	25.0%	1	25.0%
2	5.0%	2	50.0%	2	50.0%
3	0.0%	3	25.0%	3	25.0%
4	0.0%	4	0.0%	4	0.0%

Exhibit 5
Sheet 3

Indicated Profit
Sensitivity to Payout Pattern

Scenario		1=Base	2	3	4	5	6	7
Loss Pattern	Year							
	0	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	1	25.00%	100.00%	50.00%	0.00%	0.00%	0.00%	0.00%
	2	50.00%	0.00%	50.00%	100.00%	50.00%	0.00%	0.00%
	3	25.00%	0.00%	0.00%	0.00%	50.00%	100.00%	50.00%
	4	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	50.00%
Surplus % of PV Unpaid Loss		31.50%	62.00%	41.76%	31.08%	25.03%	20.77%	17.90%
Resulting Growth Model P/S		2.53	2.53	2.53	2.53	2.53	2.53	2.53
Indicated Profit Margins								
IRR Method		-0.90%	2.96%	1.02%	-0.97%	-2.88%	-4.85%	-6.72%
PVI/PVE Method		-0.90%	2.96%	1.02%	-0.97%	-2.88%	-4.85%	-6.72%
ROE Method		-1.04%	2.94%	0.98%	-1.09%	-3.16%	-5.34%	-7.52%

Assumptions for All Scenarios

Financial

Interest Rate	6.00%
Tax Rate	35.00%
Reserve Discount Rate	0.00%
Surplus as % of PV Unpaid Loss	varies
Rate for PV Calculation	6.00%
IRR Target Return	12.00%
PVI/PVE Target Return	12.00%
Rate for PVI/PVE Discounting	12.00%
ROE Target Return	12.00%
ROE Target Growth Rate	5.00%

Assumptions for All Scenarios

Underwriting

Premium		Loss		Expense	
		Fixed	72.00	Fixed	10.00
				Variable	20.0%
Premium Payment		Loss Payout		Expense Payout	
Year	%	Year	%	Year	%
0	75.0%	0	varies	0	30.0%
1	20.0%	1	varies	1	45.0%
2	5.0%	2	varies	2	20.0%
3	0.0%	3	varies	3	5.0%
4	0.0%	4	varies	4	0.0%

Exhibit 6
Sheet 1

Results for Three Point Loss Distribution
Sensitivity to Premium Full Value Reserve

Scenario	1	2	3	Average over All Scenarios
Probability	40.00%	40.00%	20.00%	
Premium	100.00	100.00	100.00	100.00
Loss	60.00	72.00	96.00	72.00
Combined Ratio	90.00%	102.00%	126.00%	102.00%
Returns				
IRR	24.11%	10.74%	-11.63%	10.74%
PVI/PVE	23.79%	10.71%	-15.45%	10.71%

Results by Year	Equity	Income	Equity Flow	Equity	Income	Equity Flow	Equity	Income	Equity Flow	Equity	Income	Equity Flow
Year												
0	38.20	0.00	-38.20	38.20	0.00	-38.20	38.20	0.00	-38.20	38.20	0.00	-38.20
1	15.74	10.56	33.02	15.74	2.76	25.22	15.74	-12.84	9.62	15.74	2.76	25.22
2	5.35	2.47	12.86	5.35	2.82	13.21	5.35	3.52	13.91	5.35	2.82	13.21
3	0.00	0.85	6.20	0.00	0.97	6.32	0.00	1.20	6.55	0.00	0.97	6.32

Exhibit 6
Sheet 2

Single Policy Company **Loss Scenario 1**

E[Loss]	72.0

UW Assumptions			Financial Assumptions		IRR and PVI/PVE Results	
	Amount	Ratio	Interest Rate	6.00%	IRR	24.11%
Premium	100.0	100.0%	Tax Rate	35.00%	PVI/PVE Discount Rate	12.00%
Loss	60.0	60.0%	Rsv Discount Rate	0.00%	PVI	13.45
Expense	30.0	30.0%	S as % of E[PV Unpaid Loss]	31.50%	PVE	56.52
Combined	90.0	90.0%	PV Loss Discount for S Calc	6.00%	PVI/PVE	23.79%

Year	Earned Premium	Incurred Loss	Stat Incurred Expense	Stat UW Income	Paid Premium	Paid Loss	Paid Expense	UW Cash Flow
0	0.0	0.0	18.0	-18.0	75.0	0.0	9.0	66.0
1	100.0	60.0	12.0	28.0	20.0	15.0	13.5	-8.5
2	0.0	0.0	0.0	0.0	5.0	30.0	6.0	-31.0
3	0.0	0.0	0.0	0.0	0.0	15.0	1.5	-16.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	100.0	60.0	30.0	10.0	100.0	60.0	30.0	10.0

Year	Unearned Premium	Loss Reserve	Expected PV Unpaid Loss	Stat Expense Reserve	Total Stat Reserves	Surplus	Assets	Receivables	Invested Assets	Inv Income
0	100.0	0.0	64.1	9.0	109.0	20.2	129.2	25.0	104.2	
1	0.0	45.0	50.0	7.5	52.5	15.7	68.2	5.0	63.2	6.3
2	0.0	15.0	17.0	1.5	16.5	5.3	21.8	0.0	21.8	3.8
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	DAC	GAAP Equity	GAAP Incurred Expense	GAAP UW Income	GAAP Pre-tax Income	Income Tax	Income	Change in Equity	Equity Flow
0	18.0	38.2	0.0	0.0	0.0	0.0	0.0	38.2	-38.2
1	0.0	15.7	30.0	10.0	16.3	5.7	10.6	-22.5	33.0
2	0.0	5.3	0.0	0.0	3.8	1.3	2.5	-10.4	12.9
3	0.0	0.0	0.0	0.0	1.3	0.5	0.9	-5.3	6.2
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total			30.0	10.0	21.4	7.5	13.9	0.0	13.9

Single Policy Company **Loss Scenario 2**

E[Loss]	72.0

UW Assumptions	Amount	Ratio	Financial Assumptions		IRR and PVI/PVE Results	
Premium	100.0	100.0%	Interest Rate	6.00%	IRR	10.74%
Loss	72.0	72.0%	Tax Rate	35.00%	PVI/PVE Discount Rate	12.00%
Expense	30.0	30.0%	Rsv Discount Rate	0.00%	PVI	6.05
Combined	102.0	102.0%	S as % of E[PV Unpaid Loss]	31.50%	PVE	56.52
			PV Loss Discount for S Calc	6.00%	PVI/PVE	10.71%

Year	Earned Premium	Incurred Loss	Stat Incurred Expense	Stat UW Income	Paid Premium	Paid Loss	Paid Expense	UW Cash Flow
0	0.0	0.0	18.0	-18.0	75.0	0.0	9.0	66.0
1	100.0	72.0	12.0	16.0	20.0	18.0	13.5	-11.5
2	0.0	0.0	0.0	0.0	5.0	36.0	6.0	-37.0
3	0.0	0.0	0.0	0.0	0.0	18.0	1.5	-19.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	100.0	72.0	30.0	-2.0	100.0	72.0	30.0	-2.0

Year	Unearned Premium	Loss Reserve	Expected PV Unpaid Loss	Stat Expense Reserve	Total Stat Reserves	Surplus	Assets	Receivables	Invested Assets	Inv Income
0	100.0	0.0	64.1	9.0	109.0	20.2	129.2	25.0	104.2	
1	0.0	54.0	50.0	7.5	61.5	15.7	77.2	5.0	72.2	6.3
2	0.0	18.0	17.0	1.5	19.5	5.3	24.8	0.0	24.8	4.3
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	DAC	GAAP Equity	GAAP Incurred Expense	GAAP UW Income	GAAP Pre-tax Income	Income Tax	Income	Change in Equity	Equity Flow
0	18.0	38.2	0.0	0.0	0.0	0.0	0.0	38.2	-38.2
1	0.0	15.7	30.0	-2.0	4.3	1.5	2.8	-22.5	25.2
2	0.0	5.3	0.0	0.0	4.3	1.5	2.8	-10.4	13.2
3	0.0	0.0	0.0	0.0	1.5	0.5	1.0	-5.3	6.3
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total			30.0	-2.0	10.1	3.5	6.6	0.0	6.6

Exhibit 6
Sheet 4

Single Policy Company **Loss Scenario 3**

E[Loss]	72.0

UW Assumptions			Financial Assumptions		IRR and PVI/PVE Results	
	Amount	Ratio	Interest Rate	6.00%	IRR	-11.63%
Premium	100.0	100.0%	Tax Rate	35.00%	PVI/PVE Discount Rate	12.00%
Loss	96.0	96.0%	Rsv Discount Rate	0.00%	PVI	-8.73
Expense	30.0	30.0%	S as % of E[PV Unpaid Loss]	31.50%	PVE	56.52
Combined	126.0	126.0%	PV Loss Discount for S Calc	6.00%	PVI/PVE	-15.45%

Year	Earned Premium	Incurred Loss	Stat Incurred Expense	Stat UW Income	Paid Premium	Paid Loss	Paid Expense	UW Cash Flow
0	0.0	0.0	18.0	-18.0	75.0	0.0	9.0	66.0
1	100.0	96.0	12.0	-8.0	20.0	24.0	13.5	-17.5
2	0.0	0.0	0.0	0.0	5.0	48.0	6.0	-49.0
3	0.0	0.0	0.0	0.0	0.0	24.0	1.5	-25.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	100.0	96.0	30.0	-26.0	100.0	96.0	30.0	-26.0

Year	Unearned Premium	Loss Reserve	Expected PV Unpaid Loss	Stat Expense Reserve	Total Stat Reserves	Surplus	Assets	Receivables	Invested Assets	Inv Income
0	100.0	0.0	64.1	9.0	109.0	20.2	129.2	25.0	104.2	
1	0.0	72.0	50.0	7.5	79.5	15.7	95.2	5.0	90.2	6.3
2	0.0	24.0	17.0	1.5	25.5	5.3	30.8	0.0	30.8	5.4
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Year	DAC	GAAP Equity	GAAP Incurred Expense	GAAP UW Income	GAAP Pre-tax Income	Income Tax	Income	Change in Equity	Equity Flow
0	18.0	38.2	0.0	0.0	0.0	0.0	0.0	38.2	-38.2
1	0.0	15.7	30.0	-26.0	-19.7	-6.9	-12.8	-22.5	9.6
2	0.0	5.3	0.0	0.0	5.4	1.9	3.5	-10.4	13.9
3	0.0	0.0	0.0	0.0	1.9	0.6	1.2	-5.3	6.6
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total			30.0	-26.0	-12.5	-4.4	-8.1	0.0	-8.1

GLM Basic Modeling: Avoiding Common Pitfalls

Geoff Werner, FCAS, MAAA, and Serhat Guven, FCAS, MAAA

Abstract Starting in the 1990's many of the larger US personal lines carriers began to implement predictive modeling techniques in the form of generalized linear modeling (GLM). Because of the early success realized by those companies, the vast majority of companies are now rushing to employ these techniques too. In their haste to keep up with competitors, many companies are making mistakes and not getting the full benefit possible.

The following are some of the most common mistakes made by companies beginning to build GLMs:

- Failing to get full buy-in from key stakeholders.
- Relying too heavily on pre-analysis.
- Using loss ratio analysis.
- Modeling raw pure premiums for all coverages directly rather than modeling at the component level.
- Restricting analysis to variables and groupings in the current rating algorithm.
- Misusing offsets.
- Treating the predictive model as a "black box".
- Limiting the use of GLMs to risk models.

This paper will address each of these pitfalls in turn. By being aware of these pitfalls, companies can hopefully minimize the transition period and achieve the full benefits of multivariate pricing as quickly as possible.

BACKGROUND

Predictive modeling has been standard practice for insurance ratemaking in the highly advanced UK marketplace for many years. While a few US companies have been doing predictive modeling for some time, it has not been until the last five years that there has been widespread acceptance of these techniques in the US marketplace.

What Is Predictive Modeling?

Essentially, predictive modeling involves using historical data to construct a statistical model that will be predictive of the future. Each observation in the historical dataset contains information or data elements that are essential in building a predictive model. Figure 1 is a visual representation of predictive modeling.

There are three types of elements in the historical dataset. First, there will be a dependent or response variable which is what the practitioner is trying to predict. For example, when modeling severity, the dependent variable is the loss amount. Second, there will be a weight associated with each observation. When modeling severity, the weight is the number of claims associated with the loss amount of the observation. Finally, there are independent variables or predictors. These are the characteristics of each observation that are being studied to ascertain whether a variable has any predictive power.

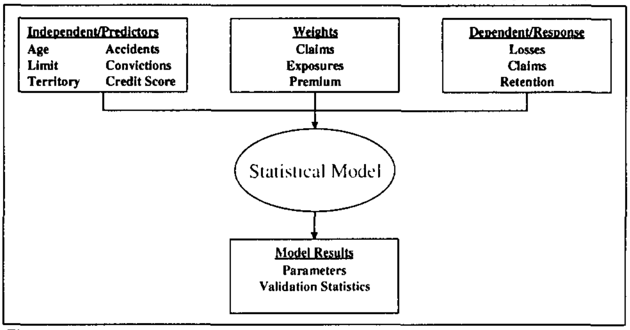


Figure 1. Overview of Predictive Modeling

The practitioner uses the historical data to build a statistical model. The output of the model is a set of parameters and validation statistics. The parameters represent the actual results; for example, when performing class plan analysis, the parameters will be the indicated relativities.[1] The validation statistics provide the practitioner with an understanding of the effectiveness of the model.

The majority of companies using predictive modeling are building such models to identify new rating variables and to better quantify the relationships between these new variables and existing rating variables.

[1] To be precise, practitioners generally use a log link function for class plan analysis. When doing so, the indicated relativities are really calculated as exp(relevant parameters).

Does Predictive Modeling Make a Difference?

Figure 2 shows a comparison of the relative predictions for automobile theft by age derived from a one-way analysis (square markers) and a simple generalized linear model (circle markers). As can be seen, the difference in these lines is quite significant for some age categories; more specifically, the percentage difference between the one-way prediction and the prediction based on the generalized linear model ranges from -38% to +6%. As this is only the difference for one factor for one cause of loss, it is apparent that the differences could be even more significant when the differences in other factors and other causes of loss are compounded.

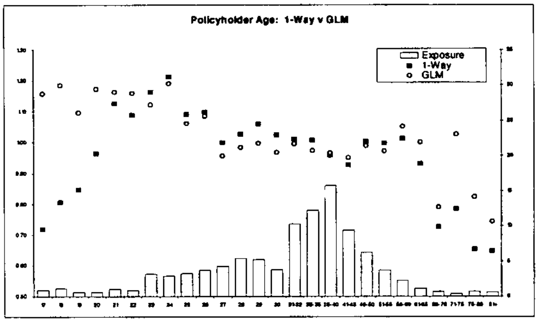


Figure 2. One-way v GLM

While the mere fact that the results are different does not prove that the multivariate results are superior to the results based on the univariate analysis, it is commonly accepted that multivariate analysis corrects for methodological flaws inherent in one-way analysis and is more accurate.[2] Thus, the companies who employ multivariate techniques will be able to better predict loss costs and develop more accurate pricing structures. Companies who fail to employ these techniques will not have accurate prices and will be susceptible to adverse selection.

What are companies doing?

The number of US companies using multivariate analysis has increased dramatically over the past five years and virtually all of the companies in the top 20 are doing some form of multivariate analysis to gain a competitive advantage with respect to classification and factor analysis. The most commonly used predictive modeling technique is generalized linear modeling (GLM). The popularity of GLMs is likely due to several key advantages of GLMs as compared to traditional and other predictive modeling techniques:

- GLMs can readily adjust for both exposure and response correlations that cause one-way analyses to fail.
- Traditional statistics (e.g., loss ratios) include a systematic and unsystematic component. Like other predictive modeling techniques, GLMs allow the model to separate the components to

[2] In the example shown, the extreme difference between the indications at the youthful ages is caused by distributional biases between age and other variables (e.g., limits and cost of vehicle). A full discussion of the reasons generalized linear model results are more accurate than those produced from traditional analysis is outside the scope of this paper. For more information on this refer to "Something Old, Something New in Classification Ratemaking With Novel Use of GLMs for Credit Insurance" written by Keith Holler, David Sommer, and Geoff Trahir and published in the 1999 CAS Winter Forum.

remove unsystematic variation or the "noise" in the data and identify systematic variation or the "signal" in the data.
- Because GLM is a predictive modeling technique, it allows the user to do more with less data than traditional techniques which require significant amounts of data in each cell for "full credibility". GLMs tend to be more robust than other predictive modeling techniques and are less susceptible to over-fitting (e.g., CART or MARS) that may occur with small data sets.
- GLMs provide the modeler with a battery of diagnostics that allow for decision-making in the context of a solid statistical framework.
- GLMs allow the modeler to assume the process being modeled follows any distribution within the exponential family. The exponential family includes common distributions like Poisson and Gamma that are generally accepted as appropriate for modeling insurance data.
- GLMs are not "black box" models. Unlike some of its predictive modeling counterparts (e.g., neural nets), a GLM is easy to interpret and allows the analyst to clearly understand how each of the predictors are influencing the prediction.

Are Companies Being as Effective as Possible?

Because predictive modeling does make a difference, companies are rushing to employ these techniques. In their haste to keep up with competitors, companies are not always taking the time to perform the analysis appropriately. Consequently, many companies are making mistakes and not getting the full benefit possible. This paper is intended to address some of the most common problems companies encounter when moving from traditional techniques to multivariate analysis, such as:
- Failing to get full buy-in from key stakeholders.
- Relying too heavily on pre-analysis.
- Using loss ratio analysis.
- Modeling raw pure premiums for all coverages directly rather than modeling at the component level.
- Restricting analysis to variables and groupings in the current rating algorithm.
- Misusing offsets.
- Treating the predictive model as a "black box".
- Limiting the use of GLMs to risk models.

This paper will address each of these pitfalls in turn. By being aware of these pitfalls, companies can hopefully minimize the transition period and achieve the full benefits of multivariate pricing as quickly as possible.

KEY STAKEHOLDER BUY-IN

It is important for all key stakeholders to support any major change to pricing techniques. The main reason that companies using multivariate techniques are improving their results is because these techniques more accurately predict the risk. Thus, by definition, the multivariate results are different—and in some cases significantly different—than the univariate results. While all of the key stakeholders do not need to fully comprehend the mathematics associated with GLMs, it is imperative they recognize the benefits of multivariate analysis and that they expect the results to be different than those from prior reviews.

The best way to communicate information of this sort is to provide simple examples based on company data that highlight distortions caused by one way analysis. To do so, the practitioner should build a very simple GLM, look for a variable for which the difference between the one-way and GLM result is materially different, and examine the correlation statistics to find a variable that is likely contributing to the difference. Those two variables can be used as a simple case study of how one-way analysis can lead to inappropriate conclusions. As this represents a major shift in mindset, it is likely that that practitioner will have to go through this process more than once.

Once management is convinced that the multivariate results are superior, the modeler can compare the results of a "quick and dirty" predictive model to the current rates to determine an estimate of the subsidization inherent in the current rates. (If available, indications based on traditional analysis can also be included in the comparison to highlight differences.) Figure 3 is an example of how this type of information can be displayed for non-technical audiences. For each individual observation, the modeler calculates Current Premium/Indicated Premium – 1.00 and categorizes each observation in the range corresponding to the inadequacy or excessiveness associated with that observation. For example, the XXX exposures in the 5%-10% bucket are risks whose current premium is 5% to 10% above the indicated premium. The cross-hatched bars represent risks whose current premium is below the indicated premium (i.e., risks being subsidized) and the solid bars represent risks whose current premium is above the indicated premium (i.e., risks subsidizing). If all risks were being charged the right rate, every observation would be in the bucket containing 0%. If the bars are spread out as in Figure 3, then there is considerable subsidization present in the rating plan. That subsidization represents opportunity for improvement.

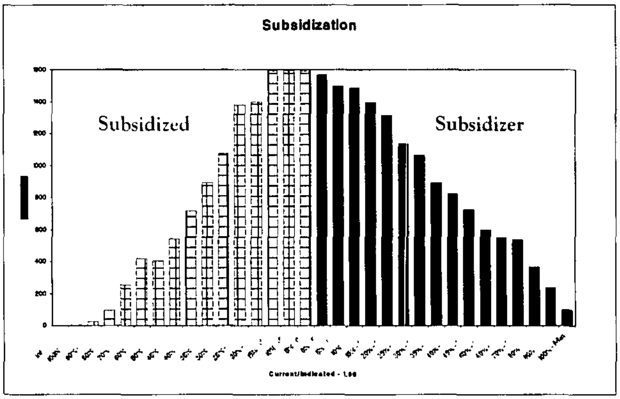


Figure 3. Subsidization Histogram

Unfortunately, in many companies the group responsible for producing the indicated relativities makes the mistake of not doing the appropriate up front communication. Failure to do so invites significant resistance when the indicated multivariate relativities are not in line with the traditional univariate results on which the others within the company have been basing decisions for years. This resistance usually leads to undesirable compromises. For example, senior management may choose to only implement one aspect of the multivariate analysis. In the best case, this weakens the effectiveness of the plan; in the worst case, it can actually result in implementation of plan that is inferior to the current plan.[3] Interestingly, gaining company-wide acceptance is usually most difficult for companies who were the most successful previously. In such companies, senior management is often very reluctant to abandon methodologies that were used to achieve the success.

PRE-ANALYSIS

Performing pre-analysis can be helpful to get a feeling for the data. Pre-analysis oftentimes includes traditional one-way data, volume measures, and correlation statistics.

Traditional one-way data includes items like raw frequencies, severities, pure premiums, and loss ratios. Examining these ratios can help the practitioner in three ways. First, it can help the modeler spot items that may distort the analysis (e.g., extraordinary losses) if left unadjusted. Second, it will highlight what others in the company may be examining, so the practitioner will be better prepared for difficult discussions when the multivariate results are different than the univariate results. Third, it allows the modeler to build a priori expectations that can help when applying judgment during the modeling. The mistake some practitioners make is putting too much emphasis on the univariate results. It is important to recall that the whole point of using multivariate analysis is to correct for the flaws inherent in univariate analysis; therefore, the results will have some differences and it is imperative the modeler does not allow the univariate results to bias judgment and ultimately limit the benefit.

Volume analysis usually includes an examination of the distribution of exposures, claim counts, and premiums. Volumes can help the practitioner decide the appropriate number of years of data necessary and whether there is enough data to do test/training analysis. The mistake some modelers make is to use these volumes to calculate traditional estimates of credibility. Traditional estimates of credibility are not necessary within the context of predictive modeling and are replaced with better diagnostics that indicate the amount of reliance that should be given to individual estimates.

Finally, correlation statistics (e.g., Cramer's V) inform the practitioner which independent predictors have a high degree of exposure correlation. The mistake too many companies make is using this information to eliminate correlated variables from the modeling process before it even starts. The practitioner should not eliminate variables at this stage, but rather note the high correlations and be aware that the inclusion/exclusion or offsetting of one of the factors will have an impact on the other factor. By not ruling out the variable before the analysis starts, the practitioner can test the various variables within the multivariate framework to determine whether only one or a combination of the highly correlated variables should be included.

Basically, if the intent is to do multivariate analysis, then the practitioner should perform multivariate analysis and avoid the temptation to make decisions during the pre-analysis stage. By making decisions before the multivariate analysis begins, the practitioner is only limiting the potential benefit.

---

[3] The careful use of offsets can help minimize the adverse impact of implementing GLM results on a piecemeal basis. Offsets are discussed in detail in a later section.

PURE PREMIUM V. LOSS RATIO ANALYSIS

When it comes to risk modeling, many companies are doing loss ratio modeling, rather than frequency/severity modeling. There are both practical and theoretical reasons that modeling loss ratios is less preferable.

If loss ratios are being modeled, it is imperative that the loss ratios be calculated on current rate level; failure to do so will make the resulting relativities inappropriate. Because the practitioner is performing multivariate analysis, it is not sufficient to use an on-level approximation method that applies an average current rate level factor to a diverse set of observations. Instead, the loss ratio for each observation must be put on the correct rate level which is done via extension of exposures (i.e., re-rating) and should be "re-underwritten" if underwriting rules were changed. This can be a very difficult--and for many companies --an impossible task. When modeling frequency and severity, the modeler uses exposures rather than premium. Thus, no current rate level adjustment is necessary.

It is widely accepted that there are standard distributions for frequency and severity. Generally speaking, a Poisson error structure is appropriate for frequency modeling and a Gamma error structure is appropriate for severity modeling. Loss ratios, on the other hand, do not follow a typical error structure as the distribution will be highly dependent on the rating structure of the individual company. This unnecessarily adds another level of uncertainty to the modeling process.

Judgment is an important part of the modeling process; therefore, it is helpful if the modeler is able to formulate some a priori expectation. When modeling with a pure premium approach, the practitioner can break the data into frequency and severity components and use their knowledge to formulate reasonable expectations. For example, when modeling auto collision frequency, the modeler may expect the age curve to decrease from youthful to adult and increase again for the most mature drivers.[4] Figure

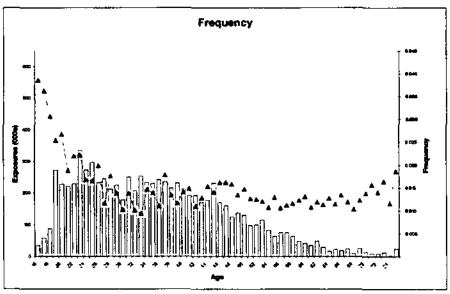


Figure 4. Raw Frequencies By Age

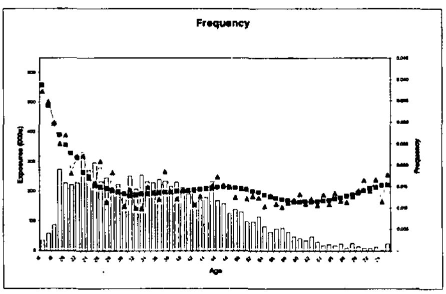


Figure 5. Raw and Modeled Frequencies By Age

[4] Depending on assignment rules, there may be a hump in the middle around the age that teenage drivers are added to the policy.

4 shows an example of raw frequency data by age of driver. To the extent that the pattern is erratic, the modeler will be able to use appropriate techniques (e.g., fit a curve or group levels) and knowledge about insurance to build a model that is captures the signal in the data (see Figure 5). If, on the other hand, the modeler is modeling loss ratios, the only expectation is that the loss ratios will be the same if the rates are perfect. Given that the current rates are probably not perfect, the modeler cannot know whether a resulting erratic pattern in the age results is due to random noise or is a very real pattern due to the underlying rates. Figure 6 shows the raw loss ratios by age for the same dataset. It is clear that it would be difficult to distinguish the noise from the true signal in the data.

Another practical issue is that loss ratio models become obsolete as soon as any change is implemented. Thus, the loss ratio models built during one review cannot be used as a starting point for any subsequent reviews. In contrast, the frequencies and severities of individual observations do not change just because a rate adjustment is made. Thus, the frequency and severity models built with one review should be a very good starting point for the next review.

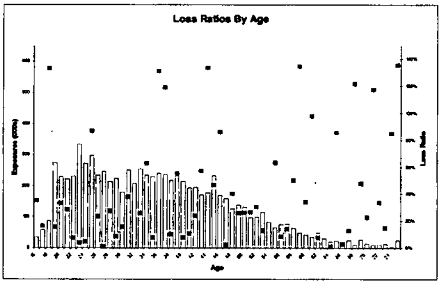


Figure 6. Raw Loss Ratios By Age

If loss ratio modeling has these issues, why do companies do it? There are three basic reasons that practitioners may model loss ratios.

1. Premiums may be readily available and exposures are not. As it is generally easier to obtain exposures than premiums this situation is rare.
2. When modeling pure premiums it is important to include all the variables in the modeling dataset.[5] In some cases, the practitioner may not be able to get all of the variables on the database and uses the premiums to try to account for some of the variation in the missing variables. This, of course, assumes the premiums reflect the true indication, which is often not the case. Additionally, this is only appropriate if the premiums are on level at the granular level.
3. The most common reason seems to be precedent. Historically speaking, companies were performing univariate analysis techniques. With univariate techniques, using loss ratios is more accurate than using pure premiums as loss ratio analysis does a better job of coping with distributional biases in the univariate world. Because of this, companies seem to have gotten into the habit of working with loss ratios. Now that companies are performing multivariate analysis, the reasons that loss ratios outperformed pure premiums in the univariate world are no longer applicable.

So, for these practical and theoretical reasons, loss ratio modeling should only be employed out of necessity. Instead, companies should pursue pure premium modeling at the component level as it will lead to better models and ultimately increased benefits.

[5] At a minimum the practitioner should include all variables that have significant correlation with other independent variables.

COMPONENT LEVEL MODELING

Once a company understands the advantages associated with pure premium modeling relative to loss ratio modeling, the next question is whether the company is going to model frequency and severity separately by cause of loss.

For years, analysts have been performing traditional loss trend analysis. When the data is available, traditional analysis is typically done by coverage or peril and for frequency and severity separately as that is the most effective way to discern underlying trends in the data. Despite that precedent, many companies try to build predictive models on a combined level.

Many companies, especially personal lines homeowners and commercial lines insurers, are tempted to model all coverages or perils combined to save time. Most homeowners insurers can separate the data by peril and should do so. While commercial lines carriers may not have the option of completely separating the data, workers compensation carriers have medical versus indemnity readily available and general liability carriers know property versus liability losses. If a practitioner wants to model on a combined basis, a quick analysis of the residuals highlights whether or not the different perils or coverages can effectively be combined. Figure 7 is an example of a plot of the residual (i.e., actual – predicted) for every observation. The x-axis represents the fitted value and the y-axis represents the magnitude of the residual. In this example, the plot has two separate concentrations. This appearance is typically seen when multiple causes of loss are included and the model is not effectively handling them. If this pattern is seen, the data should be separated and the coverages/perils should be modeled individually. If it is not possible to separate the data, then the practitioner should consider employing dispersion modeling techniques. Dispersion modeling is an advanced topic that is beyond the scope of this paper.

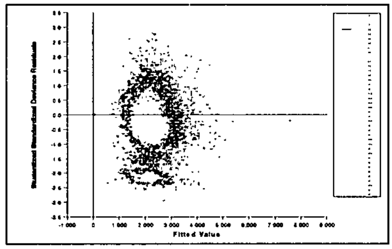


Figure 7. Residual Plot

A more common shortcut companies attempt is to model raw pure premiums directly rather than modeling frequency and severity separately. Interestingly, the anticipated time savings is usually not achieved. When modeling separately, severity modeling is usually very straight-forward and takes little time to find simple trends amongst the noise. More time is usually spent building a good model on the stable frequency data. Practitioners may believe there is time savings based on an assumption that there is only one model being built, but if raw pure premium modeling is done properly, that is not the case. In order to properly reflect the bimodal nature of pure premiums, the analyst should also build a dispersion model that coincides with the Tweedie model.[6] So, whether modeling frequency and severity separately or modeling pure premiums, the analyst is still building multiple models. But with raw pure

[6] A discussion of dispersion modeling is outside the scope of this paper. For more information on this refer to "Fitting Tweedite's Compound Poisson Model to Insurance Claims Data: Dispersion Modelling" written by Gordon Smyth and Brent Jorgensen and published in Astin Bulletin Volume 32, Number 1.

premiums, the practitioner has two challenges. First, the "noise" created by combining frequency and severity makes it more difficult to spot trends in the raw pure premiums. Second, the practitioner has the added complexity of modeling and interpreting the dispersion parameter.

LIMITATIONS ON THE ANALYSIS

When performing a class plan analysis, companies are often tempted to restrict the model to include only those variables that are currently in their rating algorithm. Additionally, for selected variables, companies often group the data to be consistent with the current rating algorithm. For example, if the rating algorithm charges the same rate for ages 30-49, companies will often automatically group ages 30-49 in the model.

The best approach for this type of modeling is to make use of all available data for the initial modeling. This includes importing all the current and potential rating variables, all available underwriting data, and any external data, if available. Some companies will interrogate 200-300 variables. These are the companies that are most likely to find the next really predictive variable.

Figure 8 represents the process when the data is being fully interrogated. The individual frequency and severity models should be built without regard to the current rating algorithm. In other words, the analyst's goal should be to use all available data to build the most predictive frequency and severity models possible. It is not necessary that the frequency and severity models be consistent. In fact, in all likelihood, the model structures for frequency and severity will be different.

After the frequency and severity models are built, the resulting frequency and severity predictions can be combined on an observation by observation basis to form modeled pure premiums. If the underlying component models are built correctly, the unsystematic variation will be removed and the modeled pure premiums will represent the systematic variation in the historical data. The practitioner then builds a pure premium or constraint model using the modeled pure premiums. It is at this stage that the practitioner limits the variables to those that will be used in the proposed rating algorithm, incorporates restrictions on rating variables, and develops underwriting rules that compliment the rates.

By limiting the data to existing variables, companies are limiting their opportunities for major improvement. The goal should not only be to improve the accuracy of existing rating structures, but also to find that new variable that is predictive and can provide a real competitive advantage.

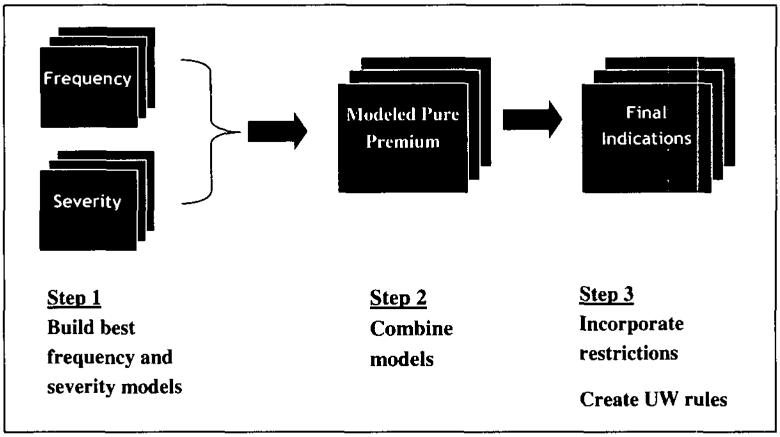


Figure 8. Modeling

INTELLIGENT USE OF OFFSETS

Offsets can be used to specify known relationships between levels of a specific factor or factors within the GLM; consequently, offsets are frequently used when the practitioner wants to fix the relativities of one or more rating factors due to some internal or external constraints. For example, a company may not want to change the multi-policy discount, so the practitioner may use an offset to force the GLM relativities to be consistent with the current discount for that factor. When an offset is used the data is adjusted so as to force the resulting parameters for that variable or variables to the desired values. The parameters for the other variables will change to try to "make up" for the difference and avoid any double-counting. Variables that are highly correlated with the offset variable(s) will change the most.

Companies who do not truly understand the implications of using offsets fail to consider there may be situations where offsets are inappropriate and, therefore, end up with unwanted results. In reality, there are situations when using offsets may be desirable and there are situations when using offsets may be undesirable. It is reasonable to assume that amount of insurance (AOI) and territory are highly correlated (i.e., homes in a particular area tend to have similar AOIs). In light of this, consider the following two examples.

1. Consider the case that a systems constraint forces the analyst to cap the relativity on homes with AOIs over $500K. This represents an undesirable subsidy that the practitioner may want to minimize. The practitioner has two options. On one hand, the practitioner can calculate the indicated GLM relativities without any offsets and cap the indicated relativity for homes over $500K. The impact of this is that the base rate will need to be adjusted to make up for the shortfall. Alternatively, the practitioner can use an offset term. By doing so, the other variables adjust to try to make up for the shortfall. The most significant changes will occur in variables that are highly correlated with the variable being constrained. In this example, the relativities in territories with a relatively heavy concentration of high-valued homes will increase to make up

for the shortfall. Note, as no variable will be perfectly correlated, a minor base rate adjustment will likely be required.

2. Assume the company makes a decision to target high-valued homes. In an effort to increase market share, the practitioner is instructed to implement relativities lower than those indicated for homes $500K and over. In this case, the practitioner does not want to use offsets. As discussed in the preceding paragraph, using offsets increases the relativities in the territories with high-value homes. Since it is immaterial whether the premiums are high due to the AOI curve or the territorial relativity, using an offset simply undoes the desired subsidy. So, if the practitioner is trying to implement a desirable subsidy, then the free-fitted relativities should be left in the model and the desired relativities should be changed outside the modeling process.

The impact of offsets is particularly important to consider in the initial stages of multivariate analysis implementation. At the onset, there are usually significant differences between the current rating plan and that indicated by the multivariate analysis. Due to regulatory constraints and renewal impact considerations, it may be difficult for a company to make all of the changes at once. If that is the case, the company needs to decide whether or not to use offsets for the variables that are not going to be changed. The following example is intended to illustrate the impact of using offsets.

The results of a full GLM analysis will be indicated relativities for all factors. Figure 9 represents the indicated and current relativities for 0, 1, 2, 3, and 4+ years of claims free driving[7] and figure 10 depicts

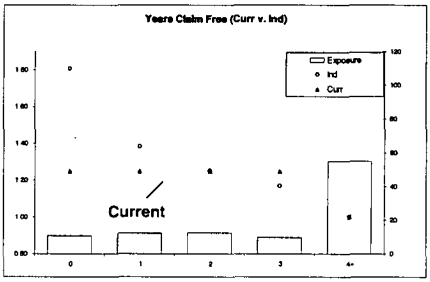


Figure 9. Indicated Years Claim Free Relativities

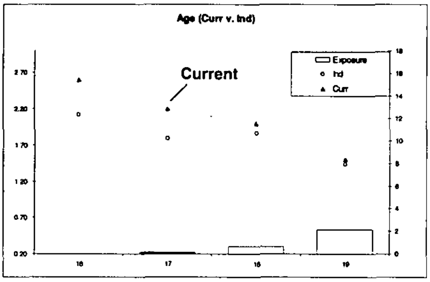


Figure 10. Indicated Age Relativities

the indicated relativities for drivers age 16-19. Assume that the company wants to implement a change to the youthful relativities with this review, but wants to keep the same relativities for number of years of claims free driving until the next review.

The modeler must decide whether to use an offset to minimize the subsidization introduced by maintaining the current relativities for years of claims free driving; if an offset is used, the other factors will adjust to make up for the subsidy. The amount the other factors will adjust is dependent on how correlated the other factors are to the years of claims free driving. Figure 11 is a graphical

[7] In this example, 4+ years of claim free driving is the base (i.e., a factor of 1.00). Thus, there are surcharges (i.e., relativities above 1.00) for 0, 1, 2, and 3+ years of claims free driving.

representation of the exposure distribution by operator age (x-axis) and years claim free (y-axis). Each of the bars stack to 100% with each segment representing a different number of years of claim free driving. If the variables are not correlated, then the segments are consistent for each bar. As can be seen by figure 11, age and years claim free are very highly correlated. More specifically, the younger ages (left side) have a high concentration of drivers with 0 and 1 years of claim free driving and the older ages (right side) have a high concentration of drivers with 4+ years of claim free driving. Thus, the inclusion of an offset for claims free driving will likely have an impact on the age relativities.

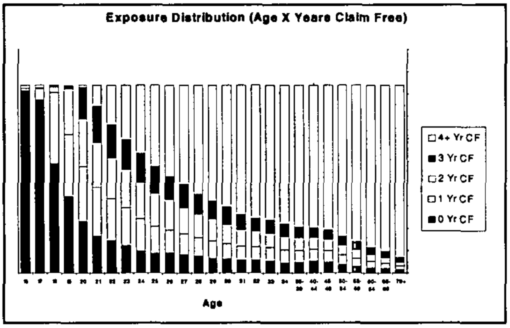


Figure 11. Distribution by Age and Years Claim Free

Figure 12 is the same chart as figure 10 with the addition of the indicated age relativities after the inclusion of the offset for years of claims free driving. In this case, the indicated relativities for ages 16-19 increased significantly to "make up" for the shortfall caused by not implementing the fully indicated surcharge for only 0 and 1 years of claim free driving. This example actually highlights the interesting case where the current is actually between the indicated and indicated with offset. Assuming the company looks at both indications, the company will realize there is an important business decision to be made. First, the company can move toward the indicated with offset as that will minimize the inequity in the rating plan. The con associated with that approach is that if the claims free driving relativities are changed with the next review, the age relativities will need to be lowered to be equitable. This leads to an increase in the age relativities in one year and a decrease in the next, which increases the chances of big premium swings for individual risks. On the other hand, the company can maintain the current age relativities or even move toward the indicated relativities without the offset. This has the benefit of reducing the potential for big premium swings, but does not correct the short run inequities in the rating plan.

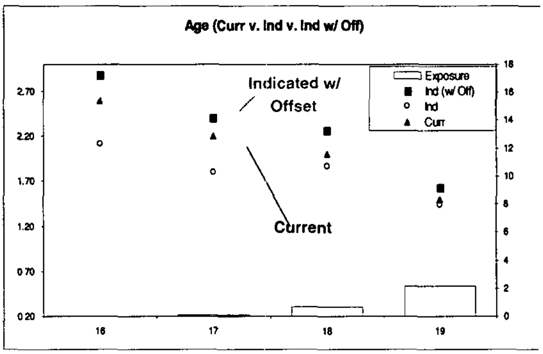


Figure 12. Ages 16-19

The point of these examples is not to encourage companies to avoid offsets. Offsets are a very important aspect of modeling with GLMs. Rather, it should be noted that offsets need to be used with an understanding of the effects so that the practitioner does not get unintended consequences. One way to prevent unintended consequences is to view indicated relativities with and without the offsets.

GLMS ARE NOT A BLACK BOX

One of the advantages of GLMs as compared to other predictive modeling techniques (e.g., neural nets) is that GLMs are not black boxes. Unfortunately, too many practitioners make the mistake of treating GLMs like black boxes. This results in models that are inappropriate and blind implementation of results that may be counterintuitive.

Figure 13 depicts the iterative nature of modeling with GLMs[8]. While certain tasks can be automated, it is important to understand that--when built properly--GLMs require involvement by the analyst in the process. The very process of building GLMs provides the analyst with valuable insights into the data and an ability to provide judgment where necessary.

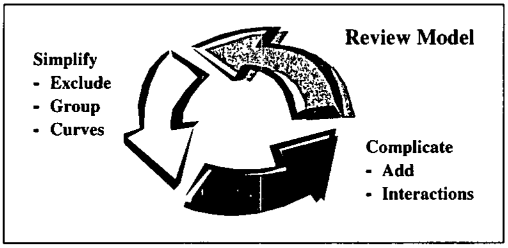


Figure 13. Iterative Modeling Process

One common mistake that companies make is to overcomplicate the models. In other words, the company correctly includes as much data as possible in the model process, but is not appropriately judicious in building the models. A quick review of companies' rating plans highlights that many companies have incredibly complex rating algorithms. It is not uncommon to see variables interacted with many different variables throughout the algorithm. Interestingly, when this is present the relativities included in the various tables tend to all be close to 1.00 suggesting the table was probably an unnecessary complication. This is an area where we should take a lesson from our counterparts in the UK who have been performing this analysis for years. Despite the fact that they have significantly more rating freedom, their rating algorithms are generally much simpler than ours. By examining proper diagnostics (i.e., standard errors, consistency tests, and type III tests) and supplying appropriate judgment throughout the modeling process, the modeler will not introduce unnecessary complications and the resulting rating algorithm will be more predictive and more manageable. Thus, the practitioner will be in a better position to make changes quickly in the future.

In contrast to the companies discussed in the preceding paragraph, there are a significant number of companies who do not introduce enough complexity in the model. A common example of oversimplification is over-smoothing. For example, a company may fit a single curve to smooth out the auto frequency data by age when multiple curves would be a better representation of the true signal. Again, GLM is not a black box. Instead, it is an iterative tool that requires human intervention. While it

[8] The picture shows that the process of building a GLMs involves determining an initial model and then testing to determine what simplifications and complications can be made to improve the performance of the model. Simplifications involve excluding variables that are not predictive, grouping levels within a variable (e.g., ages 75+) that do not add any additional predictive value, and fitting curves to continuous variables. Complications include adding new variables that can help explain variation and incorporating interactions which allow the relationship between levels of one variable to vary by level for another variable (e.g., the relationship between males and females varies by age). The circular nature is intended to show that it is an iterative procedure. As it is multivariate analysis, decisions made at any stage can impact previously made decisions.

may be wise to start slow, the practitioner should ultimately progress to the point where he/she is using the modeling as an exploratory opportunity to identify new material patterns in the data.

The other major mistake that companies make in this regard is blindly implementing results without applying appropriate judgment. A quick review of rating pages of various companies will often uncover patterns in relativities that seemingly make no sense. Figure 14 shows a comparison of indicated and current model year relativities. The current relativities are consistent with expectations as they decrease as the model year gets older. Interestingly, the indications suggest the true risk is better represented by a hump-shaped curve with the highest relativities being in the middle. This, of course, is counterintuitive and the practitioner should dig deeper into the true cause of this pattern. Since these indicated pure premium relativities were derived from the combination of frequency and severity models, the practitioner will gain insights and can work with experts from various areas within the company to determine if there are operational processes that are driving the pattern. If the overall pattern is a frequency phenomenon, then it is prudent to start with the underwriting function. For example, a pattern like this could result from underwriting rules that are especially strict for newer vehicles. If, on the other hand, the overall pattern is really a severity phenomenon, then it is prudent to start with the claims function. For example, a pattern like this could result from claims adjusters who are being relatively generous when settling claims for middle aged vehicles. By understanding the results and having cross-functional discussions about the results, the company can address the issue using the most appropriate lever (pricing, underwriting, or claims).

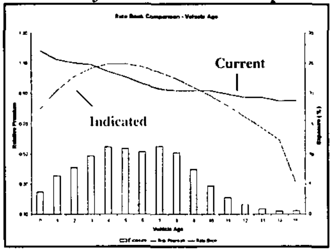


Figure 14. Indicated v Current

GLMS ARE A BUSINESS TOOL

As GLMs are generally being championed by actuaries, most of the focus has been on using GLMs to determine relativities for rating structures. Even in this paper, the examples have focused on risk modeling. However, the benefits of GLMs are clearly not restricted to the application of pricing. The following are a few of the other applications for which companies are already using GLMs:

- Practitioners are using GLMs to reduce a variety of risk variables into one score. This has obvious application in regards to creating underwriting tiers, credit scores, fire protection scores, vehicle symbols, etc.
- Many companies have begun to perform elasticity modeling. By building elasticity models for new and renewal business, companies can predict the impact of various actions on market share. A few companies are already linking the profitability and elasticity models to find the optimal pricing decision.
- Claims handlers are starting to see the advantages of GLMs and are using them to help set more accurate reserves and to provide early identification of claims that may be fraudulent or are most likely to end up in a lawsuit.
- Competitive analysis units are using GLMs to reverse-engineer competitors' rates given a large sample of rating quotes.

These are just a few of the other ways that the more advanced companies are using GLMs. Companies should not fall into the trap of thinking GLMs are only for pricing. Instead, companies should realize that GLMs can be used in a variety of circumstances given a historical database with multiple predictors and a response they would like to predict.

CONCLUSION

Predictive modeling is a very powerful tool that has been used effectively by insurance practitioners in other countries for many years. Starting in the 1990's many of the larger US personal lines carriers began to implement predictive modeling techniques in the form of generalized linear modeling (GLM). The early results from the US show that predictive modeling paid large dividends to those companies who embraced it.

Due largely to the success of the first US companies, the US is experiencing a push by many companies who want to implement predictive modeling before they are left behind. With that push companies are already getting benefits. Unfortunately, in a haste to get going, many companies have taken some shortcuts or simply made mistakes that have kept them from realizing the full benefits. By taking a step back and reconsidering some of the prior processes and decisions, companies will be able to maximize the benefits of these new techniques.

Survey of Data Management and Data Quality Texts

CAS Data Management and Information Educational Materials Working Party

Abstract:
Motivation. Recent focus on corporate governance (e.g., Sarbanes-Oxley) in the United States and the use of predictive modeling techniques in the property/casualty insurance industry have raised the profile of data management and data quality issues in the actuarial profession.
Method. Representatives of the Insurance Data Management Association (IDMA) identified seven data management texts they felt would be most helpful for actuaries. Two additional texts were added to fill out the data quality perspective.
Results. Actuaries reviewed each of the recommended texts from an actuarial perspective.
Conclusions. The working party hopes that this paper will be a resource for actuaries dealing with data management and/or data quality issues. By looking at the summary information in the tables of section 4, readers may be able to narrow down candidate books to those that will best meet their needs and then read the specific reviews in section 3.
Keywords. Data Quality; Data Administration, Warehousing and Design; Actuarial Systems; Data Collection and Statistical Reporting; Software Testing.

1. INTRODUCTION

Recent focus on corporate governance (e.g., Sarbanes-Oxley) in the United States and the use of predictive modeling techniques in the property/casualty insurance industry have raised the profile of data management and data quality issues in the actuarial profession. Actuaries have a unique role with respect to data quality because they typically understand the process and pitfalls better than management and at the same time they understand the business meaning and impact of errors better than data and systems professionals. For example Francis [1] points out that 80% or more of time spent on large predictive modeling projects is spent on data issues. Also, in December 2004, the Actuarial Standards Board updated their standard of practice on data quality (Actuarial Standard of Practice No. 23) [2].

This paper provides an overview of several resources on information quality by surveying seven non-actuarial data quality and data management textbooks recommended by the Insurance Data Management Association. Two additional texts recommended by a working party member are also reviewed. A discussion of data quality is incomplete without reference to related data management topics such as data structure, data storage, metadata, and software errors. In this paper we will employ the term "information quality" to refer to the broader set of data management topics related to data quality.

Thus, this paper provides resources for actuaries with data management or data quality questions and these resources may provide suggestions to improve data quality. Hopefully we will motivate our readers to pursue further education on information quality using one or more of the books surveyed. In addition, within our review of the literature, we present an overview of some of the key concepts of information quality.

1.1 Research Context

The actuarial literature on data quality and data management is relatively sparse. The Actuarial Standard Board (ASB) Standard of Practice No. 23 on data quality [2] provides a number of guidelines to actuaries when selecting data, relying on data supplied by others, reviewing and using data, and making disclosures about data quality. The guidelines advise actuaries to review data for reasonableness and consistency. The actuary is also advised to obtain a definition of data elements in the data, to identify questionable values and to compare data to the data used in a prior analysis. The actuary is also advised to judge whether the data is adequate for the analysis, requires enhancement or correction, requires subjective adjustment, or is so inadequate that the analysis cannot be performed.

The Casualty Actuarial Society (CAS) Committee on Management Data and Information and the Insurance Data Management Association (IDMA) also produced a white paper on data quality [3]. The white paper states that evaluating the quality of data consists of examining the data for:

- Validity,
- Accuracy, including concepts of absolute accuracy , effective accuracy and relative accuracy,
- Reasonableness, and
- Completeness.

The CAS Committee on Management Data and Information also promotes periodic calls for papers on data management and data quality which are published in the CAS *Forum*. Among the papers on information quality submitted to the program are Francis [1] and Popelyukhin [4]. Francis's focus is mainly on techniques from exploratory data analysis that can be applied by actuaries to detect glitches and other data quality issues in data supplied for an actuarial analysis. Popelyukhin describes data quality issues encountered by actuaries when relying on data supplied by non-actuaries and external data suppliers, such as that supplied by third party administrators and presents the data quality shield as a solution to insurance data quality problems.

The subject of data quality is also of interest internationally. A working party of the U.K. General Insurance Research Organization (GIRO) developed recommendations for improving the quality of reserve estimates. The Reserving (GRIT) working party report [5] recommended more focus on data quality and suggested that U.K. professional guidance notes incorporate standards from U.S. Actuarial Standards of Practice (ASOP) No. 23. Furthermore the GRIT survey found that many respondents expressed concern over data quality.

Note that none of these references specifically addresses data management.

1.2 Objective

The objective of this paper is to address gaps in actuaries' knowledge of information quality. The current CAS literature provides a basic introduction to information quality issues. However, there is very little available for those wishing a more advanced knowledge of the subject or for those who have responsibilities involving data management and data validation. Moreover, the current state of actuarial literature on information quality does not equip actuaries to become active advocates for information quality; i.e., to advise management on systems and protocols for improving information quality. This paper attempts to narrow this gap by reviewing several recommended books from an actuarial point of view.

1.3 Disclaimer

While this paper is the product of a CAS working party, its findings do not represent the official view of the Casualty Actuarial Society or the employers of the Working Party members. Moreover, while we believe the textbooks reviewed here are good sources of educational material on data management and data quality issues, we do not claim they are the only appropriate ones.

1.4 Outline

The remainder of the paper proceeds as follows. Section 2 will discuss the increasing importance of data management and data quality to actuaries, as well as how the reading list was developed. Each subsection of section 3 is a book review of one text. Section 4 summarizes and compares the working party's evaluations of the textbooks on five star rating scales.

2. BACKGROUND AND METHODS

2.1 Motivation

Information quality issues have come to forefront recently due to several key developments:

- **(Unprecedented) level of detail.** Computerization and cheap data storage along with changes in regulatory requirements have led to extraordinary amounts of data being captured, stored and provided to actuaries. Consequently, enormous amounts of data can amass enormous numbers of errors and inconsistencies.
- **Availability of new tools.** Recent years have seen the proliferation of powerful data analysis packages and technologies: from XML-enhanced data exchange to object-oriented databases to servers enabled with On-Line Analytical Processing.
- **Competition.** Competition encourages pricing techniques to be more and more precise – every percent counts. The precision of estimates is heavily dependent on the quality of the data used in the analyses. In this environment, requirements for quality of data used in pricing algorithms grow immeasurably.
- **Quality of actuaries.** Modern actuaries are more technically prepared for the challenges of dealing with huge amounts of data using contemporary tools and techniques. Prepared with the appropriate information, they should be able to tackle data quality issues with aplomb.

2.2 The Reading List

To address these issues, the CAS Committee on Data Management and Information created the Data Management Educational Materials Working Party. A casual search will reveal dozens, if not hundreds, of books on data management. The Insurance Data Management Association (www.idma.org) promotes insurance data management in multiple ways, including accreditation, online courses, information available on their website, seminars, and co-sponsoring forums. Knowing this, the working party asked the IDMA to develop the party's reading list. IDMA representatives narrowed down their syllabus to the texts they felt would be most helpful for actuaries. Louise Francis, four time winner of the CAS Data Management Call Paper program, suggested two additional texts to fill out the data quality perspective.

3. THE BOOK REVIEWS

Each of the following sections is the review of one book. The sections are ordered as the underlying texts' focus moves from data quality (3.1, 3.2) to data management (3.3 to 3.7) to special topics (3.8 and 3.9). Some of these reviews have already been published in the *Actuarial Review*. For *Corporate Information Factory* (section 3.5), the text has been altered slightly from that in the *Actuarial Review*. The texts are compared in section 4, so readers may find it helpful to skip to section 4 to determine which text(s) best address their issue(s).

3.1 Data Quality: The Accuracy Dimension

Data Quality: the Accuracy Dimension [6] by Jack E. Olson (ISBN 1-55860-891-7) focuses on data accuracy, which the author sees as the foundation for the measurement of the quality of data. The author has spent the last 36 years developing commercial software and is an expert in the field of data management systems. This background enables him to address the topic of data quality and accuracy from a practical viewpoint.

There are three parts to this book. The first part defines inaccurate data and shows that many significant business problems arise from inaccurate data. The second part focuses on how a data quality assurance program is constructed using the "inside-out" approach. The last part introduces data-intensive analytical techniques such as data profiling (the use of analytical techniques to discover the true content, structure and quality of data), along with some real world examples of profiling applications.

The author begins the first part, "Understanding Data Accuracy," by introducing real world data quality problems and the concept of data quality assurance technology. The author identifies the essential elements of this technology: experts, educational materials, methodologies, and software tools. In order to define data accuracy in the larger picture of data quality, data is defined as "having quality if it satisfies the requirements of its intended use." Some examples are used to illustrate key aspects of data quality:

- **Accuracy:** An 85% accurate database containing names, address, and phone numbers of physicians in a state would be considered poor quality for notifying physicians of a new law whereas it would be considered high data quality for a new surgical device firm to find potential customers.
- **Timeliness:** A dataset containing monthly sales information which is slow to become complete at the end of each month is poor when it is used to compute

sales bonus in that month whereas it is excellent when it is to be used for historical trend analysis.

- **Relevance:** A dataset without relevant information is of poor data quality for its intended use.
- **Completeness:** A database with 5% of information missing is probably a good quality database for general assessment but is considered to be low quality for evaluation.
- **Understood:** Dataset has to be understood for its intended purpose. *Metadata* is a term used by data management professionals for information about the data such as definitions, a description of permissible values and business relationships that define the data in a database. Comprehensive metadata is a prerequisite for good information quality.
- **Trusted:** Only trusted datasets should be used.

Data accuracy, "the most visible and dramatic dimension of data quality," is then introduced and explained. Data accuracy "refers to whether the data values stored for an object are the correct values." "To be correct, a data value must be the right value and must be represented in a consistent and unambiguous form."

The second part of the book outlines the structure of a data quality program built for identifying inaccurate data and taking actions to improve its accuracy. "A data quality assurance program is an explicit combination of organization, methodologies, and activities that exists for the purpose of reaching and maintaining high levels of data quality." An **inside-out** methodology is believed to be the best way to address accuracy. This method works from a complete and correct set of rules that define data accuracy for a particular dataset. The author defines "inaccurate data evidence" as a collection of facts which are aggregated into issues. The facts might include tabulations of the number of invalid values for variables in the data, totals of the number of missing values, etc. This evidence is produced by the data profiling process defined above. The issues are then analyzed to determine the external impact.

The second approach, the **outside-in** method, looks for issues in the business rather than looking at data. "It identifies facts that suggest that data quality problems are having an impact on the business." The facts are then examined to determine the degree of culpability attributable to defects in the data and if the data has inaccuracies that contribute to the problem.

Summarizing the two approaches to data quality programs (page 72, *fig.* 4.3):

Inside-out work flow
Data → Issues → External impacts and data entry processes

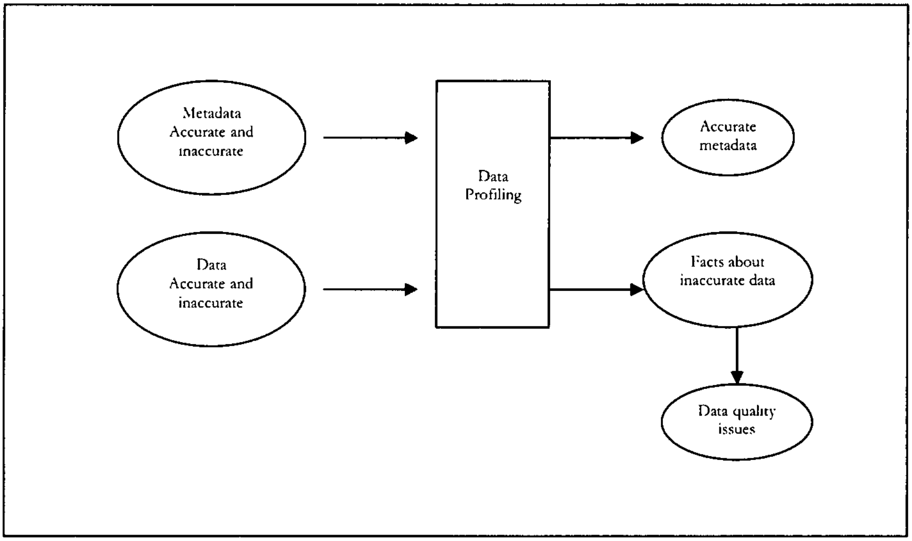
Outside-in work flow
External evidence → Issues → Data and data entry processes

The data quality assurance program also requires an assurance team to decide how it will engage the corporation to bring about improvements and return value for their effort. The author advocates that team members should only be assigned to the data quality assurance team, i.e., this is their full time job -- not a project.

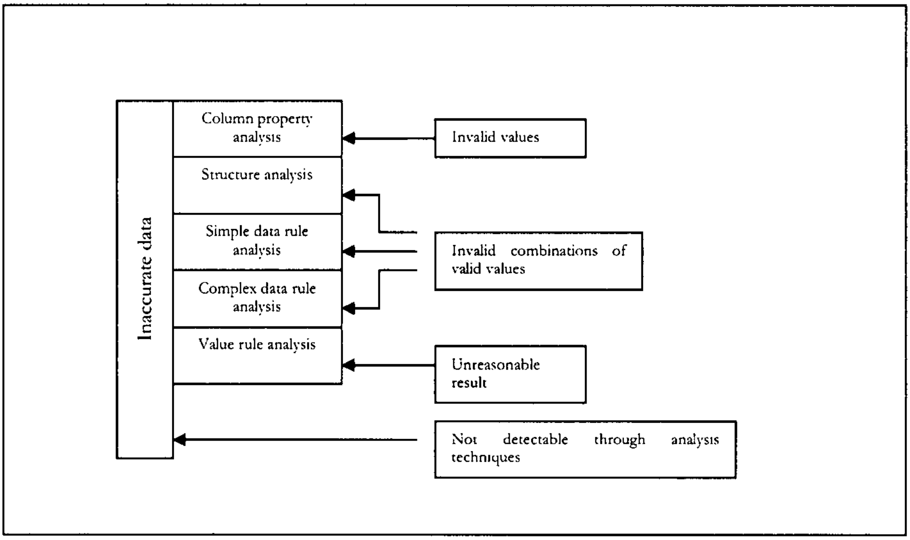
Some of the key technologies used to create and maintain an effective data quality assurance program are:

- **Metadata repositories:** metadata should define what constitutes accurate data. It is essential for determining inaccuracies in data profiling.
- **Data cleaning:** identifying and cleaning up data after data problems have been discovered. It is valuable to clean up data before moving to the next step of data profiling to avoid distortions in the discovery processes of later steps.
- **Data profiling:** the use of analytical techniques to discover the true structure, content, and quality of a collection of data.
- Data filtering
- **Data monitoring:** looking at individual transactions before they cause database changes or looking at the entire database periodically to find issues.

"Data profiling is a new technology that has emerged in the last few years." It uses any known metadata and the data itself to discover the presence of inaccuracies within a database. The general model of a data profiling process can be shown as follows (page 123, *fig.* 7.1):



Data profiling uses a bottom-up approach. It starts at the most basic level of the data and then goes to progressively higher levels of structure. The following diagram (page 131, *fig.* 7.2) illustrates how the major steps of data profiling (in the middle column) can address data issues (in the right hand column):



Within each data profiling step there can be processes for discovery, assertion testing, or value inspection. The outputs of these processes are used to make decisions. The author discusses each step in a separate chapter with real world examples of the rules and the types of investigative thought required to be effective. The author believes data profiling is probably the single most effective technology for improving the accuracy of data in corporate databases.

Overall, the book provides a thorough introduction to data accuracy and the data profiling technology that could significantly improve data quality. A reader could probably develop a data quality assurance program including data profiling after reading the text, although there is not much on statistical methodologies commonly used to detect data problems. However it does serve as a good reference for data quality structures and concepts.

3.2 Exploratory Data Mining and Data Cleaning

The primary topic of the book *Exploratory Data Mining and Data Cleaning* [7] by Tamraprni Dasu and Theodore Johnson is data quality. In data mining circles this book is the reference of choice on data quality and its authors are invited to speak on the topic at many conferences. It combines a review of the most common methods used for screening data for quality with some novel approaches developed by the authors. It also provides a review of key data quality concepts along with some data management concepts relevant to data quality.

An overview chapter summarizes the topics covered in the rest of the book and presents the authors' philosophy towards data quality. The authors lay out the methods of exploratory data mining they will be using: These include parametric summaries (measures of central tendency, dispersion and skewness), as well as non-parametric summaries such as quantiles, histograms and OLAP cubes. The authors believe in "end-to-end-data-quality," that is there are many stages in the data assembly process where data quality needs to be monitored and improved, such as during data gathering, data storage, data analysis and data integration. Their equation:

DATA + ANALYSIS = RESULTS

reflects in equation form the well known adage "garbage in – garbage out." The authors are also proponents of measuring data quality in order to promote data quality improvement.

The book has a chapter on "Exploratory Data Mining" that presents graphical and statistical techniques largely from the exploratory data analysis literature. The methods of exploratory data analysis were pioneered and the practice given its name by John Tukey (see *exploratory data analysis* at www.wikipedia.org). Its methods are widely accepted in the statistical community as a key activity within any statistical project and its methods are widely implemented in statistical software. Exploratory data mining is an application of exploratory data analysis to large databases that can be used to understand the structure of a database and to detect outliers (data glitches are often found by examining outliers). In this chapter, the authors introduce the novel concept of data depth. Data depth provides a measure of how far a record is from the center of the data or from typical data values. In order to construct such a measure, one needs a way to quantify the notion of "center" and the notion of "distance" from the center. The authors provide the Mahalanobis depth as one way to measure data depth.

In the chapter "Partitions and Piecewise Models" the authors discuss data cubes as a mechanism for exploring data. Data cubes are single or multidimensional tabular summaries of data. Statisticians have long used cross-tabulations, or slicing and dicing of data to develop a high level understanding of the structure of databases. Among practicing actuaries, pivot tables are a common example of data cubes. In this chapter, the authors introduce the concept of data pyramids for comparing two databases for changes. Unfortunately, this concept was a little difficult to follow, even after a couple of readings of the material. The authors also introduce two data mining methods which can be used to model nonlinearities and other data complexities in this chapter: piecewise regression and naïve Bayes.

In their chapter on Data Quality, the authors detail all the mishaps affecting data that create quality problems. Some of the sources of data quality problems are: unreported changes in layout, unreported changes in measurement, temporary reversion to defaults, missing and default values and gaps in time series. Being mindful of the sources of data errors, one can detect, remediate and most importantly, prevent them.

In the Data Quality chapter the authors are strong proponents of implementing data quality measures. The authors believe that in order to motivate improvements in data quality, it is imperative that data quality be measured, even when the measures are somewhat subjective. In developing their measurement approach, both static and dynamic constraints are described. Some of the metrics quantify traditional data quality components such as accuracy, consistency, uniqueness, timeliness and completeness. Others capture other features of data quality such as extent of automation (sample some transactions, follow them through the database creation processes and tabulate the number of manual interventions), successful completion of end-to-end processes (count the number of instances in a sample that, when followed through the entire process have the desired outcome), and glitches in analysis (measure the number of times and severity in a sample that data quality errors cause errors in analyses). The different metrics are weighted together into an overall data quality index using business considerations and the analysts' goals to develop weights.

The book provides a wrap-up chapter that applies the authors' quantitative techniques to the detection, correction and prevention of data quality problems. In the chapter, methods for detecting and correcting glitches are illustrated. For instance, to address the missing value problem, the authors present techniques (including data imputation) that can be used to create values that substitute for the missing data. The chapter presents an introduction to techniques for joining different data sets, including approximate joining techniques when exact matches are not found between the key fields of two databases. Finally and most importantly, the authors also stress the crucial role of metadata, the information describing the data, and discuss ways of creating good metadata.

Overall, the book provides a thorough introduction to data quality at a level that can be understood by the practicing actuary (with the exception of the material mentioned above on data pyramids).

3.3 Improving Data Warehouse and Business Information Quality

Improving Data Warehouse and Business Information Quality [8] (ISBN: 0-47125-383-9), by Larry P. English, is a complete detailed treatment of information quality for any type of business. The main theme in this book is that data is a material for informational product and (like in manufacturing) the quality of the product is determined by customer satisfaction. According to the book, *everyone* in the organization has a role in establishing and maintaining information quality to deliver a quality product to the customer. Thus actuaries, as consumers and producers of information, should establish data quality standards and communicate their data quality requirements to the stewards of all their data sources.

The book is multifaceted; it is "a concept book, a textbook, a reference book, and a practitioner's guide." It is generic enough to cover a lot of ground (scenarios, situations, setups) while detailed enough to serve as a step-by-step guide full of relevant examples. Throughout the book the author consistently uses a 4-part template for every proposed step (Input, Output, Techniques & Tools and Process Description) which makes the text immensely useful.

The book is divided into three sections: "Principles of Information Quality Improvement," "Processes for Improving Information Quality" and "Establishing the Information Quality Environment."

In section one, "Principles of Information Quality," the author lays the ground work by defining what data is, what quality is and is not and why we should be interested in information quality in the first place. He then builds upon this foundation work with detailed discussions about the high cost of low data quality and how to measure data quality with detailed examples. He continues with a discussion of quality principles applied to information as a product and each stakeholder's role in producing, planning, controlling, leading, funding, and continuously improving information.

Section two uses many flow diagrams to demonstrate the various process steps for improving information quality. For example, there are diagrams to show the steps in measuring non-quality information costs, establishing the information quality environment, establishing data quality definitions, and assessing data quality. The chapter on data definition and information architecture quality is particularly detailed as the author provides instructions on how to construct data names, build metadata repositories, and provide guidelines for quality business rules. The chapter on information quality assessment shows how to determine sample size and also includes numerous quality assessment templates to show different ways quality measurements and customer satisfaction can be presented. The author places great emphasis on data defect prevention through the process of continuous improvement as "the cost to react to quality problems can be 5 to 10 times as much as the cost of prevention."

Section three shows how "Deming's 14 points of quality" can be applied to the information product. It describes the roles and accountabilities of everyone in the organization, from information producer to executive management, as stewards of information quality. The author points out that management commitment is essential to having a quality improvement environment. He then describes how to start implementing it step by step, including: "creating a vision and objectives, identifying critical success factors, managing change, conducting an information customer survey, selecting a small manageable pilot project, defining the business problem, and assessing the systemic barriers." You clearly get the idea that this is not just about data but about managing processes and people.

With time the book has acquired the flavor of a cautionary tale about obsolete systems. If in 1999 the book was considered to be mostly about cleansing legacy systems and converting them into new shiny-bright data warehouses, nowadays it can be read as a powerful reminder of how to keep systems current and relevant in a constantly changing environment in order to avoid their transformation into "legacy" systems. According to the book, maintaining data definitions and business rules will make long strides into keeping information from becoming legacy data in need of remediation.

The book's content translates directly to the actuarial situation: actuaries rely on many pieces of data (loss runs, premiums bordereaux, claims classification, etc.), which may be quite imperfect. The caveat is that actuaries rarely (if at all) have control over their data, while the book implicitly assumes that the reader can perform the suggested data cleansing and transformation procedures. Nevertheless, the book is very useful: actuaries would definitely benefit from knowing which data defects may cause problems and of what size. Actuaries should determine the types of potential data errors with the largest impact and presumably should be able to estimate the effects they may have on their data. Ideally, actuaries would use data quality assessment reports to calculate the level of data accuracy.

The book is an extremely valuable source of information for anyone potentially affected by data quality. It can be read as a textbook, as a practitioner's guide, as a cautionary tale, or as an inspirational book. Indeed, learning about data quality problems at source level may even inspire actuaries to incorporate an estimate of data uncertainty into their methods. In summary, even though this is a very long book it does contain a wealth of ideas and techniques that can be used by everyone in the information value chain in carrying out their information quality stewardship responsibilities.

3.4 Enterprise Knowledge Management

The purpose of *Enterprise Knowledge Management* [9] (ISBN: 0-12455-840-2) by David Loshin is to provide an enterprise-wide framework for data quality. The author likens the flow of data within an organization to the assembly process in a manufacturing plant, often referring to an organization's data production as "the information factory." The author uses many quality control ideas from the world of manufacturing and applies them to the process of manufacturing information in an enterprise.

The book is divided into chapters each of which outlines one building block of an enterprise data quality program. The book is at once both technically detailed and conceptually rich.

Technical data quality concepts are illustrated by a number of real world data examples. The data examples are not insurance specific, but rather generic, typically using universal business elements such as name, address, location, and phone number. Nevertheless the concepts are universal and especially applicable in an industry like insurance, where data drives the business. The actuary will recognize many of these concepts, described generically in the text, as applicable to the actuarial applications of ratemaking, reserving, or modeling.

While containing some technical details, the text is curiously abstract, relying mostly on high-level conceptual material. It resembles an Actuarial Standard of Practice in that for each topic a list of conceptual considerations and best practices are given, but with few concrete recommendations as to which are most important. That determination is left up to the practitioner's judgment. The text is oriented towards professionals who oversee information flow within an organization: the CIO, the systems manager, or the actuary who oversees information infrastructure.

The author begins with a section on how to build support for data quality management within an organization. The first step is to get senior management buy-in for the program. Start with a small but visible data quality issue. In choosing an initial task, the author invokes the Pareto or "80-20" rule, which states that 80% of the impact is usually generated by 20% of the cases. Quantify both the soft and hard costs of allowing the issue to linger. The author recommends using a process known as COLDQ (cost of low data quality) that maps the information chain, and then builds a Data Quality Scorecard to identify potential problem nodes in the information manufacturing chain.

For instance, if the issue is faulty customer addresses, the associated costs might include hard impacts like the cost to repair data and increased customer service expense; but also soft impacts like increased customer attrition or delay in analysis and initiative implementation dependent on the data. In an insurance setting, these "soft" costs might be manifest in the inability to analyze catastrophe data or to reorganize rating territories, for example. Next, to gain buy-in, demonstrate to management the operational benefit and rate of return associated with fixing the issue. Once the issue is addressed, celebrate the solution and thereby build support and enthusiasm to address further data quality issues. A key component of the solution is to establish a data ownership policy. The author gives many different paradigms for "who should own the information" in various settings, but it should always be formalized and agreed upon.

The author discusses various dimensions of data quality, e.g., completeness, flexibility, robustness, essentialness, granularity and precision, among others, as they relate to data models, data values, information domains, information presentation, and even the corporate information policy itself. One or two indices are given as guidelines for how to compute each measure of data quality. Foe instance, a complete database is one that contains all of the data required for an analysis while the analyst may request additional data be added to an incomplete database. To measure completeness one might chart the number of requests to add new data fields over time.

Once data quality measures and thresholds have been established, they can be measured either statically or dynamically. Static measurement involves collecting and analyzing past data, usually after the end of a time cycle, and is useful for identifying chronic data quality issues. Dynamic measurement involves inserting data probes into the information chain and measuring output in real time. This is useful for identifying acute data quality issues. Data quality measurement is often implemented via a rules-engine containing data and business rules, and acceptable tolerance thresholds for each. The author spends a fair amount of time in listing considerations when evaluating different rules-based systems and products. Often the choice of a particular rules engine will depend upon whether measurements are primarily static or dynamic.

The author then devotes several chapters to data cleansing. Data cleansing is the act of "fixing" data, i.e., appending, supplementing, or overwriting data whose quality has tested low. Often data quality problems arise when merging data from two different data sources. The author describes techniques used to determine if two different data fields' members come from the same domain. The concepts of overlap, agreement, and disagreement are discussed and a formula given for computing the degree of each between two data sets.

If a data domain is unknown (this usually occurs in string fields housed in legacy mainframe data systems), a number of domain discovery techniques are given; among them agglomerative, divisive, hierarchical, and K-means clustering. Each of these clustering-based methods relies on a notion of distance between data points. Distance rules are typically Euclidean (d = ((x1–x2)^2 + (y1– y2)^2)^0.5), city block (d = |x1-x2| + |y1-y2|), or Exact Match. "Exact match" distance rules are used to compare the distance between strings and are extremely helpful in data clustering, data cleansing, spelling, and address checking routines.

One distance rule used to compare strings is "edit distance" or the minimum number of basic operations (i.e. insert, delete, transpose) needed to transform a candidate string to a target string. For example the edit distance between "intermural" and "intramural" is 3.

The author also gives a number of approximate matching techniques to match like strings using the notion of distance in combination with various word and phonetic coding schemes such as: Soundex, New York State Identification and Intelligence System (NYSIIS), Metaphone, and N-grammimg. Each of these methods attempts to simplify the phonetic representation of a word (by omitting vowels, coding like sounds, etc.) and then uses the above notions of distance on the coded entries to identify approximate string matches.

As an aid to these matching and clustering techniques, the author enumerates a number of common error paradigms and their causal conditions. For example a data format that is too strict, e.g., insisting on a middle initial for every name entry, will tend to generate erroneous, "placeholder" data entries. These are redundant records added to a database as a result of an erroneous match, or more appropriately, not finding the correct match due to inconsistencies in the fields used to join the data from two datasets.

In a specific data cleansing case study, the author describes a technique for standardizing residential and business addresses based on data rules established by the US Post Office. The author then proceeds to describe a number of general data cleansing and enhancement tools including: date/time, contextual, geographic, demographic, psychographic, and inferential data enhancement. An example of an inferential enhancement might be to assign a "primary decision maker" field to a household database based on the most frequent credit card user within the household.

Finally, the text summarizes each of the chapters as building blocks needed to build data quality practices for an enterprise. This book is a good primer on data quality concepts. It lists, in a systematic and formal way, many of the things that an actuary knows to look for intuitively in their work, but may not know how to articulate formally. While it is a long book, it is not an especially difficult read. It could be put to good use in constructing a checklist of data quality best practices that one would run through when building or implementing a new database or system architecture.

3.5 Corporate Information Factory

Corporate Information Factory [10] (ISBN 0-471-39961-2) provides an overview of information technology architecture for modern corporations. Its authors, Inmon (described by many in the industry as the father of the data warehouse), Imhoff and Sousa, describe a way of thinking about various technologies available today to give the reader a structure to incorporate them in their company's systems. The authors feel their proposed approach is "the best way to meet the long-term goals of the information processing company." Two clear strengths of their approach are that it can be implemented incrementally and it is designed to be flexible to adapt to changing business needs.

The book is divided into four parts. The first two chapters summarize the evolution of the "corporate information factory." Chapters 3 through 14 review each element of the architecture and how they are combined. Chapters 15 to 17 discuss constructing and managing the corporate information factory. Finally, the appendix contains guidelines for examining and assessing a particular corporate information factory.

The authors write: "Three fundamental business pressures are fueling the evolution of the information ecosystem: growing consumer demand, increased competition and complexity, and continued demands for improvements in operating efficiencies..." The corporate information factory can help corporations respond to these pressures by aiding them in:

- **Business operations:** running the day-to-day business,
- **Business intelligence:** helping companies understand what drives their business and the likely impact of decisions, and
- **Business management:** "If business intelligence helps companies understand what makes the wheels of the corporation turn, business management helps direct the wheels as the business landscape changes."

The authors see the big picture as follows. "The alpha and omega of the corporate information factory is the external world in which business is transacted." Information flows from the external world to the data acquisition applications of the corporate information factory. From there it can be condensed into operational reports or transformed and integrated with other data before being forwarded to primary storage management. Primary storage management includes the operational data store (ODS) and the data warehouse including historical data. The final phase, data delivery, can include data marts, decision support services and an exploration warehouse or a data mining warehouse. Managing metadata (information about the data) embraces and integrates across all three phases of the corporate information factory: data acquisition, primary storage, and data delivery.

The authors look at each of the dozen components from several points of view:

- What is the purpose or function of the component?
- What is its structure?
- How does information flow?
- What types of data does it work with?
- What types of users use it?
- What is the level of centralization versus decentralization in processing?
- How does this component interface with others?

The concepts of ecology and evolution are used frequently in the book. Corporate information systems are like an ecosystem where raw energy in the form of data is transformed by organisms, i.e., the various component information systems, into "food" or output which is then recycled into other "organisms" or information systems. These systems are never static, but evolve over time as circumstances and requirements of the users change.

Corporate information resides in a number of different data stores. These include data warehouses, data marts and operational data stores, each with their own role in meeting corporate information needs. The data warehouse is a big data repository of much of the company's data that is needed to run business intelligence systems. A data mart is a subset data repository, used for specific functions and applications containing smaller data subsets and aggregations of data. It is needed for efficiently running applications. Both play an important role in managing corporate information needs. Finally, an ODS "is a collection of detailed data that satisfies the collective, integrated operational needs of the corporation." The focus of the ODS is on information for operations, so it only contains current detailed information, not a data warehouse's multiple snapshots and summaries.

Each of the various corporate data stores has a development life cycle involving requirement gathering, analysis, design, programming, testing, and implementation. The book discusses a general database management strategy, as well as the strengths and weaknesses of various software and hardware solutions.

The different kinds of data storage may have different management requirements. For instance, the data warehouse needs are "characterized by volumes of data and unpredictable workload" (p. 251). The authors discuss the various needs and how they are addressed, how the various data stores are integrated as well as the security needs associated with the different databases.

A further consideration of corporate information systems is archival of stale data. The authors discuss how long management should wait before archiving data and what the best mechanism is for archiving it

The authors also include a discussion of multiple data warehouses and the integration of data from multiple systems. Integration of data from separate systems can be necessitated by corporate mergers and acquisitions or by the need to do more advanced analysis. Such integrations contain their own challenges, such as who owns the data, who creates and manages the new database, what types of data the database contains, and the nature of the sharing that will occur.

The book is something like the "Cliff Notes" of information management. It is a concise summary of current theory and practice with respect to developing and maintaining information systems, but it is not weighted down with a lot of technical detail. As such, the book provides an easy-to-read introduction for those not working in the area but might be too simplistic for those with deep experience in data management and information systems. The text is clearly written, but because of the multi-faceted approach sometimes it is difficult to tell where the authors are going with a discussion. Acronyms often appear in the book and their frequency sometimes becomes annoying. Also, a lot of the diagrams are trivial: they don't really illustrate their point any better than the text. Finally, despite all the points of view, there does not seem to be a lot of actionable information: this is a good text for learning about concepts, but not for implementing them.

Corporate Information Factory provides a good introduction to the broad world of information technology. This book can help actuaries better understand IT structure, concepts, issues, and goals to better frame their interactions with IT. If you are interested in a quick introduction to the topic that covers the key concepts and techniques, this book will meet your goals. If you need a more substantial introduction to information management, reference another book, perhaps one of the many books referenced by the authors.

3.6 Data Quality, the Field Guide

The focus of the book *Data Quality, the Field Guide* [11] (ISBN: 1-55558-251-6), by

Thomas C. Redman, Ph.D., is on data quality programs and efforts inside organizations. This book provides many constructive approaches to establishing or improving the data quality programs in businesses. It is a "how to" manual for those new to data management and a great refresher to those who have been in the field for a while. As the quality of the work product that an actuary produces depends so much on the quality of the input, with data being one of the key ingredients, the topic of this book should be of high interest to actuaries.

The book first reinforces that all disciplines and levels in an organization have a stake in quality data. The author presents the viewpoints of various stakeholders from the CEO to the customers of the organization.

For an actuary who has had the responsibility for data management and/or data quality in their organization, this book brings no surprises. But the nice feature for the experienced data manager – actuary or non-actuary – is the well-organized presentation of the issues with many charts and logical pictorial diagrams of data quality concepts and data quality processes to illustrate the author's points.

For those actuaries who regularly encounter quality issues in data supplied from internal or external systems and who are starting out in the area of data management and data quality, this book should be required reading. It quickly presents many concepts that a new data manager needs to know and the author presents them succinctly on a high level. Again, the illustrations and diagrams will help solidify the concepts quickly and can be adapted by readers to their own situations. Your adaptations of his charts and diagrams to a business case plan for data quality improvements will lend an authoritative flavor to your plans. It is a book worth reading for actuaries who have interactions with those responsible for data in their organizations. The knowledge and insights gained by the reader will help put them on equal footing with those who are responsible for the data.

One very important point that the author makes is that clean-ups of a database do not scale; you need to fix the source or cause of the data quality issue. "Organizations must recognize that finding and fixing data errors is time consuming, expensive, non-value-added work." Otherwise resources will be forever dedicated to cleaning a database and the problem will never go away. As the author says, "any form of clean-up without prevention is wrong-headed." In Section E of the book, the author describes the elements of a successful data quality system and how those tasks are accomplished. While you may not want to follow his solution or methods exactly, the book does provide a lot of ideas to consider as you work to improve data quality in your organization. The author makes the point that you need to consider all costs of errors – the immediate cost and the cost to those downstream; and you need to know where errors occur.

A "statistical control process" is described in Chapter 23. The process as presented is ever vigilant, focusing on continuous improvement and bottom line impact.

Another point made by the author is that I. (information or data) is not IT (information technology). It is clear from the book that one should not use IT to automate a poorly designed information chain. An information chain, as defined by the author, is an "end-to-end process that starts with original data sources creates 'information products' and continues through to the use of data in operations, decision-making, and planning." First improve the information chain and then automate using IT to reduce cost and to free up people for other tasks; IT plays a subordinate role.

The author also presents other concepts in the book that may be more applicable to the data manager such as a business case plan for data quality; the competitive advantage derived from quality data; techniques for cleaning a database; and the common elements of a successful data quality program. Reading through these sections should help practicing actuaries improve their communications with the data managers in their organizations.

Throughout the book, the author presents seventy-one "tips." For ease of reference sixteen of the most important tips are repeated at the end of the book and reorganized according to several subjects. A glossary of terms is also provided at the end of the book.

Overall the book is a quick read and presents many concepts in an easy to understand fashion for the practicing actuary.

3.7 Data Management: Databases and Organization

Data Management: Databases and Organization, fifth edition [12] (ISBN 0-47171-536-0) by Richard T. Watson is an introductory data management text. It focuses on the core skill of data modeling using SQL (structured query language) to implement the data models. It also covers such topics as the managerial perspective of data management, database architecture, emerging technologies, and data integrity.

Overall, this text is very well written. The topics are self contained, although the concepts of data modeling and SQL run throughout, so those sections should not be skipped. For actuaries, it is probably best to use the text as a reference book on particular topics because of the length (approximately 600 pages). Watson divides his book into five sections, and a brief synopsis of each follows.

Section 1, "The Managerial Perspective," defines the concept of organizational memory which includes not only computers, but also people, paper files, manuals, reports, etc. He also draws distinctions between data, information, and knowledge. According to Watson, "data are raw, unsummarized, and unanalyzed facts," while "information is data that have been processed into a meaningful form." Finally he states that "knowledge is the capacity to use information." Watson makes the interesting point that the preceding perspectives on data and information are relative. One person's information is another person's data.

In section 2, "Data Modeling and SQL," Watson considers data modeling and SQL skills as fundamental to data management. As such he devotes approximately half of the book to this topic. The style of this section is very straightforward and should be accessible to any actuary with some exposure to relational databases, such as Microsoft Access, SQL Server, Oracle, etc. He goes through in detail the basic building blocks of data modeling: modeling a single entity, one-to-many relationships, many-to-many relationships, one-to-one relationships, and recursive relationships.

The author repeatedly uses the same approach to explain new concepts, thus making the text easy to follow. First, he builds his examples using a standard data modeling diagramming syntax., Second, as each new modeling concept is introduced, a model is developed and then implemented in SQL. This is an effective technique for both data modeling and SQL since the concepts reinforce each other.

Watson also uses examples from standard relational databases such as Access and Oracle. While the book is not an Access reference and many advanced SQL features are not supported in Access, the text does give a good indication of the theoretical underpinnings about how a relational database product such as Access should be used. The text is filled with numerous exercises on both data modeling and SQL. It is a good primer for those actuaries that are interested in moving beyond Access or doing advanced database work using the macro programming capabilities of Access.

The author thoroughly illustrates the concept of normalization as a method for increasing the quality of a database design. He goes through the development of six normal forms and describes the issues that these normal forms resolve. This is perhaps a little advanced for most actuaries, but it is interesting reading if one is willing to devote the effort.

Finally, Watson provides an "SQL playbook" that contains 61 sample queries that should handle most of the data manipulation tasks that an actuary may encounter.

Section 3, "Database Architectures and Implementations," deals with more of the technical aspects of data management such as data structures and storage. It also provides an introductory background on data processing architectures such as client/server technology. If nothing else, this section and Section 4 define much of the terminology that is used in many IT shops today. This is of great use to actuaries that need to understand the key concepts of various technologies to liaise with their IT departments.

Watson devotes a chapter in this section to object-oriented (OO) data management. He does a good job of describing the object-oriented paradigm and then contrasting it with the relational-paradigm. Since the relational model is primarily used in data management, and the OO model is used primarily in software engineering, Watson posits that it is important to be able to translate between the two. Among the differences that he cites between the two paradigms is that the OO paradigm has its basis in the software engineering principles of coupling, cohesion, and encapsulation, while the relational paradigm is based on the mathematical concepts of set theory.

Section 4, "Organizational Memory Technologies," covers a potpourri of technologies. Watson devotes a chapter in this section that touches on data warehousing, data mining, and the multi-dimensional database (MDDB) or cube environment. Given that MDDB is (arguably) the best storage arrangement for actuarial triangles, this section should be of great interest to actuaries. Unfortunately, it barely scratches the surface on data warehousing and data mining. He also devotes two chapters to the Web and provides some extensive examples on how to use SQL within Java. Finally, he closes the section with a good treatment of XML (extensible markup language) and its emerging use as a data management standard.

The final section "Managing Organizational Memory" covers two topics that most actuaries should find of interest: data integrity and data administration. In this time when actuaries are being asked to become advocates for data quality, it is important for them to understand what data quality really means. Watson states that maintaining data integrity involves three goals:

1. Protecting the existence of the data so it is available whenever it is needed;
2. Maintaining the quality of the data so that it is accurate, complete, and current; and

3. Ensuring confidentiality of data so that only those authorized can access it.

He then describes many techniques to achieve these goals.

The author also covers what he calls the 18 dimensions of data quality. As an example, let's look at three of the dimensions—Accuracy, Timeliness, and Accessibility—and what conditions Watson sets for high quality (see Table 1).

Table 1

Dimension	**Conditions for high quality data**
Accuracy	Data values agree with known correct values.
Timeliness	A value's recentness matches the needs of the most time critical application requiring it.
Accessibility	Authorized users can readily access data values through a variety of devices from a variety of locations.

These three dimensions, as well as the other 15 dimensions outlined in the book, are an ongoing pursuit and not a destination. It is worthwhile for actuaries to look at all 18 dimensions and see how each of their organization's data stacks up against them.

Overall, I would highly recommend Data Management: Databases and Organization to those actuaries that are interested in learning more about the principles and challenges of data management.

3.8 Software Testing in the Real World

Edward Kit's main goal in this book [13] (ISBN 0-20187-756-2) is to prove to software companies that they need dedicated testing departments at least as big as their development departments. Given that actuaries are not in the business of making shrink-wrapped software packages for numerous outside customers, the "real world" in the title practically never intersects with the actuarial universe.

Some of the main thoughts of the book, however, will be of interest to actuaries. Considering that actuaries implement their models in software, this activity could be conceivably called "software development." Thus some notions of testing should not be fully foreign to actuaries; they just have to be adapted to the actuarial situation.

Testing according to the book should start from the "specifications" and end with the "final product" evaluation, and should be performed by an "outsider." Testing techniques range from verification to validation, i.e., from checking the "code" to examining "final product" outcomes. In the actuarial paradigm, the final product could be an Excel spreadsheet, Mathematica notebook, or Oracle stored procedure. Correspondingly, specifications could be a reserve test or pricing method, and the "code" would be formulae in cells, VBA subroutines, or SQL statements. Evidently, checking everything from methods and assumptions to auditing spreadsheet formulae and query results makes perfect sense.

The content of Kit's book is broken into 4 parts. Part I includes chapters 1 through 3. The material in these chapters is somewhat esoteric. There is a lot of discussion about what is needed to get started on the testing of software and the history of software testing. These chapters would not be applicable to the actuarial science field.

Chapters 4 through 6, which form Part II of the book, establish a framework for conducting tests on software. This section establishes some decent terminology that one could use to test a student's familiarity with testing procedures. The question we need to ask is will everyone in the industry adhere to the same terminology? For example, in the 4th chapter, there are several terms used to establish a general failure in the software code. Such terms include: "mistake, fault, failure, [or] error." Could we get some of these terms generally accepted in the actuarial industry? There are several examples of these types of definitions of principles within this section. There is one principal in particular that could prove to be useful in the actuarial science field: "the purpose of testing is to discover errors." It is a nice short and sweet principal. Chapter 5 seems to be getting to some substance. One question that it attempts to answer is when a tester should be giving special attention to the testing process. Discussions about verification (checking the code) and validation (testing the program) are also discussed in Chapter 5. Chapter 6 is not very helpful to actuaries. This chapter seems to be regurgitating different top-down methods on how to approach testing. This is probably more useful to software engineers than to actuaries. At this stage of the book, some examples would have proven to be helpful. Several lists of questions are developed for testing methods but none of the questions are ever answered. It is unclear if we are supposed to be learning how to ask or how to develop questions. More testing standards are talked about in a theoretical sense but no lists of standard questions are given. The section on "Testware" (a collection of software tools for testing) is somewhat useful. This section discusses what is actually used to test software and calls for maintaining the best testware tools beyond the testing of a single product.

Part III, which includes chapters 7 through 12, provides several different testing methods. Some of the material can be applied to what we do in actuarial science. For example, the methods used for verification could become a basis for technical reviews of an actuary's work. Still, the text lacks examples and exercises for the reader to follow. There do not appear to be definitive methods to apply to specific circumstances. The recommendations at the end of the chapters contain many phrases such as: "usually it is better to do..." or "there's a real trade-off when you do..." A decisive recommendation on a method to use in a particular situation would have been more helpful. There is a relevant exercise given on p. 67 of the book. It refers to documents in Appendices B and C. The exercise shows how verification testing can produce gains on developing software for a minimal amount of effort. Also, the section on how a tester should report an author's mistakes (in Chapter 7) is useful.

Part IV includes topics on structural designs for testing software, practices used by software engineers in testing, and getting gains from software testing. This section would not be applicable to the field of actuarial science.

The appendices follow these 4 parts, and are clearly the most useful part of the book to actuaries. There appears to be much more order and less theory in this section of the book. Appendix A gives lists of Software Testing Standards. This section may be very useful when a tester has to present results to a management team or to a group of people within the industry. For testing actuarial work, one could refer to similar standards much like we do for reserving and valuation methods. Appendix B gives good verification checklists. It is ironic that there is a functional design checklist which has a requirement to look out for designs "without examples or examples that are too few." The author could have taken this requirement and applied it to the earlier chapters in the book. Appendix B has a good deal of sample checklists which would be useful. Appendices C and D contain verification and validation exercises (respectively) and solutions that seem very useful, however extensions would be needed to translate the exercises into practical advice for Excel "developers." Appendix E contains a bibliography which is a good reference for guides on software testing. Appendix F gives source information on conferences, journals, and newsletters which may be useful for someone desiring more information on software testing. Appendix G gives a list of software technology used to check software. Appendix H contains a list of improvements in the area of terminology, product requirements, tools used for testing, and documentation which should be considered. The text should have referred to the lists and information in the appendices much more frequently.

In conclusion, actuarial practitioners who are heavily involved in spreadsheet design may occasionally find some useful tidbits in this book. However there simply are not enough examples or case studies to make any of the testing methods easy to implement. Therefore actuaries not heavily involved in systems development should probably pass on this text and wait for a more directly applicable book or article on the subject. Please note that do not wish to minimize the importance of software errors to actuaries. However this book may not be the appropriate reference for the kinds of software development projects that actuaries encounter.

3.9 Insurance Data Collection and Reporting, eighth edition

This book [14] (ISBN 1-877796-27-1), edited by Rose Castro, is the first in a series of eight books published by the Insurance Data Management Association (IDMA) designed to educate data managers. As a textbook, it is well written and quite easy to follow. There are 10 chapters in total.

The first three chapters introduce underwriting and actuarial ratemaking, highlighting the necessity of high quality insurance data that underlie these functions. As the author rightly points out in the first chapter, both line underwriters and staff underwriters need data to perform their daily jobs. Moreover, actuaries rely heavily on data to analyze loss reserves and conduct rate level experience reviews. Chapter 2 discusses general ratemaking procedures widely used by property/casualty actuaries. These procedures include pure premium method, loss ratio method and distribution of an overall indication to territories/classes. Workers Compensation ratemaking, a different animal as usual, is elaborated in the third chapter. NCCI has three types of systems to perform ratemaking functions: the administered pricing system, the advisory rate system, and the loss cost system.

Chapters 4 to 9 focus on various types of statistical agents such as ISO, NAII, and NCCI. Chapter 4 gives a general background of insurance regulation and statistical reporting. Two important court decisions (Paul v. Virginia and South-Eastern Underwriters Association) and two laws (McCarren-Ferguson Act and All-Industry Rating Bills) are cited. These help readers understand the historical context in which insurance regulation has evolved. Chapter 4 also gives a high-level review of statistical agents. Chapter 5 summarizes various statistical agent reports and three basic report designs (annual statistic compilations, Fast Track Monitoring System, and accelerated reports). Chapter 6 gives a detailed description of ISO. Besides highlighting ISO's statistical plans, the author also touches upon the process that ISO goes through after receiving data. In chapter 7 and 8, the NAII and NCCI statistical plans are described in detail. Chapter 9 identifies organizations specializing in data collection which do not fall into the above categories: mostly involuntary pools.

Chapter 10 focuses on state insurance departments including the history of insurance regulation regarding insurance data and state data needs.

Overall, this book provides excellent study material for data managers to get a good understanding of insurance data collection/reporting. Actuaries have learned most of the contents of this book through CAS exams. For them, this book not only gives a good review but also helps to piece together an understanding of data management to the insurance enterprise.

4. CONCLUSIONS

There is an actuarial standard of practice with respect to data quality and some actuaries have data management responsibilities, but there is almost nothing in actuaries' formal training to prepare them for these tasks. Furthermore, current CAS literature is comparatively cursory in its coverage of information quality topics. To fill this gap, these nine texts have been recommended to actuaries seeking more information on data quality or data management.

To help identify the best text for a specific situation, the texts are compared below in three ways. The first table (Table 2) describes the subjects covered in each book and should be helpful in determining which books are most appropriate for particular data quality and data management goals. In this table, five solid circles mean the particular topic is excellently covered in a way readily accessible to actuaries. Conversely, five empty circles mean the subject is either barely covered or addressed from a point of view that is of limited use to actuaries. Finally, a blank rating means the particular subject is not covered at all in the particular text.

Table 2: Coverage of Topics

Author	Section	Data Quality	Principles of Data Quality	Metadata	Exploratory Data Analysis	Data Audits
Olsen	3.1	●●●●○	●●●●○	●●●○○	●●○○○	●●●○○
Dasu	3.2	●●●●○	●●●○○	●●●○○	●●●●○	●○○○○
English	3.3	●●○○○	●●●◐○	●●●◐○		
Loshin	3.4	●●●○○	●●●○○	●●●○○	●●●○○	●●○○○
Inmon	3.5	●○○○○	●●○○○	●●●○○		
Redman	3.6	●●●○○	●●○○○		●●●○○	
Watson	3.7	●●●●○	●●●●○	○○○○○		
Kit	3.8	○○○○○			○○○○○	○○○○○
IDMA	3.9	○○○○○	○○○○○		○○○○○	○○○○○

Author	Section	Processing Quality	Presentation Quality	Measuring Data Quality	Data Quality Improvement Strategies	Data Management	Statistical Plans
Olsen	3.1	●●●●○		●●●●○	●●●●○	●●○○○	
Dasu	3.2	●●○○○		●●●○○	●●●○○	●●○○○	
English	3.3		○○○○○	●●○○○	●●○○○		
Loshin	3.4	●●○○○	●○○○○	●●●○○	●●○○○	●●●●○	○○○○○
Inmon	3.5				●○○○○	●●○○○	
Redman	3.6		●●○○○	●●○○○	●●●○○	●●●○○	
Watson	3.7			●○○○○	●○○○○	●●●●○	
Kit	3.8	●●○○○		○○○○○	○○○○○		
IDMA	3.9	○○○○○		○○○○○	○○○○○	●●○○○	●●●●○

Table 3: Definitions of Topics

Topic	Definition / Description
Data Quality	What is it? Why does it matter? How to achieve it?
Principles of DQ	Key attributes of "quality data"
Metadata	Information about data, e.g. business rules
EDA	Statistical and graphical tests to identify suspicious values in a data set
Data Audits	Reconcile the data intended for use to its original source(s)
Processing Quality	Ensuring quality in models and software through design, implementation and testing

Presentation Quality	Clear, correct, consistent presentation of results
Measuring DQ	Statistics to track key attributes of quality
DQ Improvement Strategies	What should an organization do to determine the level of quality required and how to achieve it?
Data Management	The bridge between those who are responsible for the collection and repository of data and those who will use the data in analyses
Statistical Plans	Examples, motivation and uses of mandated data (not detailed instructions on specific statistical plans)

The next table contains brief summaries of the dominant characteristics of each book.

Table 4: Synopsis

Author	Section	Comment
Olsen	3.1	Well written, easy to follow data quality program for companies
Dasu	3.2	Good introduction to use of exploratory data analysis in data quality
English	3.3	Complete data quality guide aimed at IT and management rather than at actuaries
Loshin	3.4	Good generic data management text. Not specific to actuarial science, but covering many thorny data issues which actuaries may encounter.
Inmon	3.5	An easy to read introduction to concepts and systems architecture
Redman	3.6	Easy to follow book with concepts an actuary can easily pick up on
Watson	3.7	Very good text on data modeling and SQL
Kit	3.8	This book should only be used by actuaries who are involved with designing software. Even so, adaptation of any of the material will be needed prior to use. The appendices and the last few chapters are the most applicable to actuaries.
IDMA	3.9	Good introduction to data collection and various agencies

The final table contains summary ratings by text. The ratings assigned provide an assessment of how suitable each book is at covering topics and what audience it is best suited for. For instance, each book is rated on whether it is geared towards beginners in information quality or at a more advanced audience that is already familiar with some of the literature. As another example, since insurance applications were not a focus of any of the books, each book is rated on its relevance to actuaries. Whereas Table 2 focuses on the topics covered in the book (such as data quality or metadata) Table 5 focuses primarily on qualities of the book as a whole that determine what audience it is best suited for (such as beginner/advanced, those wanting a more theoretical as opposed to practical knowledge, etc.). The book reviews in section 3 also contain information on the technical level of the book and the audience it is written for. Note: although many of the books were written for an audience that is actively involved in data management or data quality, they also provide a good introduction to the topic for a general audience that interacts regularly with data supplied by others.

Table 5: General Characteristics

Author	Section	(a) Actuarial Relevance	(b) Beginner / Advanced	(c) Practical / Theoretical	(d) Micro or Macro Focus	(e) Overall
Olsen	3.1	●●●○○	●●○○○	●●●○○	●●●○○	●●●○○
Dasu	3.2	●●●○○	●●●●○	●●●●○	●●○○○	●●●○○
English	3.3	●○○○○	●●●○○	●●○○○	●●○○○	●◖○○○
Loshin	3.4	●○○○○	○○○○○	●●○○○	●●○○○	●●◖○○
Inmon	3.5	●○○○○	●●●●○	○○○○○	●○○○○	●●○○○
Redman	3.6	●●○○○	●●●○○	●●○○○	●○○○○	●●●○○
Watson	3.7	●●◖○○	●◖○○○	●●○○○	●●○○○	●●●○○
Kit	3.8	○○○○○	◖○○○○	●◖○○○	●●○○○	●○○○○
IDMA	3.9	●○○○○	●●●○○	●○○○○	●●○○○	●●○○○

Notes: Generally, five solid circles is most relevant to an actuarial analyst

(a) 5 solid circles = text is written for actuaries. 5 empty circles = need to modify or extend ideas in the text before an actuary could use them.

(b) 5 solid circles = beginner: no prior IT knowledge required. 5 empty circles = advanced, e.g. reader should have worked in the field.

(c) 5 solid circles = purely practical, e.g. a tip sheet with no reasoning behind the tips. 5 empty circles = purely theoretical.

(d) 5 solid circles = hands-on analyst advice such as a book of C programs. 5 empty circles = only high-level advice, e.g. strictly executive issues.

(e) 5 solid circles = a "must-read" for all actuaries. 4 solid circles = a "must-read" for actuaries with data management responsibilities. 5 empty circles = the information is not worth the time it takes to read it.

By reviewing the summary information in these tables, the reader may be able to identify candidate books that will best meet his or her needs.

The working party hopes that this paper will be a resource for actuaries dealing with data management and/or data quality issues. More information on these issues can be found at the idma.org web site.

Acknowledgment

The working party thanks IDMA for narrowing the field of data quality and data management texts to those they felt would be most relevant to actuaries. We also thank our IDMA liaisons for their feedback and support throughout this project.

The working party also thanks the Insurance Services Office, Inc., for the use of excerpts from their homeowners module of the ISO personal lines statistical plan (other than auto).

5. REFERENCES

[1] Francis, Louise A. "Dancing with Dirty Data: Methods for Exploring and Cleaning Data." CAS Forum Winter 2005: 198-254.

[2] Actuarial Standards Board of the American Academy of Actuaries. *Actuarial Standard of Practice No. 23: Data Quality,* revised edition. Schaumburg, Illinois: American Academy of Actuaries, 2004.

[3] CAS Committee on Management Data and Information. "White Paper on Data Quality." CAS Forum Winter 1997: 145-168.

[4] Popelyukhin, Aleksey. "Watch Your TPA: A Practical Introduction to Actuarial Data Quality Management". CAS Forum Winter 1999: 239-254.

[5] Copeman, P., Gibson, L, Jones, T., Line, N, Lowe, J, Martin, P., Mathews, P., Powell, D., "A Change Agenda for Reserving: A Report of the General Insurance Reserving Issues Task Force" 2006, www.actuaries.org.uk

[6] Olson, Jack E. *Data Quality: the Accuracy Dimension.* Morgan Kaufman, 2003.

[7] Dasu, Tamraprni and Theodore Johnson. *Exploratory Data Mining and Data Cleaning.* Wiley, 2003.

[8] English, Larry P. *Improving Data Warehouse and Business Information Quality.* New York: Wiley, 1999.

[9] Loshin, David. *Enterprise Knowledge Management.* Morgan Kaufman, 2001.

[10] Inmon, William and Claudia Imhoff and Ryan Sousa. *Corporate Information Factory, second edition.* New Jersey: Wiley, 2000.

[11] Redman, Thomas C. *Data Quality, the Field Guide.* Boston: Digital Press, 2001.

[12] Watson, Richard T. *Data Management: Databases and Organization, fifth edition.* New Jersey: Wiley, 2005.

[13] Kit, Edward. *Software Testing in the Real World.* New York: Addison-Wesley, 1995.

[14] Castro, Rose. *Insurance Data Collection and Reporting, eighth edition.* New Jersey: IDMA, 2005.

Abbreviations and notations

Collect here in alphabetical order all abbreviations and notations used in the paper

ASB, Actuarial Standard Board	MDDB, Multi-dimensional Database
ASOP, Actuarial Standard of Practice	NAII, National Association of Independent Insurers
CAS, Casualty Actuarial Society	NCCI, National Council on Compensation Insurance
CIO, Chief Information Officer	NYSIIS, New York State Identification and Intelligence System
COLDQ, Cost of Low Data Quality	ODS, Operational Data Store
GIRO, General Insurance Research Organization	OLAP, On-Line Analytical Processing
GRIT, General insurance Reserving Issues Taskforce	OO, Object Oriented
IDMA, Insurance Data Management Association	SQL, Structured Query Language
ISO, Insurance Service Organization?	VBA, Microsoft Visual Basic Application
IT, Information Technology	XML, Extensible Markup language

Biographies of Working Party Contributors

Robert Campbell is Director, Commercial Lines Actuarial at Lombard Canada in Toronto, Canada. He has a Bachelor of Mathematics in Business Administration from the University of Waterloo. He is a Fellow of the CAS and a Fellow of the Canadian Institute of Actuaries. He is chair of the Data Management Educational Materials working party, participates on the CAS Committee on Data Management and Information, and was a participant on the 2006 GIRO Data Quality working party.

Lijuan Zhang is senior actuarial analyst at Insurance Service Office in New Jersey. She is responsible for ZIP Code based territory revising analysis and pricing. She has a degree in Economics from Youngstown State University. She is a Fellow of the CAS and a Member of the American Academy of Actuaries.

Louise Francis is a Consulting Principal at Francis Analytics and Actuarial Data Mining, Inc. She is involved in data mining projects as well as conventional actuarial analyses. She has a BA degree from William Smith College and an MS in Health Sciences from SUNY at Stony Brook. She is a Fellow of the CAS and a Member of the American Academy of Actuaries. She is serves on several CAS committees /working parties and is a frequent presenter at actuarial and industry symposia. She is a four-time winner of the Data Quality, Management and Technology call paper prize including one for "Dancing with Dirty Data: Methods for Exploring and Cleaning Data (2005)."

Rudy Palenik is the Commercial Actuary at Westfield Insurance Group in Westfield Center, Ohio. He is responsible for the development of rates for all the commercial lines of business. He has a degree in Math from Marquette University in Milwaukee, Wisconsin and is a Fellow of the Casualty Actuarial Society and a member of the American Academy of Actuaries. Rudy participates on a number of CAS committees including: Data Management and Information, Actuarial Education and Research Foundation, Research Paper Classifier and University Liaison.

Aleksey Popelyukhin is a Vice-President of Information Systems with the 2 Wings Risk Services and a Head of Quantitative Analytics Group with the Wall Street North Consulting in Stamford, Connecticut. He holds a Ph.D. in Mathematics and Mathematical Statistics from Moscow University (1989). Aleksey actively participates in CAS research and is frequent presenter on CAS conferences. CAS recognized Aleksey's contributions by awarding him the very first prize in "Data Management" papers competition and inviting him to the very first Working Party (on presentation of DFA/DRM results). In addition to numerous publications Aleksey helps to advance actuarial science by building convenient software tools for actuaries such as Triangle Maker®, Affinity and Actuarial Toolchest™. For those actuaries having troubles explaining statistics to the management Aleksey built a DRM presentation template available from CAS website. And for those who have troubles fitting clean models to dirty data Aleksey developed advanced data quality service called Data Quality ShieldSM. Aleksey is currently developing an integrated pricing/reserving/DRM computer system for reinsurance called "SimActuary" and also an action/adventure computer game tentatively called "Actuarial Judgement."

Gregory Scruton is Senior Vice President, Actuarial and Planning at Middlesex Mutual Assurance Company in Middletown, CT. He is responsible for pricing, reserving, planning and management information. He has a degree in Mathematics from Renssalaer Polytechnic Institute in Troy, NY. He is a Fellow of the CAS and a Member of the American Academy of Actuaries. He has participated on the CAS examination committee, and currently participates on the CAS Committee on Data Management and Information.

Virginia R. Prevosto is Principal, Consulting at Insurance Services Office, Inc. Ms. Prevosto is a Phi Beta Kappa graduate of the State University at Albany with a Bachelor of Science degree in Mathematics, *summa cum laude*. She is a Fellow of the CAS and a Member of the American Academy of Actuaries. She serves as General Officer of the CAS Examination Committee and as liaison to various other CAS admission committees. She also serves on the CAS Committee on Management Data and Information. In the past Ms. Prevosto also served on the Data Quality Task Force of the Specialty Committee of the Actuarial Standards Board that wrote the first data quality standard of practice for actuaries. Virginia has been a speaker at the Casualty Loss Reserve

Seminar on the data quality standard and to various insurance departments on data management and data quality issues. Ms. Prevosto authored the paper "Statistical Plans for Property/Casualty Insurer" and "Study Note: ISO Statistical Plans" and co-authored "For Want of a Nail the Kingdom was Lost – Mother Goose was right: Profit by Best (Data Quality) Practices" for the IAIDQ.

Dave Hudson is an Actuary for St. Paul Travelers in Hartford, CT. He has a MS degree in Mathematics from Washington State University in Pullman, WA. He is a Fellow of the CAS and a Member of the American Academy of Actuaries. He is also a member of the CAS Committee on Data Management and Information.

Keith Allen is the associate actuary for United Educators and is responsible for underwriting duties within the public school sector and general corporate actuarial issues. Allen has 13 years of experience in the insurance industry as an underwriter, claims adjuster, and actuary. Keith previously worked for Tillinghast-Towers Perrin as an actuarial specialist where he did reserving, pricing, and forecasting for various public and private entities. Prior to that, Allen worked as a claims adjuster and underwriter for State Farm Insurance where he helped develop the "Reinspection Program" used to assess coastal risks. Before joining the insurance industry, Allen was a teacher at Bellaire High School in Houston, TX. Allen holds a bachelor's degree in mathematics from the University of Texas and is an Associate of the Casualty Actuarial Society.

Shiwen Jiang is assistant vice president/actuary at Arch Insurance Group in New York. He is responsible for technical pricing. Prior to this, he has held various positions in Arch Insurance Group, The Hartford and W.R. Berkely Corp. He has a M.S. degree in Statistics from the Pennsylvania State University. He is a Fellow of the CAS and a Member of the American Academy of Actuaries. He participates on the CAS Committee on Data Management and Information and Asian Regional Committee.

USING A CLAIM SIMULATION MODEL FOR RESERVING & LOSS FORECASTING FOR MEDICAL PROFESSIONAL LIABILITY

Abstract

Various recent papers have included criticisms related to the use of link-ratio techniques for estimating ultimate losses. While this paper does not review these criticisms, it does outline development characteristics of medical professional liability1 losses that would lead the actuary to believe that link-ratio techniques may not always be the best available option for projecting ultimate losses. The paper then proceeds to provide a model that addresses these weaknesses and then extends this model for loss forecasting applications.

More specifically, this paper provides a framework for evaluating medical professional liability[1] loss exposures. The concepts used in this model are more fully discussed and described in other statistical textbooks or refereed actuarial journals. This paper is intended to provide a synthesis of existing distinct processes. Rather than repeat those discussions, a bibliographical reference is provided. The bibliography included, therefore, should be considered a critical section of this paper.

Specifics of the modeling within the framework presented are the responsibility of the actuary implementing the model. While the paper does include alternatives that may be considered within the framework, it is not intended to be a comprehensive listing of these alternatives.

This model has been developed in recognition of data availability issues for self-insured healthcare facilities. However, this model may easily be expanded for use in an insurance company context or for evaluating other medical professional liability exposures.

I. Motivation and Rationale

A recent paper published in the CAS Forum included the following statement:

[1] This model may be extended to other general liability or professional liability exposures. This model is not appropriate for coverages subject to partial payments such as workers compensation.

> "...for most, if not all, cumulative arrays the assumptions made by the standard link ratio techniques are not satisfied by the data, ..."[2] [1]

Without providing a statistical analysis to prove this statement, there are a number of intuitive reasons why we would expect this statement to be true for medical professional liability losses.

(1) Claims-made medical professional liability is generally considered a "short-tailed" coverage in comparison to other liability coverages. Occurrence coverage is generally considered "long-tailed." This leads to the natural conclusion that on an occurrence basis, the majority of loss development may be attributed to claims that are incurred but not reported ("IBNR"). Consistent with common actuarial usage, this type of development is referred to as "pure IBNR" emergence. This development should be distinguished from the development on known claims, which will be referred to as bulk development. Link ratio techniques assume future development is a function of prior cumulative experience. This is inconsistent with the understanding developed that future development is actually due to newly reported claims. These newly reported claims do not necessarily have any relationship to past claims.

This relationship between pure IBNR and bulk development may be driven by the fact that healthcare institutions, in general, are conservative by their very nature. In the aggregate, case reserves established by these conservative institutions tend to be reasonably adequate.

(2) A model is defined as "a simplified mathematical description" [3] of a more complicated process. Loss development approaches would not appear to satisfy this definition since future development is not entirely a function of cumulative losses. Therefore the "mathematical description" is not consistent with the process being modeled.

(3) Link ratio techniques are generally based on the analysis and review of loss development triangles. Given the long-tail nature of occurrence coverage the predictive ability of loss development triangles is severely compromised by inflation. Emergence in the 10th calendar period for the

[2] That same paper was later published in the Proceedings with softer language: "Most loss arrays don't satisfy the assumptions of standard link ratio techniques."

10th prior accident period is likely to differ from the for the current accident period due to inflationary pressures.

(4) Loss development data for self-insureds may be subject to various limits and deductibles. This further compounds the inflation problem. Even if we obtain triangles at constant limits and deductibles on a nominal basis, the development for each period will be different on a real, or inflation-adjusted, basis.

In addition, if actuaries choose (as they often do) to select a single development pattern and apply it to every exposure period[3], they are making the implicit assumption that trend acts in one direction – across exposure periods. Any other "direction" of trend, i.e. across settlement period, report period or maturity, would be inconsistent with the development patterns as they are used in general practice.

These factors would compromise estimates using traditional Bornhuetter-Ferguson ("B-F") or additive techniques as well as link-ratio methods.

(5) When information is aggregated, information is lost. Fundamentally, by aggregating loss information into somewhat arbitrary accident year groupings, information provided by individual claim detail is lost. It is also critical to recognize that loss development is a statistical model. In this model, parameters (loss development factors) are estimated using data (loss triangles).

The framework described herein is based on multiple underlying stochastic models and is likely to be more robust. This is because we are estimating fewer parameters with more information. However, there may be residual uncertainties that cannot be eliminated.

From a practical perspective, these methods also suffer from the following problems:

(1) To many users of actuarial information, risk (deviation from the point-estimate) is just as important as, if not more important than, the point-estimate itself. We may be able to develop statistical measures of the uncertainty involved in the selection of loss development factors - also

[3] "exposure period" is intended as a generalized term for accident period, policy, period, report period, etc.

known as parameter risk. However, we have not been able to determine the model specification risk. This risk can be quite large since the assumption of link ratio models may not be consistent with the underlying cause of loss development.

(2) Oftentimes loss triangles are simply not available. This is particularly the case for self-insureds. Self-insured entities often keep a current loss database and generally do not track aggregate loss development.

(3) For many self-insureds, excess insurance is only available on a claims-made basis. However, accruals need to be made on an occurrence basis. This would require that actuarial analyses recognize differences in limits and retentions that are dependant on the report date of the claim. Link-ratio methods do not easily recognize these differences.

The goal of this paper is to present a model that overcomes these limitations. Specifically, we present a model that is adaptable, accounts for inflation, estimates risk, and is easily extendable for loss forecasting applications.

This model has the following benefits relative to link-ratio models because:

(1) This model reduces model specification error. This is due to that fact that this model that attempts to replicate claims process. The model includes the following phases of the actual claim life cycle: an accident occurs, the claim is reported, and the claim is settled for some amount. It would be naïve to believe that each and every driver of the claim process is (or can be) included in the model presented herein. However, the model better satisfies the definition of a "model" as stated above.

(2) The model is "unified" and easily adaptable to provide consistent estimation of pure IBNR, bulk reserves and prospective loss forecasts.

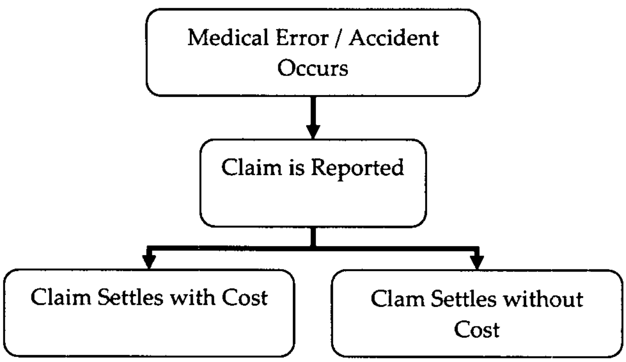
(3) The model is specifically designed to be used in a simulation environment. Given that insurance involves bearing risk – a reserve or loss forecast model should measure that risk. Use of simulation techniques is necessary in analyzing these exposures to provide an estimate of variability. Insureds retaining risk require this information as they are quite concerned with the variability in the point estimate. A model that yields only a point estimate does not accomplish this goal.

A claim level simulation model allows for the evaluation various per-occurrence and aggregate coverage alternatives in a prospective loss forecast. This information is particularly useful for insureds considering changes in their insurance program.

II. Model Overview

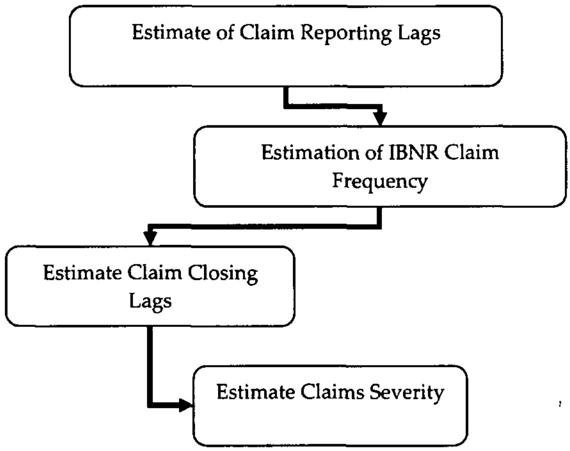
The loss-reserving model estimates indemnity and expense reserves for pure IBNR separately from bulk reserves. Each reserve component is estimated using a frequency × severity methodology. Model specification error is reduced with a frequency × severity model as it attempts to replicate the claims process. This is a benefit relative to link-ratio, additive or B-F methods, which only provide <u>proxy</u> models for loss movements in aggregate. The claim process replicated by the frequency × severity may be illustrated as follows:

FIGURE 1



Based on the diagram above, the model is specified in the following order:

FIGURE 2



The model used to estimate required bulk reserves or prospective loss forecasts are simply "special cases" of the model used to estimate pure IBNR. For this reason, the model for pure IBNR is presented first and the special cases follow.

III. The General Model – Evaluation of Pure IBNR

Claims Reporting Lag

Recognizing usual self-insured data limitations, the model employs an approach that does not require claim triangles. In typical self-insured loss runs, observed report lag for each claim will be available. These lags are calculated as the difference between the report date and the accident date for each claim. A statistical distribution may then be fit to these observed report lags.

In doing so, it must be recognized that the observed report lags have a problem similar to that found with deductible claim data[4]. With deductible claim data, our observations will not include losses below the deductible – i.e. claims are said to be truncated from below. With observed report lag data, our observations will not include claims that have not yet been reported – i.e. the observations are truncated from above. The observations, *W*, will follow a conditional distribution with the following density function:

[4] The parallel to the deductible loss data may be understood by reviewing Hogg & Klugman [6] (p. 129 –130).

$$f_w(x) = f_X(x) / F_X(M) \text{ where M = (valuation date - accident date)}$$
$$= 0 \; \textit{for } x > \textit{(valuation date – accident date)}$$

Using this density function, we can solve for the parameters of the lag distribution using maximum likelihood techniques and spreadsheet optimization tools. The likelihood function for individual claim data may be written as:

$$\begin{aligned} L(\theta) &= L(\theta; w_1, w_2, w_3, w_4, \ldots, w_n) \\ &= \prod_j f_w(x_j;\theta) \end{aligned}$$

Taking logarithms, we have:

$$\ln(L(\theta)) = \sum_j \ln(f_w(x_j;\theta))$$

where θ represents the parameter(s) of the selected distribution[5]

Using spreadsheet optimization tools, we can solve for the parameter(s), θ, which maximize the likelihood function. This analysis is presented on Exhibit 1. Column (4) of this exhibit shows the calculated observed lag (in days). The maximum report lag is calculated as shown in Column (5). The conditional likelihood and conditional log likelihood are calculated in Columns (6) and (7), respectively. The mathematics of this lag model are described in Weissner [2].

The cumulative distribution function provides our claims reporting pattern. The stochastic distribution also is easily used in a simulation analyses.

Determining Pure IBNR Claim Frequency

The report lag model is then used to estimate pure IBNR claims frequency. IBNR claims are estimated using a B-F approach. This estimation is illustrated on Exhibit 2.

For the B-F calculation, the claims reporting pattern is provided by the cumulative distribution function of the report lag model and the a priori ultimate claim estimate is determined using the average of development method estimates of ultimate claims of the mature claims periods (Column (9)).

[5] In general, we tend to use Rayleigh or Weibull distributions for lags.

The "Percent Reported" (Column (8)) is calculated using the cumulative distribution function of the lag model as determined on Exhibit 1. The function is evaluated at the difference between the data evaluation date and the midpoint of the accident period.

This is not the only approach available for estimating IBNR claims; however, based on the understanding that future claims are unrelated from prior claims, approaches, such as B-F, where IBNR claims are estimated (largely) independently of reported claim count are desirable.

Determination of Claims Closing Lag Model

The development of a conditional closing lag distribution is identical to the development of the claim reporting lag with one fundamental difference. Since the closing lag in this model represents the difference between the report date and the date of closing, the observed closing lags are truncated (from above) at the difference between the valuation date and the report date. As the reader will recall, the observed reporting lags are truncated (from above) at the difference between valuation date and the accident date. As this process is identical to that for reporting lags, the calculations underlying a closing lag model are not included in this paper.

Claims Settlement Model

Professional liability claims will settle with one of the following outcomes: (1) no payment, (2) indemnity and expense, (3) indemnity only, or (4) expense only. Depending on the quantity of available data, it may be necessary to collapse settlement outcomes into: (1) "no payment" and (2) "with cost" outcomes. The first step in determining our claim settlement model is to estimate the probability of each of these possible claim settlement outcomes.

It is recommended that the distribution of claim settlements be reviewed based on both closing year and accident year bases. Closing years are preferred as they better capture changes in claims settlement practices. Changes in claims settlement practices tend to apply to claims closed after a given date regardless of the accident date of the claim. When reviewing accident year distributions of settlements, only accident years that are completely or nearly completely closed should be considered. Consideration of immature periods may bias results towards the more quickly closed "no payment" or "expense only" settlement types. This type of review is presented on Exhibit 3.

Claims Cost Models

There are numerous previously published papers and texts describing methods to estimate stochastic claims cost models. The development of claims cost models is beyond the scope of this paper. The following is a partial listing of relevant papers and texts on this topic that the interested reader should review.

» *Klugmanm Panjer, & Wilmot – Loss Models* [3]
» *Keatinge – Modeling Losses with the Mixed Exponential Distribution* [4]
» *Philbrick – A Practical Guide to Single Parameter Pareto Distribution* [5]

In practice, mixed distribution models appear to best describe claims severity and are easily adaptable to the simulation process. Each component of the mixture represents a "type of claim." For example, a mixture of two lognormals, a point mass and a uniform distribution many be used to describe "small normal claims" (first lognormal), "large normal claims" (second lognormal), "losses clustered at the limit of insurance" (point mass) and "shock claims" (uniform distribution), respectively. An example of this model is presented in Exhibit 4 and is used later in this paper.

Credibility

Credibility of the loss data used in the estimation of the model parameters is an issue with every model. However, relative to development or B-F models, credibility should be less of an issue for the model presented herein. Credibility becomes an issue as actuaries attempt to estimate more parameters with fewer data. Relative to other models, this model estimates fewer parameters from more data.

Actuaries should recognize that, in estimating development patterns, each selected link-ratio is a parameter that is estimated from the observation in a column of observed link ratios. In addition to these link ratios, development methods or B-F methods may add other estimated parameters such as *a priori* loss estimates. Even with a 10-year development triangle, this may require the actuary to estimate more than 10 parameters. The model presented herein should generally require the estimation of fewer than 10 parameters. In addition, because this model relies on claim level detail, we have significantly more data - or information - relative to models in which claim level detail is collapsed into accident years – thereby destroying information.

Oftentimes, when credibility is an issue with B-F or development models, actuaries will simply rely on "industry data" or some other external source. This of course has a (generally unstated) credibility problem caused by lack of homogeneity.

An additional advantage of the model is that parameter uncertainty can be explicitly considered in this model. The following is a listing of relevant papers that the author has reviewed that provide uncertainty models that may be incorporated into the model framework:

» *Heckman, Philip E.; Meyers, Glenn G.- The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions* [7]
» *Kreps, Rodney - Parameter Uncertainty in (Log) Normal Distributions* [8]

Since we tend to employ lognormal claim cost models, we tend to use the derivation of parameter uncertainty as described by Kreps [8].

Claims Simulation

We now have all the elements necessary to simulate the pure IBNR reserve. Our simulation proceeds according to the claims process illustrated in Figure 1.

Step 1. Simulation of the number of claims – the number of claims is simulated using a stochastic model with a mean equal to the estimated IBNR claims frequency discussed previously. A Poisson model is often used to simulate the number of claims; however there is no requirement to do so. (We can also use a prior distribution for the Poisson parameter to incorporate parameter risk – which, as is well known, results in a negative binomial model.)

Step 2. Simulation of accident year – For each of the simulated claims, the accident year is simulated. The distribution of claims by accident year is based on the discrete distribution of IBNR claims by accident year derived in our estimation of IBNR claims frequency.

Step 3. Simulation of accident date – the accident date is simulated using a uniform distribution between the inception and expiration of the accident year simulated in Step 2. It is recognized that accident dates are not uniformly distributed throughout the accident year. That is, it is expected that there are more IBNR claims resulting from accidents occurring later within the simulated accident year. However, this is not considered to be a material weakness in the model and consideration of this nuance is a further area of development left to the interested actuary.

Step 4. Simulation of report date – the report date is simulated by adding a simulated report lag to the simulated accident date from Step 3. In this simulation, it must be recognized that the domain of possible report lags has a minimum value of the difference between the data evaluation date and simulated accident date. (For pure IBNR claims, the report date of the claim must, by definition, be greater than the valuation date of the data.) Therefore the report lag should be simulated using a truncated distribution.

Step 5. Simulation of claim closing date – the claims closing date is simulated by adding a simulated closing lag to the simulated report date from Step 4.

Step 6. Determination of present value factor – the present value factor applicable to each claim is calculated using the claim closing date. In general, for professional liability coverage, all indemnity is paid at claims closing and partial payments are not an issue. The calculated present value factor should consider that expenses are paid in advance of the claim closing date. It would not be overly difficult to develop a pattern for the payment of expenses; however, this consideration is beyond the scope of this paper.

Step 7. Simulation of claims outcomes – using the discrete distribution of claims settlement probabilities, we simulate the outcome of each claim.

Step 8. Simulation of claim value – the final step is estimation of claims settlement values. Using a mixed distribution model, the simulation of claim cost is a two-step process. The first step is the determination of which component of mixture generates the loss. This is simulated using a discrete distribution and the weight of each component in the mixture. The second step is to determine the claim value. This is simulated directly using the parameters and model-form of the component of the mixture that generates the loss.

Using the parameter uncertainty model described in Kreps [8], the severity parameters may also be simulated in each iteration.

The parameters of stochastic distribution of claim values should be adjusted for trend between the simulated accident date and the date at which the claims severity model is evaluated. The discussion of the impact of trend on loss distributions (and their parameters) is contained in Hogg & Klugman [6]. Essentially, this allows for each and every claim to be individually adjusted for trend. Furthermore, the model allows for trend adjustment based on accident date, report date or date of closing. The ability of the model to respond to trend in this manner is a benefit relative to link-ratio methods.

This simulation is shown on Exhibit 5. While spreadsheet tools can be used to simulate all these values, simulation software packages (often spreadsheet add-ins) will facilitate this process.

Through this process we have now have a comprehensive listing of all IBNR claims and all necessary information on those claims to:

» assign claims to report period,
» assign limits and retentions given claims made excess coverages, and
» calculate claims at various indemnity and / or expense retentions.

IV. A Specific Case – Bulk Reserves

The estimation of bulk reserves is simply a special case of the IBNR simulation model. Bulk reserves are simply the difference between case reserves and ultimate claims values. The process described above may also be applied to known claims with the following exceptions:

» The number of reported open claims is known and therefore does not need to be simulated.

» The accident dates and report dates of report claims are known and do not need to be simulated.

» In this simulation of closing lags it must be recognized that the domain of possible closing lags has a minimum value of the difference between the data valuation date and the actual report date. Therefore the closing lag should be simulated using a truncated distribution.

» Consideration may be given to the reserved value of the claim. If no consideration is given, the actuary (implicitly) assumes that the case reserve provides no predictive information. This may be a valid assumption for immature claims. For more mature claims, the reported value of the case should be considered in the simulation. Generally, this consideration results in "shifting" of the weights of the components of the mixed severity model or truncation of the severity distribution. Further research in this area is left to the interested reader.

» The simulated severity model should be truncated from below as the paid-to-date value of the claim. This may be a conservative adjustment as it will not allow claims to settle at their current value. An alternative would be to not consider paid amounts and allow individual claims to simulate at less then the paid value (most optimistic) or censor the resulting claim values to the paid value.

The reader should notice that these exceptions simply change the parameters of the simulation of ultimate values on known claims. The basic framework is identical to that used for the simulation of pure IBNR claims.

V. A Specific Case – Prospective Loss Forecast

The estimation of loss forecasts is also simply a special case of the IBNR simulation model. Loss forecasting and pure IBNR estimation are almost identical since no information is known about either claim type. The model is adjusted as follows to simulate prospective losses:

» The *a priori* estimate used in the B-F calculation is used as the mean estimate of prospective claims.[6]

» All claims occur within a single accident period. Therefore, the accident year need not be simulated.

[6] Essentially, this procedure is identical to that employed for pure IBNR frequency. The percentage of claims reported for a prospective period is by definition 0% and the estimated ultimate number of claims using a Bornhuetter –Ferguson model would be identical to the *a priori* frequency.

VI. Simulating From Truncated Distributions

Many aspects of this model require simulation from truncated distributions. Many simulation software packages allow for truncated distributions. If the actuary is either using spreadsheet software to perform the simulations or using simulation software that can not accommodate truncated distributions, the actuary can use inverted distributions to sample from a truncated distribution. Specifically,

- » Calculate the CDF at the truncation points. For example, for distribution truncated from below at a, calculate F(a).
- » Sample from a uniform distribution between the truncation points. In this example, the sampling would be between F(a) and 1.00 (= F(infinity)). We will designate the sampled value as U.
- » Calculate the value at which the CDF is equal U. In this example, the value would be equal to $F^{-1}(U)$ and provides our sampled value from a truncated distribution.

VII. Conclusions and Areas for Further Research

The framework of the model presented herein provides a model that is adaptable, accounts for inflation, estimates risk (both process and parameter), and is easily extendable for loss forecasting applications. Finally, this model allows for consistency in the estimation of loss forecasts and loss reserves. The model attempts to replicate the claims process rather than representing a proxy model for future emergence.

The goal of this paper is to present a model framework. However, as with all actuarial models, this model remains a "work-in-progress." Several areas for model enhancement are listed below and left to the practioner:

- » Relationships between lags and claim costs: In the current model, claims severity is independent of report lag and closing lags are independent of claims severity. The prevailing theory is that larger claims are reported later and take longer to settle.

- » Additional methods to incorporate claim information on known claims: The model provides one method by which known claim information (specifically, case reserve values) can be considered in the calculation of bulk reserves. In the evaluation of a large number of known open claims, no consideration may be necessary as all claim types would be assumed to by represented in the sample. However, for situations involving the evaluation of a smaller number of open claims, it would be desirable to develop models under which the simulated severity considered as much of the information regarding these claims as possible.

Essentially because the model is simulating the measurable aspects of the claims process, it allows the actuary an almost limitless opportunity to study various relationships.

BIBLIOGRAPHY

[1] Barnett, Glen; and Zehnwirth, Ben, *"Best Estimates for Reserves,"* Casualty Actuarial Society Forum, 1998 Vol: Fall, Proceedings of the Casualty Actuarial Society 1978 Vol: LXV Page(s): 1-9, Casualty Actuarial Society: Arlington, Virginia.

[2] Weissner, Edward W., *"Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood,"* Proceedings of the Casualty Actuarial Society 1978 Vol: LXVPage(s): 1-9, Casualty Actuarial Society: Arlington, Virginia.

[3] Klugman, S.A.; Panjer, H.H.; and Willmot, G.E., *Loss Models: From Data to Decisions*, 1998, John Wiley and Sons: New York.

[4] Keatinge, Clive L., *"Modeling Losses with the Mixed Exponential Distribution,"* Proceedings of the Casualty Actuarial Society 1999 Vol: LXXXVI Page(s): 654-698, Casualty Actuarial Society: Arlington, Virginia.

[5] Philbrick, Stephen W., *"A Practical Guide to Single Parameter Pareto Distribution,"* Proceedings of the Casualty Actuarial Society 1985 Vol: LXXII Page(s): 44-84, Casualty Actuarial Society: Arlington, Virginia.

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

[7] Heckman, Philip E.; Meyers, Glenn G., *"The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions;"* Proceedings of the Casualty Actuarial Society 1983 Vol: LXX Page(s): 22-61, Casualty Actuarial Society: Arlington, Virginia.

[8] Kreps, Rodney E., *"Parameter Uncertainty in (Log) Normal Distributions,"* Proceedings of the Casualty Actuarial Society 1997 Vol: LXXXIV Page(s): 553-580, Casualty Actuarial Society: Arlington, Virginia.

Hospital for Injured Actuarial Students

Estimation of Losses and Expense as of 9/30/2002

Estimation of Report Lag

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Observations			(2) - (1)	(3) - (1)	f (x) / F (M)	ln (6)
Incident Date	Report Date	Valuation Date	Report Lag (Days) (x)	Maximum Report Lag for Claim (M)	Conditional Likelihood (L)	Conditional Log-Likelihood ln(L)
01/06/97	01/31/97	09/30/02	25	2,093	1.305E-04	-8.9439
01/09/97	01/31/97	09/30/02	22	2,090	1.149E-04	-9.0714
01/06/97	01/31/97	09/30/02	25	2,093	1.305E-04	-8.9439
01/06/97	01/31/97	09/30/02	25	2,093	1.305E-04	-8.9439
01/07/97	01/31/97	09/30/02	24	2,092	1.253E-04	-8.9846
04/01/97	04/23/97	09/30/02	22	2,008	1.149E-04	-9.0714
01/04/97	06/02/97	09/30/02	149	2,095	7.353E-04	-7.2153
01/04/97	06/02/97	09/30/02	149	2,095	7.353E-04	-7.2153
02/28/97	06/22/97	09/30/02	114	2,040	5.763E-04	-7.4590
02/28/97	06/22/97	09/30/02	114	2,040	5.763E-04	-7.4590
02/28/97	06/22/97	09/30/02	114	2,040	5.763E-04	-7.4590
04/11/97	07/14/97	09/30/02	94	1,998	4.804E-04	-7.6410
04/13/97	07/14/97	09/30/02	92	1,996	4 706E-04	-7 6615
01/05/97	07/24/97	09/30/02	200	2,094	9 420E-04	-6 9675
04/06/97	07/28/97	09/30/02	113	2,003	5 715E-04	-7.4672
07/08/97	07/31/97	09/30/02	23	1,910	1 201E-04	-9 0270
07/09/97	07/31/97	09/30/02	22	1,909	1 149E-04	-9.0713
03/22/97	08/18/97	09/30/02	149	2,018	7.353E-04	-7.2153
03/22/97	08/18/97	09/30/02	149	2,018	7.353E-04	-7.2153
05/19/97	09/05/97	09/30/02	109	1,960	5.526E-04	-7.5009
06/30/97	09/26/97	09/30/02	88	1,918	4 510E-04	-7.7040
06/27/97	09/26/97	09/30/02	91	1,921	4.657E-04	-7.6719
02/20/97	10/09/97	09/30/02	231	2,048	1.051E-03	-6 8583
02/15/97	11/03/97	09/30/02	261	2,053	1.142E-03	-6.7748
09/14/97	11/05/97	09/30/02	52	1,842	2.701E-04	-8.2169
07/14/97	11/12/97	09/30/02	121	1,904	6 090E-04	-7.4036
07/14/97	11/12/97	09/30/02	121	1,904	6.090E-04	-7.4036
04/01/97	12/10/97	09/30/02	253	2,008	1 119E-03	-6.7951
11/01/01	08/29/02	09/30/02	301	333	4.935E-03	-5.3115
11/01/01	08/29/02	09/30/02	301	333	4.935E-03	-5.3115
11/20/01	09/10/02	09/30/02	294	314	5.397E-03	-5.2219
11/20/01	09/10/02	09/30/02	294	314	5.397E-03	-5.2219
11/20/01	09/10/02	09/30/02	294	314	5.397E-03	-5.2219
07/12/01	09/11/02	09/30/02	426	445	3.430E-03	-5.6753
01/19/00	09/12/02	09/30/02	967	985	4.762E-04	-7 6496
09/20/01	09/23/02	09/30/02	368	375	4.390E-03	-5 4285
09/20/01	09/23/02	09/30/02	368	375	4.390E-03	-5 4285
⋮	⋮	⋮	⋮	⋮	⋮	⋮
01/13/01	09/25/02	09/30/02	620	625	1 854E-03	-6 2901
09/10/01	09/25/02	09/30/02	380	385	4.240E-03	-5.4632
08/11/02	09/25/02	09/30/02	45	50	3.593E-02	-3.3263
09/10/01	09/25/02	09/30/02	380	385	4 240E-03	-5.4632
02/03/00	09/25/02	09/30/02	965	970	4.834E-04	-7.6347
01/09/00	09/27/02	09/30/02	992	995	4.280E-04	-7.7565
10/01/01	10/02/02	09/30/02	366	364	4 605E-03	-5.3806
08/08/01	10/04/02	09/30/02	422	418	3.777E-03	-5.5787
04/07/01	10/10/02	09/30/02	551	541	2.436E-03	-6 0174
10/03/01	10/15/02	09/30/02	377	362	4.686E-03	-5.3631
10/03/01	10/15/02	09/30/02	377	362	4.686E-03	-5.3631
10/07/01	10/15/02	09/30/02	373	358	4.761E-03	-5.3472
08/24/00	10/21/02	09/30/02	788	767	1.035E-03	-6.8735
10/17/01	10/21/02	09/30/02	369	348	4.980E-03	-5 3023
04/13/00	10/23/01	09/30/02	558	900	1 470E-03	-6.5228
04/13/00	10/23/01	03/31/02	558	717	1 749E-03	-6 3489
			Conditional Log- Likelihood ln(L)			-9150 5841

Report Lag Model	Model	Rayleigh
	Parameters	
	b =	437 290
	Estimation	MLE
	$\Sigma \ln(L)$	-9,150.58

Exhibit 1

Hospital for Injured Actuarial Students

Estimation of Losses and Expense as of 09/30/2002

Estimation of IBNR Claims

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Calendar Period		Avg of (1) & (2)		(4) - (3)			CDF ((5))	(7) / (8)	(9) / ((6))	A.*(6)*(1-(8))
Inception	Expiry	Average Accident Date	Valuation Date	Time Available to Report (X)	Exposure	Claims Reported to Date	Expected Percent Reported	Estimated Ultimate Claims	Frequency	Estimated IBNR Claims at 09/30/2002
1/1/1997	12/31/1997	07/02/97	09/30/2002	1,916	2,741	294	100%	294.02	0.107	0.02
1/1/1998	12/31/1998	07/02/98	09/30/2002	1,551	2,838	324	100%	324.60	0.114	0.56
1/1/1999	12/31/1999	07/02/99	09/30/2002	1,186	2,843	322	97%	330.35	0.116	7.65
1/1/2000	12/31/2000	07/01/00	09/30/2002	821	2,929	278	83%	335.75	0.115	53.63
1/1/2001	12/31/2001	07/02/01	09/30/2002	455	3,144	105	42%	251.18	0.080	194.83
1/1/2002	12/31/2002	07/02/02	09/30/2002	90	3,322	8	2%	381.74	0.115	346.26
				Total	17,816	1,331				602.95
						A.	Selected *a priori* frequency		0.106	

Exhibit 2

Hospital for Injured Actuarial Students

Estimation of Losses and Expense as of 09/30/2002

Claims Settlement Model

(1)	(2)	(3)	(4)	(5)	(6) (4) / ((3)+(4)+(5))	(7) (3) / ((3)+(4)+(5))	(8) (5) / ((3)+(4)+(5))
Period		Closed w/ Indemnity	Expense Only Claim	Closed with no payment	% of Expense Only	% of Closed with Indemnity	% of Closed with no payment
Total by Occurrence Year							
Inception	Expiry						
1/1/1997	12/31/1997	77	103	71	41.0%	30.7%	28.3%
1/1/1998	12/31/1998	63	98	72	42.1%	27.0%	30.9%
1/1/1999	12/31/1999	44	83	75	41.1%	21.8%	37.1%
1/1/2000	12/31/2000	13	42	71	33.3%	10.3%	56.3%
1/1/2001	12/31/2001	1	4	23	14.3%	3.6%	82.1%
1/1/2002	12/31/2002	0	0	2	0.0%	0.0%	100.0%
Total		198	330	314	39.2%	23.5%	37.3%
Total by Closing Year							
Inception	Expiry						
1/1/1997	12/31/1997	0	0	0			
1/1/1998	12/31/1998	2	5	4	45.5%	18.2%	36.4%
1/1/1999	12/31/1999	12	37	19	54.4%	17.6%	27.9%
1/1/2000	12/31/2000	43	58	23	46.8%	34.7%	18.5%
1/1/2001	12/31/2001	68	105	60	45.1%	29.2%	25.8%
1/1/2002	12/31/2002	72	114	65	45.4%	28.7%	25.9%
not coded		1	11	143	7.1%	0.6%	92.3%
Total		198	330	314	39.2%	23.5%	37.3%
		Selected Distribution			40.0%	24.0%	36.0%

Exhibit 3

Hospital for Injured Actuarial Students

Estimation of Losses and Expense as of 09/30/2002

Severity Model

Model: Mixed		Normal Small	Normal Large		Limit Loss		Shock
Claim Type -->		Normal Small	Normal Large		Limit Loss		Shock
Component Model -->		Lognormal	Lognormal		Point Mass		Uniform
	$\mu_i =$	10.204	11.542	Mean	1,000,000	Minimum	1,000,000
	$\sigma_i =$	0.932	1.187	Std Dev.	-	Maximum	6,000,000
Truncation Point - Maximum Possible Claim		1,000,000	1,000,000		None		None
$w_i =$		0.566	0.378		0.047		0.009

Exhibit 4

Hospital for Injured Actuarial Students

Estimation of Losses and Expense as of 09/30/2002

Simulation of True IBNR Claims
Sample Iteration

(1)	(2)	(3)	(4)	(5)	(8)
Claim No.	Accident Year	Date of Occurrence	Report Date	Type of Claim	Indemnity & Expense
1	2001	01/19/01	12/22/03	Normal Large	68,854
2	2001	04/30/01	01/08/04	Normal Small	63,381
3	2002	03/19/02	07/15/04	Normal Small	8,323
4	2000	09/23/00	01/17/03	Normal Small	14,616
5	2002	07/06/02	05/07/04	Shock Loss	2,004,453
6	2002	09/16/02	11/02/03	Normal Small	44,159
7	2002	07/12/02	09/13/03	Normal Large	32,163
8	2001	11/15/01	12/31/03	Normal Small	41,512
9	2001	12/26/01	03/19/05	Normal Small	93,695
10	2002	08/16/02	10/07/03	Normal Small	79,486
11	2002	01/08/02	10/28/02	Normal Small	41,606
12	2002	03/04/02	09/22/03	Limit Loss	1,000,000
13	2001	03/18/01	11/06/02	Normal Small	37,381
14	2002	06/01/02	04/03/04	Normal Large	294,091
⋮	⋮	⋮	⋮	⋮	⋮

Exhibit 5

Comparison of Risk Allocation Methods – Bohra-Weist DFAIC Distributions

By Trent R. Vaughn, FCAS

Abstract This study compare the results of several risk allocation methods for a realistic insurance company example.

INTRODUCTION

This purpose of this study is to compare the results of several risk allocation methods for a realistic insurance company example. The basis for the study is the fitted loss distributions of Bohra and Weist (2001), which were derived from the hypothetical data for DFA Insurance Company (DFAIC). This hypothetical data was distributed by the Casualty Actuarial Society's Committee on Dynamic Financial Analysis, as part of its 2001 call for papers.

In addition, Ruhm and Mango (2003) utilized these fitted distributions to produce 2000 simulated loss scenarios for DFAIC. This detailed simulation data was included in a spreadsheet that accompanied their paper. The Ruhm-Mango simulation data is also an important source of input for this study.

All of the data, analysis and results for the study are shown on the accompanying Excel Workbook ("bohra-weist data.xls"). The actual Bohra-Weist fitted distributions are shown on the "Data for Study" sheet. This sheet also provides some explanatory notes regarding the Ruhm-Mango simulation data. The actual Ruhm-Mango simulation data, sorted in ascending order, is shown on several different sheets, including the "RMK Capital Consumption" sheet.

In general, the calculations for each of the individual methods are displayed on a separate sheet of the Workbook. There is also a "summary" sheet that summarizes the resulting allocation and pricing for each method. In order to focus on differences in the allocation results for the various methods, each of the methods has been "calibrated" to the same overall corporate premium level. This overall premium amount is $1,242,777, which represents a total risk loading of $100,000. This total corporate risk load could be based on a financial pricing model (such as the Fama-French 3-Factor Model), or it could simply be based on a judgmental ROE or combined ratio goal that has been set by the Board of Directors.

In this study, we have also ignored complications caused by existing reserves, loss discounting, and long-tailed payouts.[1] In other words, we are assuming that these loss distributions apply to a start-up insurance company with no existing reserves. We are also implicitly ignoring differences in the duration of payments for the various lines.

Some of the methods in this study also require a specific value for the policyholders' surplus of DFAIC. We have assumed that surplus at the time of writing is $900,000. This implies a premium-to-surplus ratio of $1,242,777 / $900,000 = 1.38. The expected return-on-equity (ROE) for the company, ignoring investment income, is $100,000 / $900,000 = 11.1%.

Many of the methods in this study[2] directly determine the capital cost allocation for each of the subject lines of business. For these methods, the resulting premium by line is then determined by the following formula:

Premium = Expected Loss + Pro-Rata Allocation of Risk Load x $100,000

Note that the expected loss in this formula is undiscounted. Also, there is no provision made for underwriting and loss adjustment expenses.

The remaining methods in this study[3] directly determine the premium for each of the subject lines of business. For these methods, the total premium is "calibrated" at $1,242,777. The corresponding capital cost allocation by line is then determined according to the following formula:

Capital Cost Allocation = (Premium – Expected Loss) / $100,000

The remainder of this paper will provide explanatory notes for each of the methods, followed by short summary of the observations and results of the study.

MYERS-READ METHOD

Myers-Read (2001) proposed a capital allocation method that is based on Option Pricing Theory (OPT). Myers/Read provided a separate version of their formula for both lognormal and normal underlying loss distributions. The calculations on the "Myers-Read"

[1] See Venter (2002) for a discussion of these issues.
[2] Namely, the following methods: Myers-Read, RMK with Capital Consumption, RMK with Variance, Covariance, XTVaR99, and XTVaR with Expected Loss Cutoff.
[3] The following methods: Variance Load, Standard Deviation Load, RCR, Wang Transform, PH Transform

sheet of this workbook are based on their lognormal model, which in turn utilizes the lognormal version of Margrabe's formula.[4]

Technically, this formula requires that the distribution of aggregate (i.e. all lines combined) losses and asset values is joint lognormal. Myers-Read point out that "if each line's future loss is lognormal, then the overall loss cannot be lognormal." However, they also state: "The following derivations of default values and surplus allocations assume that total losses (the sum of all lines' losses) and asset values are joint lognormal. The authors believe that this is a reasonable approximation even when individual lines' losses are also lognormal."

Thus, the Myers-Read lognormal model is well-suited to the lognormal fitted distributions of Bohra-Weist. However, there is an important caveat. On the top of p. 556, Myers-Read provide a formula for approximating the variance of the "lognormal" aggregate losses from the individual line data.[5] But the authors point out that this formula only provides a close approximation when "the line-by-line loss volatilities are not large." For this reason, we can't utilize the lognormal line data of Bohra-Weist directly in the Myers-Read formula – because the volatility of the HO-xCat line is extremely large relative to the other lines, and the approximating formula on p. 556 will not work. Thus, in the Myers-Read sheet, the HO-xCat and HO-Cat lines of Bohra-Weist have been combined into a single "Homeowners" line. The mean and standard deviation for this combined Homeowners line is based on the sample mean and sample standard deviation for Homeowners in the Ruhm/Mango simulations.[6]

The actual Myers-Read sheet in this study is set up exactly like the tables in the Myers-Read paper. In fact, all of the headings and labels are exactly the same.[7] The "standard deviation" in column D is expressed as a percentage of the mean, as in the Myers-Read tables. Thus, in more precise terms, it is actually the "coefficient of variation" of the lognormal line distributions.

The Myers-Read formula also requires additional assumptions regarding the standard deviation of asset returns and correlations between asset values and the individual line losses. These assumptions are shown in cells D15 through I15.

[4] See Margrabe (1978) for details.
[5] This formula is cell J14 on the "Myers-Read" sheet.
[6] For details, see the explanatory notes on the "Data for Study" sheet.
[7] We have, however, omitted the normal distribution results shown on the Myers-Read tables, since the Bohra-Weist data assumes lognormality.

RMK METHODS

In the previous section, the actual Bohra-Weist fitted lognormal distributions were utilized to determine the Myers-Read allocation. RMK procedures generally do not utilize fitted loss distributions. Instead, these fitted distributions are used to generate a large number of simulated scenarios; the actual RMK calculations are then performed on the simulated scenarios.

Ruhm/Mango (2003) generated a set of 2000 simulated scenarios from the Bohra-Weist fitted lognormal distributions. Additional notes on the simulation data are contained on the "Data for Study" sheet. The remaining methods in this study use the simulation data, as opposed to the actual fitted distributions.

Clark (2005) discusses various practical applications of the RMK methodology, including the allocation of risk load by component. In general, the RMK procedure requires a "riskiness leverage ratio" as a benchmark for the measurement of risk. Clark provides RMK risk allocation examples that are based on both a capital consumption and a variance risk measure.

The "RMK Capital Consumption" sheet determines the DFAIC allocation corresponding to Clark's Exhibit 3a. The raw Ruhm-Mango simulation data, sorted in ascending order, is shown in columns A through H. The riskiness leverage ratio (column R) is based on capital consumed (column O). In this procedure, the maximum amount of capital consumed is equal to the available surplus of $900,000. This capping reflects Kreps' (2005) statement that "once you are buried it doesn't matter how much dirt is on top."

The "RMK Variance" sheet determines the allocation corresponding to Clark's Exhibit 3b. In this case, the riskiness leverage ratio is based on variance, instead of capital consumption. Clark points out that RMK with variance is equivalent to an allocation by the covariance method. In order to verify this assertion, I have also provided an explicit covariance-based allocation on the "covariance" sheet of the workbook.

VARIANCE AND STANDARD DEVIATION LOAD

Feldblum (1990) describes the variance and standard deviation load methods. Bault (1995) clarifies the underlying assumptions behind each of these two methods. The standard deviation load method sets the premium for each line according to the following formula:

Premium = Expected Loss + T x Standard Deviation

The formula for the variance load method is similar in form:

Premium = Expected Loss + T x Variance

The "variance and stdev" sheet displays the DFAIC premium calculation by line for each of these two methods. The calculations are based on the Ruhm-Mango simulation data, as opposed to the Bohra-Weist fitted distributions.[8] For each method, the "T" value (cells C2014 and C2015) is calibrated to the desired overall premium goal, and this T value is then used to determine the premium for each of the lines. The corresponding capital cost allocation is then determined according to the formula shown in the first section of this paper (see cells I9 through J14 on the "Summary" sheet for details).

EXCESS TAIL VALUE AT RISK (XTVaR)

Venter, Major, and Kreps (2005) discuss risk allocations that are based on the XTVaR risk measure. In order to utilize this method, one must select a "cutoff point" for the tail. Venter, Major, and Kreps provide the following guidance for this selection:

> *One possibility for establishing a cutoff probability for tail risk measures would be to use the probability of having any loss of capital at all. Then XTVaR would be the average loss of capital when there is a loss of capital. Another possible choice is the probability that capital is exhausted. The former is arguably more relevant to capital allocation, in that it charges for any use of capital rather than focusing on the shortfalls upon its depletion.*
>
> *On the other hand, policyholders tend to be sensitive to default. Studies suggest that they demand premium reductions one or two orders of magnitude greater than the expected value of the default cost in order to accept less than certain recovery. This is in part due to undiversified purchases of insurance. Thus the value of default has meaningful pricing effects, and policyholder concerns become quite relevant to shareholders as well.*

In this study, we have included an XTVaR calculation that is based on average loss of capital, as well as a version that is based roughly on default or insolvency. The "XTVaR Expected Loss" sheet utilizes a cutoff point that is equal to aggregate expected losses of $1,140,291.[9] As such, this version focuses on average loss of capital. The "XTVaR99" sheet utilizes a cutoff point that is equal to the 99th percentile of the aggregate loss distribution; hence, this version focuses on default outcomes.

---

[8] The calculations for these methods could also be easily performed on the fitted distributions. Provided that we have obtained a sufficiently large number of simulations, the premium and allocation results should be very similar.

[9] As in the previous section, the XTVaR calculations are based on the simulation data, as opposed to the Bohra-Weist distributions. In theory, the XTVaR formulas could also be applied directly to the fitted distributions.

Both the XTVaR with Expected Loss Cutoff and the RMK with Capital Consumption methods focus on the measurement of average capital consumed. Not surprisingly, the two methods also produce very similar capital cost allocations, as shown on the "summary" sheet. In fact, the only difference between these two methods lies in the RMK's capping procedure; that is, capital consumption in the XTVaR calculation is allowed to exceed the available surplus of $900,000. In other words, if we were to eliminate the capping procedure on the "RMK Capital Consumption" sheet, the two methods would produce an identical result.[10]

WANG TRANSFORM

Wang (2002) provides a practical discussion and description of the Wang transform method.[11] On the "Wang Transform" sheet, we have utilized this method to determine the DFAIC premiums and capital allocation.

The calculations are based on the Ruhm/Mango simulation data, shown in columns A through H. Column J assigns objective probabilities f(x) = 1/2000 to each scenario. Column K adds up the objective probabilities f(x) to get the cumulative probabilities F(x). Column L applies the Wang transform to F(x) to get the adjusted F*(x); the Sharpe ratio for this Wang transform is shown in cell 2007. Column M determines the risk-adjusted probability weights f*(x) from F*(x).

The Sharpe ratio is calibrated to produce a risk-adjusted mean of $1,242,777 for the total aggregate loss distribution. This Sharpe ratio is then used to determine the risk-adjusted means for each of the individual lines.

THE PROPORTIONAL HAZARDS (PH) TRANSFORM

Wang (1998) describes the Proportional Hazards (PH) transform. The "PH Transform" sheet applies this methodology to the DFAIC data. As with the other methods, the PH Transform has been calibrated to produce an overall risk load of $100,000; this results in a value for "r" of roughly 5/8 (see cell C2007). This r-value is then utilized to determine the PH-mean for the individual lines.

[10] This can be verfied by changing cell L2013 on the "RMK Capital Consumption" sheet to some very high value, say $10,000,000.
[11] For a more rigorous discussion and derivation, see Wang (2000).

RISK COVERAGE RATIO (RCR) METHOD

Ruhm (2001) describes the application of the risk coverage ratio (RCR) to insurance ratemaking and capital allocation. The "RCR" sheet applies the RCR method to the DFAIC data. Columns J through O display the ROE's by line and total for each of the 2000 simulated scenarios. Note that these ROE's are allowed to fall below -100% for very unfavorable scenarios.

The surplus allocation is required as an input to this method, as shown in cells C2011 to H2011. Premiums are then set (in cells C2012 through G2012) so that each line has the same risk coverage ratio. Provided that each of the lines is assigned a surplus amount greater than $0, the resulting RCR premiums by line are independent of the selected surplus allocation.[12] This can be easily verified by changing the surplus allocation in cells C2011 through F2011.

The Risk-to-Reward (or "R2R") Method is conceptually very similar to the RCR method. In fact, the two methods have been shown to produce identical results. This is verified in the "R2R" sheet of the Excel workbook.

MANGO CAPITAL CONSUMPTION

Mango (2003) introduced the capital consumption approach to capital loads. The "Mango Capital Consumption" sheet in this workbook was actually supplied by Don Mango. The methodology requires a key exogenous parameter, or utility "exponent" in cell R5. Mango provided the following explanatory remarks:[13]

> *[This method] uses the .... utility-type function to actually calculate scenario-level capital costs based on total UW loss. Costs are then spread back to LOB based on LOB UW loss only. This is different from [Clark's] RMK Capital Consumption approach, where "winners" are rewarded -- that is, LOB UW gains as well as UW losses @ scenario level are factored in, with the gains actually serving to reduce allocated capital (or cost). The approach I have done here mirrors the appendix of my Capital Consumption paper.*

Obviously, the exponent is cell R5 is a key input. Mango notes that his original choice for that exponent produced an allocation result that was very close to the Myers/Read allocation. The exponent was then adjusted to approximate the Myers/Read result as closely as possible.

---

[12] See Ruhm (2001) for an algebraic proof of this statement. Although not discussed it Ruhm, it is worth noting that this statement is only true if we allow for ROE values that are less than -100%.
[13] Source: Correspondence in private listserv group.

However, note that changes in this exponent parameter can have a big impact on the resulting allocation. Specifically, as the exponent parameter is increased (starting from a baseline of zero), more capital is allocated to Homeowners, and capital for all other lines is reduced (monotonically). In the actual "Mango Capital Consumption" sheet in the workbook, I have "calibrated" this exogenous parameter to produce the desired total risk load of $100,000 (as shown in cell S2009).[14]

SUMMARY AND CONCLUSIONS

The "summary" sheet displays the allocation results and resulting premiums for each of the methods in this study. The most significant difference between the various allocation results involves the relative allocation between the Homeowners line (which is highly cat prone) and the remaining lines. The standard deviation method allocates 59.4% of surplus to Homeowners, and the variance method allocates 88.9%. Many of the other methods fall within this s.d.-to-variance "band"; in fact, the allocations produced by the PH Transform, Covariance, Myers/Read, RMK with Variance, Mango Capital Consumption, and XTVaR99 are all remarkably similar. However, both the Wang Transform and RCR/R2R methods resulted in an Homeowners allocation that was lower than the s.d.-to-variance "band". And both the XTVaR with Expected Loss Cutoff and D. Clark's RMK with Capital Consumption resulted in an extremely low Homeowners allocation, relative to the other methods.

In general, the conceptual dichotomy between "cat prone" and "non cat prone" lines represents an important business management issue in our industry. This is especially true in the wake of Hurricanes Katrina and Rita, as most of the popular catastrophe models are now producing much higher PML's for the industry. As such, the decision regarding the percentage of capital and capital costs that should be allocated to cat-prone lines has critical implications on pricing, marketing and reinsurance purchases. Moreover, depending on the capital allocation model that you use, you'll get a much different answer to this problem.

[14] This is consistent with the approach taken in this paper for other methods that require an exogenous parameter (for example, the Sharpe ratio of the Wang Transform or the "r" of the PH-Transform). That is, the exogenous parameter has been "calibrated" to the same overall target risk load or premium.

BIBLIOGRAPHY AND REFERENCES

[1] Bault, Todd, Discussion of Feldblum, *PCAS* 1995, pp. 78-96.

[2] Bohra, Raju, and Weist, Thomas, "Preliminary Due Diligence of DFA Insurance Company," Casualty Actuarial Society *DFA Call for Papers*, 2001, pp. 25-58.

[3] Clark, David R., "Reinsurance Applications for the RMK Framework," Casualty Actuarial Society, 2005 Spring *Forum*, pp. 353-366.

[4] Feldblum, Sholom, "Risk Loads for Insurers," *PCAS*, 1990, pp. 160-195.

[5] Kreps, Rodney, "Riskiness Leverage Models," *CAS Proceedings* 2005, forthcoming.

[6] Mango, Donald F., "Capital Consumption: An Alternative Methodology for Pricing Reinsurance," Casualty Actuarial Society *Forum*, Winter 2003, pp. 351-378.

[7] Margrabe, W., "The Value of an Option to Exchange One Risky Asset for Another," *Journal of Finance*, 1978, 33:177-186.

[8] Myers, Stewart C., and Read, James A., "Capital Allocation for Insurance Companies," *Journal of Risk and Insurance*, 2001, Vol. 68, No. 4, pp. 545-580.

[9] Ruhm, David, "Risk Coverage Ratio: A Leverage-Independent Method of Pricing based on Distribution of Return," presented to 2001 ASTIN Colloquium.

[10] Ruhm, David, and Mango, Donald, "A Risk Charge Based on Conditional Probability," presented at 2003 Bowles Symposium.

[11] Venter, Gary G., Major, John A, and Kreps, Rodney E., "Additive Marginal Allocation of Risk Measures," unpublished working paper, 2005.

[12] Venter, Gary G., "Allocating Surplus – Not!" *Actuarial Review*, February, 2002.

[13] Wang, Shaun S., "A Class of Distortion Operators for Pricing Financial and Insurance Risks," *The Journal of Risk and Insurance*, 67, 1, March 2000, pp. 15-36.

[14] Wang, Shaun S., "A Universal Framework for Pricing Financial and Insurance Risks," *ASTIN Bulletin*, November, 2002.

[15] Wang, Shaun S, "Implementation of Proportional Hazards Transforms in Ratemaking," *PCAS*, 1998, pp. 940-979.

The Path of the Ultimate Loss Ratio Estimate

Michael G. Wacek, FCAS, MAAA

Abstract

This paper presents a framework for stochastically modeling the path of the ultimate loss ratio estimate through time from the inception of exposure to the payment of all claims. The framework is illustrated using Hayne's lognormal loss development model, but the approach can be used with other stochastic loss development models. The behavior of chain ladder and Bornhuetter-Ferguson estimates consistent with the assumptions of Hayne's model is examined. The general framework has application to the quantification of the uncertainty in loss ratio estimates used in reserving and pricing as well as to the evaluation of risk-based capital requirements for solvency and underwriting analysis.

Keywords: Stochastic model, diffusion process, loss development, loss ratio estimation, lognormal, Student's *t*, log *t*, parameter uncertainty

1. INTRODUCTION

Ultimate loss ratio estimates change over time. The initial loss ratio estimate that emerges from the pricing analysis for a tranche of policies soon gives way to a new estimate as time passes and claims begin to emerge (or not). By the time all claims have been paid, the loss ratio is likely to have been re-estimated many times. The focus of this paper is on how to model the future revisions of these ultimate loss ratio estimates. We illustrate the approach using loss ratio estimates based on chain ladder and Bornhuetter-Ferguson methods underpinned by a simple stochastic model described by Hayne [1].

There appears to be little, if any, actuarial literature on the subject of behavior of a ultimate loss ratio estimate *between* the time when it is made and the time when its final value becomes known, i.e., the point at which all claims have been paid. Various authors have sought to address uncertainty in the ultimate loss ratio estimate, but generally from the perspective of a single point in time.

For example, Hayne [1] proposed a lognormal model of loss development that supports the construction of confidence intervals around the ultimate loss ratio estimate[1]. Kelly [2] and Kreps [3] also used a lognormal framework to explore issues of parameter estimation and parameter uncertainty, respectively. Hodes, Feldblum and Blumsohn [4] used a slightly different lognormal development model to quantify the uncertainty in workers compensation reserves. Mack, Venter and Zehnwirth have all written extensively about

[1] Conscious that the confidence intervals he derived were dependent on the lognormal model being the correct choice, he cautiously described his results as providing a "range of reasonableness."

stochastic modeling of the loss development process[2]. Others, including Van Kampen [11], Wacek [12] and the American Academy of Actuaries Property & Casualty Risk-Based Capital Task Force [13], have sought to quantify the uncertainty in the ultimate loss ratio estimate used in pricing and reserving applications directly, without reference to the loss development process. The question on which all of these authors focused their attention is the potential variation in the final loss ratio at ultimate compared to the current ultimate loss ratio estimate, with no reference to how the ultimate loss ratio estimate might vary at intermediate points in time.

In contrast, in his acclaimed paper on solvency measurement Butsic [14] observed that loss estimates change in their march through time. He recognized that they, like stock prices, are governed by a diffusion process, a type of continuous stochastic process with a time-dependent probability structure. However, he did not propose a model of this stochastic process.

How ultimate loss ratio estimates change in the future depends in part on the method used to make the estimates. In this paper we assume that loss ratio estimates are derived from a consistently applied estimation process with minimal subjective overriding of the indicated result. We model the behavior of loss ratio estimates using stochastic versions of two loss development methods: the chain ladder method and the Bornhuetter-Ferguson method, both using paid development data. To model chain ladder estimates, we combine Hayne's and Butsic's ideas to synthesize a lognormal diffusion model for the path of the ultimate loss ratio. Then we adapt that model to the Bornhuetter-Ferguson method.

This conceptual framework, which could easily be adapted to handle other loss development models, provides actuaries with the means to give their clients more information about how much their loss ratio or reserve estimates may fluctuate from period to period. As such, it can be a useful tool for managing expectations about the variability of loss reserve estimates. It also has potential application in a number of other areas of actuarial analysis, as we will discuss later.

1.1 Organization of the Paper

The paper comprises six sections, the first being this introduction. In Section 2 we outline Hayne's lognormal model of chain ladder loss development and illustrate its application using industry Private Passenger Auto Liability data from the 2004 Schedule P. We illustrate the main benefit of a stochastic model for loss development, namely, the ability

---

[2] For example, see Mack [5], [6], Venter [7], [8] and Zehnwirth [9], [10] (the last co-authored with Barnett).

to measure the uncertainty in loss development factors and in the ultimate loss ratio estimate.

In Section 3 we discuss the effect of future loss emergence on future ultimate loss ratio estimates. We show how to use information implicit in Hayne's model to determine the distribution of future estimates derived from our stochastic versions of the chain ladder and Bornhuetter-Ferguson methods, with particular attention to the loss ratio estimate one year out. We again use industry Private Passenger Auto Liability data to illustrate the process.

In Section 4 we adjust Hayne's model to allow for parameter uncertainty, and illustrate the effect. Because the adjusted distribution does not have the multiplicative properties of the lognormal, we illustrate the use of Monte Carlo simulation to model the distribution of future ultimate loss ratio estimates.

In Section 5 we conclude with an outline of potential applications of the framework for future ultimate loss ratio estimates in loss reserving and risk-based capital applications.

2. HAYNE'S LOGNORMAL LOSS DEVELOPMENT MODEL

Hayne presented two models of chain ladder loss development: one that assumed that development is independent from one period to the next, and a second one that relaxed the independence assumption. We will adopt the first model (and henceforth refer to it simply as "Hayne's model"). Kelly [2] argued that independence is more plausible for paid loss development than for case incurred development. Therefore, we will use paid losses as the basis of our framework.

Hayne's model is quite simple. He assumed that age-to-age development factors are lognormally distributed. The product of independent lognormal random variables is also lognormal, which implies that age-to-ultimate loss development factors are lognormal. Because the product of a constant and a lognormal random variable is lognormal, if we are given the cumulative paid loss ratio at any age and the estimated parameters of the matching age-to-ultimate factor, we can determine the parameter estimates of the ultimate loss ratio. Using these parameters we can estimate the expected loss ratio (which we will take as the "best" estimate) as well as confidence intervals around that estimate.

The lognormal parameters μ and σ of the age-to-age factors can be estimated by a variety of methods. Hayne used (and we also prefer) the unbiased estimators $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n} ln(x_i)$ and $s^2 = \sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n-1}$ for μ and σ^2, respectively, where

$x_1, x_2, x_3, \ldots, x_n$ are the observed age-to-age factors[3]. $\bar{y}$ is also a maximum likelihood estimator.

2.1 Illustration of Model Parameter Estimation

We illustrate the parameter estimation for Hayne's model using the real loss development data presented in Exhibits 1 and 2. Exhibit 1 shows industry aggregate Schedule P net paid loss development data for Private Passenger Auto Liability for accident years 1995 through 2004 from the 2004 Annual Statement[4] together with the associated paid loss age-to-age development factors. The paid loss ratios at age one year are also included in the development factor exhibit. Exhibit 2 shows the natural logarithms of the age-to-age factors and the age one year paid loss ratios. The rows labeled "Mean" and "S.D." in Exhibit 2 show the unbiased estimators for μ and σ, respectively, given the data in the body of the column[5].

For example, in Exhibit 2 the mean and standard deviation of the natural logarithms of the observed age 1 to 2 development factors are 0.569 and 0.016, respectively. If we set $\mu = 0.569$ and $\sigma = 0.016$[6], these parameter estimates for prospective age 1 to 2 development imply a lognormal mean, defined as $E(x) = e^{\mu + 0.5\sigma^2}$, of 1.767, which matches the mean loss development factor calculated by the traditional method in Exhibit 1. The same is true for all of the other age-to-age factors. Similarly, the parameter estimates for the age one paid loss ratio are -1.246 and 0.069 for μ and σ, respectively, which imply a lognormal mean of 28.8%. This, too, matches the mean age one paid loss ratio shown in Exhibit 1.

The parameter estimates for the prospective age-to-age factors can be combined using the multiplicative property of lognormal distributions to determine parameter estimates for prospective age-to-ultimate factors. The product of n lognormal random variables with respective parameter sets (μ_1, σ_1), (μ_2, σ_2), $(\mu_3, \sigma_3), \ldots, (\mu_n, \sigma_n)$ is a lognormal random variable with parameters $\mu = \sum_{i=1}^{n} \mu_i$ and $\sigma = \left(\sum_{i=1}^{n} \sigma_i^2 \right)^{\frac{1}{2}}$. For example, treating age 10 as ultimate, in Exhibit 2 the μ parameter estimate for the age 7 to ultimate development factor is the sum of the mean age-to-age factors for ages 7 to 8, 8 to 9 and 9 to 10: 0.005 + 0.003 +

[3] We used unweighted estimators throughout this paper. For formulas for estimators using unequal weights for the observations, see Section 5.5 of [12].

[4] Source: Highline Data LLC as reported in the statutory filings (OneSource).

[5] Note that the standard deviation for the age 9 to 10 development factor, which is undefined, has been selected to be equal to that of the age 8 to 9 development factor in both Exhibit 1 and Exhibit 2.

[6] These parameters define the lognormal distribution that best fits the data, using unbiasedness as the criterion for "best." However, there is uncertainty about whether those parameters are correct. We address the issue of parameter uncertainty later in the paper.

0.001 = 0.009. The corresponding σ parameter is the square root of the sum of the variances of the same age-to-age factors: $\sqrt{0.000^2+0.001^2+0.001^2}=0.001$. Note that the lognormal means (labeled "LN Fit LDFs" in Exhibit 2) implied by these age-to-ultimate parameters match the age-to-ultimate development factors shown in Exhibit 1.

The ultimate chain ladder loss ratio estimates indicated by this analysis as of the end of 2004 for accident years 1995 through 2004 are summarized in Exhibit 3. In this example, the lognormal loss development model produces the same loss ratio estimates as the traditional deterministic chain ladder loss development method. If we were interested only in these mean estimates, the traditional approach would suffice. However, we also want to measure the uncertainty in the loss ratio estimates, and for that purpose the richer lognormal model is superior.

2.2 Measurement of Loss Development Uncertainty

If we assume $\mu=\bar{y}$ and $\sigma=s$ based on the data for each age-to-age development period, we can calculate the lower and upper bounds of a two-sided 95% confidence interval for prospective age-to-age factors as $e^{\bar{y}-N^{-1}(97.5\%)\cdot s}$ and $e^{\bar{y}+N^{-1}(97.5\%)\cdot s}$, respectively, where $N^{-1}(97.5\%)$ is the value of the standard normal cdf corresponding to a cumulative probability of 97.5%[7]. Similarly, using the parameter estimates for the age-to-ultimate factors we can also determine confidence intervals for age-to-ultimate factors. We have tabulated these 95% confidence intervals based on the industry Private Passenger Auto Liability Schedule P data as of the end of 2004 in Exhibit 4[8].

Exhibit 4 indicates that the age 1 to 2 development factor, which has an estimated mean of 1.767, should fall within a range of 1.710 to 1.824 95% of the time. The age 1 to ultimate development factor, which has an estimated mean of 2.508, can be expected to fall within a range of 2.423 to 2.595 95% of the time. Given the accident year 2004 paid loss ratio of 26.6% at age 1, these confidence intervals imply a paid loss ratio range at age 2 of 45.5% to 48.5% ($47.0\%\pm1.5\%$) and an ultimate loss ratio range of 64.4% to 69.0% ($66.7\%\pm2.3\%$)[9].

As we would expect, the development factors for more mature accident years have tighter confidence intervals. For example, the age 5 to 6 factor, which in a year end 2004 analysis

---

[7] $N^{-1}(97.5\%)$ is replicated in Excel by *NORMSINV*(0.975).

[8] Bear in mind that these confidence intervals are premised on the parameter estimates being correct and are narrower than confidence intervals that incorporate parameter uncertainty.

[9] While the lognormal is a skewed distribution, the skewness is imperceptible for small values of σ and the confidence intervals are, for most practical purposes, symmetrical. In this example with σ = 0.016 the skewness coefficient is 0.05. In contrast, in the case of σ = 1 it is 6.18.

would be applicable to accident year 2000, has an estimated mean of 1.020 and a 95% confidence range of 1.018 to 1.022. That implies that, 95% of the time, the accident year 2000 paid loss ratio of 76.7% as of the end of 2004 will develop to a paid loss ratio of 78.1% to 78.4% by the end of 2005, a range of 0.3 points. The 95% confidence interval for the age 5 to ultimate factor, which has an estimated mean of 1.039, is a range of 1.034 to 1.043. That implies an ultimate loss ratio range of 79.3% to 80.0%, or 0.7 points.

All of these development factor, loss ratio and confidence interval estimates are as of the end of 2004. They are all subject to change as new information in the form of actual future loss emergence becomes available. In the next section we will show how to use information implicit in Hayne's approach to model the effect of future loss emergence on these estimates.

3. A MODEL FOR FUTURE ULTIMATE LOSS RATIO ESTIMATES

Any estimate of the ultimate loss ratio for a particular accident year is quickly made obsolete by subsequent actual loss emergence. Because of this rapid obsolescence, the ultimate loss ratio must be re-estimated periodically in light of the loss development in the period since the previous evaluation. That loss development affects the new estimate in two ways.

3.1 Sources of Variation in Future Loss Ratio Estimates

First, the actual accident year loss emergence replaces the expected emergence in the loss ratio projection. For example, in Exhibit 3 the Private Passenger Auto Liability accident year 2004 ultimate loss ratio of 66.7%, estimated as of the end of 2004, was determined by applying an age-to-ultimate factor of 2.508 to the paid loss ratio of 26.6%. That age-to-ultimate factor reflected an *expected* age 1 to 2 development factor of 1.767 combined with an age 2 to ultimate factor of 1.420.

It is likely that *actual* age 1 to 2 loss development will vary from the expected. If, for example, the actual accident year 2004 emergence during 2005 (from age 1 to 2) corresponds to a development factor of 1.75, then in the ultimate loss ratio analysis conducted at the end of 2005 this actual development factor will replace the expected development factor of 1.767. If the age 2 to ultimate factor remains unchanged at 1.420, the revised chain ladder estimate of the ultimate loss ratio will become 26.6% x 1.75 x 1.42 = 66.1%. The revised

Bornhuetter-Ferguson loss ratio estimate will become 26.6% x (1.75 – 1.767) + 26.6% x 1.767 x 1.42 = 66.1%[10].

Of course, loss emergence with respect to older accident years might cause a revision in the prospective age 2 to ultimate factor. This potential for tail factor revision is a second source of uncertainty. For example, suppose the actual age 2 to 3 development on accident year 2003 during 2005 corresponds to a factor of 1.210. If that factor is averaged with the previous eight-point mean of 1.198 determined in Exhibit 1 (using loss development data through 2004), the result is a revised age 2 to 3 development factor of 1.199. Assuming the same process is repeated for the other development periods, a revised age 2 to ultimate factor will be obtained. If the resulting age 2 to ultimate factor is 1.425, the revised chain ladder ultimate loss ratio estimate is given by 26.6% x 1.75 x 1.425 = 66.3%, a reduction of 0.4% from the year end 2004 ultimate loss ratio estimate of 66.7%. The revised Bornhuetter-Ferguson estimate in this case is given by 26.6% x (1.75 – 1.767) + 26.6% x 1.767 x 1.425 = 66.5%.

The foregoing is an illustration of just one scenario of the loss development that might occur in 2005 and its effect on the ultimate loss ratio estimate. We can use information developed in Hayne's framework to model these two effects generally.

3.2 Modeling the First Source of Variation – Accident Year Development

The first effect is captured by the lognormal random variable estimated for the next year of development with respect to the accident year under review. For example, for accident year 2004, which at the end of 2004 is age 1, the lognormal distribution with $\mu = 0.569$ and $\sigma = 0.016$ models age 1 to 2 paid development. Then, since the age 1 paid loss ratio is 26.6%, the paid loss ratio distribution at age 2 is lognormal with parameters $\mu = ln\ 26.6\% + 0.569 = -0.756$ and $\sigma = 0.016$, implying a mean of 47.0%.

If the mean age 2 to ultimate factor (the tail factor) of 1.42 does not change, then the distribution of the revised chain ladder ultimate loss ratio estimate at age 2 (i.e., one year out) has lognormal parameters $\mu = ln\ 26.6\% + 0.569 + ln\ 1.42 = -0.406$ and $\sigma = 0.016$. The random variable for this chain ladder estimate x_{CL} can be expressed as a function of the paid loss ratio random variable x_P and the expected value of the mean tail factor:

$$x_{CL} = x_P \cdot E(tail) \tag{3.1}$$

[10] Assume the Bornhuetter-Ferguson expected loss ratio is 66.7%. In general, we will assume the Bornhuetter-Ferguson expected loss ratio for each accident year as of the end of 2004 is equal to the chain ladder ultimate shown in Exhibit 3, allowing us to treat the year end 2004 loss ratio estimates as identical from both methods.

The random variable x_{BF} for the comparable Bornhuetter-Ferguson estimate is a shifted version of the random variable for the age 2 paid loss ratio:

$$x_{BF} = x_P - E(x_P) + E(x_P) \cdot E(tail) \tag{3.2}$$

As defined by Formulas 3.1 and 3.2, both x_{CL} and x_{BF} reflect the uncertain impact of accident year 2004 development during 2005 on the updated ultimate loss ratio estimate that will be made at the end of 2005, but do not reflect the potential impact of tail factor revision.

3.3 Modeling the Second Source of Variation – Tail Factor Revision

The second effect, due to tail factor revision, is captured by measuring the effect of the lognormal loss development modeled for the next year on the existing mean age-to-age and age-to-ultimate factors. For example, the mean age 2 to 3 development factor shown in Exhibit 1 is 1.198. This is a mean of eight data points. What will be the effect on the mean of adding a ninth data point (representing 2005 development on accident year 2003), given that it will arise from a lognormal distribution with parameters $\mu = 0.181$ and $\sigma = 0.005$ (and mean of 1.198)? The uncertain ninth data point will contribute one-ninth weight to the revised mean age-to-age factor. There is no uncertainty about the existing mean age 2 to 3 factor – it is a constant. Therefore, the σ parameter of the distribution of the revised mean age 2 to 3 factor one year out, given an additional year of actual development, is given by $\sqrt{(\frac{8}{9} \cdot 0)^2 + (\frac{1}{9} \cdot 0.005)^2} = 0.001$. The μ parameter is given by $ln\, 1.198 - 0.5 \cdot 0.001^2 = 0.181$. We can use the same process to estimate μ and σ parameters for the comparable distributions of mean age-to-age factors one year out for all such factors comprising the development tail[11]. We can then combine the revised mean age-to-age factor parameters to determine the parameters of the revised mean age-to-ultimate factor distributions. See Exhibit 5 for a tabulation of the parameters of these revised mean age-to-age and age-to-ultimate distributions for all ages. The σ of the distributions of revised factors for age 3 to 4 and beyond is less than 0.0005 (and thus displayed as 0.000 in Exhibit 5), indicating that for Private Passenger Auto Liability, the uncertainty arising from the potential for tail factor revision is very small. This is confirmed by the very narrow confidences intervals.

[11] Bear in mind that these parameters refer to distributions of the *mean* age-to-age development factor one year out and not to distributions of the development factor itself. We are interested in the distribution of the mean development factor because changes in the mean directly affect the ultimate loss ratio estimate (which is also a mean).

3.4 Modeling the Revised Loss Ratio Estimate One Year Out

We can now combine these two effects to determine the distribution of the revised ultimate loss ratio estimate that will be determined in one year's time based on the updated loss development experience that will then be available.

To determine the distribution of the revised chain ladder estimate, we start with the actual accident year paid loss ratio, which we then multiply by the lognormal random variables for 1) the age-to-age factor for the next year of development (obtaining the random variable x_p of the paid loss ratio one year out) and 2) the revised age-to-ultimate factor beyond the next year of development. Using accident year 2004 as an example, as of the end of 2004 the ultimate loss ratio estimate is 66.7%, which has been determined by multiplying the paid loss ratio of 26.6% first by an age 1 to 2 factor of 1.767 and then by an age 2 to ultimate factor of 1.420. In order to model the ultimate loss ratio estimate one year later, at the end of 2005, we replace the constant age 1 to 2 factor of 1.767 with the lognormal random variable with parameters $\mu_1 = 0.569$ and $\sigma_1 = 0.016$. In addition, we replace the constant age-to-ultimate factor of 1.420 with the lognormal random variable with parameters $\mu_2 = 0.350$ and $\sigma_2 = 0.001$. The expected values of these two lognormal random variables are 1.767 and 1.420, respectively. The product of the paid loss ratio (a constant) and these two lognormal random variables is lognormal with parameters $\mu = lnP + \mu_1 + \mu_2$ and $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$, where P represents the actual paid loss ratio at the end of 2004, which in this example, implies $\mu = -1.325 + 0.569 + 0.350 = -0.406$ and $\sigma = \sqrt{1.016^2 + 0.001^2} = 0.017$.

Generally, we can express the random variable x_{CL} as the product of the two lognormal random variables x_p and *tail*, representing the paid loss ratio one year out and the mean tail factor:

$$x_{CL} = x_p \cdot tail \tag{3.3}$$

Now we are in a position to determine confidence intervals for the revised chain ladder ultimate loss ratio estimate at the end of 2005. The endpoints of the two-sided 95% confidence interval are given by $e^{\mu - N^{-1}(97.5\%)\cdot\sigma}$ and $e^{\mu + N^{-1}(97.5\%)\cdot\sigma}$, which imply an estimated loss ratio range one year out for accident year 2004 of 64.5% to 68.8%, or approximately $66.7\% \pm 2.1\%$. Confidence intervals for ultimate loss ratio estimates one year out for the other accident years can be estimated in the same way, and are tabulated together with those for accident year 2004 in Exhibit 6.

To determine the distribution of the comparable revised Bornhuetter-Ferguson estimate, we replace the constant $E(tail)$ in Formula 3.2 with the random variable *tail*:

$$x_{BF} = x_P - E(x_P) + E(x_P) \cdot tail \qquad (3.4)$$

We can also determine confidence intervals for the revised Bornhuetter-Ferguson loss ratio estimate at the end of 2005. However, because the sum of two lognormal random variables, in this case x_P and *tail*, is not expressible in closed distributional form, the confidence intervals must be estimated using Monte Carlo simulation. The results of a simulation involving 10,000 trials are shown in Exhibit 7. For each of the trials we randomly selected observations from the distributions of x_P and *tail*, assuming independence, and combined them according to Formula 3.4 to arrive at a simulated Bornhuetter-Ferguson estimate. After tabulating the results of 10,000 such trials, we determined the lower and upper bounds of the 95% confidence interval of the loss ratio estimate by identifying the 2.5 percentile and the 97.5 percentile of the trial values. Not surprisingly, the 95% confidence intervals for the revised Bornhuetter-Ferguson estimates are narrower in every case than the revised chain ladder estimates.

3.5 Modeling the Revised Loss Ratio Estimate – Other Time Horizons

We can extend this process to longer time horizons and determine the distribution of the ultimate loss ratio estimate two years out, three years out and so on, until the time horizon encompasses the point when all claims are expected to have been settled. The modeling is conducted in essentially the same way as for the one year time horizon. For example, in the case of a two-year horizon, the first source of uncertainty (accident year development) is modeled using the distribution of the age j to $j+2$ development factor, where j is the age in years of the accident year under review. The second source of uncertainty (potential tail factor revision) is modeled by reference to the potential effect of two additional development data points on the mean tail factor for age $j+2$ to ultimate development. The analysis of a three year time horizon focuses on accident year development from age j to $j+3$ and the tail factor from $j+3$ to ultimate, but is otherwise identical to that for the one year and two year time horizons. The analysis of the ultimate loss ratio estimate at points further in the future proceeds in the same way.

Alternatively, we can model the path of the ultimate loss ratio estimate as a succession of annual revaluations. Exhibit 8 illustrates this by plotting the results of one simulation of the path of the accident year 2004 loss ratio estimates through time for estimates determined from both chain ladder and Bornhuetter-Ferguson methods. It represents just one path among many possibilities. The simulation was performed from the vantage point of the end of 2004. As such it incorporates everything we know about actual loss development through that time as well as what we can infer about the structure of future development. We started with the accident year 2004 loss ratio estimate as of the end of 2004, which was 66.7%. Then, based on one random simulation of loss development during calendar year 2005, we made new chain ladder and Bornhuetter-Ferguson estimates of the ultimate loss ratio as of the end of 2005. We repeated the process for calendar years 2006 through 2013 using the simulated cumulative loss development through each valuation date. Exhibit 8 is a plot of the results. A complete description of the probability structure of the path can be built up from a simulation involving a large number of random trials, or in the chain ladder case, directly from the properties of the lognormal distribution.

In practice, there might not be much benefit in determining the distribution of the chain ladder ultimate loss ratio estimate for time horizons between one year and the ultimate horizon (when all claims have been settled), at least for Private Passenger Auto Liability[12]. We see this in Exhibit 9, the top half of which compares the 95% confidence intervals for the accident year 1995 through 2004 chain ladder loss ratio estimates one year out with confidence intervals for the accident year loss ratio estimates over the ultimate time horizon. If we contrast the 95% confidence interval for accident year 2004 for the one year horizon with the 95% confidence interval for the chain ladder loss ratio estimate over the ultimate time horizon, we can see that the contribution from the out years is dwarfed by the contribution from the next twelve months. The 95% confidence interval for the ultimate time horizon indicates a range for the accident year 2004 loss ratio of $66.7\% \pm 2.3\%$, which is barely wider than the range for just one year out. This is true not only for accident year 2004, but also holds for accident years 1995 through 2003.

For example, the accident year 2003 confidence interval of approximately $67.8\% \pm 0.7\%$ for a one year time horizon is almost as wide as that for the time horizon to ultimate of $67.8\% \pm 0.8\%$. For all of the older accident years, the first year of future development accounts for more than half of the variation associated with the ultimate time horizon.

This phenomenon is not confined to loss ratio estimates over short vs. longer time horizons. The same effect is also seen in other situations not related to insurance, where variability is a function of time. For example, given the common assumption that future stock price movements are lognormally distributed and independent, the 95% confidence interval for a stock price one year out, given constant annualized volatility of $\sigma = 20\%$ and an expected value of \$66.70, is \$45.07 to \$98.71, a range of \$53.64. Assuming the same expected value of \$66.70, the 95% confidence interval for the stock price two years out is \$38.22 to \$116.11, a range of \$77.80. The confidence interval range for the one-year time

[12] There might be value in doing so for other lines that display more loss development variability.

horizon stock price is 69% of the price range for the two-year time horizon. The reason for the disproportionate impact of the first period is that the confidence interval is not a linear function of σ but rather of $\sigma\sqrt{t}$, where t represents the time lag in years. In the case of chain ladder ultimate loss ratio estimation, where the age-to-age σ typically declines as the accident year ages, this effect can be even more pronounced.

Turning now to the Bornhuetter-Ferguson estimates, which are inherently less variable, the effect is smaller but still evident. The bottom half of Exhibit 9 compares the 95% confidence intervals for accident year 1995 through 2005 loss ratio estimates one year out with the confidence intervals for the loss ratio estimates over the ultimate time horizon. In the Private Passenger Auto Liability example considered here, the 95% confidence interval for the accident year 2004 loss ratio estimate is approximately $66.7\% \pm 1.6\%$, which is about two-thirds of the range of the confidence interval for estimates at the ultimate time horizon. For all of the older accident years, as in the case of the chain ladder estimates, the first year of future development accounts for more than half of the variation associated with the ultimate time horizon.

3.6 Modeling the Loss Ratio Estimate at Inception

Up to this point we have focused on modeling the distribution of the ultimate loss ratio after losses have begun to emerge. However, there is no reason why we cannot extend essentially the same procedure backward to the inception of loss exposure at age 0. Indeed, the benefit of doing so is that we can obtain a complete model of the path of the ultimate loss ratio from inception to ultimate.

The main difference in the procedure is that the lognormal model for loss emergence between age 0 and 1 describes the behavior of the paid loss ratio rather than an age-to-age factor. The rest of the analysis is merely an application of Formula 3.3.

For example, assume for the sake of illustration that the age 1 paid loss ratios in Exhibit 1 are lognormally distributed and reflect "on level" adjustments to the accident year 2005 level. The mean age 1 paid loss ratio is 28.8%, which we can take as an estimate of the 2005 "on level" age 1 paid loss ratio. The unbiased estimates of the parameters of the lognormal distribution representing the paid loss ratio at age 1 are $\mu = -1.246$ and $\sigma = 0.069$. These parameters imply a lognormal mean paid loss ratio of 28.8% that matches the sample mean. The age 1 to ultimate development factor of 2.508 implies an ultimate loss ratio estimate at inception of 72.3%.

Applying the lognormal multiplicative rule described in Section 2, the parameters of the lognormally distributed ultimate loss ratio (at the ultimate time horizon) are $\mu = -1.246 + 0.919 = -0.327$ and $\sigma = \sqrt{0.069^2 + 0.018^2} = 0.071$, implying a 95% confidence interval of 62.8% to 82.9%, a range of 20.1%. The parameters of the ultimate loss ratio one year out are $\mu = -1.246 + 0.919 = -0.327$ and $\sigma = \sqrt{0.069^2 + 0.002^2} = 0.069$. The indicated 95% confidence interval is 63.0% to 82.6%, a range of 19.6%. These calculations are summarized in Exhibit 6.

The comparable Bornhuetter-Ferguson estimate can be determined by applying Formula 3.4. Exhibit 7 shows that the 95% confidence interval for the revised Bornhuetter-Ferguson estimate of the accident year 2005 loss ratio one year out is 68.6% to 76.4%, a range of 7.8%.

4. ADJUSTING THE MODEL FOR PARAMETER UNCERTAINTY

In Section 2 we explained that, given the observations $x_1, x_2, x_3, \ldots, x_n$ arising from a lognormal process and the natural logarithms of the same observations $y_1, y_2, y_3, \ldots, y_n$ (where $y_i = \ln x_i$), the mean $\bar{y}$ and standard deviation s of the log-transformed sample are unbiased estimators of the lognormal process parameters μ and σ, respectively. The parameter selections $\mu = \bar{y}$ and $\sigma = s$ define the lognormal distribution $f(x \mid \mu, \sigma)$ that best fits the data, using unbiasedness as the criterion for "best."

However, while these are good estimates of the parameters, there is uncertainty about their true values. Fortunately, by combining information contained in the sample with results from sampling theory, it is possible to determine the mixed distribution $f(x)$ that reflects the probability weighted contribution of all of the potential parameter values. Wacek [12] showed that $f(x)$ defines a "log t" distribution[13] and, in particular that the random variable $y = \ln x$ is Student's t with $n-1$ degrees of freedom, mean $\bar{y}$ and variance $s^2 \cdot \frac{n+1}{n} \cdot \frac{n-1}{n-3}$.

4.1 Log t Confidence Intervals

The bounds of the two-sided log t 95% confidence interval are given by $e^{\bar{y} - T_{n-1}^{-1}(97.5\%) \cdot s \cdot \sqrt{(n+1)/n}}$ and $e^{\bar{y} + T_{n-1}^{-1}(97.5\%) \cdot s \cdot \sqrt{(n+1)/n}}$, respectively, where $T_{n-1}^{-1}(97.5\%)$ is the value of the standard Student's t cdf with $n-1$ degrees of freedom corresponding to a cumulative probability of 97.5%[14]. Two-sided 95% confidence intervals for Private Passenger Auto

[13] The log t bears the same relationship to the Student's t distribution that the lognormal bears to the normal.

[14] $T_{n-1}^{-1}(97.5\%)$ is replicated in Excel by *TINV*(0.05, n-1).

Liability age-to-age factors, based on the log *t* distribution, are shown in Exhibit 10. Unfortunately, the log *t* distribution does not share the multiplicative property of the lognormal. As a result, we cannot specify the distribution of age-to-ultimate development factors in closed form. Instead the age-to-ultimate factor distributions and related confidence intervals must be estimated using a Monte Carlo simulation procedure that determines the age-to-ultimate factor from the underlying age-to-age factors for each random trial.

In the top section of Exhibit 10, we have tabulated the indicated log *t* 95% confidence intervals for age-to-age factors based on the industry Private Passenger Auto Liability 2004 Schedule P data, together with the ratios of these confidence interval bounds to the lognormal confidence interval bounds given in Exhibit 4. In addition, we have tabulated the sample size for each development period as well as $T_{n-1}^{-1}(97.5\%)$ and the degrees of freedom used in the calculations. At the risk of being seen as statistically less than rigorous, we set a minimum degrees of freedom value of 3 for purposes of calculating the confidence intervals to avoid using log *t* distributions with an undefined variance.

The log *t* confidence intervals shown in Exhibit 10 for age-to-age factors are very close to the lognormal confidence intervals given in Exhibit 4. The largest difference is in the age 1 to 2 factor, where upper bound of the log *t* interval is 1.839, which is only 0.8% larger than the lognormal upper bound of 1.824. The percentage differences for the other age-to-age factors are smaller.

In the lower section of Exhibit 10, we have tabulated the 95% confidence intervals for age-to-ultimate factors indicated by a Monte Carlo simulation involving 10,000 trials. As was the case with the age-to-age factors, the differences between the log *t* confidence intervals and lognormal confidences intervals for the age-to-ultimate factors are quite small. For example, the largest difference is in the age 1 to ultimate confidence interval, where the upper bound of the log *t* interval is 2.619. This is only 0.9% larger than the lognormal upper bound of 2.595. The percentage differences for the other age-to-ultimate factors are smaller. This suggests that, at least for Private Passenger Auto Liability, the effect of parameter uncertainty is small enough that it can be ignored. However, it is important to bear in mind that this might not be the case for other lines of business.

4.2 Log *t* Simulation of Development Factors

In the Monte Carlo simulation of age-to-ultimate factors, for each trial we randomly selected one age-to-age factor from each of the log *t* distributions representing development from age 1 to 2, age 2 to 3, ..., age 9 to 10. Treating age 10 as ultimate, we then multiplied these age-to-age factors in the usual way to determine a set of age-to-ultimate factors for that trial. After the results of the 10,000 trials were tabulated, we determined the lower and upper bounds of the 95% confidence interval for each age-to-ultimate factor (age 1 to ultimate, age 2 to ultimate, etc.) by identifying the 2.5 percentile and the 97.5 percentile of the 10,000 trial values.

To make the random age-to-age factor selections, we started with a random draw R from the uniform distribution defined on the interval [0, 1]. Because R has a value between 0 and 1, it can be treated as though it is a cumulative probability. The number $T_{n-1}^{-1}(R)$ that corresponds to a standard Student's t cumulative probability of R is a random number from the standard Student's t distribution with $n-1$ degrees of freedom, which has a mean of zero and a variance of $\frac{n-1}{n-3}$. More generally, the corresponding random number from the Student's t distribution with $n-1$ degrees of freedom, mean M and variance $C^2 \cdot \frac{n-1}{n-3}$ is given by $M + T_{n-1}^{-1}(R) \cdot C$, which corresponds to a random number of $e^{M+T_{n-1}^{-1}(R) \cdot C}$ from the related log t distribution. Substituting the appropriate values of $\bar{y}$ for M and $s\sqrt{(n+1)/n}$ for C, we obtain $e^{\bar{y}+T_{n-1}^{-1}(R) \cdot s\sqrt{(n+1)/n}}$ as the value of a randomly selected age-to-age factor.

Putting some numbers to it, a draw of R = 0.873 implies a random age 1 to 2 development factor from the corresponding log t with 8 degrees of freedom of $exp(0.569 + 1.229 \cdot 0.016\sqrt{10/9}) = 1.803$[15]. If the next draw is R = 0.239, then the random age 2 to 3 factor, drawn from the corresponding log t with 7 degrees of freedom, is $exp(0.181 + (-0.749) \cdot 0.005\sqrt{9/8}) = 1.194$. Random numbers corresponding to the other development periods are similarly obtained. Then the age 1 to ultimate factor, the age 2 to ultimate factor, age 3 to ultimate factor, and so on, are obtained by multiplication. Tabulation of these results completes the first trial. The process is repeated in the same way for 10,000 trials.

4.3 Log *t* Simulation of Future Loss Ratio Estimates

Under conditions of parameter uncertainty the distribution of future loss ratio estimates must also be modeled using Monte Carlo simulation. Each of the lognormal age-to-age

[15] $T_{n-1}^{-1}(R)$ is replicated in Excel by $TINV(2(1-R), n-1)$ if R > 0.5, and $-TINV(2R, n-1)$, if R ≤ 0.5. $TINV$ assumes users are interested in two-tailed applications and therefore takes as its first argument the total two-tail probability. It returns values only from the right half of the distribution.

development components identified in Section 3 must be replaced with corresponding log *t* components.

For example, to estimate the distribution of the updated chain ladder estimate of the accident year 2004 ultimate loss ratio at the end of 2005, given the year end 2004 estimate of 66.7%, we tabulated 10,000 randomly obtained year end 2005 loss ratio estimates. To determine each loss ratio estimate, we randomly selected from the log *t* distributions that represent the factors that contribute to the uncertainty in that estimate. For each trial we randomly selected one factor from the distribution of accident year 2004 development during 2005 and one factor from each of the age-to-age factor distributions that contribute to the revised tail factor. Then we multiplied all of these factors and the paid loss ratio as of year end 2004 to arrive at the ultimate loss ratio estimate for that trial.

This is illustrated in detail in Exhibit 11 for one trial, where the simulated actual accident year 2004 age 1 to 2 development factor is 1.727 (compared to an expected factor of 1.767) and the revised tail factor is 1.418 (compared to an expected 1.420). The product of the year end 2004 paid loss ratio and these two factors is the revised estimated ultimate loss ratio for accident year 2004 as of the end of 2005.

To arrive at approximate distributions of revised chain ladder ultimate loss ratio estimates for all of the accident years 1995 through 2004 as of the end of 2005, the process described in the preceding paragraph was repeated 10,000 times for each accident year. The results of this process are summarized in Exhibit 12, which, as the log *t* version of Exhibit 9, compares the 95% confidence intervals for the accident year 1995-2004 loss ratio estimates one year out with the confidence intervals for the estimates over the ultimate time horizon. The chain ladder estimates are summarized in the top half of the exhibit and the Bornhuetter-Ferguson estimates in the bottom half. As we observed in the lognormal case, much of the potential variation in the ultimate loss ratio estimates that is expected over the time horizon to ultimate is encompassed in the variation expected over a one-year time horizon. For example, the log *t* 95% confidence interval for the chain ladder estimate of the accident year 2004 loss ratio one year out of $66.7\% \pm 2.7\%$ is nearly as wide as the 95% confidence interval of $66.7\% \pm 2.9\%$ for the same loss ratio over the ultimate time horizon. Similarly, the accident year 2003 confidence interval for the chain ladder estimate of approximately $66.7\% \pm 0.9\%$ for a one year time horizon is also nearly as wide as that for the time horizon to ultimate of $67.8\% \pm 1.1\%$. For the older accident years, the proportion of the variation associated with the ultimate time horizon accounted for by the first year of future development is somewhat smaller, but the absolute size and significance of the confidence intervals for those years is much smaller.

Note that the log t confidence intervals are at least as wide in every case as the comparable lognormal confidence intervals shown in Exhibit 7.. In fact, in the case of the chain ladder estimates, for every accident year 1995-2004 the log t confidence intervals for the one-year time horizon are at least as wide as the lognormal confidence intervals for the ultimate time horizon!

5. CONCLUSIONS

There are a number of potential applications of the framework we have described for modeling future estimates of the ultimate loss ratio, ranging from loss reserving to pricing to analysis of risk-based capital. While a detailed discussion of these applications is beyond the scope of this paper, we will touch briefly on some examples.

5.1 Loss Reserving

The framework presented in this paper gives reserve actuaries a way to manage their clients' expectations. Reserve clients don't like surprises and often express frustration that loss ratio or reserve estimates change significantly from one period to the next. We have shown in this paper that a large proportion of the potential variation in ultimate estimates can be present in the first year of future development. As we saw in the Private Passenger Auto Liability example we presented, this phenomenon is particularly pronounced when the estimates are determined using the chain ladder method, but it can also be present if the estimates are derived from the Bornhuetter-Ferguson approach. It seems likely that most reserve clients do not understand this phenomenon. Actuaries have done a good job in getting clients to understand that ultimate loss estimates are subject to large potential variation, but many clients seem to expect that variation to emerge only in the distant future, if at all.

We suggest that the uncertainty in loss ratio and reserve estimates be framed in terms of how these estimates might change at the *next* valuation by presenting the ultimate estimates together with confidence intervals consistent with the valuation time horizon. For example, if the next valuation will be in one year, then the results would be presented with one-year time horizon confidence intervals. Then, because the potential variation has been explained to them in advance, clients might be better able to accept the revised estimates produced at the next valuation. This framework also naturally facilitates the explanation of the reasons for estimate revisions in terms of the sources of variation. For example, how much of the revision is due to actual accident year development and how much is due to a tail factor revision caused by loss emergence on the older accident years?

While we have focused much of our discussion on historical accident years and thus implicitly on reserving, we can easily extend this framework to encompass certain aspects of the pricing and underwriting, which can be used to assess risk load requirements, reinsurance risk transfer characteristics as well as to establish expectations for paid loss emergence during the first year after inception.

5.2 Risk-Based Capital

The framework described can also be applied to analysis of the issues outlined by Butsic [14] in his paper on solvency measurement in risk-based capital applications. He advocated the use of a common time horizon for measurement of all kinds of risks on both sides of the balance sheet. He showed how long term solvency protection could be achieved by periodic assessment and adjustment of risk-based capital using a short time horizon, e.g., one year. In particular, Butsic proposed that the risk-based capital charge at the beginning of each period be calibrated to a suitably small Expected Policyholder Deficit (EPD)[16] expressed as a ratio to expected unpaid losses. The capital charge would be reset at the beginning of each new period based on asset and/or liability changes during the period just ended. While he illustrated his approach with numerical examples, he did not describe a model for how claim liabilities change from one period to the next. The model presented in this paper, using parameters determined from Schedule P data, could be used together with Butsic's approach to test and refine the capital charges employed in the NAIC and rating agency risk-based capital models[17]. Moreover, to the extent that these risk-based capital charges imply the minimum amount of capital needed by an underwriter to assume risk, the model potentially has application to the problem of capital allocation for pricing applications as well.

5.3 Other Stochastic Loss Development Models

We have used Hayne's simple lognormal model to illustrate how to model the future behavior of loss ratio estimates. However, the same conceptual approach can be used with other stochastic models. If ultimate loss ratios are estimated using a different stochastic model, the path of future revisions to those ultimate loss ratio estimates can be determined using the ideas presented in this paper.

---

[16] The EPD is defined as the expectation of losses exceeding available assets. It can be viewed as the expected value of the proportion of policyholder claims that will be unrecoverable because of insurer insolvency.

[17] For stress testing these solvency models it may make sense to use the chain ladder model, which produces more variable loss ratio estimates, rather than the Bornhuetter-Ferguson model.

7. REFERENCES

[1] Hayne, Roger M., "An Estimate of Statistical Variation in Development Factor Methods", *PCAS LXXII*, **1985**, 25-43, http://www.casact.org/pubs/proceed/proceed85/85025.pdf
[2] Kelly, Mary V., "Practical Loss Reserving Method with Stochastic Development Factors", *Casualty Actuarial Society Discussion Paper Program*, May **1992** (Volume 1), 355-381, http://www.casact.org/pubs/dpp/dpp92/92dpp355.pdf
[3] Kreps, Rodney E., "Parameter Uncertainty in (Log)Normal Distributions", *PCAS* LXXXIV, **1997**, 553-580, http://www.casact.org/pubs/proceed/proceed97/97553.pdf
[4] Hodes, Douglas M., Sholom Feldblum and Gary Blumsohn, "Workers Compensation Reserve Uncertainty", *Casualty Actuarial Society Forum*, Volume: Summer, **1996**, 61-150, http://www.casact.org/pubs/forum/96sforum/96sf061.pdf
[5] Mack, Thomas, "Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates", *ASTIN Bulletin*, Volume: 23:2, **1993**, 213-225, http://www.casact.org/library/astin/vol23no2/213.pdf
[6] Mack, Thomas, "Measuring the Variability of Chain Ladder Reserve Estimates", *Casualty Actuarial Society Forum*, Volume: Spring (Volume 1), **1994**, 101-182, http://www.casact.org/pubs/forum/94spforum/94spf101.pdf
[7] Venter, Gary G., "Introduction to Selected Papers from the Variability in Reserves Prize Program", *Casualty Actuarial Society Forum*, Volume: Spring (Volume 1), **1994**, 91-100, http://www.casact.org/pubs/forum/94spforum/94spf091.pdf
[8] Venter, Gary G., "Testing the Assumptions of Age-to-Age Factors", *PCAS* LXXXV, **1998**, 807-847, http://www.casact.org/pubs/proceed/proceed98/980807.pdf
[9] Zehnwirth, Benjamin, "Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals and Risk Based Capital", *Casualty Actuarial Society Forum*, Volume: Spring (Volume 2), **1994**, 447-606, http://www.casact.org/pubs/forum/94spforum/94spf447.pdf
[10] Barnett, Glen and Benjamin Zehnwirth, "Best Estimates for Reserves", *PCAS* LXXXVII, **2000**, 245-321, http://www.casact.org/pubs/proceed/proceed00/00245.pdf
[11] Van Kampen, Charles E., "Estimating the Parameter Risk of a Loss Ratio Distribution", *Casualty Actuarial Society Forum*, Volume: Spring, **2003**, 177-213, http://www.casact.org/pubs/forum/03spforum/03spf177.pdf
[12] Wacek, Michael G., "Parameter Uncertainty in Loss Ratio Distributions and its Implications", *Casualty Actuarial Society Forum*, Volume: Fall, **2005**, 165-202, http://casact.org/pubs/forum/05fforum/05f165.pdf
[13] American Academy of Actuaries Property/Casualty Risk-Based Capital Task Force, "Report on Reserve and Underwriting Risk Factors", *Casualty Actuarial Society Forum*, Volume: Summer, **1993**, 105-171, http://casact.org/pubs/forum/93sforum/93sf105.pdf
[14] Butsic, Robert P., "Solvency Measurement for Property-Liability Risk-Based Capital Applications", *Casualty Actuarial Society Discussion Paper Program*, May **1992** (Volume 1), 311-354, http://www.casact.org/pubs/dpp/dpp92/92dpp311.pdf

Abbreviations and notations

μ, first parameter of lognormal, $E(\bar{y})=\mu$

σ, second parameter of lognormal, $E(s^2)=\sigma^2$

EPD, expected policyholder deficit

$f(x\,|\,\mu,\sigma)$, distribution of x, given known parameters μ,σ

$f(x)$, distribution of x (unknown parameters)

n, number of points in sample

N^{-1}(prob), standard normal inverse distribution function

P, actual paid loss ratio

R, random number from unit uniform distribution

s, standard deviation of log-transformed sample

T_{n-1}^{-1}(prob), standard normal inverse distribution function

$tail$, random variable for mean tail factor one year out

x_1, x_2, x_3,..., x_n, lognormal sample

x_{BF}, Bornhuetter-Ferguson estimate of ultimate loss ratio

x_{CL}, chain ladder estimate of ultimate loss ratio

x_p, cumulative paid loss ratio

y_1, y_2, y_3,..., y_n log-transformed sample

$\bar{y}$, mean of log-transformed sample

Biography of the Author

Michael Wacek is President of Odyssey America Reinsurance Corporation based in Stamford, CT. Over the course of more than 25 years in the industry, including nine years in the London Market, Mike has seen the business from the vantage point of a primary insurer, reinsurance broker and reinsurer. He has a BA from Macalester College (Math, Economics), is a Fellow of the Casualty Actuarial Society and a Member of the American Academy of Actuaries. He has authored a number of papers.

EXHIBIT 1

ANNUAL STATEMENT FOR THE YEAR 2004 OF THE * INDUSTRY AGGREGATE *
SCHEDULE P - PART 3B - PRIVATE PASSENGER AUTO LIABILITY/MEDICAL

CUMULATIVE PAID NET LOSSES AND DEFENSE AND COST CONTAINMENT EXPENSES REPORTED AT ANNUAL INTERVALS ($000,000 OMITTED)

AY	1	2	3	4	5	6	7	8	9	10
1995	17,674	32,062	38,619	42,035	43,829	44,723	45,162	45,375	45,483	45,540
1996	18,315	32,791	39,271	42,933	44,950	45,917	46,392	46,600	46,753	
1997	18,606	32,942	39,634	43,411	45,428	46,357	46,681	46,921		
1998	18,816	33,667	40,575	44,446	46,476	47,350	47,809			
1999	20,649	36,515	43,724	47,684	49,753	50,716				
2000	22,327	39,312	46,848	51,065	53,242					
2001	23,141	40,527	48,284	52,661						
2002	24,301	42,168	50,356							
2003	24,210	41,640								
2004	24,468									

PAID AGE-TO-AGE LOSS DEVELOPMENT FACTORS

AY	1 Loss Ratio	1-2 LDF	2-3 LDF	3-4 LDF	4-5 LDF	5-6 LDF	6-7 LDF	7-8 LDF	8-9 LDF	9-10 LDF
1995	28.0%	1.814	1.205	1.088	1.043	1.020	1.010	1.005	1.002	1.001
1996	27.7%	1.790	1.198	1.093	1.047	1.022	1.010	1.004	1.003	
1997	27.1%	1.771	1.203	1.095	1.046	1.020	1.007	1.005		
1998	27.1%	1.789	1.205	1.095	1.046	1.019	1.010			
1999	29.8%	1.768	1.197	1.091	1.043	1.019				
2000	32.2%	1.761	1.192	1.090	1.043					
2001	31.6%	1.751	1.191	1.091						
2002	30.4%	1.735	1.194							
2003	27.7%	1.720								
2004	26.6%									
Mean	28.8%	1.767	1.198	1.092	1.045	1.020	1.009	1.005	1.003	1.001
S.D.	2.0%	0.029	0.006	0.003	0.002	0.001	0.002	0.000	0.001	0.000
C.V.	7.0%	0.016	0.005	0.002	0.002	0.001	0.002	0.000	0.001	0.000
Cum Mean		2.508	1.420	1.185	1.085	1.039	1.018	1.009	1.004	1.001

Source: Highline Data LLC as reported in the statutory filings (OneSource)

EXHIBIT 2

ANNUAL STATEMENT FOR THE YEAR 2004 OF THE * INDUSTRY AGGREGATE *
SCHEDULE P - PART 3B - PRIVATE PASSENGER AUTO LIABILITY/MEDICAL

NATURAL LOGARITHMS OF PAID AGE-TO-AGE LOSS DEVELOPMENT FACTORS IN EXHIBIT 1

AY	1 Loss Ratio	1-2 LDF	2-3 LDF	3-4 LDF	4-5 LDF	5-6 LDF	6-7 LDF	7-8 LDF	8-9 LDF	9-10 LDF
1995	-1.274	0.596	0.186	0.085	0.042	0.020	0.010	0.005	0.002	0.001
1996	-1.282	0.582	0.180	0.089	0.046	0.021	0.010	0.004	0.003	
1997	-1.307	0.571	0.185	0.091	0.045	0.020	0.007	0.005		
1998	-1.304	0.582	0.187	0.091	0.045	0.019	0.010			
1999	-1.210	0.570	0.180	0.087	0.042	0.019				
2000	-1.135	0.566	0.175	0.086	0.042					
2001	-1.151	0.560	0.175	0.087						
2002	-1.191	0.551	0.177							
2003	-1.282	0.542								
2004	-1.325									
Mean	-1.246	0.569	0.181	0.088	0.044	0.020	0.009	0.005	0.003	0.001
S.D.	0.069	0.016	0.005	0.002	0.002	0.001	0.002	0.000	0.001	0.001
LN Fit LDFs	28.8%	1.767	1.198	1.092	1.045	1.020	1.009	1.005	1.003	1.001
Cum Mean	-0.327	0.919	0.350	0.170	0.082	0.038	0.018	0.009	0.004	0.001
Cum S.D.	0.071	0.018	0.006	0.004	0.003	0.002	0.002	0.001	0.001	0.001
LN Fit LDFs	72.3%	2.508	1.420	1.185	1.085	1.039	1.018	1.009	1.004	1.001

Source: Highline Data LLC as reported in the statutory filings (OneSource)

EXHIBIT 3

SUMMARY OF ESTIMATED ULTIMATE LOSS RATIOS
FROM PAID LOSS DEVELOPMENT ANALYSIS

PRIVATE PASSENGER AUTO LIABILITY

INDUSTRY AGGREGATE EXPERIENCE

Accident Year	Net Earned Premiums	Net Paid Losses	Net Paid Loss Ratio	Age-to-Ult Factor	Estimated Ultimate Loss Ratio
1995	63,183	45,540	72.1%	1.000	72.1%
1996	66,006	46,753	70.8%	1.001	70.9%
1997	68,764	46,921	68.2%	1.004	68.5%
1998	69,343	47,809	68.9%	1.009	69.6%
1999	69,231	50,716	73.3%	1.018	74.6%
2000	69,444	53,242	76.7%	1.039	79.6%
2001	73,143	52,661	72.0%	1.085	78.1%
2002	79,922	50,356	63.0%	1.185	74.6%
2003	87,242	41,640	47.7%	1.420	67.8%
2004	92,064	24,468	26.6%	2.508	66.7%

EXHIBIT 4

SUMMARY OF PAID LOSS DEVELOPMENT FACTORS
WITH ASSOCIATED LOGNORMAL 95% CONFIDENCE INTERVALS

PRIVATE PASSENGER AUTO LIABILITY

BASED ON INDUSTRY AGGREGATE EXPERIENCE

AGE-TO-AGE FACTORS

Paid Loss Development Period	Est μ for LDF	Est σ for LDF	Mean LDF	Lognormal 95% Confidence LDF Lower Bound	Lognormal 95% Confidence LDF Upper Bound
9-Ult*	0.001	0.001	1.001	1.000	1.002
8-9	0.003	0.001	1.003	1.002	1.004
7-8	0.005	0.000	1.005	1.004	1.005
6-7	0.009	0.002	1.009	1.006	1.012
5-6	0.020	0.001	1.020	1.018	1.022
4-5	0.044	0.002	1.045	1.041	1.048
3-4	0.088	0.002	1.092	1.087	1.097
2-3	0.181	0.005	1.198	1.187	1.209
1-2	0.569	0.016	1.767	1.710	1.824

AGE-TO-ULTIMATE FACTORS

Paid Loss Development Period	Est μ for LDF	Est σ for LDF	Mean LDF	Lognormal 95% Confidence LDF Lower Bound	Lognormal 95% Confidence LDF Upper Bound
9 - Ult*	0.001	0.001	1.001	1.000	1.002
8 - Ult	0.004	0.001	1.004	1.002	1.006
7 - Ult	0.009	0.001	1.009	1.007	1.011
6 - Ult	0.018	0.002	1.018	1.015	1.022
5 - Ult	0.038	0.002	1.039	1.034	1.043
4 - Ult	0.082	0.003	1.085	1.079	1.091
3 - Ult	0.170	0.004	1.185	1.176	1.193
2 - Ult	0.350	0.006	1.420	1.403	1.436
1 - Ult	0.919	0.018	2.508	2.423	2.595

* Age 10 deemed to be ultimate

EXHIBIT 5

SUMMARY OF REVISED MEAN PAID LOSS DEVELOPMENT FACTORS ONE YEAR OUT

PRIVATE PASSENGER AUTO LIABILITY

BASED ON INDUSTRY AGGREGATE EXPERIENCE

MEAN AGE-TO-AGE FACTORS ONE YEAR OUT

						Lognormal 95% Confidence	
Paid Loss Development Period	Est σ for Actual LDF	Actual LDF Weight	Est σ for Revised Mean LDF	Est μ for Revised Mean LDF	Revised Mean LDF	Revised Mean LDF (Lower Bound)	Revised Mean LDF (Upper Bound)
9-Ult*	0.001	1/2	0.000	0.001	1.001	1.001	1.002
8-9	0.001	1/3	0.000	0.003	1.003	1.002	1.003
7-8	0.000	1/4	0.000	0.005	1.005	1.005	1.005
6-7	0.002	1/5	0.000	0.009	1.009	1.009	1.010
5-6	0.001	1/6	0.000	0.020	1.020	1.020	1.020
4-5	0.002	1/7	0.000	0.044	1.045	1.044	1.045
3-4	0.002	1/8	0.000	0.088	1.092	1.091	1.093
2-3	0.005	1/9	0.001	0.181	1.198	1.197	1.199
1-2	0.016	1/10	0.002	0.569	1.767	1.761	1.772

MEAN AGE-TO-ULTIMATE FACTORS ONE YEAR OUT

				Lognormal 95% Confidence	
Paid Loss Development Period	Est σ for Revised Mean LDF	Est μ for Revised Mean LDF	Revised Mean LDF	Revised Mean LDF (Lower Bound)	Revised Mean LDF (Upper Bound)
9 - Ult*	0.000	0.001	1.001	1.001	1.002
8 - Ult	0.000	0.004	1.004	1.003	1.005
7 - Ult	0.000	0.009	1.009	1.008	1.010
6 - Ult	0.000	0.018	1.018	1.017	1.019
5 - Ult	0.001	0.038	1.039	1.038	1.040
4 - Ult	0.001	0.082	1.085	1.084	1.086
3 - Ult	0.001	0.170	1.185	1.183	1.186
2 - Ult	0.001	0.350	1.420	1.417	1.422
1 - Ult	0.002	0.919	2.508	2.499	2.517

* Age 10 deemed to be ultimate

EXHIBIT 6

ANALYSIS OF ESTIMATED ULTIMATE LOSS RATIOS (CHAIN LADDER) ONE YEAR OUT

PRIVATE PASSENGER AUTO LIABILITY

BASED ON INDUSTRY AGGREGATE EXPERIENCE

			First Effect Accident Year Devt		Second Effect Tail Factor Revision					Lognormal 95% Confidence	
Accident Year	Devt Age	Net Paid Loss Ratio	Est μ for Actual 1-Yr LDF	Est σ for Actual 1-Yr LDF	Est μ for Revised Mean LDF	Est σ for Revised Mean LDF	Est μ for Est Ult L/R 1 Yr Out	Est σ for Est Ult L/R 1 Yr Out	Estimated Ultimate Loss Ratio 1 Yr Out	Est L/R 1 Yr Out (Lower Bound)	Est L/R 1 Yr Out (Upper Bound)
1995	10	72.1%	0.000	0.000	0.000	0.000	-0.327	0.000	72.1%	72.1%	72.1%
1996	9	70.8%	0.001	0.001	0.000	0.000	-0.344	0.001	70.9%	70.8%	71.0%
1997	8	68.2%	0.003	0.001	0.001	0.000	-0.378	0.001	68.5%	68.4%	68.6%
1998	7	68.9%	0.005	0.000	0.004	0.000	-0.363	0.001	69.6%	69.5%	69.6%
1999	6	73.3%	0.009	0.002	0.009	0.000	-0.293	0.002	74.6%	74.4%	74.8%
2000	5	76.7%	0.020	0.001	0.018	0.000	-0.228	0.001	79.6%	79.5%	79.8%
2001	4	72.0%	0.044	0.002	0.038	0.001	-0.247	0.002	78.1%	77.8%	78.4%
2002	3	63.0%	0.088	0.002	0.082	0.001	-0.292	0.003	74.6%	74.3%	75.0%
2003	2	47.7%	0.181	0.005	0.170	0.001	-0.389	0.005	67.8%	67.1%	68.4%
2004	1	26.6%	0.569	0.016	0.350	0.001	-0.406	0.017	66.7%	64.5%	68.8%
2005	0	0.0%	-1.246	0.069	0.919	0.002	-0.327	0.069	72.3%	63.0%	82.6%

EXHIBIT 7

ANALYSIS OF ESTIMATED ULTIMATE LOSS RATIOS (BORNHUETTER-FERGUSON)
ONE YEAR OUT

PRIVATE PASSENGER AUTO LIABILITY

BASED ON INDUSTRY AGGREGATE EXPERIENCE

			First Effect Accident Year Devt			Second Effect Tail Factor Revision				Lognormal 95% Confidence *	
Accident Year	Devt Age	Net Paid Loss Ratio	Est μ for Actual 1-Yr LDF	Est σ for Actual 1-Yr LDF	Expected Paid L/R 1 Yr Out	Est μ for Revised Mean LDF	Est σ for Revised Mean LDF	Expected Mean Tail 1 Yr Out	Estimated Ultimate Loss Ratio 1 Yr Out	Est L/R 1 Yr Out (Lower Bound)	Est L/R 1 Yr Out (Upper Bound)
1995	10	72.1%	0.000	0.000	72.1%	0.000	0.000	1.000	72.1%	72.1%	72.1%
1996	9	70.8%	0.001	0.001	70.9%	0.000	0.000	1.000	70.9%	70.8%	71.0%
1997	8	68.2%	0.003	0.001	68.4%	0.001	0.000	1.001	68.5%	68.4%	68.6%
1998	7	68.9%	0.005	0.000	69.3%	0.004	0.000	1.004	69.6%	69.5%	69.6%
1999	6	73.3%	0.009	0.002	73.9%	0.009	0.000	1.009	74.6%	74.4%	74.8%
2000	5	76.7%	0.020	0.001	78.2%	0.018	0.000	1.018	79.6%	79.5%	79.8%
2001	4	72.0%	0.044	0.002	75.2%	0.038	0.001	1.039	78.1%	77.8%	78.4%
2002	3	63.0%	0.088	0.002	68.8%	0.082	0.001	1.085	74.6%	74.3%	75.0%
2003	2	47.7%	0.181	0.005	57.2%	0.170	0.001	1.185	67.8%	67.2%	68.3%
2004	1	26.6%	0.569	0.016	47.0%	0.350	0.001	1.420	66.7%	65.1%	68.2%
2005	0	0.0%	-1.246	0.069	28.8%	0.919	0.002	2.508	72.3%	68.6%	76.3%

* Based on Monte Carlo simulation of $x_{BF} = x_P - E(x_P) + E(x_P) \cdot tail$

EXHIBIT 8

One Path of the Accident Year 2004 Ultimate Loss Ratio Estimate

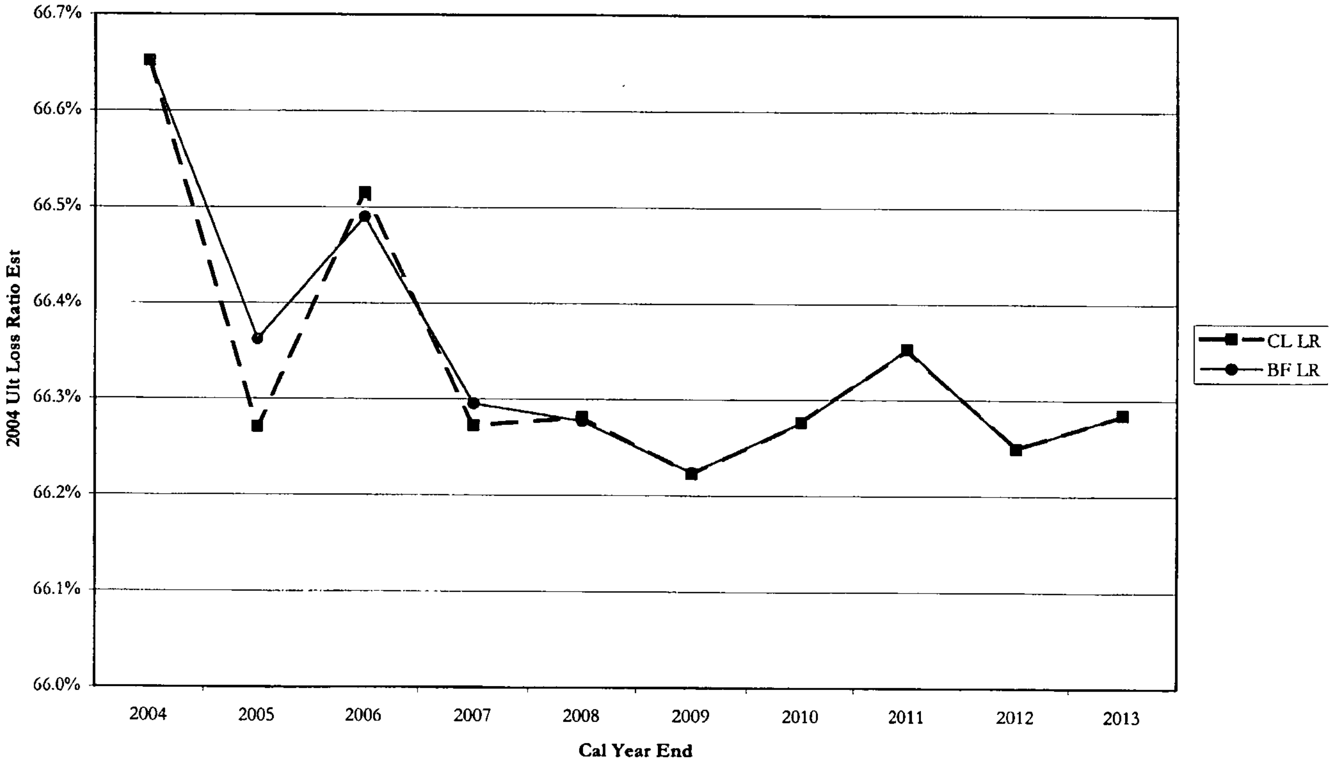


EXHIBIT 9

LOGNORMAL CONFIDENCE INTERVALS - ULTIMATE LOSS RATIOS
ONE YEAR VS. ULTIMATE TIME HORIZONS

PRIVATE PASSENGER AUTO LIABILITY

BASED ON INDUSTRY AGGREGATE EXPERIENCE

		95% Confidence Intervals - Chain Ladder Estimates			
Accident Year	Dec 2004	One Year Horizon		Ultimate Horizon	
1995	72.1%	72.1%	72.1%	72.1%	72.1%
1996	70.9%	70.8%	71.0%	70.8%	71.0%
1997	68.5%	68.4%	68.6%	68.4%	68.6%
1998	69.6%	69.5%	69.6%	69.4%	69.7%
1999	74.6%	74.4%	74.8%	74.3%	74.8%
2000	79.6%	79.5%	79.8%	79.3%	80.0%
2001	78.1%	77.8%	78.4%	77.7%	78.5%
2002	74.6%	74.3%	75.0%	74.1%	75.2%
2003	67.8%	67.1%	68.4%	67.0%	68.6%
2004	66.7%	64.5%	68.8%	64.4%	69.0%

		95% Confidence Intervals - B-F Estimates			
Accident Year	Dec 2004	One Year Horizon		Ultimate Horizon	
1995	72.1%	72.1%	72.1%	72.1%	72.1%
1996	70.9%	70.8%	71.0%	70.8%	71.0%
1997	68.5%	68.4%	68.6%	68.4%	68.6%
1998	69.6%	69.5%	69.6%	69.4%	69.7%
1999	74.6%	74.4%	74.8%	74.3%	74.8%
2000	79.6%	79.5%	79.8%	79.3%	80.0%
2001	78.1%	77.8%	78.4%	77.7%	78.5%
2002	74.6%	74.3%	75.0%	74.1%	75.2%
2003	67.8%	67.2%	68.3%	67.0%	68.6%
2004	66.7%	65.1%	68.2%	64.4%	69.0%

EXHIBIT 10

LOG *t* CONFIDENCE INTERVALS FOR PAID LOSS DEVELOPMENT FACTORS REFLECTING PARAMETER UNCERTAINTY

PRIVATE PASSENGER AUTO LIABILITY

BASED ON INDUSTRY AGGREGATE EXPERIENCE

AGE-TO-AGE FACTORS

	Supporting Information for Confidence Interval Determination			Log *t* 95% Confidence			Log *t* Ratio to Lognormal	
Paid Loss Development Period	Sample Size n	Degrees of Freedom n-1 **	$T_{n-1}^{-1}(97.5\%)$	LDF Lower Bound	LDF Mean	LDF Upper Bound	At Lower Bound	At Upper Bound
9-Ult*	1	3	3.182	0.998	1.001	1.004	0.998	1.002
8-9	2	3	3.182	1.000	1.003	1.005	0.999	1.001
7-8	3	3	3.182	1.004	1.005	1.006	0.999	1.001
6-7	4	3	3.182	1.004	1.009	1.015	0.998	1.002
5-6	5	4	2.776	1.017	1.020	1.023	0.999	1.001
4-5	6	5	2.571	1.039	1.045	1.050	0.998	1.002
3-4	7	6	2.447	1.085	1.092	1.099	0.998	1.002
2-3	8	7	2.365	1.184	1.198	1.212	0.997	1.003
1-2	9	8	2.306	1.697	1.767	1.839	0.992	1.008

AGE-TO-ULTIMATE FACTORS***

	Log *t* 95% Confidence			Log *t* Ratio to Lognormal	
Paid Loss Development Period	LDF Lower Bound	LDF Mean	LDF Upper Bound	At Lower Bound	At Upper Bound
9 - Ult*	0.998	1.001	1.004	0.998	1.002
8 - Ult	1.000	1.004	1.008	0.998	1.002
7 - Ult	1.005	1.009	1.013	0.998	1.002
6 - Ult	1.011	1.018	1.025	0.997	1.003
5 - Ult	1.031	1.039	1.047	0.997	1.004
4 - Ult	1.075	1.085	1.095	0.996	1.004
3 - Ult	1.171	1.185	1.198	0.996	1.004
2 - Ult	1.397	1.420	1.443	0.996	1.005
1 - Ult	2.401	2.508	2.619	0.991	1.009

* Age 10 deemed to be ultimate

** Judgmentally limited to a minimum of 3. (Variance not defined, if d.f. < 3.)

*** From Monte Carlo simulation (10,000 trials)

EXHIBIT 11

MONTE CARLO SIMULATION OF ESTIMATED ULTIMATE LOSS RATIO FOR ACCIDENT YEAR 2004
ONE YEAR OUT

ILLUSTRATION OF ONE RANDOM TRIAL - REFLECTING PARAMETER UNCERTAINTY

PRIVATE PASSENGER AUTO LIABILITY

BASED ON INDUSTRY AGGREGATE EXPERIENCE

Devt Period	Expected LDF	Sample Size n	Degrees of Freedom n-1 **	Uniform Random Number R	$T_{n-1}^{-1}(R)$	$\bar{y}$	s	$\sqrt{\frac{n+1}{n}}$	Random LDF: Accident Yr Devt *	Random LDF: Revised Tail *
9-Ult***	1.001	2	3	0.561	0.167	0.001	0.000	1.225		1.001
8-9	1.003	3	3	0.074	-1.938	0.003	0.000	1.155		1.002
7-8	1.005	4	3	0.084	-1.810	0.005	0.000	1.118		1.005
6-7	1.009	5	4	0.484	-0.043	0.009	0.000	1.095		1.009
5-6	1.020	6	5	0.899	1.468	0.020	0.000	1.080		1.020
4-5	1.045	7	6	0.128	-1.255	0.044	0.000	1.069		1.044
3-4	1.092	8	7	0.131	-1.220	0.088	0.000	1.061		1.092
2-3	1.198	9	8	0.396	-0.273	0.181	0.001	1.054		1.198
1-2	1.767	9	8	0.116	-1.293	0.569	0.016	1.054	1.727	
									1.727	1.418

Revised **Chain Ladder** Loss Ratio Estimate One Year Out = Paid Loss Ratio x Actual Acc Year Devt x Revised Tail Factor

= 26.6% x 1.727 x 1.418 = 65.1%

Revised **B - F** L/R Estimate One Year Out = Actual Paid L/R - Expected Paid L/R + Expected Paid L/R x Revised Tail Factor

= (26.6% x 1.727) - (26.6% x 1.767) + (26.6% x 1.767) x 1.418 = 65.6%

* = $\exp(\bar{y} + T_{n-1}^{-1}(R) \cdot s\sqrt{(n+1)/n})$

** Judgmentally limited to a minimum of 3. (Variance not defined, if d.f. < 3.)

*** Age 10 deemed to be ultimate

EXHIBIT 12

LOG t CONFIDENCE INTERVALS - ULTIMATE LOSS RATIOS
ONE YEAR VS. ULTIMATE TIME HORIZONS

PRIVATE PASSENGER AUTO LIABILITY

BASED ON INDUSTRY AGGREGATE EXPERIENCE

95% Confidence Intervals - Chain Ladder Estimates

Accident Year	Dec 2004	One Year Horizon		Ultimate Horizon	
1995	72.1%	72.1%	72.1%	72.1%	72.1%
1996	70.9%	70.7%	71.1%	70.7%	71.1%
1997	68.5%	68.3%	68.7%	68.3%	68.8%
1998	69.6%	69.4%	69.7%	69.3%	69.8%
1999	74.6%	74.2%	75.0%	74.1%	75.1%
2000	79.6%	79.3%	79.9%	79.0%	80.3%
2001	78.1%	77.7%	78.5%	77.4%	78.9%
2002	74.6%	74.1%	75.1%	73.8%	75.5%
2003	67.8%	66.9%	68.6%	66.7%	68.9%
2004	66.7%	64.0%	69.4%	63.8%	69.6%

95% Confidence Intervals - B-F Estimates

Accident Year	Dec 2004	One Year Horizon		Ultimate Horizon	
1995	72.1%	72.1%	72.1%	72.1%	72.1%
1996	70.9%	70.7%	71.1%	70.7%	71.1%
1997	68.5%	68.3%	68.7%	68.3%	68.8%
1998	69.6%	69.4%	69.7%	69.3%	69.8%
1999	74.6%	74.2%	75.0%	74.1%	75.1%
2000	79.6%	79.3%	79.9%	79.0%	80.3%
2001	78.1%	77.7%	78.5%	77.4%	78.9%
2002	74.6%	74.2%	75.1%	73.8%	75.5%
2003	67.8%	67.1%	68.4%	66.7%	68.9%
2004	66.7%	64.8%	68.5%	63.8%	69.6%

A Test of Clinical Judgment vs. Statistical Prediction in Loss Reserving for Commercial Auto Liability

Michael G. Wacek, FCAS, MAAA

Abstract

This paper is a case study of the quality of clinical judgment in loss reserving for Commercial Auto Liability in the U.S. for accident years 1995 through 2001. Research on clinical vs. statistical prediction in non-insurance fields indicates that relatively simple models frequently produce better results than human experts with access to the same information. To test the quality of clinical judgment vs. statistical prediction in the Commercial Auto Liability loss reserving process, we compared the ultimate loss ratios actually booked by the U.S. insurance industry for these accident years at twelve, twenty-four and thirty-six months of development to comparable loss ratio estimates generated by mechanical application of several basic loss development methods. The booked ultimate loss ratios differed significantly from those indicated by the mechanical application of chain ladder and Bornhuetter-Ferguson methods, implying that the booked ultimate loss ratios were not determined using those methods, at least not without significant adjustment. We then compared all of these booked and estimated loss ratios to the ultimate loss ratios booked as of the end of 2004, which we treated as proxies for the true ultimate loss ratios. In most cases, the mechanically generated ultimate loss ratio estimates were closer to the booked estimates as of the end of 2004 than were the earlier booked loss ratios. The conclusion must be that, either the booked ultimate loss ratios were based on other methods that are inferior to the chain ladder and Bornhuetter-Ferguson or judgmental adjustments were made to the indicated ultimate loss ratios that reduced the quality of the final selections. Further research would be required to determine whether this is a general loss reserving phenomenon or one confined to Commercial Auto Liability during the time period studied.

Keywords: loss reserving, commercial auto liability, chain ladder, Bornhuetter-Ferguson

1. INTRODUCTION

Research on clinical vs. statistical prediction in non-insurance fields indicates that relatively simple quantitative models often produce better results than human experts with access to the same information. "Clinical prediction" refers to the conclusion reached by an expert when presented with a set of facts about a problem of a type with which he or she has experience. "Statistical prediction" refers to the conclusion indicated by a quantitative or statistical formula or model using a set of quantifiable facts about a problem. Clinical prediction does not preclude the use of statistical methods, but where they are employed they are augmented by consideration of other information and the judgment of the expert. For further background on this research see Snijders, Tazelaar and Batenburg [1].

The process of establishing the loss reserve liability to be carried on an insurer's balance sheet generally meets the definition of clinical rather than statistical prediction. Quantitative methods are used to make estimates of ultimate losses, but the estimate of the required loss reserve that is selected for booking on the balance sheet is almost never the unadjusted output of a formula. Typically, the loss reserve actuary makes adjustments to formula output before making recommendations to executive management. Those recommendations frequently take the form of a range of reasonable estimates. Ultimately, the loss reserve liability selected to be booked on the balance sheet reflects both statistical information and the judgment of the actuary and management.

In this paper we describe the results of a test of the quality of clinical vs. statistical prediction with respect to Commercial Auto Liability ultimate loss ratio estimates for accident years 1995 through 2001 for the U.S. industry in total using Schedule P data reported in Best's Aggregates & Averages. We expected to find insignificant differences in booked ultimate loss ratios from those indicated by the chain ladder and Bornhuetter-Ferguson methods, which we classify as statistical prediction methods[1]. To the extent that there were differences, we expected the judgmentally selected loss ratios to be superior. The chain ladder and Bornhuetter-Ferguson methods are relatively crude approaches that do not and cannot incorporate all of the quantitative and qualitative information available about emerging claims. It should be possible to improve upon the estimates that emerge from these methods. Indeed, much of the recent actuarial literature on loss reserving has focused on methods that are statistically, if not qualitatively, superior to the chain ladder and Bornhuetter-Ferguson.

2. SUMMARY OF FINDINGS

In this section we describe the results of our comparison of the industry's booked ultimate loss ratio estimates with statistically predicted ultimate loss ratios. Our purpose was first to determine whether the booked results appear to be based on any of the statistical methods and then to determine whether the booked loss ratios, which were based at least to some extent on clinical judgment, were better or worse than statistically predicted ones.

2.1 Comparison of Clinically and Statistically Predicted Loss Ratios

To test the proposition that the booked ultimate loss ratios for accident years 1995 through 2001 were consistent with estimates indicated by statistical loss development analysis methods, we compared the ultimate loss ratios actually booked by the U.S. insurance industry for these accident years at twelve, twenty-four and thirty-six months of

[1] In fact, our initial purpose in studying Commercial Auto Liability ultimate loss ratios from this period was to determine whether their behavior over time conformed to the model described by Wacek [2], which assumes that selected ultimate loss ratios are largely derived from the loss development models with relatively little injection of judgment.

development to comparable loss ratio estimates generated by mechanical application of the chain ladder and Bornhuetter-Ferguson methods using both paid and case incurred loss data as well as the average of all four of these methods[2].

Figures A, B and C show comparisons of the clinical and statistical predictions for loss ratio valuations as of twelve, twenty-four and thirty-six months, respectively, after inception of the accident year. In Figure A, which shows the ultimate loss ratio estimates as of twelve months, we see that the booked loss ratio estimates (represented by the dashed line) were almost always the lowest of all of the methods[3]. If the booked estimates were based on one or more of the statistical methods, we would expect to see the booked loss ratio estimates within the cluster of statistical estimates and not at the edge or outside of it, as they are here. Between 1995 and 1997, the booked estimates seem to track the Bornhuetter-Ferguson case incurred indications, but after that they diverge sharply downward. In Figure B, which compares the ultimate loss ratio estimates as of twenty-four months, we see the same pattern as at twelve months, but it is even clearer. The statistical method estimates were clustered more closely together than at twelve months and this tighter clustering accentuates the divergence of the booked and statistical estimates. For each year from 1999 through 2001 the distance of the booked estimate from the closest statistical estimate was greater than the *range* of the five statistical predictions! We see the same pattern again in the thirty-six-months comparison shown in Figure C, which further reinforces the conclusion that the booked ultimate loss ratios must have been determined by a different process.

Exhibits 1, 2 and 3 make the same comparisons as Figures A, B and C in tabular form. For example, referring to Exhibit 2, we see that the range of statistical ultimate loss ratio estimates for accident year 1999 at twenty-four months was 88.4% to 91.7%, a range of 3.3 loss ratio points. The clinical prediction, represented by the booked loss ratio, was 83.6%, which is nearly five points below the lowest of the statistical estimates. The divergence is even more striking at thirty-six months, where the statistical estimates range from 91.9% to 92.7%, a range of 0.8% points. The booked ultimate loss ratio was 87.7%, again five points below the lowest statistical estimate and more than six times the size of the range of the statistical estimates! The pattern is similar for accident years 2000 and 2001.

The booked loss ratio estimates were so different from those produced by the chain ladder and Bornhuetter-Ferguson methods and their average that we concluded that the

---

[2] For a detailed explanation of the methods and data used to determine these estimates, see Appendix A.

[3] The estimate for accident year 1995 is the notable exception. The 1996 and 1997 booked estimates are the lowest (but essentially tied with the B-F case incurred estimates). Each of the 1999-2001 booked estimates is the lowest by a significant amount.

booked loss ratios could not have arisen directly from any of those methods, especially after 1997. To the extent those methods were used, the statistical indications were so heavily adjusted that the final loss ratio estimates selected for booking were effectively independent of those methods.

2.2 Accuracy of Clinically vs. Statistically Predicted Loss Ratios

To test the proposition that the clinically predicted booked ultimate loss ratios were better estimates than the statistical predictions, we compared the clinical and statistical predictions to the ultimate loss ratio estimates booked as of December 2004, which we treated as reasonable proxies for the true ultimate loss ratios[4].

The clinically predicted loss ratios were *not* better estimates than the purely statistical predictions. In fact, in most cases the booked ultimate loss ratios were far inferior to the mechanically generated ones in predicting the "true" ultimate loss ratios. Figures D, E and F are graphical comparisons of the prediction errors of the various ultimate loss ratio estimation methods for estimates made as of twelve months, twenty-four months and thirty-six months, respectively. A positive error implies a loss ratio projection that is higher than the "true" ultimate loss ratio. A negative error implies a loss ratio projection that is lower than the "true" ultimate loss ratio. A visual inspection of Figures D, E and F makes clear that the clinically predicted loss ratios showed prediction errors of a larger magnitude than the statistical indications for most accident years and all three valuations. Several of the methods showed a negative bias, i.e., a tendency to underestimate the "true" ultimate loss ratio, especially between 1997 and 2000, but the clinically determined booked estimates showed the most pronounced negative bias[5]. That negative bias in the booked estimates was not confined to the 1997 through 2000 period and instead was fairly persistent across accident years and at all three valuations.

For a more detailed look, see the tabular summary of the prediction errors provided in Exhibits 4, 5 and 6, which compare the clinical and statistical prediction errors at twelve

[4] Based on historical development patterns, by December 2004 the expected paid and case incurred losses for the oldest year in our accident year sample, 1995, were both more than 99% of ultimate losses. Even the youngest year, 2001, was substantially developed, with expected paid losses at more than 80% and expected case incurred losses at more than 95% of ultimate losses as of December 2004, leaving little likelihood of development surprises that would materially affect the ultimate loss ratio estimate beyond that date.

[5] At the twelve months valuation, the mean error of the statistical estimates was –2.2% in 1997; -4.9% in 1998, -6.8% in 1999 and –2.5% in 2000, an average error of –4.1% over the period. Clearly, the statistical methods did not perform well in this time period. However, the errors in the booked estimates at twelve months were much larger: -5.9% in 1997, -9.2% in 1998, -13.8% in 1999 and –10.4 in 2000, an average of –9.8% for the period. At twenty-four and thirty-six months, respectively, the mean errors of the statistical indications for 1997-2000 were –1.6% and -0.6% compared to –7.1% and –4.1% for the booked estimates.

months, twenty-four months and thirty-six months valuations, respectively. The clinically predicted booked ultimate loss ratio was the most accurate of the estimates in 1995 at all three valuations. However, for *all* other accident years at all three valuations, the clinical prediction proved to be either the least or second least accurate of the six predictive methods. It was the *least* accurate of the six methods in four of the seven accident years as of twelve months, and five of the seven years as of the twenty-four months and thirty-six months valuations. That means that two-thirds of the time *any* of the statistical methods would have been better than the clinical approach that was actually used! The clinical estimates also had by far the highest sum of squared errors of all the methods at all three valuations. Finally, the clinical estimates showed the largest bias (and that bias was negative) at all three valuations.

Statistical prediction outperformed clinical prediction for Commercial Auto Liability ultimate loss ratio estimation by a wide margin in this time period!

3. CONCLUSIONS

We do not know whether the superiority of statistical loss reserving methods that we saw here with respect to Commercial Auto Liability is confined to the circumstances of that line of business during the time period studied or whether it is a more general phenomenon. That would be an interesting question for further research. All we can say is that the industry would have set more accurate Commercial Auto Liability loss reserves for accident years 1995 through 2001, if it had simply booked the indications of any one of the five statistical methods we tested (the best of which was the simple average of the two chain ladder and two Bornhuetter-Ferguson estimates).

It is beyond the scope of this paper to explain the poor performance of clinical prediction of Commercial Auto Liability ultimate loss ratios between 1995 and 2001, but let's consider a few possibilities that may also warrant further study.

One possibility is that the negative bias we observed had a purely technical basis arising from the skewness of aggregate loss distributions. In his 1985 paper Stanard [3] made the following observation about chain ladder loss ratio indications in the small samples he studied: "...[T]he median prediction error...was usually negative...but a few large cases of over-prediction made the mean prediction error (the bias) positive." If the industry's Commercial Auto Liability experience comprised individual portfolios that displayed enough skewness to result in the effect that Stanard observed, then perhaps the negative bias we saw resulted merely from chain ladder or other over-projections being judgmentally capped. In that case the sum of the individual portfolio estimates would be biased low. We don't know whether this effect could be large enough to fully explain the phenomenon we observed.

A second possibility, one suggested by research in other fields, is that the expert judgment exercised by actuaries and management is not always so expert. Perhaps qualitative and even quantitative judgments based on "experience" are risky and even biased. Perhaps what we observed is that even highly trained and experienced insurance professionals can be fooled by "anomalies" in the data that actually are part of the fundamental statistical pattern, the "correction" of which can degrade rather than improve the result.

It is interesting that irrespective of the limitations of the chain ladder and Bornhuetter-Ferguson methods from a theoretical standpoint, they performed better than the method actually used to reserve Commercial Auto Liability from 1995 through 2001. It is a reminder that theoretical advances in loss reserving methodology will have no effect on the accuracy of booked estimates if the indications are ignored or overridden by judgment! We saw that while the chain ladder and Bornhuetter-Ferguson methods underestimated the ultimate loss ratios during the period 1997 through 2000, the addition of clinical judgment *more than doubled* that underestimation. While we must be careful not to over-generalize from this limited study, at very least it suggests that actuaries must be mindful that the exercise of judgment in loss reserving has the potential to compound rather than reduce reserving errors. That is not to say that judgment should never be exercised, but it must be exercised with great care.

FIGURE A

Commercial Auto Liability - U.S. Industry
Accident Year 1995-2001 Ultimate Loss Ratio Estimates
As of Twelve Months

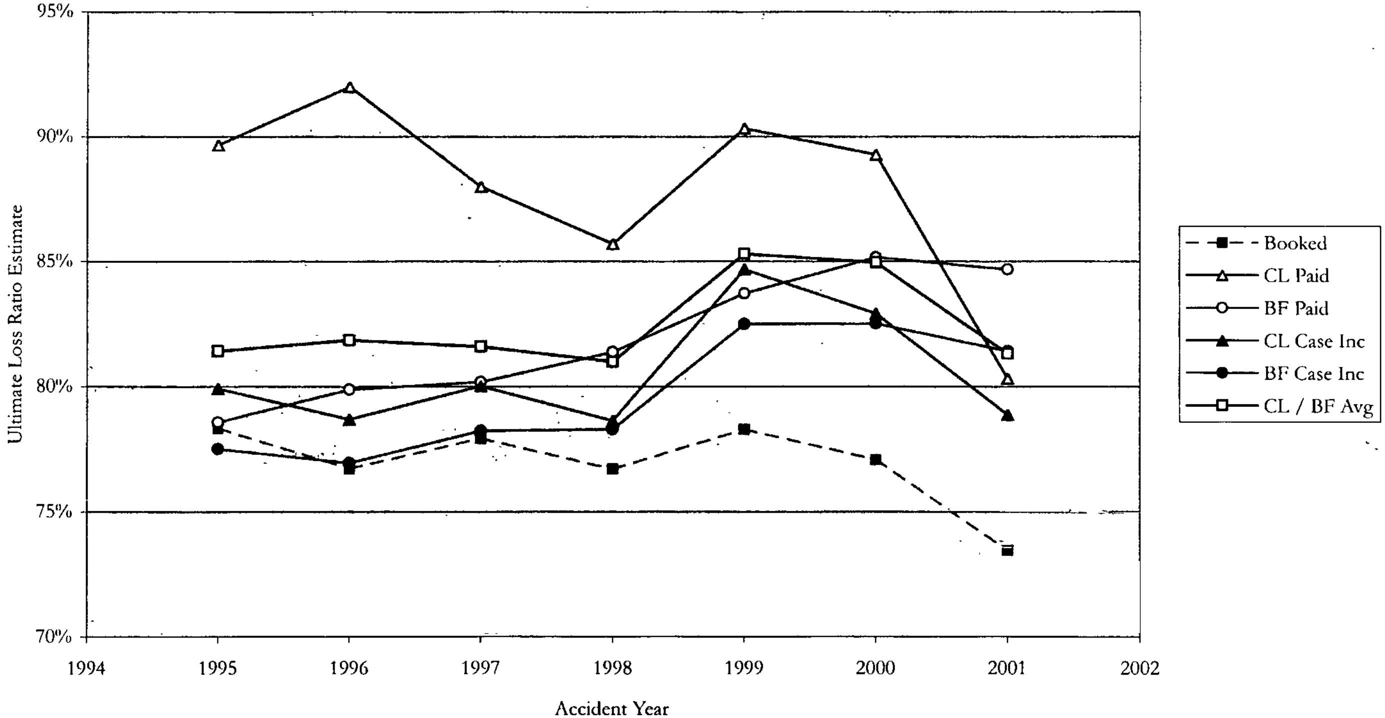


FIGURE B

Commercial Auto Liability - U.S. Industry
Accident Year 1995-2001 Ultimate Loss Ratio Estimates
As of Twenty-Four Months

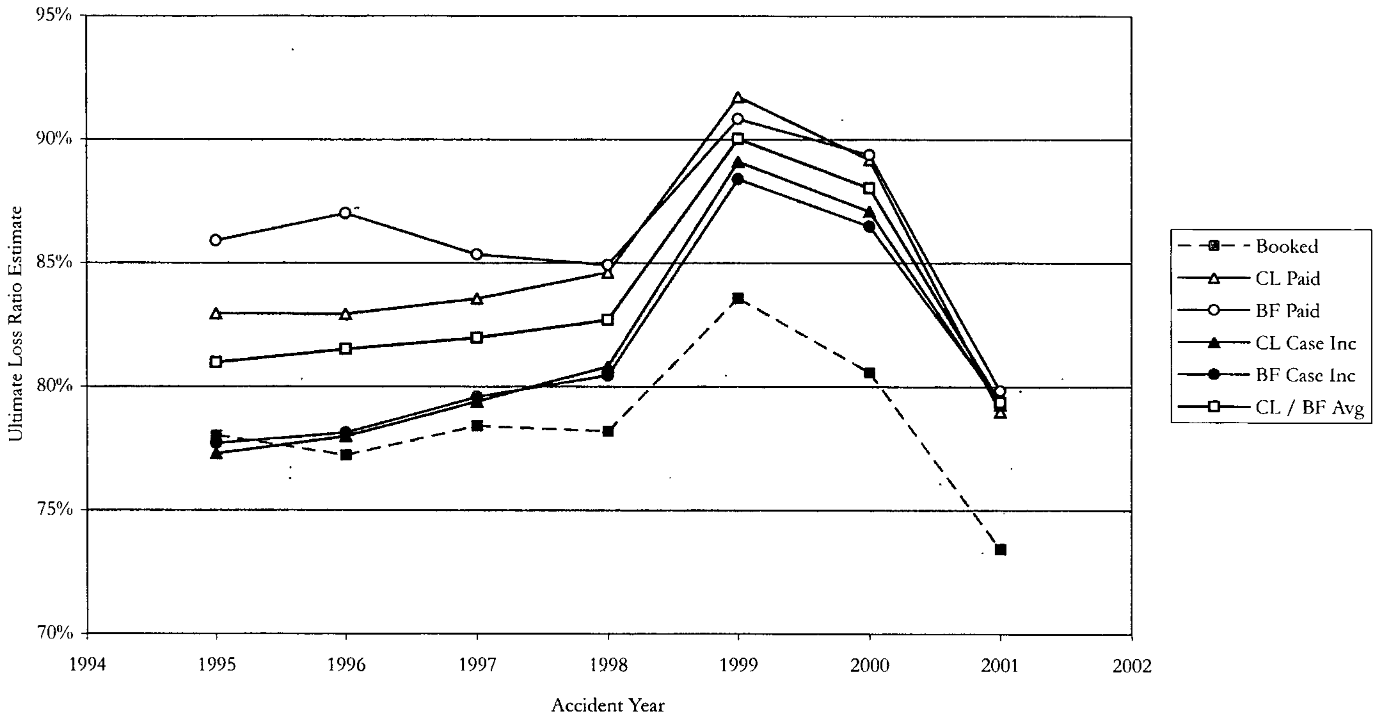


FIGURE C

Commercial Auto Liability - U.S. Industry
Accident Year 1995-2001 Ultimate Loss Ratio Estimates
As of Thirty-Six Months

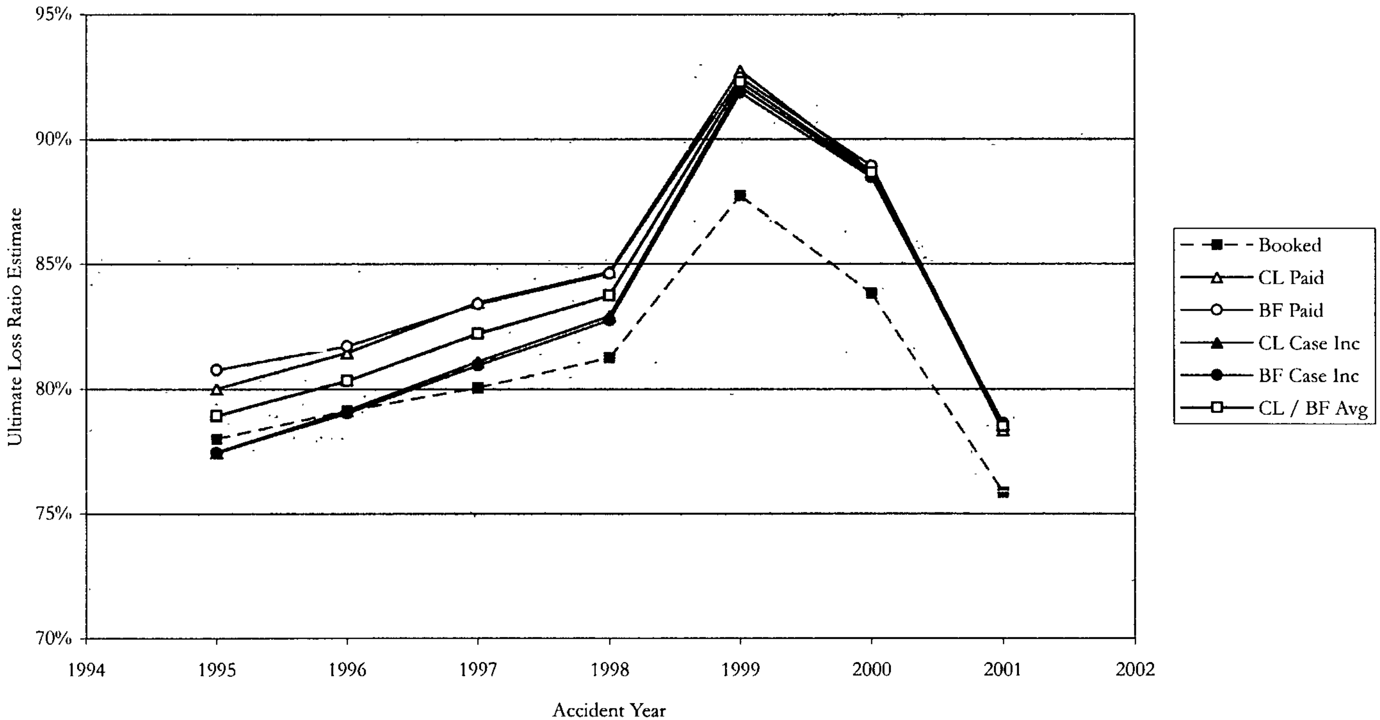


EXHIBIT 1

Comparison of Clinical and Statistical Ultimate Loss Ratio Predictions

Accident Years 1995-2001 as of 12 Months

	Statistical Predictions					Clinical Prediction
Accident Year	(1) Paid Chain Ladder	(2) Paid B-F	(3) Case Inc Chain Ladder	(4) Case Inc B-F	(5) Average CL & B-F	(6) Actual Booked
1995	89.7%	78.6%	79.9%	77.5%	81.4%	78.3%
1996	92.0%	79.9%	78.7%	76.9%	81.9%	76.7%
1997	88.0%	80.2%	80.0%	78.2%	81.6%	77.9%
1998	85.7%	81.4%	78.6%	78.3%	81.0%	76.7%
1999	90.3%	83.7%	84.7%	82.5%	85.3%	78.3%
2000	89.3%	85.2%	82.9%	82.5%	85.0%	77.1%
2001	80.3%	84.7%	78.9%	81.4%	81.3%	73.5%

Notes.

Column	Comments
(1)	See Appendix Exhibit A-3 (upper portion)
(2)	See Appendix Exhibit A-4 (upper portion)
(3)	See Appendix Exhibit A-6 (upper portion)
(4)	See Appendix Exhibit A-7 (upper portion)
(5)	Simple average of Columns (1) through (4)
(6)	See Appendix Exhibit A-1 ("12 Months" Ratios in upper portion)

EXHIBIT 2

Comparison of Clinical and Statistical Ultimate Loss Ratio Predictions

Accident Years 1995-2001 as of 24 Months

	Statistical Predictions					Clinical Prediction
Accident Year	(1) Paid Chain Ladder	(2) Paid B-F	(3) Case Inc Chain Ladder	(4) Case Inc B-F	(5) Average CL & B-F	(6) Actual Booked
1995	83.0%	85.9%	77.3%	77.7%	81.0%	78.0%
1996	82.9%	87.0%	78.0%	78.1%	81.5%	77.2%
1997	83.6%	85.4%	79.4%	79.6%	82.0%	78.4%
1998	84.6%	84.9%	80.8%	80.5%	82.7%	78.2%
1999	91.7%	90.8%	89.1%	88.4%	90.0%	83.6%
2000	89.2%	89.4%	87.1%	86.5%	88.0%	80.6%
2001	79.0%	79.8%	79.3%	79.4%	79.4%	73.4%

Notes.

Column	Comments
(1)	See Appendix Exhibit A-3 (middle portion)
(2)	See Appendix Exhibit A-4 (middle portion)
(3)	See Appendix Exhibit A-6 (middle portion)
(4)	See Appendix Exhibit A-7 (middle portion)
(5)	Simple average of Columns (1) through (4)
(6)	See Appendix Exhibit A-1 ("24 Months" Ratios in upper portion)

EXHIBIT 3

Comparison of Clinical and Statistical Ultimate Loss Ratio Predictions

Accident Years 1995-2001 as of 36 Months

	Statistical Predictions					Clinical Prediction
Accident Year	(1) Paid Chain Ladder	(2) Paid B-F	(3) Case Inc Chain Ladder	(4) Case Inc B-F	(5) Average CL & B-F	(6) Actual Booked
1995	80.0%	80.8%	77.5%	77.4%	78.9%	78.0%
1996	81.5%	81.7%	79.1%	79.0%	80.3%	79.1%
1997	83.5%	83.4%	81.1%	80.9%	82.2%	80.1%
1998	84.7%	84.6%	82.9%	82.8%	83.7%	81.3%
1999	92.7%	92.4%	92.1%	91.9%	92.3%	87.7%
2000	88.7%	88.9%	88.6%	88.5%	88.7%	83.8%
2001	78.3%	78.5%	78.6%	78.6%	78.5%	75.9%

Notes.

Column	Comments
(1)	See Appendix Exhibit A-3 (lower portion)
(2)	See Appendix Exhibit A-4 (lower portion)
(3)	See Appendix Exhibit A-6 (lower portion)
(4)	See Appendix Exhibit A-7 (lower portion)
(5)	Simple average of Columns (1) through (4)
(6)	See Appendix Exhibit A-1 ("36 Months" Ratios in upper portion)

FIGURE D

Commercial Auto Liability - U.S. Industry
Accident Year 1995-2001 Ultimate Loss Ratio Estimates
Errors in Projections as of Twelve Months

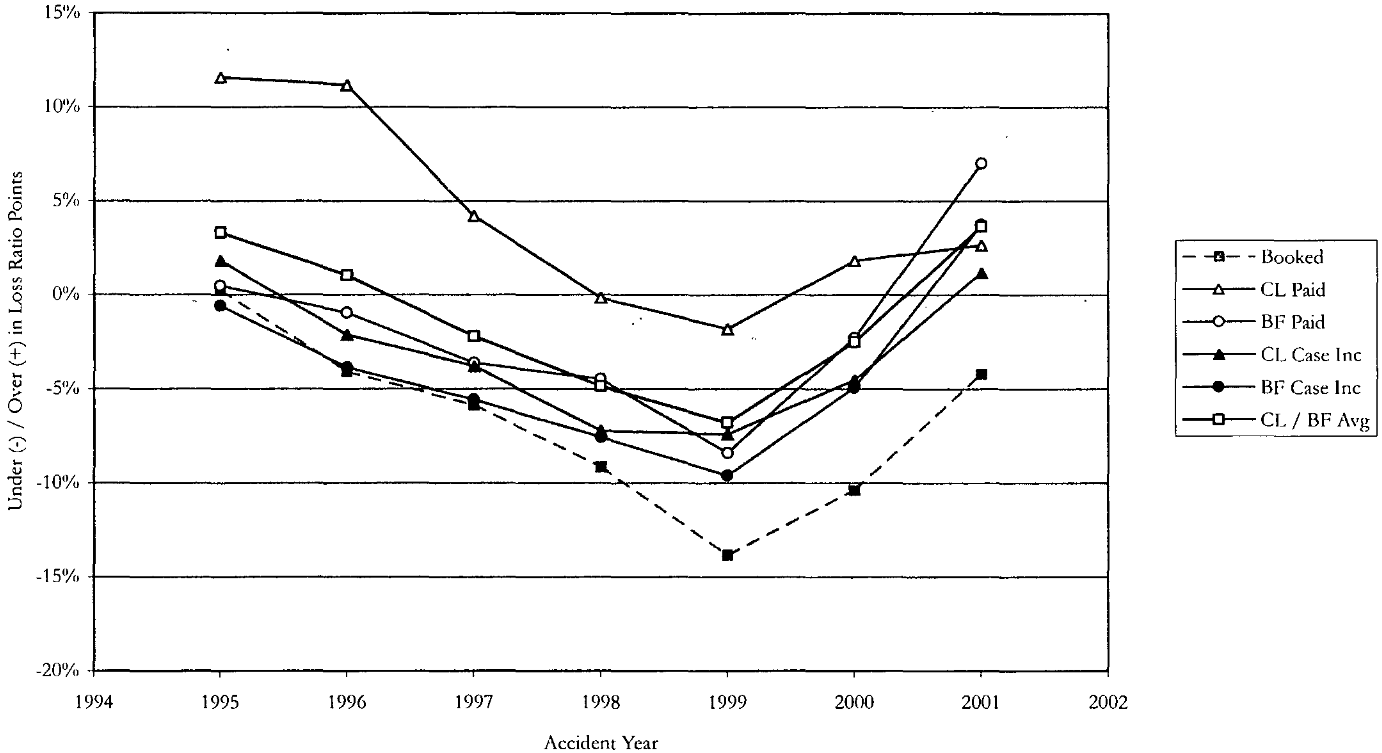


FIGURE E

Commercial Auto Liability - U.S. Industry
Accident Year 1995-2001 Ultimate Loss Ratio Estimates
Errors in Projections as of Twnty-Four Months

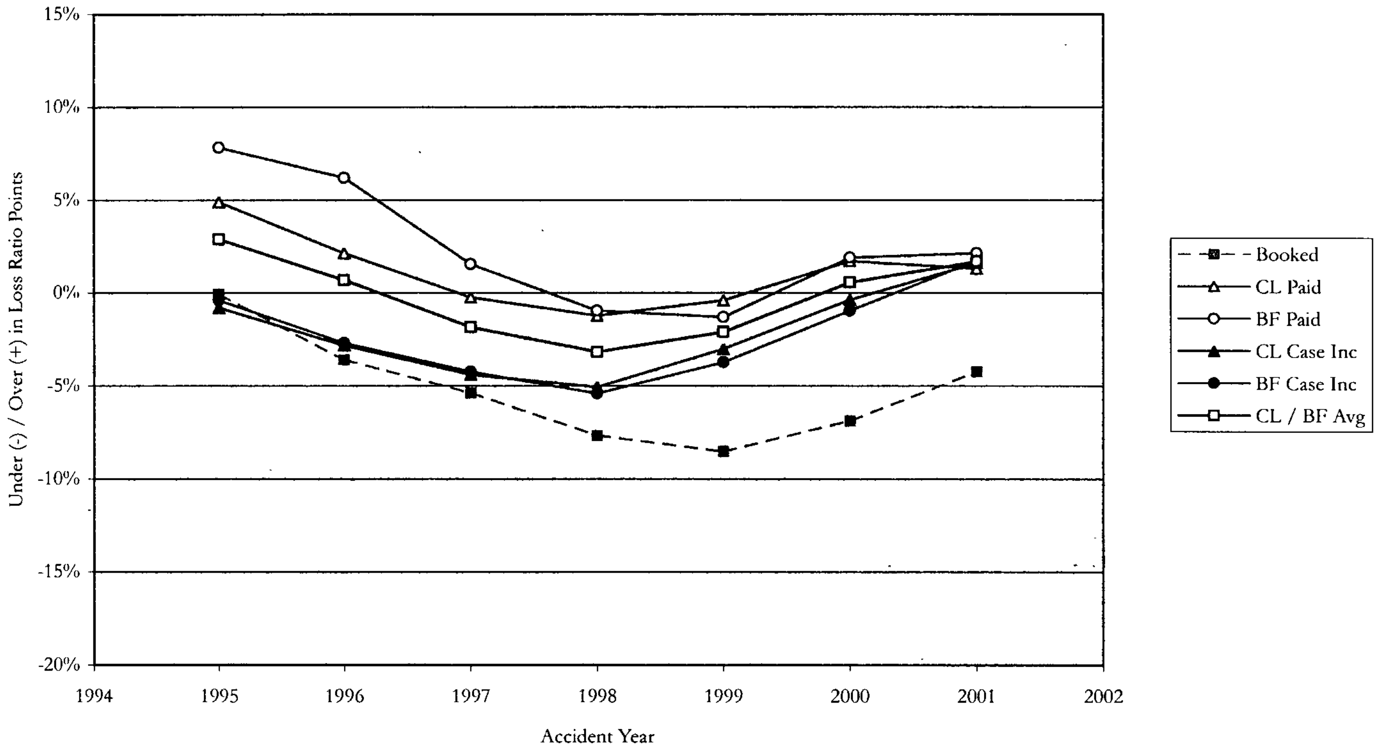


FIGURE F

Commercial Auto Liability - U.S. Industry
Accident Year 1995-2001 Ultimate Loss Ratio Estimates
Errors in Projections as of Thirty-Six Months

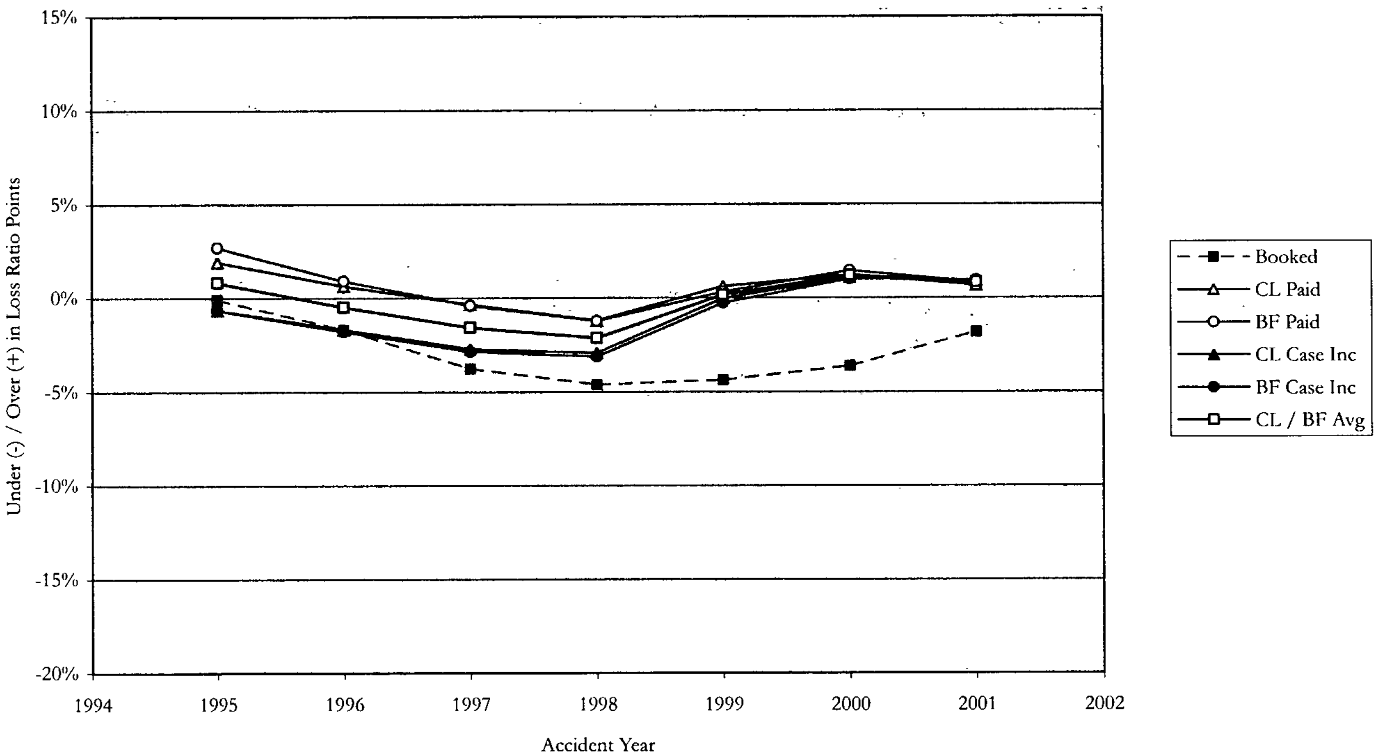


EXHIBIT 4

Comparison of Clinical and Statistical Ultimate Loss Ratio Predictions

Accident Years 1995-2001 as of 12 Months

		Statistical Predictions					Clinical Prediction
Accident Year	(1) Proxy for True Ultimate	(2) Paid Chain Ladder	(3) Paid B-F	(4) Case Inc Chain Ladder	(5) Case Inc B-F	(6) Average CL & B-F	(7) Actual Booked
1995	78.1%	11.6% *	0.5%	1.8%	-0.6%	3.3%	0.2% +
1996	80.8%	11.2% *	-0.9% +	-2.1%	-3.9%	1.1%	-4.1%
1997	83.8%	4.2%	-3.6%	-3.8%	-5.6%	-2.2% +	-5.9%
1998	85.9%	-0.2% +	-4.5%	-7.2%	-7.6%	-4.9%	-9.2% *
1999	92.1%	-1.8% +	-8.4%	-7.4%	-9.6%	-6.8%	-13.8% *
2000	87.5%	1.8% +	-2.3%	-4.5%	-4.9%	-2.5%	-10.4% *
2001	77.7%	2.7%	7.0% *	1.2% +	3.8%	3.7%	-4.2%
Mean Error		4.2%	-1.7%	-3.2%	-4.1%	-1.2% +	-6.8% *
Sum of Errors^2		2.90%	1.59%	1.52%	2.35%	1.07% +	4.52% *
Number of Best (+)		3	1	1	0	1	1
Number of Worst (*)		2	1	0	0	0	4

Notes.

Column	Comments
(1)	See Appendix Exhibit A-1 ("December 2004" Ratios in upper portion)
(2)	Exhibit 1, Column (1) minus Exhibit 4, Column (1)
(3)	Exhibit 1, Column (2) minus Exhibit 4, Column (1)
(4)	Exhibit 1, Column (3) minus Exhibit 4, Column (1)
(5)	Exhibit 1, Column (4) minus Exhibit 4, Column (1)
(6)	Exhibit 1, Column (5) minus Exhibit 4, Column (1)
(7)	Exhibit 1, Column (6) minus Exhibit 4, Column (1)

EXHIBIT 5

Comparison of Clinical and Statistical Ultimate Loss Ratio Predictions

Accident Years 1995-2001 as of 24 Months

	Statistical Predictions						Clinical Prediction
Accident Year	(1) Proxy for True Ultimate	(2) Paid Chain Ladder	(3) Paid B-F	(4) Case Inc Chain Ladder	(5) Case Inc B-F	(6) Average CL & B-F	(7) Actual Booked
1995	78.1%	4.9%	7.8% *	-0.8%	-0.4%	2.9%	-0.1% +
1996	80.8%	2.1%	6.2% *	-2.8%	-2.7%	0.7% +	-3.6%
1997	83.8%	-0.3% +	1.5%	-4.4%	-4.2%	-1.8%	-5.4% *
1998	85.9%	-1.2%	-1.0% +	-5.1%	-5.4%	-3.2%	-7.7% *
1999	92.1%	-0.4% +	-1.3%	-3.0%	-3.7%	-2.1%	-8.5% *
2000	87.5%	1.7%	1.9%	-0.4% +	-1.0%	0.6%	-6.9% *
2001	77.7%	1.3% +	2.1%	1.6%	1.7%	1.7%	-4.2% *
Mean Error		1.2%	2.5%	-2.1%	-2.2%	-0.2% +	-5.2% *
Sum of Errors^2		0.35%	1.13%	0.65%	0.72%	0.30% +	2.39% *
Number of Best (+)		3	1	1	0	1	1
Number of Worst (*)		0	2	0	0	0	5

Notes.

Column	Comments
(1)	See Appendix Exhibit A-1 ("December 2004" Ratios in upper portion)
(2)	Exhibit 2, Column (1) minus Exhibit 4, Column (1)
(3)	Exhibit 2, Column (2) minus Exhibit 4, Column (1)
(4)	Exhibit 2, Column (3) minus Exhibit 4, Column (1)
(5)	Exhibit 2, Column (4) minus Exhibit 4, Column (1)
(6)	Exhibit 2, Column (5) minus Exhibit 4, Column (1)
(7)	Exhibit 2, Column (6) minus Exhibit 4, Column (1)

EXHIBIT 6

Comparison of Clinical and Statistical Ultimate Loss Ratio Predictions

Accident Years 1995-2001 as of 36 Months

	Statistical Predictions						Clinical Prediction
Accident Year	(1) Proxy for True Ultimate	(2) Paid Chain Ladder	(3) Paid B-F	(4) Case Inc Chain Ladder	(5) Case Inc B-F	(6) Average CL & B-F	(7) Actual Booked
1995	78.1%	1.9%	2.7% *	-0.6%	-0.6%	-0.8%	-0.1% +
1996	80.8%	0.7%	0.9%	-1.7%	-1.8% *	-0.5% +	-1.7%
1997	83.8%	-0.3% +	-0.4%	-2.7%	-2.9%	-1.6%	-3.7% *
1998	85.9%	-1.2% +	-1.2%	-2.9%	-3.1%	-2.1%	-4.6% *
1999	92.1%	0.6%	0.3%	0.0% +	-0.2%	0.2%	-4.4% *
2000	87.5%	1.3%	1.5%	1.1%	1.0% +	1.2%	-3.6% *
2001	77.7%	0.7% +	0.8%	0.9%	1.0%	0.8%	-1.8% *
Mean Error		0.5%	0.7%	-0.8%	-0.9%	-0.2% +	-2.8% *
Sum of Errors^2		0.08% +	0.13%	0.21%	0.23%	0.10%	0.74% *
Number of Best (+)		3	0	1	1	1	1
Number of Worst (*)		0	1	0	1	0	5

Notes.

Column	Comments
(1)	See Appendix Exhibit A-1 ("December 2004" Ratios in upper portion)
(2)	Exhibit 3, Column (1) minus Exhibit 4, Column (1)
(3)	Exhibit 3, Column (2) minus Exhibit 4, Column (1)
(4)	Exhibit 3, Column (3) minus Exhibit 4, Column (1)
(5)	Exhibit 3, Column (4) minus Exhibit 4, Column (1)
(6)	Exhibit 3, Column (5) minus Exhibit 4, Column (1)
(7)	Exhibit 3, Column (6) minus Exhibit 4, Column (1)

Appendix A

A.1 Sources of Data Used in Analysis

Our analysis of accident years 1995 through 2001 was based on industry aggregate Schedule P data for Commercial Auto Liability as reported in the 1995 through 2005 volumes of Best's Aggregates and Averages[6]. In particular, we used information about net earned premiums, net ultimate losses, net paid losses and net IBNR from Schedule P, Parts 1C, 2C, 3C and 4C, respectively[7]. We determined case incurred losses by subtracting net IBNR from net ultimate losses.

The loss development history for accident years 1995 through 2001 can be found in the 2005 volume of Best's Aggregates and Averages [14], which is a compilation of information from the industry's 2004 Annual Statements. We have tabulated paid, case incurred and booked ultimate loss and loss ratio information from that volume in Appendix Exhibit A-1 for accident years 1995 through 2001 as of twelve, twenty-four and thirty-six months and also as of December 2004. We used the ultimate loss ratio estimates booked as of December 2004 as proxies for the "true" ultimate loss ratios.

We turned to older volumes of Best's Aggregates & Averages for the loss development data needed to make the statistical ultimate loss ratio predictions for accident years 1995 through 2001 at twelve, twenty-four and thirty-six months. For example, we developed the initial expected loss ratio for the Bornhuetter-Ferguson analysis of accident year 1995 as of twelve months using loss development data from the 1994 Schedule P as reported in the 1995 Best's Aggregates and Averages [4]. For the first chain ladder and Bornhuetter-Ferguson analyses of accident year 1995 at twelve months, we augmented the previous loss development factor triangle available from the 1994 Schedule P with 1995 data from the 1996 volume of Best's Aggregate and Averages [5]. We computed the development factors corresponding to development between December 1994 and December 1995 from the data in the 1996 volume and added those development factors to the previous triangle[8]. Similarly, for the analysis of the later accident years and/or later valuation dates we continued to augment the triangle of loss development factors using data from later volumes of Best's Aggregates and Averages. (See [6] through [13].)

---

6 1995 [4], 1996 [5], 1997 [6], 1998 [7], 1999 [8], 2000 [9], 2001 [10], 2002 [11], 2003 [12], 2004 [13], 2005 [14]

7 All references to "net losses" should be understood to include the "defense and cost containment expenses" reported in Parts 2C, 3C and 4C of Schedule P.

8 It is more reliable to calculate development factors using data for both numerator and denominator from within a single Schedule P than to take numerator and denominator from Schedules P from successive years, because of slight differences in the companies included in Best's Aggregates and Averages from year to year.

We tabulated the paid loss ratios as of twelve months together with the age-to-age paid development factors in Appendix Exhibit A-2A. The standard format for a triangle of loss development factors shows all development factors for a given accident year in a single row. In that format, the loss development factors associated with the development observed within individual calendar years appear on the positively sloped diagonals.

Appendix Exhibit A-2A departs slightly from the standard format to show all of the development factors observed in a given *calendar year* in a single row rather than on a diagonal. In this format, the development factors associated within individual accident years appear on the negatively sloped diagonals. The advantage of this format is that the five-year average development factors, which are tabulated in the upper section of Appendix Exhibit A-2B, can be computed by reference to five rows of data rather than more complicated references to the five points in a triangular array. This is particularly helpful in this analysis, where we are projecting seven accident years at three different valuations.

The lower section of Appendix Exhibit A-2B shows the cumulative mean development factors to age ten years (which is the outer bound of our development data) and the age ten loss ratios indicated by applying the age twelve months to age ten years development factor to the trailing five-year loss ratio as of twelve months. Those loss ratios, multiplied by a tail factor, are used as initial expected loss ratios in the Bornhuetter-Ferguson analysis as of twelve months.

Appendix Exhibits A-5A and A-5B are the case incurred analogues to Appendix Exhibit A-2A and A-2B. They are tabulations of the case incurred loss ratios as of twelve months and the case incurred development factors based on the Schedule P data contained in the 1995 through 2005 volumes of Best's Aggregates & Averages.

A.2 Clinically and Statistically Predicted Loss Ratios

In this section we describe the source of the booked ultimate loss ratio estimates that we classify as clinical predictions and discuss the details underlying the five judgment-free statistical prediction methods used in our analysis.

A.2.1 Clinically Predicted Ultimate Loss Ratios

The clinical predictions of ultimate loss ratios are available from the 2004 Schedule P compilations that appear in the 2005 Best's Aggregates and Averages. We have tabulated these ultimate loss ratio estimates together with the underlying earned premium and ultimate loss dollars in the upper portion of Appendix Exhibit A-1 in the sections labeled "12

Months", "24 Months" and "36 Months." The earned premium figures are from Schedule P, Part 1C. The ultimate loss figures are from Schedule P, Part 2C.

A.2.2 Statistically Predicted Ultimate Loss Ratios

We made statistical predictions of the ultimate loss ratios for accident years 1995-2001 using the unadjusted results of five loss development methods: 1) the paid chain ladder method, 2) the paid Bornhuetter-Ferguson method, 3) the case incurred chain ladder method, 4) the case incurred Bornhuetter-Ferguson method and 5) the simple average of the results of methods 1-4. We call these statistical predictions because we used the indicated results of each of these methods in every case and injected no subjective judgment by adjusting results that might seem odd or unreasonable.

Paid Chain Ladder Ultimate Loss Ratio Estimates

We determined chain ladder ultimate loss ratio estimates by applying paid loss development factors to paid loss ratios in the standard way. For example, for the 1995 accident year projection at twelve months, we first calculated mean age-to-age factors from historical paid loss data available, using five-year simple means where possible, reflecting the development patterns observed during calendar years 1991 through 1995[9]. These mean age-to-age factors and the cumulative factors they imply out to age ten years are tabulated in Appendix Exhibit A-2B. We used these mean factors as estimates of the appropriate prospective development factors applicable to the 1995 accident year. For the tail factor (age ten years to ultimate) we used the relationship between estimated ultimate losses and paid losses as reported in the 2004 Schedule P, which yielded a factor of 1.009[10]. We then multiplied the 1995 paid loss ratio at twelve months by the age twelve months to ultimate loss development factor derived from the age-to-age factors and the tail. We used the same procedure to determine prospective development factors for use with the other accident years and valuations.

Appendix Exhibit A-3 summarizes the calculation of the paid chain ladder ultimate loss ratio estimates for accident years 1995-2001 as of twelve, twenty-four and thirty-six months.

[9] Appendix Exhibit A-2 shows that for factors corresponding to development from age seven years and beyond, fewer than five loss development factors were available for the mean calculations.

[10] Based on estimated ultimate losses of $8,916,383 and paid losses of $8,835,898 as of December 2004 as reported in Appendix Exhibit A-1.

Paid Bornhuetter-Ferguson Ultimate Loss Ratio Estimates

We determined estimates of the accident year 1995-2001 ultimate loss ratios at twelve, twenty-four and thirty-six months using the version of the Bornhuetter-Ferguson method with paid loss data described below.

First, we determined the initial expected loss ratio for each accident year to be used in the first Bornhuetter-Ferguson analysis at twelve months. Using data available at the beginning of each accident year we calculated the simple mean of the paid loss ratios as of twelve months from the five prior accident years. For example, to calculate the initial expected loss ratio for accident year 1995, we first computed the mean of the 1990-1994 paid loss ratios as of twelve months, (15.0% + 14.5% + 14.4% +15.5% + 17.6%) / 5 = 15.4%, which we took as the expected paid loss ratio for accident year 1995 at twelve months. We calculated age-to-age development factors in the same way. We then multiplied the 15.4% by the age twelve months to ultimate paid development factor (including the tail factor of 1.009 discussed in the paid chain ladder section) to arrive at 77.7% as the initial Bornhuetter-Ferguson expected loss ratio for accident year 1995. While this procedure is crude, and it is easy to think of ways to improve upon it, in the present circumstances it has the merit of being based only on data available in Schedule P. No additional data or no subjective judgment is required. For 1996 and other accident years we calculated the initial loss ratio for the twelve months valuation in the same way. See Appendix Exhibits A-2A and A-2B for compilations of the historical and average paid loss ratios and age-to-age factors together with implied cumulative development factors out to age ten years on which the initial expected loss ratios were based.

In some versions of the Bornhuetter-Ferguson method the initial expected loss ratio is used not only at the twelve-month valuation but also at all subsequent ones. We believe, however, that it is more common to update the expected loss ratio for the analysis at later valuations, and the version we used for our analysis uses updated expected loss ratios. We again sought to avoid injecting either exogenous information not available from Schedule P or subjective judgment into the analysis, so we adopted the indicated ultimate loss ratio indication from the paid chain ladder method at the previous valuation as the expected loss ratio for all valuation dates beyond twelve months.

The expected loss ratio depends on two quantifiable elements: 1) the expected development in the next twelve months and 2) the expected development beyond the next twelve months. By definition, the first element becomes obsolete in twelve months and is replaced in the estimation process by the actual development that has occurred. In contrast, in twelve months the second element continues to lie entirely in the future. However, the loss development in the tail observed during the previous twelve months has probably affected our estimate of that future development. In other words, the age-to-ultimate factor has probably been revised to reflect the most recent year of development on the older accident years.

The Bornhuetter-Ferguson ultimate loss ratio estimate typically combines the actual accident year emergence with the updated tail. This can be expressed in formula terms as follows:

$$BF\ Loss\ Ratio = Actual\ Paid\ Loss\ Ratio + ELR \times (1 - \frac{1}{LDF_{updated}})$$

Because in our formulation the expected loss ratio was explicitly constructed as the product of the expected paid loss ratio and the expected age-to-ultimate development factor ($ELR = Expected\ Paid\ Loss\ Ratio \times LDF$), we concluded that $LDF_{updated}$ should also be used to update the expected loss ratio as follows:

$$BF\ Loss\ Ratio = Actual\ Paid\ Loss\ Ratio + ELR \times \frac{LDF_{updated}}{LDF} \times (1 - \frac{1}{LDF_{updated}})$$

This adjustment has the effect of updating the expected loss ratio in light of the updated development data to: $ELR = Expected\ Paid\ Loss\ Ratio \times LDF_{updated}$. We recognize that this is not the standard Bornhuetter-Ferguson formulation. However, it is conceptually more consistent with the premise of the expected loss ratio to make this adjustment than not to make it.

The details of the paid Bornhuetter-Ferguson analysis performed for accident years 1995 through 2001 as of twelve, twenty-four and thirty-six months are presented in Appendix Exhibit A-4.

Case Incurred Chain Ladder Ultimate Loss Ratio Estimates

The ultimate loss ratio analysis using the chain ladder method with case incurred loss data paralleled the paid chain ladder analysis. The only differences were that it used case incurred loss data instead of paid loss data and the tail factor (for age ten years to ultimate) was determined from the relationship between accident year 1995 ultimate losses and case incurred losses (rather than paid losses) reported in the 2004 Schedule P. The case incurred tail factor determined in this way was 1.002[11]. See Appendix Exhibit A-5 for compilation of these historical and average case incurred loss ratios and age-to-age factors together with implied cumulative development factors out to age ten years.

The results of the case incurred chain ladder analysis for accident years 1995 through 2001 as of twelve, twenty-four and thirty-six months are summarized in Appendix Exhibit A-6.

Case Incurred B-F Ultimate Loss Ratio Estimates

Similarly, the case incurred Bornhuetter-Ferguson loss ratio analysis paralleled the paid Bornhuetter-Ferguson analysis, except that it used case incurred rather than paid data from Appendix Exhibit A-5. The results from this analysis of accident years 1995 through 2001 as of twelve, twenty-four and thirty-six months are summarized in Appendix Exhibit A-7.

Average of Incurred Chain Ladder and B-F Methods (Paid and Case Incurred)

Ultimate loss ratio selections are rarely determined from only one method. The simple average approach adopted here as a fifth statistical prediction acknowledges in a simple way the practice of combining estimates from different methods.

[11] Based on estimated ultimate losses of $8,916,383 and case incurred losses of $8,895,998 as of December 2004 as reported in Appendix Exhibit A-1.

APPENDIX EXHIBIT A-1

2004 Annual Statement (U.S. Industry)
Selected Premium and Loss Statistics

		Estimated Ultimate Net Losses and Loss Expense							
Accident Year	Net Earned Premiums	12 Months		24 Months		36 Months		December 2004	
		Dollars	Ratio	Dollars	Ratio	Dollars	Ratio	Dollars	Ratio
1995	11,419,308	8,944,478	78.3%	8,909,903	78.0%	8,907,535	78.0%	8,916,383	78.1%
1996	11,945,125	9,164,925	76.7%	9,224,673	77.2%	9,452,826	79.1%	9,660,376	80.9%
1997	12,101,165	9,430,510	77.9%	9,488,547	78.4%	9,687,547	80.1%	10,141,169	83.8%
1998	12,165,123	9,331,198	76.7%	9,512,292	78.2%	9,885,056	81.3%	10,445,429	85.9%
1999	12,053,631	9,436,430	78.3%	10,073,714	83.6%	10,575,733	87.7%	11,103,268	92.1%
2000	12,929,133	9,966,148	77.1%	10,416,697	80.6%	10,837,941	83.8%	11,037,507	85.4%
2001	14,186,157	10,420,178	73.5%	10,416,359	73.4%	10,761,679	75.9%	11,018,475	77.7%

		Case Incurred Losses and Loss Expense							
Accident Year	Net Earned Premiums	12 Months		24 Months		36 Months		December 2004	
		Dollars	Ratio	Dollars	Ratio	Dollars	Ratio	Dollars	Ratio
1995	11,419,308	5,349,752	46.8%	7,155,266	62.7%	8,035,265	70.4%	8,895,998	77.9%
1996	11,945,125	5,599,565	46.9%	7,554,912	63.2%	8,590,063	71.9%	9,624,782	80.6%
1997	12,101,165	5,810,562	48.0%	7,761,367	64.1%	8,911,313	73.6%	10,075,215	83.3%
1998	12,165,123	5,725,649	47.1%	7,899,777	64.9%	9,112,603	74.9%	10,357,940	85.1%
1999	12,053,631	6,064,094	50.3%	8,537,262	70.8%	9,923,657	82.3%	10,956,003	90.9%
2000	12,929,133	6,256,104	48.4%	8,793,340	68.0%	10,162,998	78.6%	10,788,755	83.4%
2001	14,186,157	6,350,997	44.8%	8,668,276	61.1%	9,922,085	69.9%	10,503,768	74.0%

		Net Paid Losses and Loss Expense							
Accident Year	Net Earned Premiums	12 Months		24 Months		36 Months		December 2004	
		Dollars	Ratio	Dollars	Ratio	Dollars	Ratio	Dollars	Ratio
1995	11,419,308	2,080,653	18.2%	4,400,438	38.5%	6,188,228	54.2%	8,835,898	77.4%
1996	11,945,125	2,298,993	19.2%	4,670,807	39.1%	6,642,691	55.6%	9,532,038	79.8%
1997	12,101,165	2,320,305	19.2%	4,824,751	39.9%	6,916,574	57.2%	9,936,030	82.1%
1998	12,165,123	2,334,107	19.2%	4,942,814	40.6%	7,062,840	58.1%	10,108,623	83.1%
1999	12,053,631	2,486,813	20.6%	5,329,527	44.2%	7,657,087	63.5%	10,524,675	87.3%
2000	12,929,133	2,652,474	20.5%	5,540,847	42.9%	7,840,880	60.6%	10,279,657	79.5%
2001	14,186,157	2,617,173	18.4%	5,367,450	37.8%	7,607,185	53.6%	9,122,500	64.3%

APPENDIX EXHIBIT A-2A

Accident Year Paid Loss Development Factors
By Calendar Year of Observed Development

Calendar Year	Age 1 Loss Ratio	Age 1 - 2	Age 2 - 3	Age 3 - 4	Age 4 - 5	Age 5 - 6	Age 6 - 7	Age 7 - 8	Age 8 - 9	Age 9 - 10
1990	15.0%	2.291	1.464	1.234	1.120	1.055				
1991	14.5%	2.291	1.454	1.221	1.110	1.059	1.030			
1992	14.4%	2.311	1.465	1.213	1.109	1.057	1.031	1.016		
1993	15.5%	2.242	1.451	1.210	1.105	1.054	1.030	1.015	1.007	
1994	17.6%	2.265	1.456	1.196	1.101	1.050	1.028	1.016	1.008	1.004
1995	18.2%	2.165	1.449	1.205	1.099	1.048	1.025	1.013	1.008	1.004
1996	19.2%	2.115	1.422	1.202	1.104	1.047	1.024	1.011	1.005	1.004
1997	19.2%	2.032	1.406	1.209	1.098	1.047	1.024	1.012	1.006	1.004
1998	19.2%	2.079	1.422	1.197	1.096	1.046	1.020	1.012	1.007	1.003
1999	20.6%	2.118	1.434	1.198	1.093	1.049	1.025	1.013	1.006	1.004
2000	20.5%	2.143	1.429	1.208	1.105	1.047	1.021	1.010	1.004	1.002
2001	18.4%	2.089	1.437	1.215	1.101	1.046	1.023	1.010	1.006	1.002
2002	14.3%	2.051	1.415	1.215	1.105	1.049	1.023	1.010	1.007	1.004
2003	12.8%	2.090	1.417	1.196	1.096	1.046	1.020	1.007	1.006	1.004

Data Source Legend
- 1994 Schedule P as reported in 1995 Best's Aggregates & Averages
- 1995-2003 Schedules P as reported in 1995-2004 Best's Aggregates & Averages
- 2004 Schedule P as reported in 2005 Best's Aggregates & Averages

APPENDIX EXHIBIT A-2B

Accident Year Paid Loss Development Factors

Trailing Five-Year Average Age to Age Development

Year	Age 1 Loss Ratio	Age 1 - 2	Age 2 - 3	Age 3 - 4	Age 4 - 5	Age 5 - 6	Age 6 - 7	Age 7 - 8	Age 8 - 9	Age 9 - 10
1994	15.4%	2.280	1.458	1.214	1.109	1.055	1.030	1.016	1.007	1.004
1995	16.0%	2.255	1.455	1.209	1.105	1.054	1.029	1.015	1.007	1.004
1996	17.0%	2.220	1.448	1.205	1.103	1.051	1.028	1.014	1.007	1.004
1997	17.9%	2.164	1.437	1.204	1.101	1.049	1.026	1.013	1.007	1.004
1998	18.7%	2.131	1.431	1.202	1.099	1.048	1.024	1.013	1.007	1.004
1999	19.3%	2.102	1.427	1.202	1.098	1.048	1.024	1.012	1.006	1.004
2000	19.8%	2.097	1.423	1.203	1.099	1.047	1.023	1.012	1.006	1.003
2001	19.6%	2.092	1.426	1.205	1.099	1.047	1.023	1.011	1.006	1.003
2002	18.6%	2.096	1.427	1.207	1.100	1.048	1.022	1.011	1.006	1.003
2003	17.4%	2.098	1.426	1.206	1.100	1.047	1.022	1.010	1.006	1.003

Trailing Five-Year Average Development to Age 10 Years

Year	Age 1 Loss Ratio	Age 1 - 2	Age 2 - 3	Age 3 - 4	Age 4 - 5	Age 5 - 6	Age 6 - 7	Age 7 - 8	Age 8 - 9	Age 9 - 10
1994	77.0%	5.000	2.193	1.504	1.238	1.117	1.058	1.028	1.012	1.004
1995	78.2%	4.877	2.163	1.487	1.230	1.113	1.057	1.027	1.012	1.004
1996	80.5%	4.736	2.134	1.473	1.222	1.108	1.054	1.026	1.011	1.004
1997	81.6%	4.548	2.102	1.463	1.215	1.103	1.052	1.025	1.011	1.004
1998	82.7%	4.427	2.077	1.452	1.208	1.099	1.049	1.024	1.011	1.004
1999	83.7%	4.338	2.064	1.447	1.204	1.097	1.047	1.023	1.010	1.004
2000	85.2%	4.312	2.056	1.445	1.202	1.094	1.044	1.021	1.009	1.003
2001	84.5%	4.315	2.062	1.447	1.200	1.092	1.043	1.020	1.009	1.003
2002	80.8%	4.336	2.069	1.450	1.201	1.092	1.043	1.020	1.009	1.003
2003	75.2%	4.332	2.065	1.448	1.200	1.091	1.041	1.019	1.009	1.003

APPENDIX EXHIBIT A-3

Paid Chain Ladder Ultimate Loss Estimates
As of Twelve, Twenty-Four and Thirty-Six Months

As of Twelve Months

Accident Year	Loss Ratio at 12 Mo.	Loss Development Factors To 10 Yrs	Tail	To Ult	CL Ult Loss Ratio
1995	18.2%	4.877	1.009	4.921	89.7%
1996	19.2%	4.736	1.009	4.779	92.0%
1997	19.2%	4.548	1.009	4.590	88.0%
1998	19.2%	4.427	1.009	4.467	85.7%
1999	20.6%	4.338	1.009	4.378	90.3%
2000	20.5%	4.312	1.009	4.352	89.3%
2001	18.4%	4.315	1.009	4.354	80.3%

As of Twenty-Four Months

Accident Year	Loss Ratio at 24 Mo.	Loss Development Factors To 10 Yrs	Tail	To Ult	CL Ult Loss Ratio
1995	38.5%	2.134	1.009	2.153	83.0%
1996	39.1%	2.102	1.009	2.121	82.9%
1997	39.9%	2.077	1.009	2.096	83.6%
1998	40.6%	2.064	1.009	2.083	84.6%
1999	44.2%	2.056	1.009	2.075	91.7%
2000	42.9%	2.062	1.009	2.081	89.2%
2001	37.8%	2.069	1.009	2.088	79.0%

As of Thirty-Six Months

Accident Year	Loss Ratio at 36 Mo.	Loss Development Factors To 10 Yrs	Tail	To Ult	CL Ult Loss Ratio
1995	54.2%	1.463	1.009	1.476	80.0%
1996	55.6%	1.452	1.009	1.465	81.5%
1997	57.2%	1.447	1.009	1.460	83.5%
1998	58.1%	1.445	1.009	1.459	84.7%
1999	63.5%	1.447	1.009	1.460	92.7%
2000	60.6%	1.450	1.009	1.463	88.7%
2001	53.6%	1.448	1.009	1.461	78.3%

APPENDIX EXHIBIT A-4

Paid Bornhuetter-Ferguson Ultimate Loss Estimates
As of Twelve, Twenty-Four and Thirty-Six Months

As of Twelve Months

Accident Year	Loss Ratio at 12 Mo.	BF ELR	Age to Ult LDF Current	Age to Ult LDF Prior	BF Ult Loss Ratio
1995	18.2%	77.7%	4.921	5.045	78.6%
1996	19.2%	78.9%	4.779	4.921	79.9%
1997	19.2%	81.2%	4.590	4.779	80.2%
1998	19.2%	82.3%	4.467	4.590	81.4%
1999	20.6%	83.4%	4.378	4.467	83.7%
2000	20.5%	84.5%	4.352	4.378	85.2%
2001	18.4%	85.9%	4.354	4.352	84.7%

As of Twenty-Four Months

Accident Year	Loss Ratio at 24 Mo.	BF ELR	Age to Ult LDF Current	Age to Ult LDF Prior	BF Ult Loss Ratio
1995	38.5%	89.7%	2.153	2.182	85.9%
1996	39.1%	92.0%	2.121	2.153	87.0%
1997	39.9%	88.0%	2.096	2.121	85.3%
1998	40.6%	85.7%	2.083	2.096	84.9%
1999	44.2%	90.3%	2.075	2.083	90.8%
2000	42.9%	89.3%	2.081	2.075	89.4%
2001	37.8%	80.3%	2.088	2.081	79.8%

As of Thirty-Six Months

Accident Year	Loss Ratio at 36 Mo.	BF ELR	Age to Ult LDF Current	Age to Ult LDF Prior	BF Ult Loss Ratio
1995	54.2%	83.0%	1.476	1.486	80.8%
1996	55.6%	82.9%	1.465	1.476	81.7%
1997	57.2%	83.6%	1.460	1.465	83.4%
1998	58.1%	84.6%	1.459	1.460	84.6%
1999	63.5%	91.7%	1.460	1.459	92.4%
2000	60.6%	89.2%	1.463	1.460	88.9%
2001	53.6%	79.0%	1.461	1.463	78.5%

APPENDIX EXHIBIT A-5A

Accident Year Case Incurred Loss Development Factors
By Calendar Year of Observed Development

Calendar Year	Age 1 Loss Ratio	Age 1 - 2	Age 2 - 3	Age 3 - 4	Age 4 - 5	Age 5 - 6	Age 6 - 7	Age 7 - 8	Age 8 - 9	Age 9 - 10
1990	42.6%	1.392	1.145	1.067	1.034	1.014				
1991	41.7%	1.409	1.128	1.058	1.026	1.011	1.008			
1992	42.0%	1.397	1.124	1.056	1.024	1.014	1.005	1.003		
1993	44.2%	1.345	1.114	1.058	1.025	1.010	1.007	1.003	1.001	
1994	46.7%	1.363	1.123	1.048	1.024	1.010	1.006	1.004	1.002	1.002
1995	46.8%	1.362	1.120	1.051	1.020	1.007	1.002	1.001	1.001	1.001
1996	46.9%	1.337	1.121	1.050	1.024	1.009	1.002	1.002	1.000	1.001
1997	48.0%	1.349	1.123	1.060	1.025	1.008	1.007	1.003	1.001	1.000
1998	47.1%	1.336	1.137	1.065	1.024	1.010	1.003	1.002	1.001	1.000
1999	50.3%	1.380	1.148	1.068	1.018	1.010	1.003	1.001	0.999	1.001
2000	48.4%	1.408	1.154	1.069	1.028	1.011	1.004	1.000	0.999	0.999
2001	44.8%	1.406	1.162	1.082	1.036	1.013	1.006	1.003	1.002	1.002
2002	36.9%	1.365	1.156	1.073	1.036	1.017	1.008	1.003	1.003	1.003
2003	34.2%	1.375	1.145	1.059	1.022	1.009	1.002	0.998	1.001	1.000

Data Source Legend

- 1994 Schedule P as reported in 1995 Best's Aggregates & Averages
- 1995-2003 Schedules P as reported in 1995-2004 Best's Aggregates & Averages
- 2004 Schedule P as reported in 2005 Best's Aggregates & Averages

APPENDIX EXHIBIT A-5B

Accident Year Case Incurred Loss Development Factors

Trailing Five-Year Average Age to Age Development

Year	Age 1 Loss Ratio	Age 1 - 2	Age 2 - 3	Age 3 - 4	Age 4 - 5	Age 5 - 6	Age 6 - 7	Age 7 - 8	Age 8 - 9	Age 9 - 10
1994	43.4%	1.381	1.127	1.057	1.027	1.012	1.007	1.003	1.002	1.002
1995	44.3%	1.375	1.122	1.054	1.024	1.011	1.006	1.003	1.001	1.002
1996	45.3%	1.361	1.121	1.053	1.023	1.010	1.005	1.002	1.001	1.001
1997	46.5%	1.351	1.120	1.053	1.024	1.009	1.005	1.002	1.001	1.001
1998	47.1%	1.350	1.125	1.055	1.023	1.009	1.004	1.002	1.001	1.001
1999	47.8%	1.353	1.130	1.059	1.022	1.009	1.003	1.002	1.001	1.001
2000	48.1%	1.362	1.137	1.062	1.024	1.010	1.004	1.002	1.000	1.000
2001	47.7%	1.376	1.145	1.069	1.026	1.010	1.004	1.002	1.001	1.000
2002	45.5%	1.379	1.151	1.072	1.028	1.012	1.005	1.002	1.001	1.001
2003	42.9%	1.387	1.153	1.070	1.028	1.012	1.004	1.001	1.001	1.001

Trailing Five-Year Average Development to Age 10 Years

Year	Age 1 Loss Ratio	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9
1994	75.3%	1.732	1.254	1.113	1.053	1.026	1.013	1.007	1.004	1.002
1995	75.4%	1.702	1.237	1.103	1.046	1.022	1.011	1.006	1.003	1.002
1996	75.9%	1.675	1.231	1.098	1.043	1.020	1.009	1.005	1.002	1.001
1997	77.4%	1.663	1.230	1.098	1.043	1.019	1.010	1.005	1.002	1.001
1998	78.5%	1.667	1.235	1.098	1.041	1.017	1.008	1.004	1.002	1.001
1999	80.3%	1.680	1.241	1.099	1.038	1.015	1.006	1.003	1.001	1.001
2000	82.3%	1.710	1.255	1.104	1.040	1.015	1.006	1.002	1.001	1.000
2001	83.9%	1.758	1.278	1.116	1.044	1.018	1.007	1.003	1.001	1.000
2002	81.2%	1.785	1.294	1.124	1.049	1.020	1.008	1.003	1.002	1.001
2003	76.9%	1.792	1.292	1.121	1.047	1.019	1.007	1.003	1.002	1.001

APPENDIX EXHIBIT A-6

Case Incurred Chain Ladder Ultimate Loss Estimates
As of Twelve, Twenty-Four and Thirty-Six Months

As of Twelve Months

Accident Year	Loss Ratio at 12 Mo.	Loss Development Factors To 10 Yrs	Tail	To Ult	CL Ult Loss Ratio
1995	46.8%	1.702	1.002	1.706	79.9%
1996	46.9%	1.675	1.002	1.679	78.7%
1997	48.0%	1.663	1.002	1.666	80.0%
1998	47.1%	1.667	1.002	1.671	78.6%
1999	50.3%	1.680	1.002	1.683	84.7%
2000	48.4%	1.710	1.002	1.714	82.9%
2001	44.8%	1.758	1.002	1.762	78.9%

As of Twenty-Four Months

Accident Year	Loss Ratio at 24 Mo.	Loss Development Factors To 10 Yrs	Tail	To Ult	CL Ult Loss Ratio
1995	62.7%	1.231	1.002	1.233	77.3%
1996	63.2%	1.230	1.002	1.233	78.0%
1997	64.1%	1.235	1.002	1.238	79.4%
1998	64.9%	1.241	1.002	1.244	80.8%
1999	70.8%	1.255	1.002	1.258	89.1%
2000	68.0%	1.278	1.002	1.281	87.1%
2001	61.1%	1.294	1.002	1.297	79.3%

As of Thirty-Six Months

Accident Year	Loss Ratio at 36 Mo.	Loss Development Factors To 10 Yrs	Tail	To Ult	CL Ult Loss Ratio
1995	70.4%	1.098	1.002	1.101	77.5%
1996	71.9%	1.098	1.002	1.100	79.1%
1997	73.6%	1.099	1.002	1.101	81.1%
1998	74.9%	1.104	1.002	1.107	82.9%
1999	82.3%	1.116	1.002	1.119	92.1%
2000	78.6%	1.124	1.002	1.127	88.6%
2001	69.9%	1.121	1.002	1.123	78.6%

APPENDIX EXHIBIT A-7

Case Incurred Bornhuetter-Ferguson Ultimate Loss Estimates
As of Twelve, Twenty-Four and Thirty-Six Months

As of Twelve Months

Accident Year	Loss Ratio at 12 Mo.	BF ELR	Age to Ult LDF Current	Age to Ult LDF Prior	BF Ult Loss Ratio
1995	46.8%	75.4%	1.706	1.736	77.5%
1996	46.9%	75.6%	1.679	1.706	76.9%
1997	48.0%	76.1%	1.666	1.679	78.2%
1998	47.1%	77.6%	1.671	1.666	78.3%
1999	50.3%	78.7%	1.683	1.671	82.5%
2000	48.4%	80.5%	1.714	1.683	82.5%
2001	44.8%	82.5%	1.762	1.714	81.4%

As of Twenty-Four Months

Accident Year	Loss Ratio at 24 Mo.	BF ELR	Age to Ult LDF Current	Age to Ult LDF Prior	BF Ult Loss Ratio
1995	62.7%	79.9%	1.233	1.240	77.7%
1996	63.2%	78.7%	1.233	1.233	78.1%
1997	64.1%	80.0%	1.238	1.233	79.6%
1998	64.9%	78.6%	1.244	1.238	80.5%
1999	70.8%	84.7%	1.258	1.244	88.4%
2000	68.0%	82.9%	1.281	1.258	86.5%
2001	61.1%	78.9%	1.297	1.281	79.4%

As of Thirty-Six Months

Accident Year	Loss Ratio at 36 Mo.	BF ELR	Age to Ult LDF Current	Age to Ult LDF Prior	BF Ult Loss Ratio
1995	70.4%	77.3%	1.101	1.101	77.4%
1996	71.9%	78.0%	1.100	1.101	79.0%
1997	73.6%	79.4%	1.101	1.100	80.9%
1998	74.9%	80.8%	1.107	1.101	82.8%
1999	82.3%	89.1%	1.119	1.107	91.9%
2000	78.6%	87.1%	1.127	1.119	88.5%
2001	69.9%	79.3%	1.123	1.127	78.6%

4. REFERENCES

[1] Snijders, C., F. Tazelaar and R. Batenburg, "Electronic Decision Support for Procurement Management: Evidence on Whether Computers Can Make Better Procurement Decisions", *Journal of Purchasing and Supply Management*, Volume 9, Number 5, September **2003**, 191-198, http://www.uu.nl/content/iscorepaper212.pdf

[2] Wacek, Michael G., "The Path of the Ultimate Loss Ratio Estimate", *Casualty Actuarial Society Forum,* Winter 2007.

[3] Stanard, James N., "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques", *PCAS* **1985**, Vol. LXXII, 124-148, http://www.casact.org/pubs/proceed/proceed85/85124.pdf.

[4] *Best's Aggregates & Averages (Property-Casualty) – United States,* 1995 Edition, Oldwick (NJ), A. M. Best Company, **1995**, 117, 137, 142, 147.

[5] *Best's Aggregates & Averages (Property-Casualty) – United States,* 1996 Edition, Oldwick (NJ), A. M. Best Company, **1996**, 117, 137, 142, 147.

[6] *Best's Aggregates & Averages (Property-Casualty) – United States,* 1997 Edition, Oldwick (NJ), A. M. Best Company, **1997**, 161, 181, 186, 191.

[7] *Best's Aggregates & Averages (Property-Casualty) – United States,* 1998 Edition, Oldwick (NJ), A. M. Best Company, **1998**, 169, 188, ?, 198.

[8] *Best's Aggregates & Averages (Property-Casualty) – United States,* 1999 Edition, Oldwick (NJ), A. M. Best Company, **1999**, 205, 224, 229, 234.

[9] *Best's Aggregates & Averages (Property-Casualty) – United States,* 2000 Edition, Oldwick (NJ), A. M. Best Company, **2000**, 205, 224, 229, 234.

[10] *Best's Aggregates & Averages (Property-Casualty) – United States,* 2001 Edition, Oldwick (NJ), A. M. Best Company, **2001**, 213, 232, 237, 242.

[11] *Best's Aggregates & Averages (Property-Casualty) – United States,* 2002 Edition, Oldwick (NJ), A. M. Best Company, **2002**, 213, 232, 237, 242.

[12] *Best's Aggregates & Averages (Property-Casualty) – United States,* 2003 Edition–Supplement, Oldwick (NJ), A. M. Best Company, **2003**.

[13] *Best's Aggregates & Averages (Property-Casualty) – United States,* 2004 Edition, Oldwick (NJ), A. M. Best Company, **2004**, 166, 185, 190, 195.

[14] *Best's Aggregates & Averages (Property-Casualty) – United States,* 2005 Edition, Oldwick (NJ), A. M. Best Company, **2005**, 180, 199, 204, 209.

Abbreviations and notations

BF, abbreviation for "Bornhuetter-Ferguson"
CL, abbreviation for "chain ladder"
ELR, expected loss ratio used in Bornhuetter-Ferguson analysis
LDF, loss development factor

Biography of the Author

Michael Wacek is President of Odyssey America Reinsurance Corporation based in Stamford, CT. Over the course of more than 25 years in the industry, including nine years in the London Market, Mike has seen the business from the vantage point of a primary insurer, reinsurance broker and reinsurer. He has a BA from Macalester College (Math, Economics), is a Fellow of the Casualty Actuarial Society and a Member of the American Academy of Actuaries. He has authored a number of papers.