The Casualty Actuarial Society *Forum*
Fall 2006 Edition
Including the 2006 Reserves Call Papers

To CAS Members:

This is the Fall 2006 Edition of the Casualty Actuarial Society *Forum*. It contains 12 Reserves Call Papers and five 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 2006 CAS Reserves Call Papers
Presented at the
2006 CAS Casualty Loss Reserve Seminar
September 11-12, 2006
Renaissance Waverly Hotel
Atlanta, Georgia

The Fall 2006 Edition of the CAS *Forum* is a cooperative effort between the Committee for the CAS *Forum* and the Committee on Reserves.

The CAS Committee on Reserves present for discussion 12 papers prepared in response to their 2006 call for papers.

This *Forum* includes papers that will be discussed by the authors at the 2006 CAS Casualty Loss Reserve Seminar, September 11-12, 2006, in Atlanta, Georgia.

2006 Committee on Reserves

Thomas A. Ryan, *Chairperson*

Jeffrey R. Carlson
Ron Fowler
Aaron M. Halpert
Steven C. Herman
Bertram A. Horowitz
Gloria A. Huberman
Warren H. Johnson Jr.
C.K. Stan Khury
Dale F. Ogden
Christopher Edward Olson
Susan R. Pino
Mark R. Shapland
Wendy W. Tobey

2006 CAS Reserves Call Papers

Additional Papers

Considerations Regarding Standards of Materiality in Estimates of Outstanding Liabilities

Emmanuel T. Bardis, FCAS, MAAA, PhD, MBA
Christina L. Gwilliam, FCAS, MAAA
Stephen P. Lowe, FCAS, MAAA
Atul S. Malhotra, FCAS, MAAA, MBA

Emmanuel T. Bardis
Towers Perrin
111 Huntington Avenue
8th Floor
Boston, MA 02199-7612
(617)638-3807
manolis.bardis@towersperrin.com

Christina L. Gwilliam
Towers Perrin
2107 Wilson Blvd.
Suite 500
Arlington, VA 22201-3062
(703)351-4807
christina.gwilliam@towersperrin.com

Stephen P. Lowe
Towers Perrin
Forestal Centre
175 Powder Forest Drive
Weatogue, CT 06089-9658
(860)843-7057
stephen.lowe@towersperrin.com

Atul S. Malhotra
ACE USA
455 Market Street
Suite 500
San Francisco, CA 94105
(415)547-4587
atul.malhotra@ace-ina.com

Abstract: This paper reports on our research into the issues associated with establishing standards for materiality associated with claim liability estimates. In our research we explored several alternative methods for developing benchmarks for materiality. Rather than restrict ourselves to theoretical considerations, we tested the various methods empirically using public data for individual companies and various lines of business. The empirical test results raise many practical issues that must be considered in such an exercise. This paper is meant to promote discussion on this topic and related issues.

Keywords: reserve variability; uncertainty and ranges; materiality; range of reasonable estimates; range of reasonably probable outcomes; statement of actuarial opinion

Table of Contents

1. EXECUTIVE SUMMARY

As a result of the recent accounting scandals and the stock market boom/bust, there has been an increased desire by shareholders, regulators and rating agencies for transparency in financial statements. Within the non-life / property and casualty insurance sector, the largest liability on an insurer's balance sheet is the loss reserve. There is an increased desire to better understand the uncertainty associated with estimates of unpaid claims underlying the loss reserve. A single point estimate gives no sense of the degree of certainty (or uncertainty) as to the likelihood that actual claim liabilities will ultimately be close to the estimate. Therefore actuaries are increasingly asked to supply a range of reasonably possible outcomes. In the U.S., Appointed Actuaries are required to identify significant risks and uncertainties that could result in material adverse deviation in the loss reserve, and to specify the materiality standard for the specific company. There is little guidance on how to estimate the range of reasonable estimates, or on what this materiality standard should be. This paper seeks to explore ways to measure reserve volatility and to assist the actuary in these areas. In the context of the paper we develop a framework that is designed to answer two distinct questions:

- *By what amount must two estimates of unpaid claim liabilities differ to be considered materially different from each other?*
- *What is the magnitude of the reasonably probable total deviation in actual claim liabilities from the estimate of expected claim liabilities?*

Both of these questions are related to the volatility of the claim generation process characterizing non-life / property and casualty exposures, but they focus on different issues that arise from the uncertainty the volatility creates. Note that materiality in the context of actuarial opinions has a different meaning. For actuarial opinions, materiality is related to an adverse claim liability deviation that would significantly affect the viability of a company. Our use of the term materiality is explained in our Conceptual Framework in Section 2.3.

The first question gives rise to a *Range of Reasonable Estimates,* ideally reflecting uncertainties as to the parameters and model selected to produce estimates of the expected claim liabilities. The second question gives rise to a *Range of Reasonably Probable Outcomes*, incorporating process as well as parameter and model risk. Both ranges depend on standards that must give due consideration to statistical, financial, and solvency perspectives.

In addition to providing a framework for analyzing the two questions posed above, the paper reports on our empirical research, in which we explored several alternative methods for measuring process, parameter, and model risk, and for translating the amount of measured risk into benchmark ranges. The empirical test results raise many practical measurement issues that will require further research to resolve. While the paper presents empirical results for illustration and comparison, the ranges derived are subject to substantive limitations and should therefore not be considered a recommendation.

1.1 Research Approach

We designed our study using the framework of statistical hypothesis testing. We used data from the 2003 Annual Statement of a sample of U.S. insurers for the personal auto liability, homeowners, workers compensation and other liability lines of business. To measure uncertainty in the unpaid claims we used two stochastic methods on individual lines of business: the Bootstrapping methodology of England and Verrall, and the Mack stochastic methodology[1]. The coefficients of variation resulting from our analysis provide a measure of the reserve volatility. As explained in more detail in subsequent sections of this paper, we endeavored to bifurcate total volatility into process and parameter risk. Next, we used two approaches to estimate materiality standards. The two approaches are a percentile/threshold approach and a tail value at risk (TVar) approach. Finally, these monoline results were combined to recognize the risk diversification benefits of multi-line writers. We used a Copula[2] type approach to aggregate the claim liability distributions. We note specifically that we have not used statistical hypothesis testing as our approach. Instead, we use the terminology or the framework associated with hypothesis testing to explain the results of our study for the reader's benefit.

1.2 Results

We derived indicated reserve ranges on two bases: the "range of estimation" basis, which is used to estimate the range of reasonable estimates, and the "range of outcomes" basis,

[1] The Bootstrapping and Mack methods are described in subsequent sections of the text as well as in Appendix A.
[2] Copula theory is described in subsequent sections of the text as well as in Appendix C.

which is used to estimate the range of reasonably probable outcomes. As shown in the table below, the outcome standards are higher than the estimation standards by an average of 75%.

Standards of Materiality – Mack

Line of Business	Range of Estimation: Lower Tail Test	Range of Estimation: Upper Tail Test	Range of Outcome: Lower Tail Test	Range of Outcome: Upper Tail Test
Personal Auto Liability	-5.8%	6.7%	-10.2%	12.2%
Homeowners	-9.7%	11.4%	-17.5%	21.5%
Workers Compensation	-13.6%	16.4%	-20.8%	26.2%
Other Liability	-16.4%	20.2%	-28.0%	37.7%

One reason for the difference between the two types of ranges is that outcome standards include process and parameter risk whereas estimation standards only include parameter risk.

Finally, we created a fictitious company that writes all four lines of business to see the benefit of risk diversification.

Standards of Materiality – Mack

Type	Lower Tail	Upper Tail
Range of Estimation	-12.4%	15.4%
Range of Outcomes	-14.4%	18.1%

1.3 Conclusions and Implications

The major conclusions of our studies were as follows:

- Materiality can have different implications when viewed from a statistical, financial or solvency perspective.
- Materiality standards should clearly be different in a Range of Estimation context than in a Range of Outcomes context.

- Standards of materiality should vary by line of business. Lines of business that historically exhibit higher volatility should have higher standards of materiality (i.e., wider ranges).
- Materiality standards can be arrived at using a framework of statistical hypothesis testing and applying techniques such as percentile/threshold and or TVar.
- Any approach to studying or deriving standards of materiality requires the measure of an appetite for adverse outcomes such as benchmark percentile/threshold of adverse deviation or benchmark exceedence ratio. In terms of the hypothesis testing framework, this relates to one's tolerance level for making a "Type I" or a "Type II" error. Specifically, all else being equal, a wide materiality standard range allows a higher probability of accepting the hypothesis that two reserve estimates are not materially different when in fact they are (i.e., it involves a higher probability of a Type II error). Conversely, a lower materiality standard increases the risk of a Type I error (i.e., concluding that two estimates are statistically different when in fact they are not).
- It is our recommendation that these benchmarks be derived based on combined industry data. Then materiality standards can be derived for individual companies using these benchmarks and their own implied volatility.
- The percentile/threshold and the TVar approaches used in this study yield different standards of materiality applied on the same data as they essentially measure volatility differently. The latter is a more conservative approach.
- Diversification for multi-line writers reduces overall volatility of liabilities compared to mono-line writers, requiring lower levels of surplus, and thus multi-line writers should have lowers standards of statistical and financial materiality compared to mono-line writers.
- The results of our analysis showed that financially impaired companies in general should have narrower standards of materiality compared to financially healthy companies.
- Some of the other conclusions that we reached as a by-product of our extensive use of standard stochastic methodologies are as follows:
- Standard volatility-measuring techniques overstate the volatility of the underlying loss exposure (loss generating process) when used on data without any adjustment for exogenous and endogenous factors impacting the company. For example, these methods

are influenced by trends, changes in case reserving levels, changes in claim settlement rates, and other factors. Adjustments should be carried out to scrub the triangles of these factors before these methodologies can be applied.

- The standard Mack and Bootstrapping stochastic methods usually give different measures (answers) for volatility of the underlying loss data. Our research on industry data showed that the Bootstrapping method has a tendency to overreact to sudden changes in data.
- Both the Mack and Bootstrapping stochastic methodologies give different results for volatility when applied to paid and incurred loss data of the same underlying loss exposure. Both methods apply with more confidence to paid loss development data. The results of these stochastic methods when applied to incurred loss development data, where negative development is prevalent, are not very credible
- The standard stochastic methodologies such as Mack and Bootstrapping do not perform well in differentiating between process and parameter risk. Loss data should be adjusted to a stationary basis in order to achieve a credible differentiation between process and parameter risk.

2. INTRODUCTION

2.1 Background

Actuaries today are being asked by the investment and regulatory communities not only to specify their best estimate of a property and casualty insurer's claim liabilities[3], but also to specify a range of reasonably possible outcomes around their best estimate. These requests in part are being driven by a spate of reserve increases taken by major insurers (particularly those writing U.S. business) in the last few years, which has heightened the issue of "reserve risk." In general there is a movement towards understanding the uncertainty or variability associated with estimates of claim liabilities, as the range around the estimate provides insight as to the solidity of the reserves recorded on the balance sheet (i.e., what percentile within the range of estimates does the carried reserve represent[4]?). Understanding the variability is important to the external stakeholders bearing the risk (shareholders and policyholders), and to the directors of the company who are responsible for managing its risk and capital. A single point estimate gives no sense of the degree of certainty (or lack of certainty) as to the likelihood that the actual claim liabilities will ultimately be close to the estimate.

Additionally, an issue that actuaries and directors of insurance companies often face is how to reconcile differences between alternative estimates of claim liabilities: management's estimate, internal actuarial estimates, and external actuarial estimates. In such instances, directors are faced with the difficult task of choosing a reserve to record based on one of the alternative estimates. How should they make this decision? Are these estimates different enough that one can assume that they are truly differences in opinion, or do they merely reflect differences in methods and assumptions that are within a range of reasonableness?

---

[3] Throughout this paper we refer generically to claim liabilities as being the uncertain amount that will ultimately be paid by the insurer to settle claims arising from insurance coverage that it has provided. The term is meant to be inclusive of defense, adjustment, and other settlement costs in addition to direct payments to the claimant.

[4] In this paper we do not address the issue of how an estimate of liabilities is translated into a reserve on the balance sheet. Generally the literature is vague on this subject, specifying for example that the company should record its "best estimate". While some may interpret this as implying that the reserve should be set equal to the mean estimate, others might interpret it as requiring that the reserve be set at the median, or some other percentile that includes a margin. For purposes of exposition we have therefore assumed that there is a pre-ordained mapping from the selected distribution of claim liability outcomes to an appropriate reserve; the focus of our inquiry is on the selection of the distribution itself.

Given two estimates that are different, are the differences between the two estimates material, and will the booking of reserves based on either of the estimates cause the users of the financial statement to draw different conclusions?

These issues have gained importance lately with changes to the year-end 2005 U.S. opinion process for non-life companies. The Model Law developed by the National Association of Insurance Commissioners, which has been adopted by a few states at this juncture but is expected to be adopted by most states, specifies that opinions should include an Actuarial Opinion Summary that details the opining actuary's own point estimate and range, if one was generated.

2.2 Purpose/Objective of the Paper

This paper is intended to address the following two questions, both of which arise as practical issues in actuarial practice today:

1. *By what amount must two estimates of claim liabilities differ to be considered materially different from each other?* This question often arises in the context of reserve opinions, for example when a reviewing actuary is comparing his or her estimate to management's estimate underlying the held reserve. For sufficiently small differences the conclusion should be that the two estimates are not significantly different. However, at some point the difference between the two estimates becomes sufficiently large that it is significant.

2. *What is the magnitude of the reasonably probable total deviation (adverse or favorable) in actual claim liabilities from the current estimate of expected claim liabilities?* This question arises in the context of solvency, for example when one is stress-testing the balance sheet against the possibility of adverse deviation from the expected level of claim liabilities that would have a significant impact on the company.

Both of these questions are related to the volatility embedded in the claim generation process characterizing non-life / property and casualty exposures, but they focus on different issues that arise from the uncertainty that the volatility creates. In responding to either question, actuaries need benchmark *standards for materiality*, typically expressed as a percentage of the claim liabilities[5], to guide their responses.

This paper reports on our research into the issues associated with establishing standards for materiality associated with claim liability estimates. In our research we explored several alternative methods for developing benchmarks for materiality. Rather than restrict ourselves to theoretical considerations, we tested the various methods empirically using public data for individual companies and various lines of business. The empirical test results raise many practical issues that must be considered in such an exercise.

This paper is meant to promote discussion on this topic and related issues. Our approach is not meant to be definitive, and our empirical results are subject to substantive limitations. The latter are provided for illustration and comparison, and should not be taken as a recommendation. We expect that our approach will continue to evolve with further exploration on the topic.

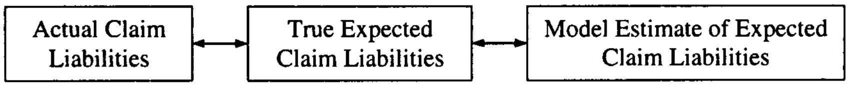
2.3 Conceptual Framework

The historical loss development data the actuary can use to estimate claim liabilities are a relatively small sample of realizations of the claim generation process. The actual claims generated in each accident or underwriting year are the result of (a) randomness, and (b) differences in environmental influences. These influences are both exogenous (the socio-economic conditions at the time) and endogenous (underwriting and claim handling procedures in place at the time). From the available data, the actuary is asked to discern the expected value of the claim liabilities, and the distribution of possible outcomes around that expectation. With imperfect knowledge, the actuary can only provide an *estimate* of the expected value and the underlying distribution, creating a second level of uncertainty above that inherent in the claim generation process itself.

Within a reserving context, actuaries attempt to estimate the true, but unknown, expected claim liabilities by applying an actuarial model to the available historical data. It helps to think about the uncertainty involved in estimating claim liabilities in terms of the following continuum:

---

[5] In certain contexts, materiality standards might also be expressed as a percentage of net income or capital.



The true expected claim liabilities could be considered as the indication from the "perfect" actuarial model where:

— there is no uncertainty associated with the models inputs; and

— all the assumptions employed by the actuarial model are correct.

The potential differences between the actual claim liabilities and the true expected claim liabilities are due to process risk while the potential differences between the true expected claim liabilities and the actuary's model estimate are due to parameter risk and model risk. A detailed description of all the risks associated with the measurement of claim liabilities follows.

- Process risk represents the fundamental uncertainty due to the presence of randomness when losses are generated. Even when an actuary can achieve a "perfect" model, the random nature within which losses are generated would prohibit that actuary from calculating the actual claim liability amount.
- Parameter risk is the uncertainty associated with the unknown parameters of statistical models, even if the selection of the model is correct (i.e., we might know with certainty that the link ratios at a certain maturity follow a log-normal distribution, but we are not sure about the correct parameters associated with that distribution); and
- Model risk is the risk associated with the uncertainty that the loss generating process is not represented correctly by the particular model selected.

Some actuarial literature separates that risk between model risk and specification risk; the former relates to the question if the selected model is correct while the latter relates to the question if the distributions employed by the model are correct). Model risk is the most difficult type of risk to measure since every stochastic model is based on the premise that its fundamental assumptions are correct. Traditional stochastic reserving models, including Mack and Bootstrapping, ignore model risk. One way of approximating model risk is hindcast testing. With hindcast testing a model employs a subset of the historical data to project losses for the remainder of the historical period and compare the actual and projected results. The resulting residuals provide a proxy for model risk.

In the context of uncertain claim liabilities, materiality must be examined from several different perspectives.

- The statistical perspective on materiality reflects the fact that one is estimating the shape and parameters of an unknown claim liability distribution.
- The financial perspective on materiality relates to the question: Would users of the financial statements draw different conclusions if the figures presented were different? This perspective draws on the other elements of the balance sheet, and the income statement.
- The solvency perspective on materiality links the uncertainties associated with the claim liabilities to the capital and claims-paying capacity of the enterprise.

Materiality questions arise most commonly in the context of alternative actuarial estimates, relating to the first question posed at the outset of our paper: Given the uncertainty in the estimation process, is the difference between one actuarial estimate of the claim liabilities and another actuarial estimate significant? In the context of this question we are concerned with the uncertainty of the *expected* liabilities (and not random variations between actual and expected, i.e., process risk); only parameter and model risk are relevant. In other words, the relevant distribution is the distribution of the estimated mean.

> The *Range of Reasonable Estimates* is the range within which alternative estimates of the expected claim liabilities would be deemed to be immaterial, in the sense that (a) the difference between the estimates is not statistically significant, and (b) the difference in the resulting reserves is not financially material. Within this range one could not say that one estimate was actuarially "better" than the other. An actuary reviewing the reserves of a company would accept the reserves if his or her own estimate were within this range.

Materiality can also arise in the context of solvency and risk management, in which one should consider the total risk embedded in the claim liability estimation process, including parameter, model and process risks. In this case we are interested in the actual liability outcomes, so we need to measure all types of risk that could have an adverse effect on a company's surplus.

> The *Range of Reasonably Probable Outcomes* is the range within which the alternative actual claim outcomes are expected to fall with reasonable confidence, in the sense that (a) the outcomes outside of the range which, while possible, have low statistical probability, and (b) for a reasonably well capitalized company, outcomes within the range would not threaten the solvency of the company.

In this paper we focus on materiality standards for the range of reasonable estimates and the range of reasonably probable outcomes. In the first context we refer to the relevant materiality as *estimation materiality*, which ideally will reflect only model and parameter risk. In developing estimation materiality standards we considered only the statistical perspective. We did not consider the financial or solvency perspective; however, as a refinement it might be appropriate to consider the financial perspective.

The latter range relates principally to the financial and capital management (or solvency) perspective on materiality and links the uncertainty associated with the actual claim liability distribution to the finances of the company. All types of risk (model, parameter, process) that could have an adverse effect on the income and capital needs of the company should be measured here. In this context we refer to the relevant materiality standard as *outcome materiality*. When measuring outcome materiality we considered the statistical and solvency perspective, but not the financial perspective.

We note that there is not a clear distinction between the concepts of Range of Reasonable Estimates and Range of Reasonably Probable Outcomes. The underlying precept of our analysis is reserve volatility, which is captured in the definition of Range of Reasonable Estimates. The Range of Reasonably Probable Outcomes is a slightly broader concept in that it tries to incorporate reserve volatility in conjunction with management input and the financial condition of the company (i.e. surplus).

Additionally, in setting materiality standards we did not consider other sources of risk, such as market, credit, operational or insurance underwriting risk.

In summary, for a given set of claim liabilities, the objective is to develop:

a) an appropriate standard for a *range of reasonable estimates*, reflecting appropriate criteria for *estimation materiality*; and

b) an appropriate standard for a *range of reasonably probable outcomes*, reflecting appropriate criteria for *outcome materiality*.

To develop these ranges, it is necessary to estimate the claim liability distribution and to separate process from parameter and model risk. As we discuss later in the Methodology section, the claim liability distributions in this paper are estimated with stochastic reserving methods, which provide distributions for both the actual claim liabilities and the estimate of the expected claim liabilities.

Once appropriate claim liability distributions have been produced, the two ranges embodying our materiality standards can be obtained from them. In the case of each distribution, this requires the selection or derivation of a threshold [5]. The threshold can be based either on a specified percentile of the distribution (generally, a VaR approach), or on a specified expected exceedence value (generally, a TVaR approach)

The percentile threshold approach is a point measure in the sense that it measures the probability of an outcome being worse than a given monetary threshold (e.g., probability of ruin). While the percentile threshold approach measures the probability that a particular value will be exceeded say once every 100 years, the expected "exceedence" threshold approach measures the expected value of the exceeded amount (every 100 years) when the threshold is exceeded. The expected exceedence threshold approach provides values higher than the percentile threshold approach, as it is influenced by the outcomes of remote loss outcomes. In the chart shown below the percentile threshold approach focuses on finding the shaded region, whereas the expected exceedence threshold approach focuses on estimating the expected value of losses exceeding the threshold, as a percentage of expected liabilities. Essentially, these two paradigms measure "tail" risk differently.

We formulate the problem of analyzing estimation materiality in the framework of statistical hypothesis testing. Although we do not actually perform hypothesis testing, this framework has the advantage of helping to explain the variables required to calculate materiality and analyze the results obtained from our analysis. The only divergence between a true statistical hypothesis testing and the methodology employed in this paper is that, while statistical hypothesis testing compares the distributions of two estimates of the mean, in this paper we compare the distribution of expected claim liabilities to an alternative point estimate of the mean that is considered to be certain. In that respect our employed approach resembles the measurement of a statistical confidence level.

Consider a distribution of expected claim liabilities where:

$C\alpha$ = the mean of the distribution

m_1 = the upper bound of the range of reasonable estimates

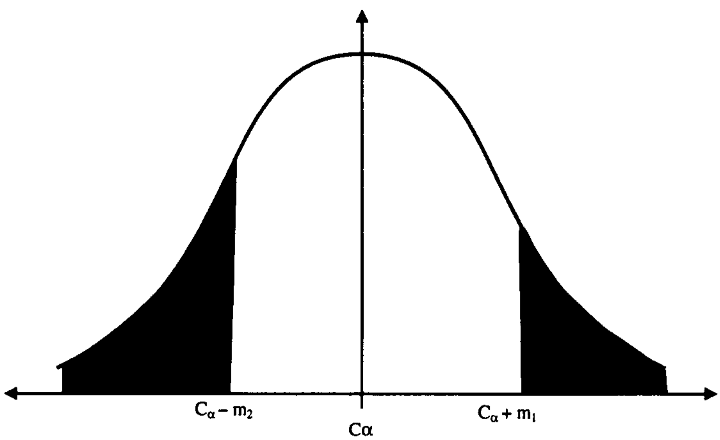
m_2 = the lower bound of the range of reasonable estimates

We can set up the problem in this framework as follows:

H_0 (Null Hypothesis): The two estimates of the expected claim liabilities are the same

H_1 (Alternate Hypothesis): The two estimates are not the same

A formulation of the problem pictorially is as follows:



Reserves in the range $C_\alpha - m_2$ to $C_\alpha + m_1$ are not considered significantly different

The Type I error in statistical hypothesis testing measures the probability of rejecting the null hypothesis when the null hypothesis is true. Typically, m1 and m2, defining the range of reasonable estimates, are determined by selecting a significance level, reflecting an acceptably low probability of a Type I error. The significance level is measured by "r" (in our paper), shown by the shaded region in the chart above. Note that as the stringency of the significance level is tightened, the range of reasonable estimates expands.

If the alternative estimate of the expected claim liabilities falls outside of the range of reasonable estimates (in other words if the alternative estimate amount falls in the shaded region in the chart above) then we can reject the null hypothesis that the original estimate underlying the reserve and the alternative estimate are essentially the same.

The formulation for analyzing outcome materiality follows a more traditional confidence level construct. However, the picture is essentially the same as that shown above. We seek to define a range of reasonably probable outcomes, such that the likelihood of actual claim liabilities being outside of that range is reasonably small. However, rather than defining the range purely from a statistical perspective, we define it with reference to a solvency perspective as well. The benchmark level of outcome materiality is based on an empirical analysis of the typical relationship of reserves to risk-based capital, and the level of adverse deviation that would cause the insurer to "ruin" by failing the risk-based capital adequacy test.

In the context of outcome materiality, a higher probability of ruin corresponds to a smaller range of reasonably probable outcomes.

For both types of ranges, we develop empirical measures of m_1 and m_2 in this paper. They may be interpreted as an explicit function of three primary variables amongst others:

$m = f(\sigma, r, \alpha)$ where:

> σ is the implied volatility of the claim liabilities for line of business under consideration, or the uncertainty of the estimated mean;
>
> r is the selected threshold. The corresponding factor in statistical hypothesis testing is the probability of Type I error; and
>
> α is the implied percentile of the carried reserves in relationship to the expected claim liabilities.

m (defining the upper or lower bound of the range) is directly proportional to σ. A more volatile book of business will require a larger allocation of surplus and thus will have a higher m. In other words, the more volatile a book of business, the greater the uncertainty associated with the claim liability estimates. As a result, the corresponding m should be greater for this line to consider the greater uncertainty of the loss process. In Step 1 of the Methodology section we outline how we calculated the implied volatility of each line of business.

α, the implied percentile of the carried reserves, is another important factor. If the carried reserves are booked at a higher percentile of the claim liability distribution then a lower standard of outcome materiality is acceptable.

As r increases m should decrease, a higher r (i.e., a larger shaded area) reflects a higher level of conservatism. A higher r also implies a higher probability of ruin (i.e., it is easier for actual claim liabilities to fall in the shaded region). A higher r also implies that it is easier to conclude that the alternative estimate of claim liabilities is different from the original estimate underlying the reserve.

Other factors that should be considered in selecting the thresholds that define the materiality standards may be the following:

- type of exposures involved
- primary / reinsurance limits
- size of reserves / expected loss / no. of exposures or claims
- average age of reserves
- expectation of parameter risk associated with the particular LOB
- probable maximum loss
- asset variability
- net income variability

In our study, we have not specifically analyzed the impact of these issues in the calculation of the standards of materiality. Generally, the impact of the above factors on the standards will depend on whether they add to or decrease the volatility of the claim liabilities, or increase or decrease the uncertainty associated with the financial and solvency status of the company.

Both of the stochastic reserving models employed in our analysis measure process and parameter risk but neither of them measures explicitly model risk. Further research is needed in the area of the measurement of model risk.

2.4 Acknowledgements

The authors would like to thank Rick Shaw, Amy Bouska, Tom Ghezzi, Jeanne Hollister, Ron Kozlowski, Kahshin Leow and Ollie Sherman for their thoughtful insights on the topic.

3. METHODOLOGY

3.1 Overview

The overall approach is as follows:

Step 1 - Obtain sample balance sheet and historical claim development data for selected companies and lines of business.

Step 2 - Apply stochastic methods to the historical claim development data to measure the distribution of the actual claim liabilities, and the distribution of the estimated expected claim liabilities for each company and each line of business.

Step 3 - Select estimation and outcome thresholds. For outcome materiality thresholds, base selections on typical balance sheet solvency impacts for selected companies.

Step 4 - Develop ranges embodying the materiality standards, based on both percentile thresholds and expected exceedence ratio thresholds.

Step 5 - Recognize risk diversification benefits among multiple lines by incorporating correlation and aggregating the individual line of business distributions to build an aggregate distribution to arrive at ranges embodying the overall materiality standards at a legal entity level.

The following sections will elaborate on each step.

3.2 Step 1 – Data and Data Limitations

U.S. insurers are required to file Annual Statements with state regulatory authorities. The required format includes income statements, balance sheets, cash flows and schedules focusing on aspects such as historical claim development (Schedule P), reinsurance recoverables (Schedule F) and investment (Schedule D). As noted previously, a Statement of Actuarial Opinion must accompany each Annual Statement. Annual Statements and Statements of Actuarial Opinion are in the public domain and can be viewed at each state's

Department of Insurance. We used an internal Annual Statement database, based on data obtained annually from A.M. Best.

We used data from various sections of the Annual Statement. The claim liability development triangles and premiums were obtained from the Schedule P for each company. Measurement of capital came from the "Five-Year Historical Data" exhibit.

We analyzed four lines of business: Personal Auto Liability, Homeowners, Workers Compensation and Other Liability-Occurrence policy forms. These lines were selected to reflect the spectrum from short-tail to long-tail, and the spread in volatility.

Within the U.S. non-life insurance sector, it is common for an insurer to operate through multiple legal entities under common management, often referred to as a group. Multiple entities within a group offer flexibility in terms of capitalization, pricing and regulatory domain. An insurer must file an Annual Statement and a Statement of Actuarial Opinion for each legal entity. Therefore our analysis is done at the legal entity, not group level.

There are often inter-company pooling arrangements whereby an insurer allocates results to entities which may or may not have written the business. The pooling percentages may vary by line and year. The pooling applies to each aspect of the Annual Statement, including the Schedule P data triangles we use, that is, the analyzed triangles may represent a percentage of a larger triangle. Therefore, when we consider the relative size of the sample entities, we need to adjust for pooling. All figures presented are adjusted to reflect the effect of pooling.

We included insurers that cover the spectrum from small single-state or regional to large national companies. For our purposes, we define size in relation to the premium earned from 1994-2003 for each line. Companies with premium below $3 billion are considered small for that line, companies above $10 billion are defined as big, and the rest are medium. For example, a large national writer such as Hartford Financial includes a legal entity, Harford Fire Insurance Company, which we consider small, medium and large depending on the line of business under consideration. (See the table below; figures in $000's)

Line of Business	Net Earned Premium for Line	Size Categorization for Line
Workers Compensation	$ 11,818,872	Big
Personal Auto Liability	$ 10,463,646	Big
Homeowners	$ 4,913,186	Medium
Other Liability	$ 2,914,352	Small

The segmentation by size is intended to group companies with expected similar reserve volatility. All things being equal, we expect claim liabilities on larger volumes of business to be less volatile than smaller books of business.

We took the data "as is," meaning that extensive cleansing of the data was not undertaken. In several instances we adjusted anomalous data, with care not to sanitize the data. Even with these adjustments, some of the data appears to be implausible; companies with implausible data were excluded from our analysis.

While we restricted our analysis to publicly available Annual Statement data, it should be noted that insurers have additional information available internally. Companies often segment their business into more homogenous groups than Annual Statement line of business. The concepts applied here on a line of business basis are illustrative and can also be utilized for different segmentation.

3.3 Step 2 – Use stochastic methods to measure volatility of unpaid claim liabilities

We used the Bootstrapping methodology as described by England and Verall [6] and the Mack Stochastic methodology [12] to estimate the volatility of claim liabilities. For a brief description of these methods please refer to Appendix A. The CV (Coefficient of Variation) is our chosen measure of volatility. The Mack method generates the first two moments of the claim liability distribution, the mean and the standard deviation, while the Bootstrapping method produces an empirical distribution of claim liabilities, so CVs are easily calculated in both cases. The input historical claim development triangles used for both methods are paid loss development triangles (including only allocated loss adjustment expense). In addition, we augmented the Bootstrapping method described in the paper to recognize development beyond the maturity of the triangle (i.e., in the tail). We have assumed a tail that extends to

10-12 years for Homeowners and Personal Auto Liability, 15-20 years for Occurrence Liability and 40-50 years for Workers Compensation. The tail was estimated by fitting an inverse power curve to the development factors for ages of 48 months and beyond, based on Richard E. Sherman's [14] approach as outlined in "Extrapolating, Smoothing, and Interpolating Development Factors." A uniform tail was selected to apply to all accident years within a company. Additionally, for the Mack method we selected standard errors associated with the tail volatility. The selection was essentially based on the empirical results of the Bootstrapping method. We compared the CVs produced by the Bootstrapping method for each line of business in our sample database with the inclusion of a tail factor and exclusion of the tail factor. The difference in the CVs including and excluding the tail factor was then selected as a measure of the standard error associated with the tail factor.

As stated above, the Bootstrapping method provides more than just the mean and variance of the claim liability distribution; it generates the entire distribution. In almost all cases the mean of the distribution generated from the Mack and Bootstrapping methods was different from the carried reserve amount, therefore we performed a linear transformation to force the mean of the distribution to be equal to the carried reserves, while preserving the CV of the distribution. When we describe the "percentile/carried reserve," we are assuming the carried reserve is the best estimate. This is an assumption, not an assertion. Readers are directed to "Management's Best Estimates of Loss Reserves" [10] by Rodney Kreps that notes the mean of the distribution is "probably not a good estimate, as it is almost surely low."

We note that the use of paid claim development data in our analysis is essentially dictated by the inherent limitation of the Bootstrapping and Mack stochastic reserving methodologies used in our analysis. These methodologies do not respond well to reported loss (case reserves + paid losses) data. Indeed both methods produce unreasonable results when used on reported loss triangles which occasionally have age-to-age loss development factors below 1.0 followed by positive development (age-to-age development factors above 1.0). Both of these methods require a somewhat smooth progression of age-to-age loss development factors from immature to mature valuations, declining from high loss development factors for immature data to low development factors for mature data.

Another limitation is that both methodologies assume an essentially stationary process, i.e., that there are no endogenous and exogenous influences on the loss generating process such as company-specific changes in operations, claim settlement rates, premium/exposure growth or changes, large settlements, evolving interpretations of liabilities in the court system, hurricanes, and so forth. Realistically, the loss development data reflected in the triangles of a company are hardly ever stationary, as they include both exogenous and endogenous influences, which cause additional volatility in the loss development triangle. As a result, stochastic reserving methodologies that rely on the volatility inherent in the loss triangle almost always overstate the volatility of the underlying loss generating process. In order to adjust for this distortion, we adjusted the volatility estimates arrived from the use of these stochastic reserving methodologies downward. The adjustment factors were calculated using industry-wide paid loss triangles from A.M. Best (27 to 30 company composite, depending on the line of business), adjusting the triangles for industry exposure changes and frequency trend and other exogenous influences. The frequency trend is applied to adjust for observed declines in claim frequency due to safer workplaces, safer cars, and so forth over the years. We then postulate that this process should create stationary triangles, absent of any exogenous and endogenous factors described above. Thus the volatility present in these stationary triangles will be the true volatility produced by the loss generating process. We measured this volatility in the industry-wide triangles, by line of business, and used it to adjust downwards our overall volatility results produced by the stochastic methods employed in this paper. While we believe these adjustments are reasonable, refinements in the techniques used to better achieve the desired stationarity can certainly improve upon them.

We also independently tested the assumption that the process risk implied by the stochastic methods employed in our analysis is overstated, by comparing the claim volatility calculated by the stochastic methods to the claim volatility obtained via hindcast testing. We employed an independent historical data set for 20 companies and measured the performance of deterministic reserving techniques as they tried to estimate the claim liabilities for these companies. We first estimated the claim liabilities using information that was available at a given point in time and then looked at the available run-off information to see what the actual claim liabilities amounts were with hindsight. The observed estimation error over time and across all companies provides a proxy for the total risk associated with the evaluation of claim liabilities. For the workers compensation line, the hindcast tests results indicate a CV of total risk equal to 8.1%. By comparison, we obtained a parameter-only risk CV of 11.0% from the Mack method. Most of the companies in the hind-cast testing were rather large, with reserves in excess of $100 million, so the associated process risk for these companies should be quite low. The fact that the total risk CV from the hind-cast testing is lower when compared to the parameter risk CV from the Mack method suggests that the process risk and parameter risk implied by our stochastic methods could be overstated. The inability of traditional stochastic reserving methods to separate the variability due to changes in endogenous and exogenous influences, from the true claim volatility due to the claim generating process, is the main reason for this presumed overstatement of process and parameter risk.

3.4 Step 3 – Select Significance Threshold Levels

For outcome materiality, we calculated threshold significance levels for both financially healthy companies and financially impaired ones. Financially impaired companies should get an earlier warning flag when something is wrong with their reserves compared to financially healthy companies, since the underlying assumption is that an adverse claim liability deviation causes a greater financial "hurt" to financially impaired companies.

We employed the "bright line test," which we understand is utilized by the NAIC, in the measurement of outcome benchmark significance levels. The bright line test measures the difference between the surplus as regards to the policyholders and the RBC (Risk Based Capital) capital amount, proposed by the NAIC, that would downgrade the company to the next lower RBC level. If the claim liabilities of a company sustain an adverse deviation greater or equal to the capital level difference mentioned above, that company would be downgraded to the next lower RBC level. That capital level difference to the next lower RBC level, given a distribution around carried reserves, provides a maximum standard of materiality for the company (i.e., the officers of that company would, at least, want to know under what adverse claim liability deviation the company would be downgraded to the next lower RBC level). They might want though to set up an earlier warning flag, based on their experience with the company's financial results, so the adverse deviation from the bright line test can serve, at least, as a maximum standard of materiality.

An assumption in the above analysis is that these companies did not experience significant changes in their distribution of exposures, by line, during the historical period of the analysis. The implied volatility from the claim liabilities for each company was calculated on an all lines combined basis considering both process and parameter risk. An outcome materiality significance level threshold was calculated for an upper tail test within the percentile threshold context and an exceedence ratio threshold was calculated within the

TVar context, given a level of volatility associated with the carried reserve. The outcome materiality significance level threshold for a lower tail test for the percentile threshold approach was calculated judgmentally based on the assumption that the magnitude of the standards of materiality should be higher for an upper tail test when compared to a lower tail test.

For the corresponding estimation materiality significance level threshold we employed a 7.5% rule of thumb benchmark for the upper tail test. That 7.5% represents the average of the 5% to 10% significance levels usually employed in statistical hypothesis testing. Interestingly, we tested the validity of this assumption (7.5%) by estimating the benchmark significance level threshold using parameter risk only from the outputs of the Mack method and found that the resulting estimation materiality benchmark significance level threshold was, on average, similar to the 7.5% that we assumed.

The resulting thresholds for financially healthy companies were as follows:

	Percentile Threshold Benchmark Significance Levels		**Tail Value at Risk Benchmark Exceedence Ratio**	
	Lower Tail Probability	Upper Tail Probability	Lower Tail Expected Excess	Upper Tail Expected Excess
Estimation materiality	10.0%	7.5%	n/a	2.0%
Outcome materiality	8.0%	6.0%	n/a	1.5%

All other things being equal, the resulting outcome materiality standards are higher from the corresponding estimation materiality standards, a logical relationship when considering the higher amount of risk associated with outcome materiality standards. For the majority of the healthy companies the resulting outcome materiality benchmark significance level, for the upper tail test, was 0.0%. This result highlights the fact that most of the healthy companies are so well capitalized that they need to suffer an adverse claim liability deviation in excess of the 99.9% percentile of their claim liability distribution in order to get downgraded into the next lower RBC level.

We also performed the above analysis on a group of 16 companies that were financially impaired. These companies were either in rehabilitation or liquidation. For arriving at the outcome materiality benchmark significance levels for the financially impaired companies we used the bright line test as well as one year adverse development from Schedule P data. As expected, we estimated much higher benchmark significance levels and benchmark exceedence ratios compared to the healthy company figures mentioned above. The outcome materiality benchmark significance levels for adverse deviation for the financially impaired companies were 18% compared to 6% for financially healthy companies.

Specifically, we performed the following steps to come up with the outcome materiality benchmark significance levels and benchmark exceedence ratios.

We employed 39 financially healthy companies from our A.M. Best data base. The calculation of an upper tail test outcome materiality benchmark significance level threshold, for the Percentile Threshold approach, followed the steps outlined below:

1) For each of these companies we measured the total risk, for all lines combined, associated with their claim liability distribution. That claim liability distribution was calculated from loss and ALAE Schedule P Part 3 triangle data using the Mack stochastic reserving method. We further assumed that the mean C_α of the stochastic distributions is equal to the carried reserves for the companies.

2) From the Bright Line Test we calculated the adverse claim liability deviation that would downgrade each company into the next lower RBC level. That adverse claim liability deviation m represents a maximum standard of materiality.

3) We then added the mean of the distribution to the adverse claim liability deviation. The area under the claim liability distribution in excess of $C_\alpha + m$ represents the upper tail outcome materiality benchmark significance level. It measures the probability of extreme claim liability outcomes that a company must experience before it gets downgraded into the next lower RBC level.

For the calculation of an outcome materiality exceedence ratio for the TVar approach the first two steps outlined above were identically repeated. As a last step we measured the average of claim liability outcomes that exceed the carried reserves by the standard of materiality. These excess losses were calculated as a ratio to the expected claim liabilities, producing the exceedence ratio threshold.

3.5 Step 4 – Estimate Materiality Standards for each Individual Line

Based on the selected outcome materiality benchmark significance levels and exceedence ratios we then calculated the outcome materiality standards for each company in our sample database. The calculation proceeds as follows:

1) For each company triangle we generate a claim liability distribution using both the Bootstrapping and the Mack method.
2) We normalize each loss reserve distribution so that the mean of the distribution is equal to the carried reserve of the company.
3) The outcome materiality standard is equal to the difference between the percentile implied by the outcome materiality benchmark significance level, as described above, and the percentile implied by the carried reserve.
4) The outcome materiality standard implied by the TVar approach is calculated as the difference between the percentile implied by the benchmark exceedence ratio and the percentile implied by the carried reserve.
5) The estimation materiality standards are calculated in a similar fashion using the estimation materiality benchmark significance levels and exceedence ratios.

3.6 Step 5 – Recognize Risk Diversification Benefits Among Multiple Lines

Few companies are monoline writers. For multi-line writers, the standards of materiality should incorporate the risk diversification associated with underwriting more than one lines of business. Aggregate claim liability distributions can be calculated from the individual line distributions. In our analysis we incorporate a Copula type of approach that performs the aggregation procedure. More information regarding the Copula approach is included in Appendix C.

The Mack and Bootstrapping stochastic reserving methods mentioned above measure the claim volatility for an individual line of business. In case where more than one lines of business are considered we need a model that aggregates the individual lines distributions. The mean of the aggregate distribution is the sum of the individual lines means. However we cannot arrive at the percentiles of the aggregate distribution by simply adding the individual lines percentiles. Straight summation makes sense only in the case of 100% correlation across all lines, a highly unlikely situation. The volatility of the aggregate distribution is influenced by two factors:

- The claim volatility for each individual line of business: The larger the claim volatility for each individual line, the larger the volatility of the aggregate distribution, all other things being equal; and
- The correlations across lines: The larger the correlation among individual lines, the larger the volatility of the aggregate distribution, all other things been equal.

Statisticians have shown that the aggregate distribution of any combination of n random variables can be written as a function of the n individual variables distributions (*Sklar theorem 1996*). This function is called Copula. We are employing one Copula model in our analysis that provides a convenient way of calculating the aggregate distribution of several lines of business. Two components are needed for the Copula model:

- The distributions of the individual lines of business; and
- The correlation coefficients among these lines.

The Copula model employed in our analysis is the Normal Copula. For the Normal Copula a correlation matrix based on the assumed correlations among the various lines must be selected. The correlation matrix for the Normal Copula should be positive-definite (i.e., invertible) for the Copula to work.

The selected correlation among the various lines is based on modeling of economic variables such as general/price inflation, wage inflation, auto inflation, and medical inflation. This is done by first building forecasting models for auto inflation and medical inflation as a function of general/price inflation. The models have an autoregressive component in that the inflationary component being modeled reverts back to it long term mean. Next we modeled the impact of each of these inflationary components on each line of business. The model used was a geometric model employed by Robert P. Butsic [3] to model the impact of different inflationary components on losses of different lines of business including social inflation. Once the impact of these inflationary components on each line of business is ascertained then we can construct a distribution of losses for each line of business by forecasting these economic variables. The correlation matrix is then estimated by empirically measuring the correlation between the simulated losses for each line of business. Our assumed correlation matrix is included in Appendix C. The advantage of such models is that correlation between the claims experience is an emergent property.

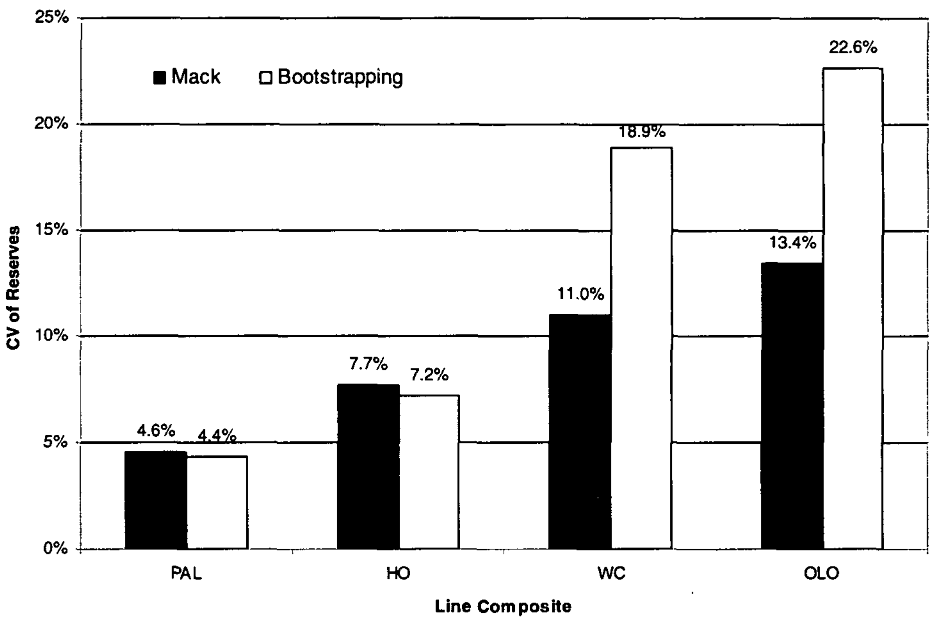
4. RESULTS and CONCLUSIONS

4.1 Reserve Volatility

As can be seen in Exhibit 4.1.1, the volatility is relatively small for Personal Auto Liability (PAL), somewhat larger for Homeowners (HO), larger still for Workers Compensation (WC), and even larger for Other Liability Occurrence (OLO.) The relative magnitude is as expected. The HO line is impacted by catastrophes and the HO claim liabilities are more volatile when compared to the OLO liabilities. OLO is impacted by some high severity claims so intuitively is more volatile. WC also has high severity claims but there is enough frequency/consistency that overall it is less volatile than OLO.

Exhibit 4.1.1

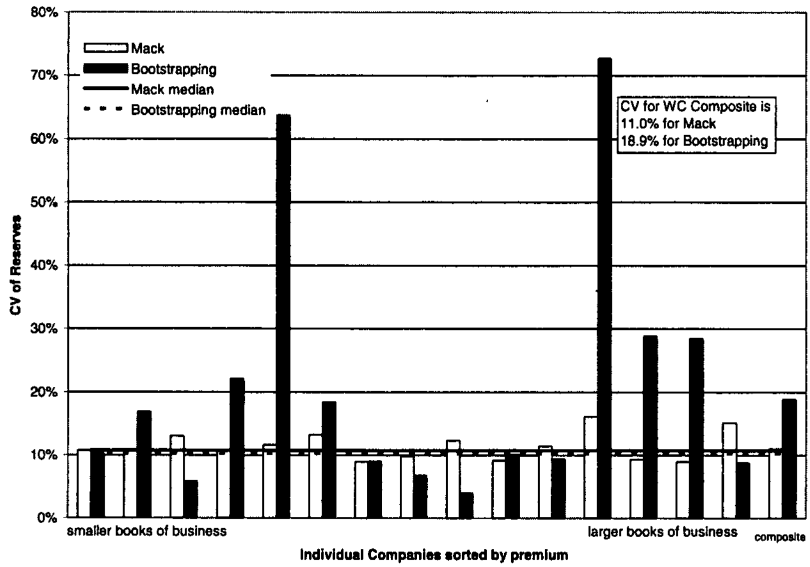
Comparison of Parameter Risk CVs from Mack and Bootstrapping Methods



The Bootstrapping method is more sensitive to outlier development factors and so it generates significantly larger CVs for some companies, as displayed in Exhibit 4.1.2.

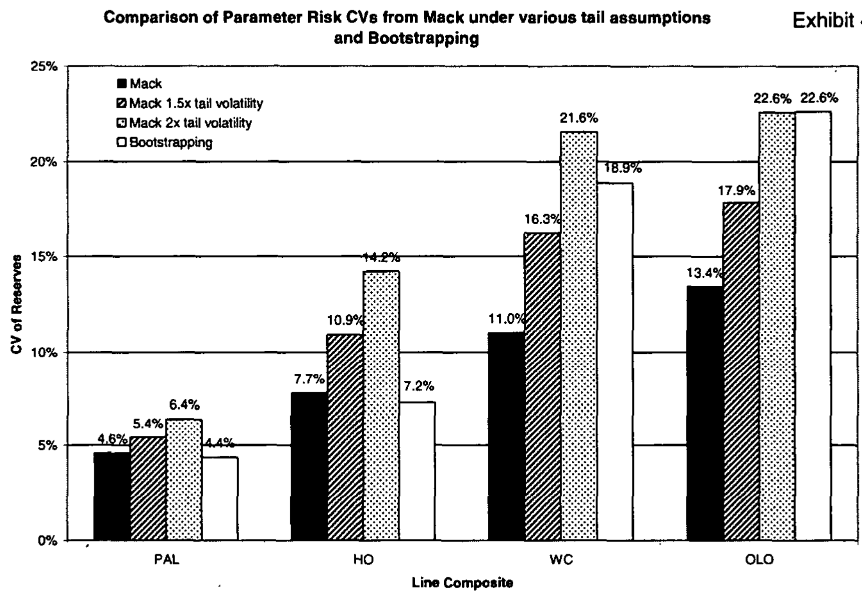
Exhibit 4.1.2

Comparison of Parameter Risk CVs from Mack and Bootstrapping Methods
Workers Compensation



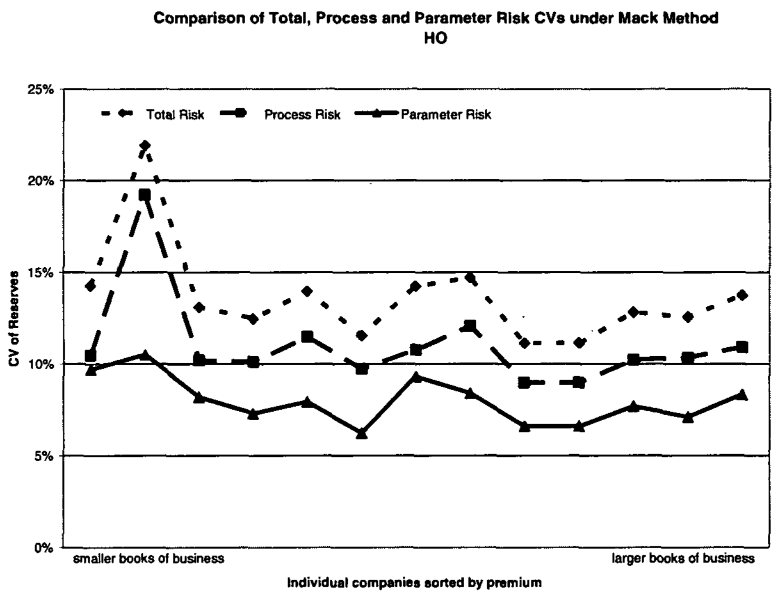
As discussed in Section 3.3, the Mack method is dependent on the assumed volatility in the tail. We tested our tail volatility assumption to determine how sensitive the analysis is to our supposition. We increased the volatility in the tail by 50% and 100%. For Workers Compensation and Other Liability Occurrence, the increased tail volatility drove the Mack CVs closer to the Bootstrapping CVs. On the other hand, the adjustment created a difference for Personal Auto and Homeowners, whereas the CVs were quite similar before increasing the tail volatility. These results are displayed in Exhibit 4.1.3 (see next page)

Exhibit 4.1.3



As shown in the Exhibit 4.1.4, total, parameter and process risk all generally follow the same

Exhibit 4.1.4

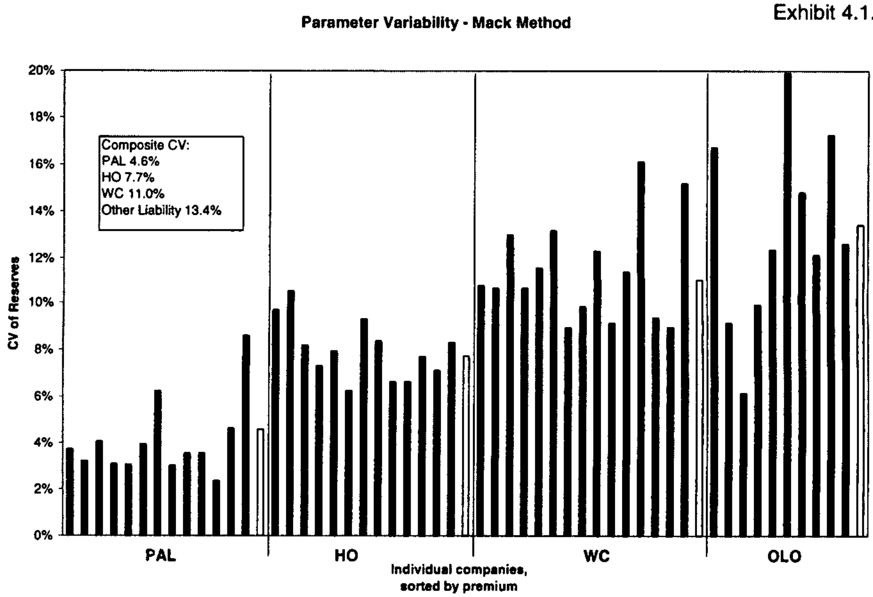


relationship: process is usually larger than parameter risk, and naturally, total risk is the largest.

We expected that parameter risk is invariant of size, while process risk should decrease by the size of the company. Our analysis calculates process risk that is independent of size. That result suggests that the stochastic methods employed in our analysis could possibly overstate process risk.

We summarize the results by size, expecting larger books of business to be less volatile, however this was not the case. For Personal Auto Liability, the most volatile companies were generally the larger ones. There was not much variation in the size of selected Workers Compensation companies with two-thirds of them categorized as small. The results were mixed with high CVs coming from both small and large companies. Each bar in Exhibit 4.1.5 represents a company and is sorted from by premium volume, with smaller companies on the left.

Exhibit 4.1.5



An analysis by size of company that employs a larger sample of companies is probably needed to draw more credible conclusions in the comparison of claim volatility by size of company.

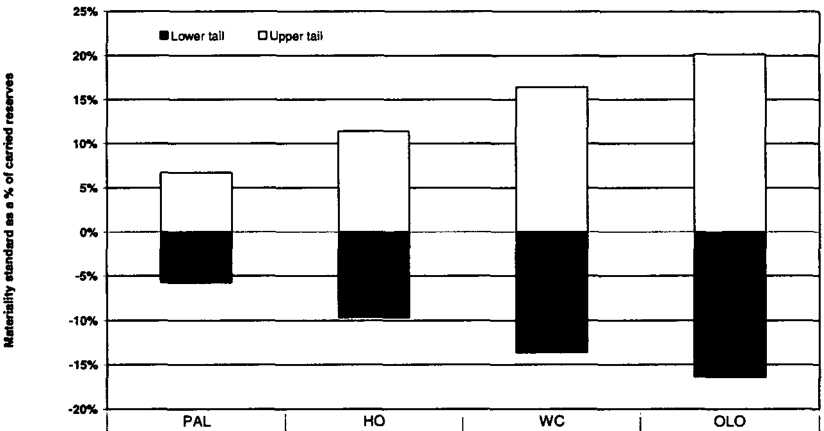
4.2 Materiality Standards

The following graphic summarizes the estimation materiality standards, based on the Mack method, for the four lines of business under consideration. All the standards shown in the remainder of the paper were calculated, unless otherwise noted, as a percentage of carried reserves and using the Percentile Threshold approach.

The resulting estimation materiality standards are higher than what actuaries are accustomed to, partly because these techniques overstate volatility unless adjustments are made for exogenous and endogenous factors.

Exhibit 4.2.1

Estimation Materiality Standards based on the Mack Method



	PAL	HO	WC	OLO
□ Upper tail	6.7%	11.4%	16.4%	20.2%
■ Lower tail	-5.8%	-9.7%	-13.6%	-16.4%

For Personal Auto the standards are close to the +/-5% range of expected that is often employed, as a rule of thumb. For the remainder of the lines the resulting standards are much higher. The calculated standards of materiality could be overstated due to the suspected overstatement of the process and parameter risk produced by the Mack method. Exhibit 4.2.1 (see previous page) graphs the upper and lower tail estimation materiality standards by company and by line. For the more volatile lines the wider range of the materiality standards is evident.

The following table compares the estimation materiality standards produced by the Mack and Bootstrapping models. The actual upper and lower tail estimation materiality standards are calculated as follows:

Estimation Standards of Materiality – Bootstrapping vs. Mack

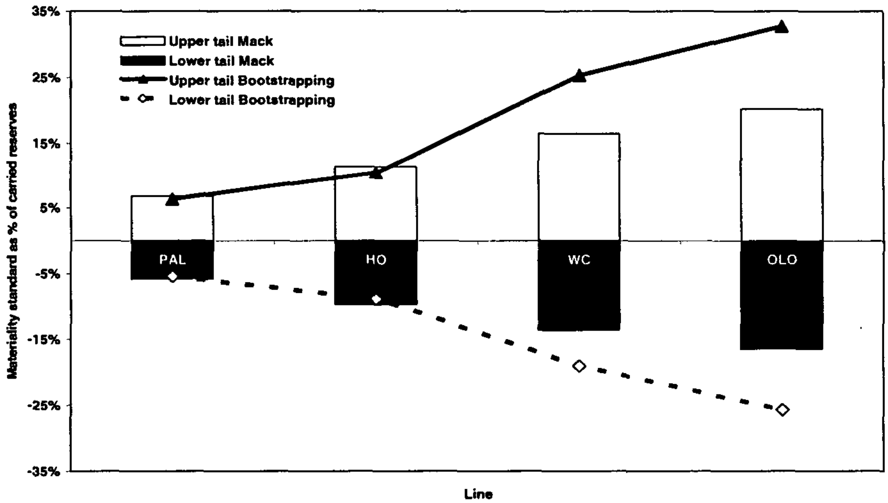
	Mack		**Bootstrapping**	
Line of Business	**Lower Tail Test**	**Upper Tail Test**	**Lower Tail Test**	**Upper Tail Test**
Personal Auto Liability	-5.8%	6.7%	-5.4%	6.3%
Homeowners	-9.7%	11.4%	-8.8%	10.5%
Workers Compensation	-13.6%	16.4%	-19.0%	25.3%
Other Liability	-16.4%	20.2%	-25.7%	32.7%

The resulting estimation materiality standards between the two methods are relatively close for the Personal Auto and Homeowners lines of business. For the two long-tail lines, workers compensation and other liability, the Bootstrapping statistical standards are 40% to 60% higher when compared to the Mack standards.

The additional tail volatility implied by the Bootstrapping method produces the higher estimation materiality standards. Exhibit 4.2.2 compares the estimation materiality standards for the two stochastic methods employed in the analysis.

Exhibit 4.2.2

Comparison of Estimation Materiality Standard - Mack vs. Bootstrapping



The comparison of estimation and outcome standards of materiality is summarized in the following table, for the Mack method, for both lower and upper tail tests:

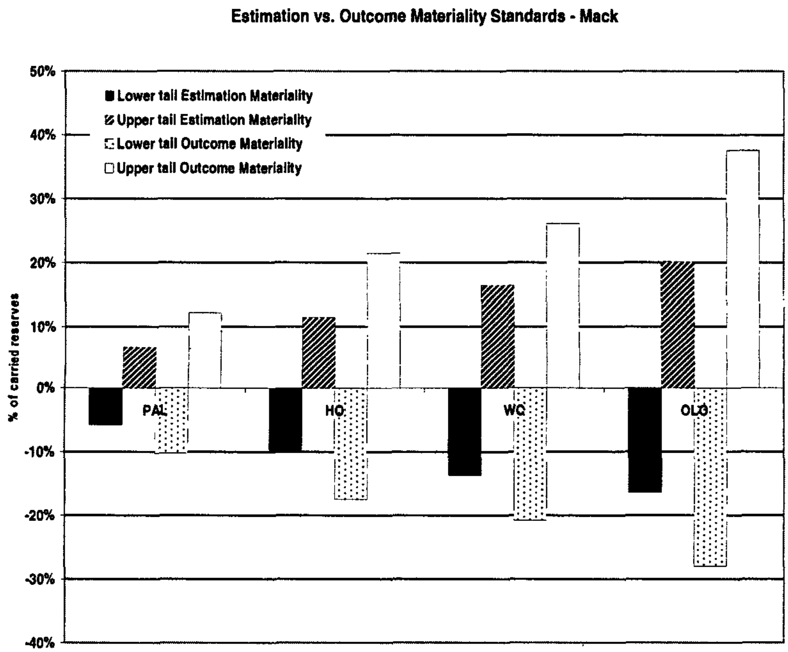
Estimation vs. Outcome Materiality Standards– Mack

	Estimation standards		Outcome standards	
Line of Business	Lower Tail Test	Upper Tail Test	Lower Tail Test	Upper Tail Test
Personal Auto Liability	-5.8%	6.7%	-10.2%	12.2%
Homeowners	-9.7%	11.4%	-17.5%	21.5%
Workers Compensation	-13.6%	16.4%	-20.8%	26.2%
Other Liability	-16.4%	20.2%	-28.0%	37.7%

The outcome materiality standards are, on average, 75% higher when compared to the estimation materiality standards. There are two reasons that explain this relationship: (1)

outcome materiality standards employ the calculation of both process and parameter risk while estimation materiality standards employ parameter risk only. The inclusion of process risk increases the outcome materiality standards; and (2) the benchmark significance level is higher for the estimation materiality standards when compared to the benchmark significance level for the outcome materiality standards. All other things been equal, the resulting outcome materiality standards should be higher since the corresponding probability of Type I error is lower. Exhibit 4.2.3 provides a comparison of the outcome and estimation materiality standards.

Exhibit 4.2.3



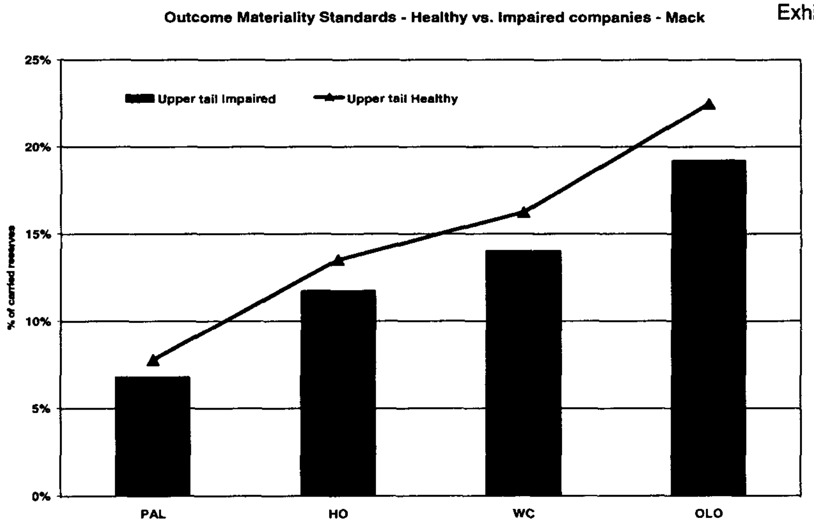
The following table compares the upper tail outcome materiality standards for financially healthy and financially impaired companies.

Outcome Materiality Standards– Upper Tail
Healthy vs. Liquidated Companies – Mack

Line of Business	Financially Healthy	Financially Impaired
Personal Auto Liability	12.2%	6.8%
Homeowners	21.5%	11.7%
Workers Compensation	26.2%	14.0%
Other Liability	37.7%	19.2%

The outcome materiality standards are much higher for the financially healthy companies when compared to the corresponding standards for the financially impaired companies. For a financially impaired company, a lower outcome materiality standard is reasonable since it provides an earlier warning flag if an adverse claim liability deviation is experienced by that company. The lower standards compensate for the greater reserve uncertainty associated with the reserves of a financially impaired company coupled by lower reserve to surplus

Exhibit 4.2.4



ratios. Moreover, our selected significance level benchmarks of 18% for financially impaired companies vs. 6% for financially healthy ones, allows for a greater probability of Type I error for the financially impaired companies, decreasing in effect their respective outcome materiality standards. Exhibit 4.2.4 (see previous page) compares the upper tail outcome materiality standards for financially healthy and financially impaired companies.

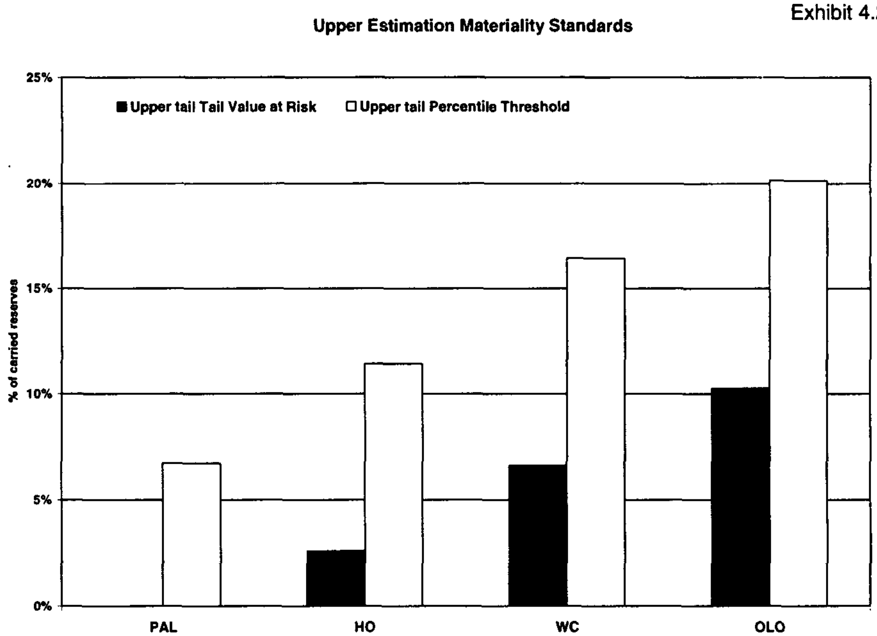
The following table compares the upper tail estimation materiality standards for the two risk measures employed in our analysis, the Percentile Threshold approach and the TVar approach.

Estimation Materiality Standards – Mack

Line of Business	Percentile Threshold	Tail Value at Risk
Personal Auto Liability	6.7%	0.0%
Homeowners	11.4%	2.6%
Workers Compensation	16.4%	6.6%
Other Liability	20.2%	10.3%

The standards implied by the TVar approach are considerably lower when compared to the standards produced by the Percentile Threshold approach. The reason of that observed difference lies on the varying fundamental assumptions of the two risk measures. The Percentile Threshold approach measures the probability that the actual claim liability amount would exceed a selected dollar threshold. It does not consider the magnitude of the resulting deficiency. A $1 reserve deficiency gets the same weight as a $1 million reserve deficiency under the Percentile Threshold approach. On the other hand, the TVar approach measures the expected risk of material adverse deviation. The higher the risk of material adverse deviation, the higher measure of risk is calculated by the TVar approach. In other words, the TVar approach penalizes a company for the probability of extreme claim liability outcomes. Since for most of the property and casualty (general non-life) insurance companies there is a small chance of very large claim liability outcomes, the TVar approach, on average, assigns more reserve risk to the companies when compared to the Percentile Threshold approach. The higher risk associated with the TVar approach results in lower standards of materiality since an earlier warning flag is more appropriate in the presence of more reserving risk. Exhibit 4.2.5 compares the upper tail estimation materiality standards for the two measures of risk employed in our analysis.

Exhibit 4.2.5



We also created a fictitious company that writes the four lines of business under consideration with a reserve distribution approximating the distribution of the whole industry. Employing a Normal Copula approach we calculated the CVs of the claim liabilities for the company. The resulting total risk CV is 11.0% while the parameter risk CV is 10.2%. The risk diversification associated with the underwriting of four, instead of one, lines of business results into combined CVs that are lower when compared to the CVs from the two long tail lines of business (workers compensation and other liability). Exhibit 4.2.6 (see next page) compares the resulting aggregate CVs from the four monoline writers to the CVs of the fictitious multi-line company.

Exhibit 4.2.6

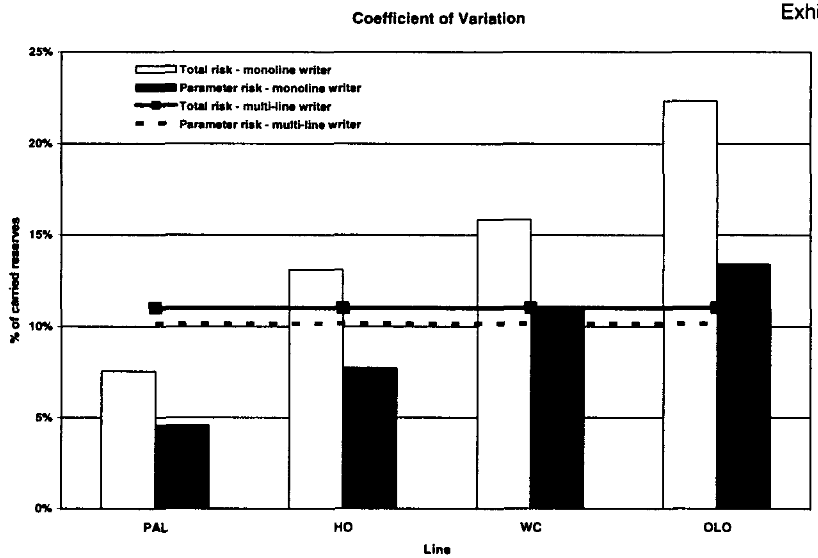
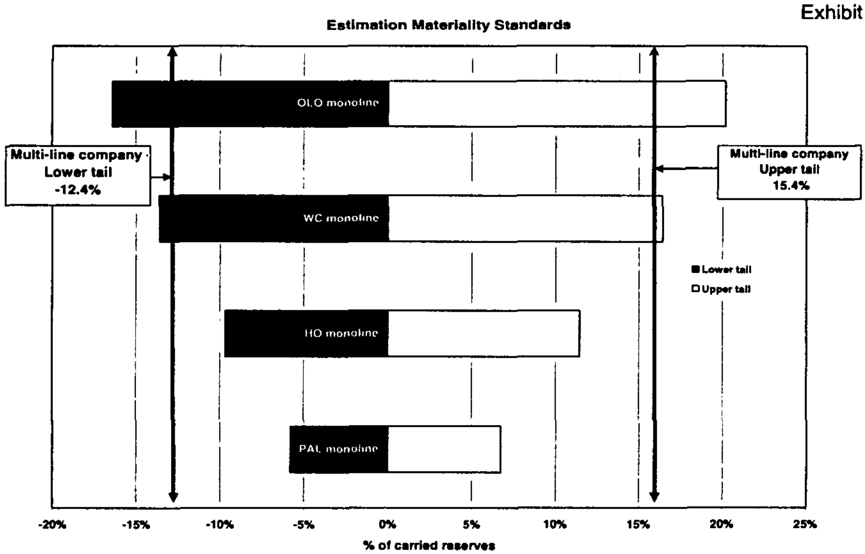


Exhibit 4.2.7 compares the estimation standards of materiality of the fictitious four lines writer to the four monoline writers. The resulting upper tail statistical standard is 15.4% while the lower tail statistical standard is 12.4%. These standards are affected by the higher weight given to the long tail lines (30% for Personal Auto Liability, 6% for Homeowners, 35% for Workers Compensation and 29% for Other Liability Occurrence.)

Exhibit 4.2.7



The following table compares the outcome and estimation materiality standards for a writer of the four lines of business.

Standards of Materiality – Mack

Type	Lower Tail	Upper Tail
Estimation materiality standards	-12.4%	15.4%
Outcome materiality standards	-14.4%	18.1%

The outcome materiality standards are, on average, 15% higher when compared to the estimation materiality standards. This relationship is reasonable in light of process risk which is considered in the outcome materiality standards but not in the estimation materiality standards.

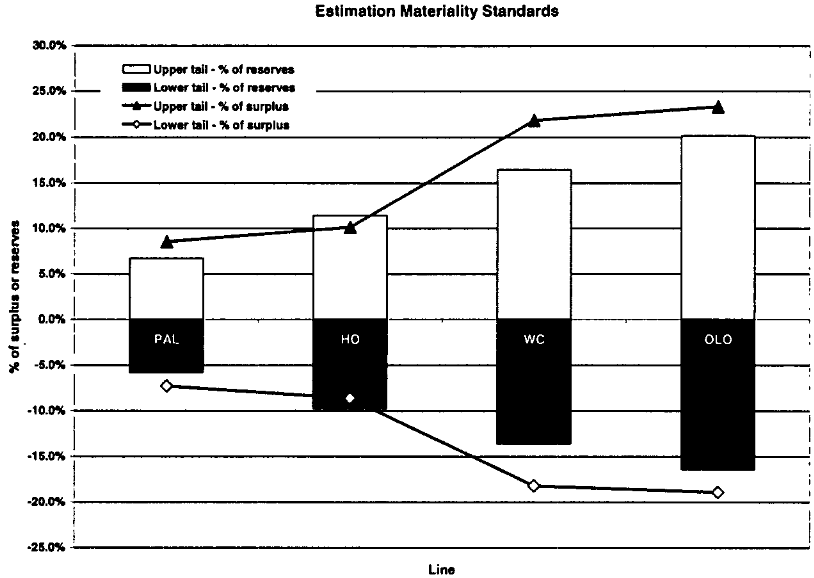
The following table summarizes the estimation materiality standards, for the four lines of business under consideration, as a percentage of individual company surplus.

Estimation Materiality Standards – Mack (as a % of surplus)

Line of Business	Lower Tail	Upper Tail
Personal Auto Liability	-7.3%	8.5%
Homeowners	-8.6%	10.1%
Workers Compensation	-18.2%	21.9%
Other Liability	-18.9%	23.4%

The resulting percentages for the upper tail test are in the area of 10% for short tail lines and in the area of 20% for long tail lines. Exhibit 4.2.8 (see next page) compares the estimation materiality standards, as a percentage of both surplus and carried reserves, for each line of business analyzed.

Exhibit 4.2.8



The following table compares the implied volatility for each line of business analyzed, measured by the coefficient of variation, to the resulting estimation materiality standards for the upper tail test in the Percentile Threshold approach.

Comparison of Parameter Risk CVs and Estimation Materiality Standards – Mack

Line of business	CV	Upper tail estimation materiality standards
Personal Auto Liability	4.6%	6.7%
Homeowners	7.7%	11.4%
Workers Compensation	11.0%	16.4%
Other Liability	13.4%	20.2%

The standards of materiality increase for the more volatile lines. The uncertainty associated with the calculation of the claim liabilities for a volatile line is quite high, and the large associated standards of materiality reflect that uncertainty. All other things being equal, two independent actuarial estimates that measure volatile claim liabilities should be given the benefit of the underlying uncertainty before considered materially different from one another. Another way to intuitively think about this result is that lines of business, or books of business, which show a high level of volatility usually have a higher percentage of total surplus allocated to them and thus have a higher cushion to absorb adverse deviation. That results in a higher standard of materiality as a percentage of reserves.

As we pointed earlier, the results quoted above might be overstated as the stochastic methods employed in this paper presumably overstate process and parameter risk. Thus both the CVs and the materiality standards derived above are overstated. We performed the adjustments to reduce the overstatement, described earlier in this paper, on three companies for three different lines of business to get an approximate impact of the overstatement of volatility by the stochastic methods. The impact of the overstatement of the CV was calculated by subtracting the CV obtained from the true volatility of the adjusted industry triangle from the CV of the unadjusted individual company triangle. Using the benchmark significance levels we calculated the adjusted standards of materiality. This procedure was performed separately for each line of business. The results are presented in the following tables:

Outcome Materiality Standards– Mack, Upper Tail or Adverse Deviation

Line of Business	Before Adjustment	After Adjustment
Personal Auto Liability	12.2%	5.7%
Workers Compensation	26.2%	18.0%
Other Liability	37.7%	16.7%

Estimation Materiality Standards– Mack, Upper Tail or Adverse Deviation

Line of Business	Before Adjustment	After Adjustment
Personal Auto Liability	6.7%	3.6%
Workers Compensation	16.4%	12.5%
Other Liability	20.2%	11.5%

As these tables show, the impact of this overstatement can be significant. To make a thorough assessment of the impact of the adjustment is beyond the scope of this paper.

APPENDICES

A. Technical Appendix – Stochastic Methods Employed

The deterministic methods provide a best estimate of the claim liabilities. In comparison, stochastic methods provide a claim liability distribution around the best estimate, in addition to the best estimate. We employed two stochastic methods in our analysis. Each of these methods represents the two families of stochastic methods described below:

"Chain Ladder" family of methods. These methods employ cumulative loss and expense triangle data and generally are based on the premise that the underlying assumptions of the chain ladder method (CLM) are correct. The Thomas Mack method is probably the best-known representative of this family. It provides the first two moments of the claim liability distribution (i.e., the mean and the variance of the distribution)

"Simulation" family of methods. These techniques provide an empirical distribution of the claim liabilities. Our representative of this family is Bootstrapping, a powerful, yet simple, technique that employs simulations and avoids the fitting of complicated analytical models.

A more detailed description of these two methods follows:

Mack method

The Mack method [12] specifies the first two moments of the claim liability distribution only. It essentially calculates the standard error of the claim liability distribution based on the inherent uncertainty of the underlying data. Our research employed the following notation: Let C_{ik} denote the cumulative loss payments for accident year i, $1 \le i \le I$ and development year k, $1 \le k \le I$, where I is the total number of accident years. The values of C_{ik} are known for $i+k \le I+1$. We want to estimate the values of C_{ik} for $i+k > I+1$. The value of the reserves for accident year i is:

$$R_i = C_{iI} - C_{i,I+1-i}, \qquad \text{(A.1)}$$

where C_{iI} represents the true ultimate loss for accident year i. The expected ultimate loss amount for accident year i is calculated by the formula:

$$\mathbf{C}_{iI} = C_{i,I+1-i} \text{ x } \mathbf{f}_{I+1-i} \text{ x...x } \mathbf{f}_{I-1}, \tag{A.2}$$

where $2 \le i \le I$ and $\mathbf{f}_k$ are the observed volume weighted ATA factors from maturity k to k+1 for $1 \le k \le I-1$. Notice the bolded figure $\mathbf{C}_{iI}$ that represents an estimate of the ultimate loss for accident year i employing historical ATA factors $\mathbf{f}_k$ for $1 \le k \le I-1$. The true value of the ultimate loss for accident year I is denoted by C_{iI} and depends on the actual ATA factors f_k whose values are currently unknown.

There are three major assumptions that form the base of this paper:

(1) $E(\frac{C_{i,k+1}}{C_{i,k}} / C_{i1},\ldots,C_{ik}) = f_k$ for $1 \le i \le I$ and $1 \le k \le I-1$, i.e. the expected value of the loss development factor $\frac{C_{i,k+1}}{C_{i,k}}$ equals f_k, where f_k is the unknown "true" development factor which is the same for all accident years. Moreover the loss development factor $\frac{C_{i,k+1}}{C_{i,k}}$ equals f_k irrespective of the prior development $C_{i1},\ldots,C_{ik}$.

(2) The variables $\{C_{i1},\ldots,C_{iI}\}$ and $\{C_{j1},\ldots,C_{jI}\}$ for different accident years $i \neq j$ are independent (i.e. the loss payments in an accident year are independent from the loss payments in another accident year). Under this assumption, the ATA estimators $\mathbf{f}_k$ are unbiased i.e. $E(\mathbf{f}_k) = f_k$.

(3) The 3rd major assumption of the paper satisfies the principle of the theory of point estimation that among all the unbiased estimators of the ATA factors preference should be given to the one with the smallest variance. This principle can be restated as:

$$\mathrm{Var}(C_{j,k+1} / C_{j1},\ldots,C_{jk}) = C_{jk} \times \alpha_k^2, \tag{A.3}$$

where $1 \le j \le I$, $1 \le k \le I-1$ with unknown proportionality constants α_k^2 for $1 \le k \le I-1$.

With the help of the previous stated assumptions, we calculated the mean squared error (mse) of the ultimate losses for accident year i. This mse of the ultimate loss is defined as:

$$\mathrm{mse}(\mathbf{C}_{iI}) = E[(C_{iI} - \mathbf{C}_{iI})^2 / C_{ik} \text{ for } i+k \le I+1]. \tag{A.4}$$

That mean square error is a conditional expectation of the actual triangle data, since Mack measures future claim volatility given a run-off triangle. It can easily be shown that the mse of the ultimate losses and the reserves for a particular accident year i are equal, i.e. $\mathrm{mse}(\mathbf{C}_{iI}) = \mathrm{mse}(\mathbf{R}_i)$. The square root of the mean squared error of the reserves is called the standard error (s.e.) of the reserves. Based on the previously stated assumptions the standard error of the reserves is calculated for every accident year i, $\mathrm{s.e.}(\mathbf{R}_i)$, and for all accident years combined, $\mathrm{s.e.}(\mathbf{R})$. The resulting formulas are as follows:

$$(\mathrm{s.e.}(\mathbf{C}_{iI}))^2 = \mathbf{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{a_k^2}{f_k^2} \left(\frac{1}{C_{ik}} + \frac{1}{\sum_{j=1}^{I-k} Cjk}\right), \text{ and} \tag{A.5}$$

$$(\mathrm{s.e.}(\mathbf{R}))^2 = \sum_{i=2}^{I} \left\{ (\mathrm{s.e.}(\mathbf{R}_i))^2 + \mathbf{C}_{iI} \left(\sum_{j=i+1}^{I} \mathbf{C}_{jI} \right) \sum_{k=I+1-i}^{I-1} \frac{2a_k^2 / f_k^2}{\sum_{n=1}^{I-k} C_{nk}} \right., \tag{A.6}$$

$$\text{where: } \boldsymbol{\alpha}_k^2 = \frac{1}{I-k-1} \sum_{J=1}^{I-k} C_{jk} \left(\frac{C_{j,k+1}}{C_{jk}} - \mathbf{f}_k \right)^2, \ 1 \le k \le I-2.$$

Using the identity that: $E(X - c)^2 = Var(X) + [E(X) - c]^2$ where c is a constant, we can re-write the mean square error as:

$$mse(\mathbf{R_i}) = Var(R_i \mid D) + [\, E(R_i \mid D) - \mathbf{R_i} \,]^2, \quad \text{(A.7)}$$

where D is the observed triangle data, (i.e., we can decompose the total claim liabilities risk into the sum of pure future random error $Var(R_i \mid D)$ and the deviation between the model estimated claim liabilities and the true Expected Claim liabilities (i.e., the parameter risk)). All the components of the mean square error can be calculated based on the implicit assumption of the Mack model that the chain ladder estimated link ratios are unbiased, minimum variance estimators of the true unknown loss development factors.

Bootstrapping method

Bootstrapping [6] is based on theory developed by England and Verrall. In some of their earlier work, they proved that the reserve estimates from the CLM are identical to reserves produced by an over-dispersed Poisson generalized linear model (GLM). As a result, the residuals produced from a chain ladder model fitted to a historical triangle can be treated as residuals of a regression model. The residuals of regressions should be approximately independent and identically distributed around zero. The Bootstrapping technique samples, with replacement, the residuals of the CLM. The resulting simulated residuals can be considered as residuals from a triangle that have approximately the same statistical characteristics as the triangle that produced the original residuals. Using appropriate residuals (the so-called Pearson residuals) we can produce new sets of incremental payments and subsequently new reserve indications from each simulation.

The Bootstrapping algorithm steps are as follows:

Using all years volume weighted loss development factors (LDFs) from the original triangle, a "fitted" triangle is calculated by applying these LDFs to the latest diagonal of the original triangle.

Fitted incremental values are compared to actual incremental values to calculate unscaled residuals. The formula for the Pearson residual is = (actual – fitted) / sqrt(fitted).

Residuals are normalized by an appropriate scale factor: $\sqrt{\frac{n}{n-p}}$, where n is the number of data point in the triangle and p is the number of the parameters in the over-dispersed Poisson GLM model. The scaling factor adjusts for the difference in the degrees of freedom between the parameter free Bootstrapping model and the over-dispersed Poisson GLM model.

The model re-samples these scaled residuals with replacement. The re-sampling is performed once per simulation. Using these re-sampled residuals, an incremental "bootstrap" loss triangle is created based on the Pearson residuals formula. These incremental losses are converted to cumulative, from which all years volume weighted LDFs are calculated. These are then used to "complete the square," by application of the LDFs to the latest diagonal. Reserves are then calculated for each simulation, and a distribution is assembled using the results of all the simulations. This step captures parameter risk only.

Process risk is introduced by treating each incremental from the bootstrap triangle as the mean of a gamma random variable with variance proportional to the mean. The subsequent steps are identical to those shown above.

Tail variability is modeled by using an inverse power curve fit (the so-called Sherman inverse power curve). The parameters of a linear regression are fitted to available age to age factors (ATA) from all accident years as follows:

$$ATA = 1 + a \times t^{-b}, \qquad \text{(A.8)}$$

where a and b are the fitted parameters while t represents the development year. The fitting procedure employs the natural logarithms of the ATA factors and the resulting formula is:

$$\ln(ATA\text{-}1) = \ln(a) - b \ln(t). \qquad \text{(A.9)}$$

With the use of a linear regression the a and b parameters are calculated based on a least square error approach. The development factors in the tail of the triangle vary at each simulation since the ATA factors from the historical years vary at each simulation too.

Separate fits are used for parameter and total risk. The length of the tail is different by line, as described in the body of the paper.

B. Technical Appendix – Financially Impaired Companies

The following companies were analyzed to establish upper bounds on the significance level. These companies were impaired in 2002 according to Best's Insolvency Study, Property/ Casualty U.S. Insurers 1969-2002.

A.M. Best #	Company name as listed in A.M. Best database	State
03627	Aberdeen Insurance Co	TX
02681	Acceptance Insurance Co	NE
03754	American Growers	NE
00685	American Professionals Insurance Co	IN
12181	Aries Insurance Co, Inc	FL
02141	Casualty Reciprocal Exchange	MO
02412	Equity Mutual Insurance Co	MO
10561	Grange Mutual Insurance	OR
02592	Highlands Casualty Co	TX
02239	Highland Insurance Co	TX
02812	Highlands Lloyds	TX
11860	Legion Indemnity Co	IL
02352	Legion Insurance Co	PA
02348	National Automobile & Casualty Insurance Co	CA
00213	NN Insurance Co	WI
10626	Oak Casualty	IL
02880	Pacific Automobile Insurance Co	CA
02376	Pacific National Insurance Co	CA
03658	PAULA Insurance Co	CA
10420	Security Indemnity Insurance Co	NJ
00858	State Capital Insurance Co	NC
02489	Statesman Insurance Co	IN
12110	Villanova Insurance Co	PA
00942	Wasatch Crest Mutual Insurance Co	UT
10630	Western Specialty Insurance Co	IL

C. Technical Appendix – Copula

Copula theory provides a convenient way to calculate the aggregate distribution of several random variables, given a predetermined correlation matrix among these variables. We started with n=4 lines of business where the mean of the claim liabilities μ, (an (nx1) vector), and the nxn correlation matrix C of the claim liabilities between lines are already given. An assumption that needs to be satisfied is that the correlation matrix C is positive definite (an nxn matrix C is positive definite if it is symmetric and if $x'Cx > 0$ for every n-dimensional column vector $x \neq 0$). In the following steps we will describe the normal copula methodology.

1. The claim liability distribution for each line of business is calculated based on the Mack or Bootstrapping methods.
2. Employing the so-called Cholesky decomposition method, we can calculate a randomly generated n-variate normal vector X with each of its vectors satisfying the predetermined correlation matrix C. The required steps for this Cholesky decomposition are as follows:
    1. Since C is a positive definite matrix we can prove, with the help of intermediate algebra, that C can be factored as follows:
       $C = L \times L'$, (where L is a lower triangular matrix from the Cholesky decomposition and L' is the transpose of L);
    2. We introduce a linear transformation X, i.e. $X = \mu + L \times z$, , where z is an nx1 vector from a standard normal distribution, i.e. $z \sim N(0,1)$;
    3. Then: $E(X) = \mu + E(L) \times E(z) = \mu$, since $E(z)=0$ &
       $Var(X) = E((X-\mu)x(X-\mu)') = E((L \times z)x(L \times z)') =$
       $E((L \times (z \times z)') \times L') = E(L \times L') = C$;
       since $Var(z)= E(z \times z') = I$ (i.e. the identity matrix) and $L \times L' = C$;

    and

    - The end result is an n-variate normal vector X, where $X \sim N(\mu,C)$, i.e. the n-variate normal vector X has the required mean μ and the required correlation matrix C.

Copula theory has been gaining acceptance among actuaries. For example [13], "Correlation and Aggregate Loss Distributions With An Emphasis On the Iman-Conover Method", written by Stephen J. Mindenhall and published in the Winter 2006 CAS Forum, explains in more detail the multivariate Normal Copula approach described above.

We assumed the following correlations between lines:

	Personal Auto Liability	Homeowners	Workers Compensation	Other Liability
Personal Auto Liability	1.00	0.40	0.38	0.60
Homeowners	0.40	1.00	0.40	0.40
Workers Compensation	0.38	0.40	1.00	0.19
Other Liability	0.60	0.40	0.19	1.00

D. Technical Appendix – Detailed Calculation of Materiality Standard

Step 1: Historical paid loss triangle, company A and application of stochastic methods

We start with a historical paid loss triangle for the legal entity A. **We assume that A is a mono-line writer.** We denote the random variable of the unpaid claim liabilities by X. Employing the Mack method we can calculate the first two moments of the claim liability distribution, i.e. the mean, E(X), and the corresponding coefficient of variation, CV(X). Using the Bootstrapping method we calculate an empirical distribution of the claim liabilities. As a byproduct of this empirical distribution we can calculate the mean and the coefficient of variation of the claim liabilities.

For the calculation of estimation materiality standards only parameter risk was considered while for the calculation of outcome materiality standards two types of risk (i.e., process and parameter risk) were considered.

Step 2: Calculation of benchmark significance levels/exceedence ratios

In all the steps of our analysis, except the second step, we employ data from individual companies in order to calculate standards of materiality. For the calculation of the benchmark significance levels and benchmark exceedence ratios, we employ a subset of the industry-wide data, not company specific data. The benchmarks were calculated separately for a group of 39 financially healthy companies and a group of 16 financially impaired companies. These benchmarks were employed in the calculation of outcome materiality standards.

For each company in our database we employed the Mack method to calculate the mean, E(X), and coefficient of variation, CV(X), of their respective total risk claim liability distribution. For the calculation of these distributions we employed loss and ALAE triangles from Schedule P, Part 3 Summary. For simplicity, we added the losses for all lines of business written by a company before calculating its claim liability distribution. We implicitly assumed that each company had not experienced any change in its exposures, among their various lines of business, over the past 10 years. A more detailed, but also more time-consuming approach for each company would be to calculate the claim liability distribution for each of their individual lines and then calculate the aggregate distribution based on the combination of these individual lines distributions.

Another underlying assumption is that each company's claim liability distribution has a log-normal form.

We then recorded the risk based capital amount (RBC) for each company, as provided in their respective annual statements, on the "Five-Year Historical Data" page. Based on the RBC amount we calculated the different NAIC-mandated regulatory, or company action levels. So for example if a company had an RBC amount of $10,000 then we have the following levels:

RBC Action Levels	**"Required Policyholder Surplus"**
No action required (>100%)	$10,000 or more
Company action required (75%-100%)	$7,500 to $10,000
Regulatory action required (50%-75%)	$5,000 to $7,500
Regulatory control authorized (35%-50%)	$3,500 to $5,000
Regulatory control mandated (<35%)	$3,500 or less

As a next step, we measured the difference between the surplus as regards to policyholders and the RBC capital amount that would downgrade each company to the next lower RBC action level. So for example, if the surplus of the company is $12,500 then the calculated difference is $2,500(=12,500-10,000), while if the surplus of the company is $6,000 then the calculated difference is $1,000(=6,000-5,000.) If the company sustained an adverse claim liability deviation greater or equal to the calculated difference it would be downgraded to the next lower RBC level. The difference indicated above can serve as a maximum standard of materiality, within the solvency perspective of materiality.

For each company under consideration we calculated a maximum standard of materiality m. We also assumed that the claim liability distribution of each company is log-normal, while the mean and variance of these distributions have already been calculated by the Mack method. Given the first two moments of a log-normal claim liability distribution we can easily calculate percentiles.

The benchmark significance level is the area in the tail of the company's claim liability distribution, in excess of the mean plus the maximum materiality standard (i.e. E(X)+m.) This area represents the probability of extreme claim liability outcomes that, if materialize, would downgrade the company to the next lower RBC level.

The benchmark exceedence ratio is equal to the expected losses in excess of E(X)+m, as a ratio to the expected claim liabilities E(X). This ratio represents the expected risk of material adverse deviation as a percentage of carried reserves that, if materialize, would downgrade the company to the next lower RBC level.

Steps 3a and 3b describe in more detail how to calculate the percentiles and expected losses, in excess of a given threshold, for a log-normal distribution.

Finally, we calculated the weighted average, across all companies, benchmark significance levels and benchmark exceedence ratios using the carried reserves of each company as a weight. For healthy companies the weighted average benchmark significance level is 6.0% while the weighted average benchmark exceedence ratio is 1.5%. These benchmarks were employed for the calculation of upper tail test outcome materiality standards. For the lower tail test outcome materiality standards we selected judgmentally a benchmark significance level equal to 8.0%. The selection of higher benchmark significance level for the lower tail test makes sure that the resulting outcome standards of materiality are higher for the upper tail test when compared to those of the lower tail test.

The benchmark significance levels and benchmark exceedence ratio for the estimation materiality standards were calculated based on judgment, as explained in the text.

Step 3a: Estimation materiality standards, Company A

Mack – Percentile Threshold approach

We have already calculated:

(a) The mean of the claim liability distribution, E(X);

(b) The coefficient of variation of the claim liability distribution CV(X); and

(c) The benchmark significance level r for estimation materiality. This is 7.5% for the upper tail test and 10.0% for the lower tail test.

We make the additional assumption that the claim liabilities follow a log-normal distribution with parameters μ and σ, i.e. $X \sim LN(\mu,\sigma)$. The logarithm of X then is normally distributed with parameters μ and σ, i.e. $\ln(X) \sim N(\mu,\sigma)$. From introductory statistical theory we can calculate μ and σ by:

$$\mu = \ln(E(X)) - \frac{\sigma^2}{2}, \text{where } \sigma = \sqrt{\ln(1+CV(X)^2)} \qquad \text{(D.1)}$$

The purpose of the percentile threshold approach is to calculate a range of reasonable estimates around the carried reserves that is outside the upper and lower tails of the distribution, as defined by the benchmark significance levels.

For the calculation of the upper tail estimation materiality standard, we subtracted the mean reserves from the 92.5th (=1-0.075) percentile implied by the benchmark significance level:

$$\textbf{Materiality standard - Upper tail} = E(X) * \exp^{\varphi(0.925)*\sigma - \sigma^2/2} - E(X), \qquad \text{(D.2)}$$

where $\phi(0.925)$ represents the 92.5^{th} percentile of the standard normal distribution function. The first component of the preceding formula represents the 92.5^{th} percentile of the log-normal distribution X of the claim liabilities.

For the calculation of the lower tail estimation materiality standard we subtract the 10^{th} percentile implied by the benchmark significance level from the mean reserves:

$$\textbf{Materiality standard - Lower tail} = E(X) - E(X) * \exp^{\varphi(0.10)*\sigma - \sigma^2/2}, \qquad \text{(D.3)}$$

where $\phi(0.10)$ represents the 10.0^{th} percentile of the standard normal distribution function. The second component of the preceding formula represents the 10.0^{th} percentile of the log-normal distribution X of the claim liabilities.

For the purpose of this analysis we employed a mean of the claim liabilities equal to the carried reserve for legal entity A.

Bootstrapping – Percentile Threshold approach

The Bootstrapping stochastic method calculates an empirical distribution of the claim liabilities. The Bootstrapping method produces a few thousand random realizations of the empirical claim liability distribution though a simulation approach. The first step is to linearly transform the stochastic claim liability distribution to make sure that the mean of that distribution is equal to the carried reserves for legal entity A. The transformed distribution has the same coefficient of variation as the original stochastic empirical distribution. The *percentile* function in excel can calculate the various percentiles of the resulting transformed distribution.

The upper tail estimation materiality standard is calculated as follows:

Materiality standard - Upper tail =

92.5^{th} percentile of simulated claim liability distribution - E(X).

The lower tail estimation materiality standard is calculated as follows:

Materiality standard - Lower tail =

E(X) - 10^{th} percentile of simulated claim liability distribution.

Mack – Expected exceedence/TVar approach

For the Mack approach we were provided with the mean, E(X), and the coefficient of variation, CV(X), of the claim liability distribution. Again we assume that the claim liabilities X follow a lognormal distribution with parameters μ and σ. The selected benchmark exceedence ratio is equal to 2.0%.

The purpose of the expected exceedence approach is to calculate a standard of materiality that when added to the carried reserves, the expected losses in excess of these carried reserves plus the materiality standard, is equal to 2.0% of the carried reserves, (for estimation materiality standards.) In other words, if the company experiences actual losses in excess of the expected losses plus the standard of materiality, then the expected material adverse deviation is equal to 2.0% of the carried reserves. A risk of material adverse deviation exists when the actual losses exceed expected losses (i.e., E(X)), by the selected materiality standard. By construction the TVar measure of risk focuses only on the upper tail of the distribution.

Available optimization routines in Excel™, such as *SOLVER*, can help us calculate the standard of materiality m. When we add this standard of materiality to the carried reserves E(X) then the expected losses in excess of E(X)+m are equal to 2% of the carried reserves.

The formula for the expected losses, in excess of the carried reserves plus the materiality standard (i.e. E(X)+m), is as follows:

$$\{1 - \frac{E[X;E(X)+m]}{E(X)}\} \times E(X), \tag{D.4}$$

where E[X;E(X)+m] represents the expected losses from the claim liability distribution limited to E(X)+m (the so called limited expected value function.)

With an assumption of a log-normal distribution for X ~ LN(μ,σ), we calculated the expected losses limited to an upper limit c as follows:

$$E[X;c] = \exp^{\mu+\sigma^2/2} \times \Phi(\frac{\ln(c)-\mu-\sigma^2}{\sigma}) + c \times [1 - \Phi(\frac{\ln(c)-\mu}{\sigma})], \tag{D.5}$$

where $\Phi(x)$ is the standard normal cumulative distribution function.

Again, for the purpose of our analysis we employed a mean of the claim liabilities equal to the carried reserve for legal entity A.

Bootstrapping – Expected exceedence/TVar approach

The empirical distribution produced by the Bootstrapping stochastic reserving method is linearly transformed, as explained in the "*Bootstrapping – Percentile Threshold approach*" section. With the help of *SOLVER*, we can calculate a standard of materiality m that when added to the mean E(X) of the claim liability distribution, the expected losses in excess of E(X)+m are equal to 2.0% of the carried reserves. Again, when a company experiences actual losses that exceed expected losses (i.e., E(X)) by the selected materiality standard amount m, then the expected risk of material adverse deviation is equal to 2.0% of the carried reserves.

The analysis proceeds as follows: We start with a few thousands simulations of the transformed empirical distribution. From each simulated value we subtract the mean of the distribution plus the materiality standard (i.e. E(X)+m.) If the difference:

Simulated value – E(X) – m,

is positive, then the difference represents a material adverse deviation, since the simulated losses exceed the expected loss amount plus the materiality standard amount. If, on the other hand, the difference is negative, then we set it equal to zero since we are interested only in material adverse deviations. We average the material adverse deviations over all the simulated values and we divide this average material adverse deviation by the expected claim liability amount. *SOLVER* ensured that we selected a standard of materiality m that would produce exactly a 2.0% expected risk of material adverse deviation, as a percentage of carried reserves.

Step 3b: Outcome materiality standards, Company A

For the calculation of the outcome materiality standards, we employ exactly the same methodologies described in step 3a for the two stochastic methods, the Mack and Bootstrapping, and the two measures of risk, the percentile threshold approach and the expected exceedence/TVar approach. The only difference is in the benchmark significance level r for outcome materiality. This is 6.0% for the upper tail test and 8.0% for the lower tail test. The outcome benchmark exceedence ratio is 1.5%.

Step 4: Outcome materiality standards, Company B

We assumed that company B was a multi-line writer. The additional analysis, compared to the mono-line company A case, relates to the calculation of the aggregate claim liability distribution from the combination of all lines written by company B.

As a first step, we calculate the claim liability distributions for each of the n lines of business written by company B. Moreover, we assume an nxn correlation matrix C that describes the correlations among these various lines. Based on the Cholesky decomposition methodology described in section "*Normal Copula theory basics*", we can calculate an n-variate normal array X that satisfies the correlation matrix C provided. As a last step, we re-sort the n lines claim liability distributions produced by the Mack and Bootstrapping methods based on the ranking of the nx1 vectors in X. This way, we can achieve the predetermined correlation among the various lines claim liability distributions. We then add all these re-sorted line distributions together to create an aggregate distribution that represents the combined all-lines liabilities for company B.

Having produced the aggregate distribution for all lines combined we then calculate estimation and outcome materiality standards for company B employing the same techniques described in steps 3a and 3b.

The following Exhibits 1 through 5 illustrate the calculation of outcome materiality standards for company A for both the Mack and Bootstrapping stochastic methods and both the Percentile Threshold and TVar measures of risk approaches.

5. REFERENCES

[1] A.M. Best Company, Best Insurance Reports, 2003.

[2] A.M. Best, "Special Report, Best's Insolvency Study, Property/Casualty U.S. Insurers 1969-2002," May 2004.

[3] Butsic, Robert P., "Solvency Measurement for Property - Liability Risk-Based Capital Applications," *Journal of Risk and Insurance*, December 1994, Vol. 61, No. 4, 656-690.

[4] CAS, "Statement of Principles Regarding Property and Casualty Loss and Loss Adjustment Expense Reserves", 1988.

[5] CAS Working Party on Quantifying Variability in Reserve Estimates, "The Analysis and Estimation of Loss & ALAE Variability: A Summary Report", *CAS Fall Forum*, 2005, 1-118.

[6] England, Peter and Richard Verrall, "Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving," *Insurance: Mathematics and Economics*, 1999, Vol. 25, 281-293.

[7] Faivre, Fabien, "Copula: A New Vision for Economic Capital and Application to a Four Line of Business Company," *ASTIN Bulletin*, May 2001, Vol. 31, No. 1, 213-249.

[8] International Accounting Standards Committee Foundation, "Draft Statement of Principles," ED 5.

[9] Kaye, Paul, "Risk Management in Insurance, A Guide To Risk Measurement, Capital Allocation And Related Decision Support Issues," *CAS Discussion Paper Program*, 2005, 1-34.

[10] Kreps, Rodney E., "Management's Best Estimate of Loss Reserves", *CAS Fall Forum*, 2002, 247-258.

[11] Levine, George M., "Considerations Regarding Materiality and Range of Reserves in Connection with Actuarial Standard of Practice #36", *CAS Fall Forum*, 2002, 259-286.

[12] Mack, Thomas, "Measuring the Variability of Chain Ladder Reserve Estimates," *CAS Spring Forum*, 1994, 101-182.

[13] Mindenhall, Stephen J., "Correlation and Aggregate Loss Distributions With An Emphasis On the Iman-Conover Method," *CAS Winter Forum*, 2006, 103-204.

[14] Sherman, Richard E., "Extrapolating, Smoothing, and Interpolating Development Factors," *CAS Proceedings*, 1984, Vol. LXXI, Nos. 135 and 136, 122-155.

Variance and Covariance Due to Inflation

David R. Clark, FCAS

Abstract

Motivation. This paper looks at the problem of measuring correlation between reserve segments. The research was motivated by the 2005 CAS Working Party on Reserve Variability.

Method. Using a random-walk time series model for inflation, we can estimate the variance of a stream of inflation-sensitive payments. The same calculations can be performed to estimate the covariance between two streams of payments.

Results. Formulas are presented for estimating and calculating the variance in reserves attributable to inflation. All of these calculations are performed analytically, without requiring simulation.

Conclusions. Covariance between reserve segments due to common sensitivity to inflation can be easily modeled. This provides a convenient and intuitive way of calculating dependence between reserve segments in order to estimate variance at a company level.

Availability. Excel spreadsheet examples of the calculations described in this paper are available from the author.

Keywords. Inflation, Reserving, Time-Series, Correlation, Covariance

1. INTRODUCTION

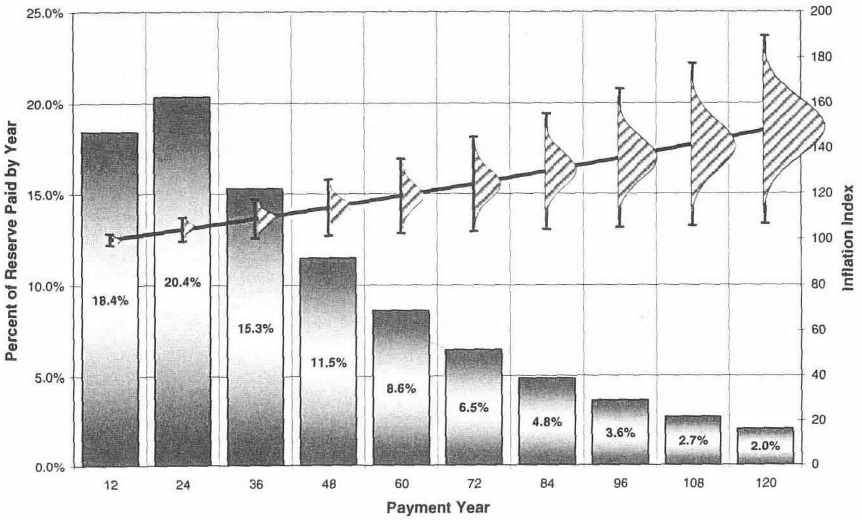
This paper addresses the question of how to estimate the correlation between the future payments in two or more different reserve segments.

The motivation for this paper was the Working Party on Reserve Variability [6], which outlined the many current approaches for estimating variability for a single reserve segment – typically based on a single development triangle. An area of research identified by the Working Party was the question of correlation between two or more reserve segments.

The approach that we will follow for evaluating correlation will be based on first principles about one of the underlying causes of correlation. That is, we begin by asking why we think that there is a correlation structure that needs to be considered. From first principles, we know that inflation has an impact on the amount of loss dollars to be paid, and that different reserve segments may be affected by the same inflation index. For example, a medical claim for an injured worker and a bodily injury claim under Auto Liability may both be dependent upon a common medical inflation driver.

This basic concept is illustrated in the graph below. The bars represent a forecast of loss payments over a ten year time horizon; the line represents the "expected" inflation index built into the forecasted payment stream. If we know the variability in the inflation index (represented by the bell curves), then we can calculate the variance of the future loss payments due to inflation[1].

Inflation Variability for Sample Loss Payout



As the bell curves around the inflation index illustrate, the variance due to inflation increases for longer time horizons. The uncertainty in the estimate of a loss payment ten years in the future is greater than the uncertainty in the estimate of a loss payment one year in the future.

The extension to correlation then follows. If we know that two or more reserve segments are affected by the same inflation index, then we know that they will be correlated with each other.

---

[1] This concept is not new: see the papers by Taylor [5], Hodes et al [4], or Brehm [2] listed in the references.

The question then turns to the source of the inflation index used in this variance calculation. The inflation index should ideally be extracted from the insurance loss data itself, but in practice insurance data is rarely stable enough to provide a reliable estimate. A reasonable alternative is to use an external source for the inflation index.

We will follow the inflation model as outlined in the research work commissioned by the Casualty Actuarial Society (see [1]). This research assumes that inflation follows a mean-reverting random walk. Briefly, this means that the inflation rate in one year is dependent on the inflation rate in the prior year, but that it will eventually "revert" to a long-run average inflation rate. More informally, a mean-reverting model allows us to talk about *periods* of high or low inflation rather than just individual years being higher or lower than average.

Because we are limiting the discussion to the variance and covariance due to inflation, we are able to produce closed-form solutions for all of the variance and covariance terms. All of this can alternatively be incorporated into a larger simulation model if that is preferred.

After describing the basic model of inflation variability (section 2) and the formulas for variance and covariance of the reserve segments (section 3), we will look at a method for refining the calculation to include different sensitivities to inflation by reserve segment (section 4), and then finally how to integrate variance due to inflation with variance from other sources (section 5).

2. BASIC MODEL

We assume that loss inflation rates follow a mean-reverting time series model. This is described using an autoregressive AR(1) model.

$$X_t = \mu \cdot (1-r) + X_{t-1} \cdot r + e_t$$

X_t logarithm of $1+i_t$ (i_t= the inflation rate at time t)

μ logarithm of the 1+long-term average inflation rate i

r factor representing the strength of the reversion (or "persistence")
$r=0$ would be a pure "random draw" model
$r=1$ would be a pure "random walk" model

e_t normally distributed error term, with variance σ^2

Because the model can be transformed into a linear relationship, the parameters can be calculated easily with linear regression.

If we select, for example, a component of the consumer price index (CPI), then the variables are:

Independent Variable (X_{t-1}): $\ln\left(\frac{CPI(2)}{CPI(1)}\right), \ln\left(\frac{CPI(3)}{CPI(2)}\right), \cdots, \ln\left(\frac{CPI(n-1)}{CPI(n-2)}\right)$

Dependent Variable (X_t): $\ln\left(\frac{CPI(3)}{CPI(2)}\right), \ln\left(\frac{CPI(4)}{CPI(3)}\right), \cdots, \ln\left(\frac{CPI(n)}{CPI(n-1)}\right)$

The slope of the regression line is the parameter *r*. We can estimate the long-run average inflation rate by the intercept/(1-*r*), though we will see that the magnitude of this average does not affect our variability calculations.

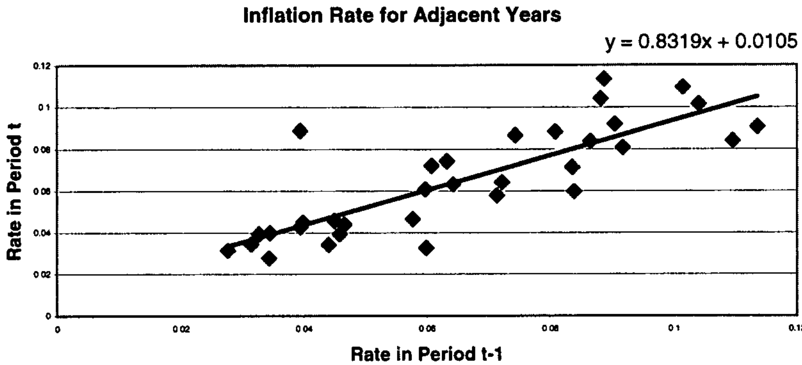
The standard error of the regression (the average deviation of the actual dependent variables from the values predicted by the fitted line) is our estimate of sigma, σ.

We will illustrate this calculation using the medical component of the CPI, though the reserving actuary is free to use any loss-inflation index deemed appropriate. Table 1 below shows this calculation based on data available through the Bureau of Labor Statistics. We calculate the logarithms of changes in the CPI, and then perform a simple linear regression on the X_t and X_{t-1} columns.

This data is, of course, meant purely for illustration and the analyst should decide carefully as to what external inflation index is most representative for the losses to be paid.

Table 1

Year (t)	CPI	Inflation %	X_t	X_{t-1}			
1970	34.0						
1971	36.1	6.18%					
1972	37.3	3.32%	0.0327	0.059932			
1973	38.8	4.02%	0.039427	0.0327			
1974	42.4	9.28%	0.088728	0.039427	X = ln(1+ Inflation %)		
1975	47.5	12.03%	0.113581	0.088728			
1976	52.0	9.47%	0.090514	0.113581			
1977	57.0	9.62%	0.091808	0.090514	Slope	0.831857	r
1978	61.8	8.42%	0.080852	0.091808	Intercept	0.010527	μ*(1-r)
1979	67.5	9.22%	0.088224	0.080852			
1980	74.9	10.96%	0.104026	0.088224	Long-Term	0.062605	μ
1981	82.9	10.68%	0.101481	0.104026			
1982	92.5	11.58%	0.109574	0.101481	Std Error	0.014738	σ
1983	100.6	8.76%	0.083944	0.109574			
1984	106.8	6.16%	0.059806	0.083944			
1985	113.5	6.27%	0.060845	0.059806			
1986	122.0	7.49%	0.072218	0.060845			
1987	130.1	6.64%	0.064282	0.072218			
1988	138.6	6.53%	0.063289	0.064282			
1989	149.3	7.72%	0.074366	0.063289			
1990	162.8	9.04%	0.086565	0.074366			
1991	177.0	8.72%	0.083627	0.086565			
1992	190.1	7.40%	0.071401	0.083627			
1993	201.4	5.94%	0.057743	0.071401			
1994	211.0	4.77%	0.046565	0.057743			
1995	220.5	4.50%	0.04404	0.046565			
1996	228.2	3.49%	0.034325	0.04404			
1997	234.6	2.80%	0.027659	0.034325			
1998	242.1	3.20%	0.031469	0.027659			
1999	250.6	3.51%	0.034507	0.031469			
2000	260.8	4.07%	0.039896	0.034507			
2001	272.8	4.60%	0.044985	0.039896			
2002	285.6	4.69%	0.045853	0.044985			
2003	297.1	4.03%	0.039477	0.045853			
2004	310.1	4.38%	0.042826	0.039477			



3. CALCULATING THE VARIANCE OF PAYMENTS

We proceed by showing the calculation of variance for a single payment and then building the model step-by-step up to the covariance between two streams of payments.

3.1 Calculating the Variance of a Single Payment

A one year inflation factor $(1+i_i)$ is lognormally distributed, which means that a loss payment one year in the future – if unaffected by random factors other than inflation – would also be lognormally distributed.

With no "mean reversion" (r =0), the coefficient of variation, CV, of the loss payment would be $\sqrt{\exp(\sigma^2)-1}$. An inflation factor two years out $CPI(2) = (1+i_1)\cdot(1+i_2)$ would also be lognormally distributed, but the CV would increase to $\sqrt{\exp(2\cdot\sigma^2)-1}$.

The simplicity of this expression is due to the assumption that i_1 and i_2 are independent and identically distributed, and also the fact that the product of two lognormal random variables is also a lognormal random variable.

If we introduce the concept of mean reversion such that $r > 0$, then the formula for the CV of the single year factor does not change, but the two-year inflation index CPI(2) becomes:

$$\text{CPI(2)} = (1+i_1)\cdot\left(\frac{(1+i_1)}{E[1+i_1]}\right)^r\cdot(1+i_2).$$

The $\text{CV}_{n=2}$ increases to become $\sqrt{\exp\{(1+(1+r)^2)\cdot\sigma^2\}-1}$.

The index for subsequent years is created in a similar manner. For n=3, we have

$$\text{CPI(3)} = (1+i_1)\cdot\left(\frac{(1+i_1)}{E[1+i_1]}\right)^r\cdot(1+i_2)\cdot\left(\left(\frac{(1+i_1)}{E[1+i_1]}\right)^r\cdot\frac{(1+i_2)}{E[1+i_2]}\right)^r\cdot(1+i_3).$$

The $\text{CV}_{n=3}$ becomes $\sqrt{\exp\{(1+(1+r)^2+(1+r+r^2)^2)\cdot\sigma^2\}-1}$.

In the special case in which $r=1$, we have a $\text{CV}_{n=3}$ of $\sqrt{\exp\{(1+2^2+3^2)\cdot\sigma^2\}-1}$.

More generally, the CV for *n* years of inflation is given by:

$$CV_n = \sqrt{\exp\{n\cdot\sigma^2\}-1} \qquad \text{for } r=0.$$

$$CV_n = \sqrt{\exp\left\{\left(\frac{n}{(1-r)^2}-\frac{2\cdot r\cdot(1-r^n)}{(1-r)^3}+\frac{r^2\cdot(1-r^{2n})}{(1-r)^2\cdot(1-r^2)}\right)\cdot\sigma^2\right\}-1} \qquad \text{for } r<1$$

or, alternatively

$$CV_n = \sqrt{\exp\left\{\frac{n \cdot (n+1) \cdot (2n+1)}{6} \cdot \sigma^2\right\} - 1} \qquad \text{for } r = 1$$

A more detailed derivation of these formulas is given in Appendix A.

We note that when the reversion term r is close to 1, changes in the inflation rate are "persistent," meaning that the inflation level will not return to its long-run average very quickly. In these cases, the variance of a loss payment in the distant future will have a much greater variance than under the "random draw" model with $r = 0$.

The table below shows the CV implied for a single payment at various points in the future using different assumptions about the reversion parameter r.

		Sigma =	0.024996	
	CV_n for Selected Reversion Parameters			
n	r = 0	r = .50	r = .80	r = 1
1	0.0250	0.0250	0.0250	0.0250
2	0.0354	0.0451	0.0515	0.0559
3	0.0433	0.0629	0.0799	0.0937
4	0.0500	0.0785	0.1090	0.1376
5	0.0559	0.0923	0.1380	0.1870

3.2 Calculating the Covariance Between Two Payments

Suppose that we have an inflation factor for a given number of years n, and a second factor for $n+k$. We quickly recognize that there must be a strong correlation since n of the $n+k$ years are common to both factors. Using the same mean reversion model, the correlation coefficient can be written[2]:

[2] The term $Cov_{n,k}$ is a "scaled" value which is the dollars of covariance divided by the means of the losses at times n and $n+k$. This is sometimes called the "coefficient of covariation" and is convenient notation because of the parallel to the coefficient of variation (CV) used earlier.

$$\rho_{n.n+k} = \frac{Cov_{n.k}}{CV_n \cdot CV_{n+k}}.$$

The term in the numerator is proportional to the covariance, and is given as follows:

$$Cov_{n.k} = \exp\left\{\left(\frac{n}{(1-r)^2} - \frac{r\cdot(1+r^k)\cdot(1-r^n)}{(1-r)^3} + \frac{r^{k+2}\cdot(1-r^{2n})}{(1-r)^2\cdot(1-r^2)}\right)\cdot\sigma^2\right\} - 1 \qquad \text{for } r < 1$$

or, alternatively

$$Cov_{n.k} = \exp\left\{\frac{n\cdot(n+1)}{2}\left(\frac{2n+1}{3}+k\right)\cdot\sigma^2\right\} - 1 \qquad \text{for } r = 1$$

Note also that $Cov_{n.k} = CV_n^2$ when k =0.

Sigma = 0.025000
Reversion = 0.500000

Matrix of Correlation Coefficients

	1	2	3	4	5
1	1	0.83188775	0.69611104	0.59742763	0.52484632
2	0.83188775	1	0.91052622	0.80678419	0.71882568
3	0.69611104	0.91052622	1	0.94009581	0.85838825
4	0.59742763	0.80678419	0.94009581	1	0.95526523
5	0.52484632	0.71882568	0.85838825	0.95526523	1

3.3 Calculating the Variance of a Stream of Payments

Given these terms, we are able to set up a matrix of correlation coefficients, or covariances, in order to calculate the variance for a sum of payments. The full correlation structure between the individual payments due to inflation is captured in this matrix.

If we have a vector of *N* loss payments, $\vec{P}$, and an *N*-by-*N* matrix of covariance terms such that $M(i,j) = Cov_{i,j-i}$ for $i < j$, then we can calculate the variance for the stream of payments as:

$$Var(P) \;=\; \vec{P} \cdot M \cdot \vec{P}^T \qquad P = \text{sum of all payments in the vector}$$

Or equivalently,

$$Var(P) \;=\; \sum_{i=1}^{N}\sum_{j=1}^{N} P(i) \cdot M(i,j) \cdot P(j)$$

3.4 Calculating the Covariance Between Two Streams of Payments

If we have two vectors of loss payments ${}_A\vec{P}$ and ${}_B\vec{P}$, both with *N* elements, then the covariance of the two sums can be calculated in a similar manner.

$$Cov({}_AP,\, {}_BP) \;=\; {}_A\vec{P} \cdot M \cdot {}_B\vec{P}^T.$$

The correlation between the two payouts will be a single number, and generally a number approaching 1.000, indicating a very strong correlation. This is because our model assumes that both payment streams are directly affected by the inflation rate, and that inflation is the only source of variability. In Section 4, we soften the first assumption by allowing different degrees of sensitivity to inflation by line of business. In Section 5, we show how to bring in other sources of variability.

4. MEASURING THE SIGNIFICANCE OF INFLATION BY SEGMENT

As mentioned above, the variance/covariance model assumes that the CPI directly affects the amount of loss payment. This may not be exactly true, and we would want the ability to control the degree to which loss development is dependent on inflation.

The degree of inflation for a given risk class (RC) will be controlled by a parameter ${}_{RC}\gamma$, which is applied as an exponent to the CPI. This parameter could be set equal to zero for the cases in which a risk class is unaffected by inflation.

Adjusted Inflation Index for Risk Class A: $CPI^{{}_{A}\gamma}$

In calculating the time-series parameters for this adjusted index, the reversion parameter r is unchanged regardless of the γ; the sigma will change to become $\sigma \rightarrow \gamma \cdot \sigma$. This adjustment is easily incorporated into the CV calculation.

$$CV_n = \sqrt{\exp\left\{\left(\frac{n}{(1-r)^2} - \frac{2 \cdot r \cdot (1-r^n)}{(1-r)^3} + \frac{r^2 \cdot (1-r^{2n})}{(1-r)^2 \cdot (1-r^2)}\right) \cdot \gamma^2 \cdot \sigma^2\right\} - 1}$$

Similarly, the covariance term, when there are two risk classes, A and B, with different degrees of dependence on inflation, is modified as below:

$$Cov_{n,k} = \exp\left\{\left(\frac{n}{(1-r)^2} - \frac{r \cdot (1+r^k) \cdot (1-r^n)}{(1-r)^3} + \frac{r^{k+2} \cdot (1-r^{2n})}{(1-r)^2 \cdot (1-r^2)}\right) \cdot {}_{A}\gamma \cdot {}_{B}\gamma \cdot \sigma^2\right\} - 1$$

We note that this expression is the same as the earlier calculation when ${}_A\gamma = {}_B\gamma = 1$, and the covariance is zero when either ${}_A\gamma$ or ${}_B\gamma$ is zero.

The next question to address is the method for estimating the parameter γ for a given business segment. We begin by defining a simple model for loss payments from a triangle. The formulas below give a model ignoring inflation:

$$c_{y,d} \approx \alpha_y \cdot \beta_d$$

Where $c_{y,d}$ = incremental loss paid in accident year y and development period d. For example, $c_{1999,3}$ would be the amount paid for accident year 1999 between 24 and 36 months.

α_y = a measure of exposure for accident year y, such as onlevel premium. This can be supplied from external sources or be estimated from the triangle itself.

β_d = a parameter representing the amount of development in development period d.

This model is introduced for simplicity only. When we combine this simple two factor (AY and development period) model with an assumption that incremental payments follow an over-dispersed Poisson distribution, then the results match an all-year weighted average chain-ladder calculation.

In order to include an inflation index in this model, we expand the expression with a term including a CPI curve.

$$c^*_{y,d} \approx \alpha_y \cdot \beta_d \cdot CPI(y+d-1)^\gamma$$

From this expanded model, we immediately notice that the no-inflation model is a special case when $\gamma = 0$, so that $c_{y,d} = c^*_{y,d}$. If payments are directly proportional to inflation, then we would expect $\gamma = 1$; and if we expect a "leveraged" effect of inflation (say, in excess layers) then $\gamma > 1$.

Given an explicit model, as above, we are then able to estimate the parameter γ that maximizes a likelihood function or minimizes some other error function. We also have available the goodness-of-fit statistics to test the value of including inflation.

To illustrate, we will work with a small triangle of [cumulative] paid data:

AY	1	2	3	4	5	6	7
1998	13,822	26,045	34,915	41,064	45,228	47,942	49,730
1999	13,710	27,104	36,777	43,309	47,266	49,501	
2000	14,409	28,805	38,328	44,772	49,022		
2001	15,120	28,945	38,692	45,169			
2002	13,344	25,970	34,922				
2003	13,506	25,926					
2004	14,765						

The incremental paid losses from this triangle are then given by:

AY	1	2	3	4	5	6	7
1998	13,822	12,223	8,870	6,149	4,164	2,714	1,788
1999	13,710	13,394	9,673	6,532	3,957	2,235	
2000	14,409	14,396	9,523	6,444	4,250		
2001	15,120	13,825	9,747	6,477			
2002	13,344	12,626	8,952				
2003	13,506	12,420					
2004	14,765						

Based on maximum likelihood estimation[3], we have the following fitted parameters:

y	α_y	d	β_d
1998	49,730	1	0.2737
1999	51,347	2	0.2573
2000	53,571	3	0.1814
2001	54,089	4	0.1227
2002	49,018	5	0.0800
2003	48,824	6	0.0490
2004	53,946	7	0.0360

These fitted values are equivalent to calculating the α's as the chain-ladder ultimates. The fitted values from this model, corresponding to the actual incremental payments, are shown in the triangle below.

AY	1	2	3	4	5	6	7
1998	13,611	12,796	9,023	6,099	3,978	2,435	1,788
1999	14,054	13,212	9,316	6,298	4,107	2,514	
2000	14,662	13,784	9,720	6,571	4,285		
2001	14,804	13,917	9,813	6,634			
2002	13,416	12,612	8,893				
2003	13,363	12,563					
2004	14,765						

The model is then expanded for the inflation adjustment.

y	α_y	d	β_d	CPI	Index	γ	Index^{γ}
1998	49,730	1	0.2761	242.1	1.000	1.655	1.000
1999	48,043	2	0.2424	250.6	1.035		1.059
2000	46,709	3	0.1593	260.8	1.077		1.131
2001	43,867	4	0.1003	272.8	1.127		1.218
2002	37,028	5	0.0609	285.6	1.180		1.315
2003	34,448	6	0.0348	297.1	1.227		1.403
2004	35,499	7	0.0239	310.1	1.281		1.506

With the fitted values including this inflation parameter are as follows:

AY	1	2	3	4	5	6	7
1998	13,732	12,764	8,962	6,075	3,981	2,430	1,788
1999	14,046	13,172	9,327	6,331	4,106	2,519	
2000	14,587	13,796	9,783	6,571	4,285		
2001	14,759	13,978	9,808	6,625			
2002	13,440	12,595	8,887				
2003	13,348	12,578					
2004	14,765						

For example, the first development period for AY 2003 has a fitted value equal to:

$$13{,}348 = 34{,}448 \times .2761 \times 1.403.$$

[3] For this calculation, we will assume that each cell follows an Over-Dispersed Poisson (ODP) distribution with a common variance/mean ratio ϕ. Appendix A gives the full details of this model.

The parameter value 1.655 acts as a "leveraging" effect on inflation, meaning that payments increase at a faster rate than the CPI would indicate. However, in this example, as with most real data sets, there is significant uncertainty in the estimate of the γ parameter. The loss development triangle simply is not a sufficient base for estimating it credibly. Informally, the reason for this is that we can pick almost any value for γ and then fit α_y and β_d vectors that reasonably approximate the historical loss development (see Appendix B for further insight as to why this is the case). It is for this reason that we recommend that the γ parameter be selected by the model user rather than via a fitted model.

The example given above shows that the parameter γ, for measuring the sensitivity to inflation, often lacks great predictive value, that is $\hat{c}^*_{y,d}$ is not much better than $\hat{c}_{y,d}$. This suggests that the use of an external inflation index in calculating variability needs to be justified on *a priori* theoretical grounds and not solely on statistical tests. As a starting assumption, $\gamma = 1$ for each risk class is most reasonable.

The difficulty in estimating the parameter γ does not mean that losses are unaffected by inflation, but merely that a triangle format is not a sufficient basis to discern what the relationship to inflation is.

5. COMBINING OTHER SOURCES OF VARIABILITY

The discussion to this point has been limited to the variability strictly due to inflation. Naturally the variability of loss payments is driven by many other sources, and we need to be able to combine these different sources into a single calculation. Some of these other sources would include:

- Changes in an injured person's condition (recovery, deterioration, death)
- Newly reported claims not originally in the triangle ("true" IBNR)
- Legal or regulatory changes impacting the coverage provided in the insurance policy

These types of variability are, arguably, independent of changes in the rate of inflation and can therefore be treated as statistically independent.

The most common method for including all types of variability is through the use of a large simulation model; however, that is not necessary if we are interested just in the means and variances of the payments.

Section 5 will follow the same logic as Section 3, by starting with a single payment and then showing step-by-step how the calculations are generalized to produce a full covariance matrix on payment streams.

5.1 Calculating the Variance of a Single Payment

Suppose that we have a random variable for the payment amount at a specific time *t*, and denote this expected amount C_t. The timing of the payment is known with certainty, and we have an estimate of its mean $E(C_t)$ and variance $Var(C_t)$ from sources other than inflation. These values may have come from a stochastic reserving model, or may have been simply selected by a reserving actuary.

The next step is to assume that we have an estimate of the inflation index at time *t*, based on the equations from Sections 3 and 4 above.

$$CV_t = \sqrt{\exp\left\{\left(\frac{t}{(1-r)^2} - \frac{2 \cdot r \cdot (1-r^t)}{(1-r)^3} + \frac{r^2 \cdot (1-r^{2t})}{(1-r)^2 \cdot (1-r^2)}\right) \cdot \gamma^2 \cdot \sigma^2\right\} - 1}$$

The inflation index will be represented by a second random variable b_t, with a mean of one $E(b_t) = 1$ and a variance of $Var(b_t) = CV_t^2$. We make the further assumption that the inflation index is statistically independent of the other sources of variance in C_t.

The variance of the product of the two random variables is then calculated as follows.

$$Var(b_t \cdot C_t) = Var(b_t) \cdot Var(C_t) + Var(b_t) \cdot E(C_t)^2 + E(b_t)^2 \cdot Var(C_t)$$

The derivation of this expression is given in Appendix C.

For the reader familiar with the literature of the Casualty Actuarial Society, the description to this point should not be surprising. In fact, the formulas are identical with what is usually referred to as "mixing" parameters, and the use of the notation "*b*" is a deliberate choice to be consistent with this idea.

The inflation index can be viewed as a "parameter variance" component with the total variance above regrouped as follows.

$$Var(b_t \cdot C_t) = \underbrace{Var(C_t) \cdot E(b_t^2)}_{\text{Process Variance}} + \underbrace{Var(b_t) \cdot E(C_t)^2}_{\text{Parameter Variance}}$$

5.2 Calculating the Covariance of a Two Payments

If we have two payments, taking place at different times, *t* and *t+k*, then the covariance between these two payments is calculated in a formula that generalizes the variance formula above.

$$\begin{aligned} Cov(b_t \cdot C_t, b_{t+k} \cdot C_{t+k}) &= Cov(b_t, b_{t+k}) \cdot Cov(C_t, C_{t+k}) \\ &+ Cov(b_t, b_{t+k}) \cdot E(C_t) \cdot E(C_{t+k}) + E(b_t) \cdot E(b_{t+k}) \cdot Cov(C_c, C_{t+k}) \end{aligned}$$

For the special case of *k=0*, this expression reduces to the variance formula above.

5.3 Calculating the Variance of a Stream of Payments

The variance of a stream of payments is a linear combination of the variance and covariance terms calculated above.

We again start with a vector of N expected loss payments, $\vec{P} = \{E(C_t)\}_{t=1}^N$. We now assume that we also know the covariance matrix from sources other than inflation, $M_C(i,j) = Cov(C_i, C_j)$.

As in Section 3.3, we also create an N-by-N matrix of covariance terms for the inflation indices corresponding to each loss payment: $M_b(i,j) = Cov(b_i, b_j)$.

The covariance matrix, representing each pair of loss payments in the payment stream $\vec{P}_N$, is calculated by applying the formula from Section 5.2 on an element-by-element basis.

$$M_{b \cdot C}(i,j) = M_b(i,j) \cdot M_C(i,j) + M_b(i,j) \cdot E(C_i) \cdot E(C_j) + M_C(i,j)$$

The variance of the sum of all payments in the stream is then calculated as the sum of all entries in this combined matrix $M_{b \cdot C}$.

Once again, this may be viewed as a combination of a matrix of expected "process variance" and a matrix of "parameter variance" elements.

$$M_{b\cdot C}(i,j) = \underbrace{M_C(i,j)\cdot\{1+M_b(i,j)\}}_{\text{Process Variance}} + \underbrace{M_b(i,j)\cdot E(C_i)\cdot E(C_j)}_{\text{Parameter Variance}}$$

We may also note that the sum of the "parameter variance" elements is identical to what we denoted $Var(P) = \vec{P}\cdot M\cdot \vec{P}^T$ in Section 3.3.

At this point the reader may have a concern about where all of these numbers come from. The matrix of covariances related to inflation M_b is created using the formulas from Section 3, but do we really have all of the covariances from other sources needed for M_C ? It may be that these are not available and a further simplification is needed.

The easiest way to simplify this process is to include an assumption that the ultimate loss C and the variance of the ultimate loss $Var(C)$ are known. We further assume that the payment pattern on a percent basis is fixed and certain. That is, the dollar amount of ultimate loss may vary, the same percent will always be paid in the first year. By this assumption, all of the C_t payments are perfectly correlated and have the same coefficient of variation (standard deviation divided by mean) CV_C. The elements of the M_C matrix are then easily defined as follows.

$$M_C(i,j) = E(C_i)\cdot E(C_j)\cdot CV_C^2$$

The overall covariance matrix then simplifies greatly.

$$M_{b\cdot C}(i,j) = M_b(i,j)\cdot M_C(i,j) + M_b(i,j)\cdot E(C_i)\cdot E(C_j) + M_C(i,j)$$

becomes

$$M_{b\cdot C}(i,j) = \left\{CV_C^2 + \left(1+CV_C^2\right)\cdot M_b(i,j)\right\}\cdot E(C_i)\cdot E(C_j)$$

5.4 Calculating the Covariance between Two Streams of Payments

The example of how to combine the variance due to inflation with variance from other sources can now be generalized to the discussion of the covariance between two reserve risk classes such as different lines of business.

If we have two risk classes A and B, each with selected payment streams such that we create an NxN matrix of covariance terms between each of the payments. As with the single payment stream example, this can be set up as a matrix.

$$M_{AB}(i,j) \;=\; Cov\left({}_AC_i, {}_BC_j\right)$$

To combine this with the variance due to inflation, we then use the following formula.

$$M_{b\cdot AB}(i,j) \;=\; M_b(i,j)\cdot M_{AB}(i,j) + M_b(i,j)\cdot E\left({}_AC_i\right)\cdot E\left({}_BC_j\right) + M_{AB}(i,j)$$

If the two reserve segments are not correlated based on any factors other than inflation, then all the elements of this matrix are zero, and no calculations are necessary.

We may also simplify the matrix if, as in the previous section, we introduce the assumption that the percent payment pattern for each risk class is fixed and known. The matrix M_{AB} then becomes a constant amount times the cross-product of the payments. The correlation coefficient ρ_{AB} for sources other than inflation is introduced.

$$M_{AB}(i,j) \;=\; E\left({}_AC_i\right)\cdot E\left({}_BC_j\right)\cdot\left\{\rho_{AB}\cdot CV_A\cdot CV_B\right\}$$

This again leads to a simpler version of the covariance matrix.

$$M_{b\cdot AB}(i,j) \;=\; \left\{\rho_{AB}\cdot CV_A\cdot CV_B + \left(1+\rho_{AB}\cdot CV_A\cdot CV_B\right)\cdot M_b(i,j)\right\}\cdot E\left({}_AC_i\right)\cdot E\left({}_BC_j\right)$$

The covariance term between the two risk classes is the sum of all of the terms in this matrix.

The correlation coefficient $\rho_{b\cdot AB}$ (including both inflation and other sources) between these two risk classes is then calculated as follows.

$$\rho_{b\cdot AB} \;=\; \frac{sum\{M_{b.AB}\}}{\sqrt{sum\{M_{b.A}\}\cdot sum\{M_{b.B}\}}} \;=\; \frac{\rho_{AB}\cdot CV_A\cdot CV_B + \left(1+\rho_{AB}\cdot CV_A\cdot CV_B\right)\cdot\Sigma_{AB}^2}{\sqrt{\left\{CV_A^2+\left(1+CV_A^2\right)\cdot\Sigma_A^2\right\}\cdot\left\{CV_B^2+\left(1+CV_B^2\right)\cdot\Sigma_B^2\right\}}}$$

$$\begin{aligned}
\text{where } \Sigma_A^2 &= sum\left\{M_b(i,j)\cdot E\left({}_AC_i\right)\cdot E\left({}_AC_j\right)\right\}/E\left({}_AC\right)^2\\
\Sigma_B^2 &= sum\left\{M_b(i,j)\cdot E\left({}_BC_i\right)\cdot E\left({}_BC_j\right)\right\}/E\left({}_BC\right)^2\\
\Sigma_{AB}^2 &= sum\left\{M_b(i,j)\cdot E\left({}_AC_i\right)\cdot E\left({}_BC_j\right)\right\}/\left\{E\left({}_AC\right)\cdot E\left({}_BC\right)\right\}
\end{aligned}$$

These expressions can also be written in matrix notation.

$$\begin{aligned}
\Sigma_A^2 &= Var\left({}_AP\right) = {}_A\vec{P}\cdot M_b\cdot{}_A\vec{P}^T\\
\Sigma_B^2 &= Var\left({}_BP\right) = {}_B\vec{P}\cdot M_b\cdot{}_B\vec{P}^T\\
\Sigma_{AB}^2 &= Cov\left({}_AP, {}_BP\right) = {}_A\vec{P}\cdot M_b\cdot{}_B\vec{P}^T
\end{aligned}$$

In these formulas, we have included the same inflation covariance matrix M_b. However, if we include adjustment factors other than ${}_A\gamma={}_B\gamma=1$, then we would need to adjust the matrices as shown in Section 4.

With this formula, we are able to combine the correlation due to inflation with correlation from other sources without having to define all of the inter-dependencies between individual payments. If the user is uncomfortable with assuming that the payout patterns do not vary, then the more general formulas can be run.

6. RESULTS AND DISCUSSION

Having completed a fairly rigorous description of the formulas for calculating covariance due to inflation, it is worthwhile showing a simplified numerical example to illustrate how this can be implemented in practice.

We begin with the inflation model defined in Section 2, in which we calculated:

Reversion parameter r = .831857

Variability Sigma σ = .014738

If both reserve risk classes A and B are directly proportional to this inflation index, such that ${}_A\gamma={}_B\gamma=1$, then we have an inflation covariance matrix M_b as show below (each element of the matrix being one calculation of the formula in Section 3.2).

Matrix of Covariance Factors M_b

0.00022	0.00040	0.00055	0.00067	0.00078	0.00086	0.00094	0.00100	0.00105	0.00109
0.00040	0.00095	0.00140	0.00178	0.00210	0.00236	0.00258	0.00276	0.00291	0.00304
0.00055	0.00140	0.00233	0.00311	0.00375	0.00429	0.00473	0.00510	0.00541	0.00567
0.00067	0.00178	0.00311	0.00443	0.00553	0.00644	0.00720	0.00784	0.00837	0.00881
0.00078	0.00210	0.00375	0.00553	0.00722	0.00864	0.00982	0.01080	0.01161	0.01229
0.00086	0.00236	0.00429	0.00644	0.00864	0.01069	0.01240	0.01382	0.01501	0.01600
0.00094	0.00258	0.00473	0.00720	0.00982	0.01240	0.01477	0.01675	0.01840	0.01977
0.00100	0.00276	0.00510	0.00784	0.01080	0.01382	0.01675	0.01941	0.02163	0.02348
0.00105	0.00291	0.00541	0.00837	0.01161	0.01501	0.01840	0.02163	0.02456	0.02699
0.00109	0.00304	0.00567	0.00881	0.01229	0.01600	0.01977	0.02348	0.02699	0.03014

We then introduce two reserve segments, having ten year payment patterns as below.

Year	Risk Class A $_AP$	Risk Class B $_BP$
1	46.40%	15.20%
2	12.10%	11.60%
3	8.40%	10.50%
4	6.80%	10.00%
5	5.70%	9.40%
6	4.90%	9.10%
7	4.50%	8.90%
8	4.00%	8.60%
9	3.70%	8.40%
10	3.50%	8.30%

From this information, we can calculate the CVs from inflation as follows:

$$\Sigma_A^2 = Var({}_AP) = {}_A\bar{P} \cdot M_b \cdot {}_A\bar{P} = .0470^2$$

$$\Sigma_B^2 = Var({}_BP) = {}_B\bar{P} \cdot M_b \cdot {}_B\bar{P} = .0803^2$$

$$\Sigma_{AB}^2 = Cov({}_AP, {}_BP) = {}_A\bar{P} \cdot M_b \cdot {}_B\bar{P} = .0610^2$$

The correlation coefficient from inflation only is then estimated as follows.

$$\frac{\Sigma_{AB}^2}{\sqrt{\Sigma_A^2 \cdot \Sigma_B^2}} = \frac{.0610^2}{.0470 \cdot .0803} = .989$$

This very significant correlation is, again, due to the fact that inflation is the only factor contributing to the variance of either reserve risk class.

We can generalize this by including variability from other sources. We will assume that the risk classes A and B have CVs from sources other than inflation of .100 and .160 respectively, and that these are independent. Further, we will include the simplifying assumption that the ultimate losses are variable but that the percentage payout patterns are fixed. The resulting correlation coefficient, reflecting all sources of variance is given below.

$$\rho_{b \cdot AB} = \frac{\Sigma_{AB}^2}{\sqrt{\left(CV_A^2 + (1+CV_A^2)\cdot\Sigma_A^2\right)\cdot\left(CV_B^2 + (1+CV_B^2)\cdot\Sigma_B^2\right)}}$$

$$= \frac{.0610^2}{\sqrt{\left(.100^2 + (1+.100^2)\cdot .0470^2\right)\cdot\left(.160^2 + (1+.160^2)\cdot .0803^2\right)}} = .188$$

All of these numbers are meant purely for illustration purposes, but they do show that the formulas produce results in reasonable ranges.

The general process for estimating variance and covariance due to inflation can be summarized in the steps below:

- Select an external index, such as a component of the CPI
- Estimate the variance (σ^2) and reversion (r) parameters for the inflation index
- Select a default inflation-sensitivity parameter γ for each risk class
- Estimate the future loss payment stream for each risk class
- Calculate the variance of each risk class due to inflation
- Calculate the covariance between each pair of risk classes

7. CONCLUSIONS

The formulas outlined in this paper provide a very simple method for estimating the sensitivity of losses and reserves to movement in inflation rates. The advantages of this approach may be summarized as below:

1) The basic idea is very easy to explain: loss payments move with inflation
2) Variability due to inflation can be linked to economic forecast models
3) The calculation of variances and correlation can be performed in an Excel spreadsheet in closed form

The chief disadvantage that is identified is that external inflation indices, such as components of the consumer price index (CPI) have not been shown to be significant explanatory variables for movement in insurance loss amounts.

In spite of the difficulty in estimating the sensitivity parameter γ, however, we have a reasonable baseline value of $\gamma = 1$. The model therefore can provide a correlation structure between reserve risk classes based on external knowledge of inflation with a minimal need for arbitrary assumptions.

Appendix A: Derivation of Key Formulas

This appendix provides a more detailed derivation of the key variance and covariance formulas given in the body of the paper.

The formulas in this paper are able to be written in a compact form by capitalizing on a useful property of the lognormal distribution; namely that the product of lognormal random variables is again a lognormal random variable. Analogously, the sum of normal (Gaussian) random variables is again a normal random variable.

The autoregressive model, AR(1), is written in a recursive linear form, after taking the logarithms of the inflation trend factors.

$$X_t \;=\; X_{t-1} \cdot r + b + \sigma \cdot e_t$$

X_t logarithm of $1+i_t$ (i_t= the inflation rate at time t)

e_t standard normal random variable, $e_t \propto Normal(0,1)$

The distribution of X_t, conditional upon a known value for X_{t-1}, is then given as

$$X_t \mid X_{t-1} \;\propto\; Normal(X_{t-1} \cdot r + b, \sigma).$$

The variance of the conditional random variable is then

$$Var(X_t \mid X_{t-1}) \;=\; \sigma^2.$$

The random variable for the logarithm of the inflation rate two or more years out is found by expanding the recursive expression:

$$X_t \mid X_{t-2} \;=\; (X_{t-2} \cdot r \;+\; b \;+\; \sigma \cdot e_{t-1}) \cdot r \;+\; b \;+\; \sigma \cdot e_t$$

$$X_t \mid X_{t-3} \;=\; ((X_{t-3} \cdot r \;+\; b \;+\; \sigma \cdot e_{t-2}) \cdot r \;+\; b \;+\; \sigma \cdot e_{t-1}) \cdot r \;+\; b \;+\; \sigma \cdot e_t$$

This expanding of the recursive formula can be generalized as

$$X_t \mid X_0 \;=\; X_0 \cdot r^t \;+\; \sum_{i=1}^{t} (b + \sigma \cdot e_i) \cdot r^{t-i}.$$

The variance for this more general form is therefore given as below.

$$X_t \mid X_0 \;\propto\; Normal\left(X_0 \cdot r^t + b \cdot \sum_{i=1}^{t} r^{t-i}, \sigma \cdot \sqrt{\sum_{i=1}^{t} r^{2 \cdot (t-i)}} \right)$$

The variance for the random variable conditional upon a point "t" years prior is then:

$$Var(X_t \mid X_0) = \sigma^2 \cdot \sum_{i=1}^{t} r^{2\cdot(t-i)} = \sigma^2 \cdot \frac{1-r^{2t}}{1-r^2} \qquad \text{if } r<1$$

or $$Var(X_t \mid X_0) = \sigma^2 \cdot t \qquad \text{if } r=1$$

These and subsequent simplifications are possible based on three fundamental identities.

$$\sum_{k=1}^{n} k = 1+2+3+\cdots+(n-1)+n = \frac{n\cdot(n+1)}{2}$$

$$\sum_{k=1}^{n} k^2 = 1^2+2^2+3^2+\cdots+(n-1)^2+n^2 = \frac{n\cdot(n+1)\cdot(2n+1)}{6}$$

$$\sum_{k=1}^{n} r^{k-1} = 1+r^1+r^2+r^3+\cdots+r^{n-2}+r^{n-1} = \frac{1-r^n}{1-r}$$

The random variable X_t represents the inflation rate "*t*" years in the future, and the expression $Var(X_t \mid X_0)$ is the variance around that rate. For our purposes, we need the variance of the inflation index at this future point; the index includes the variances of all of the annual inflation rates from the base time to the future period.

For this next step, we must remember that the inflation rate at a given point in the future is correlated with the inflation rates at subsequent points. This implies that the normal error terms e_i are included multiple times in the summation below.

$$(X_1+X_2+\cdots+X_n \mid X_0) = \sum_{j=1}^{n}(X_j \mid X_0) = \sum_{j=1}^{n}\left\{X_0\cdot r^j + \sum_{i=1}^{j}(b+\sigma\cdot e_i)\cdot r^{j-i}\right\}$$

If we make the substitution $S_n = (X_1+X_2+\cdots+X_n \mid X_0)$, then the random variable can be written more compactly as below.

$$S_n = E(S_n)+\sigma\cdot\sum_{j=1}^{n}\left\{\sum_{i=1}^{j} e_i\cdot r^{j-i}\right\} = E(S_n)+\sigma\cdot\sum_{j=1}^{n}\left\{e_{n+1-j}\cdot\sum_{i=1}^{j} r^{i-1}\right\}$$

In order to calculate the variance for this summation, we make use of the following relationships.

$$Var(S_n) = E(S_n^2) - E(S_n)^2 \qquad \text{and} \qquad E(e_i) = 0 \quad \forall i$$

The variance for the sum of these annual rates therefore requires the collapsing of the double summation.

$$Var(S_n) = \sigma^2 \cdot \sum_{j=1}^{n} \left(\sum_{i=1}^{j} r^{j-i} \right)^2 = \sigma^2 \cdot \sum_{j=1}^{n} \left(\frac{1-r^j}{1-r} \right)^2$$

This summation can be further simplified as shown below:

$$Var(S_n) = \frac{\sigma^2}{(1-r)^2} \cdot \sum_{j=1}^{n} \left(1 - 2r^j + r^{2j}\right)$$

$$= \frac{\sigma^2}{(1-r)^2} \cdot \left\{ t - 2r \cdot \left(\frac{1-r^n}{1-r} \right) + r^2 \cdot \left(\frac{1-r^{2n}}{1-r^2} \right) \right\}$$

Alternatively, for the special case in which $r = 1$, we can write

$$Var(S_n) = \sigma^2 \cdot \frac{n \cdot (n+1) \cdot (2n+1)}{6}.$$

The final step for the variance calculation is to translate the variance of the normal random variable X_t into the expression for the CV of the lognormal random variable.

We can accomplish this by making note of the following relationship[4].

$$CV(e^X)^2 = \frac{Var(e^X)}{E(e^X)^2} = e^{Var(X)} - 1$$

This provides the translation to all of the formulas given in section 3.1 of the paper.

By analogy, there is an expression for the [standardized] covariance of two random variables.

---

[4] As the reader might expect, this relationship holds when X is a normal random variable, but it is not generally true for other distributions.

$$Cov^*\left(e^X,e^Y\right) = \frac{Cov\left(e^X,e^Y\right)}{E\left(e^X\right)\cdot E\left(e^Y\right)} = e^{Cov(X,Y)}-1$$

For this covariance expression, we recall that we are looking for the relationship between two sums of random variables $(X_1+X_2+\cdots X_n)$ and $(X_1+X_2+\cdots+X_n+\cdots+X_{n+k})$, which we may again denote S_n and S_{n+k} for convenience.

$$S_n = E(S_n)+\sigma\cdot\sum_{j=1}^{n}\left\{e_{n+1-j}\cdot\sum_{i=1}^{j}r^{i-1}\right\}$$

$$S_{n+k} = E(S_{n+k})+\sigma\cdot\sum_{j=1}^{n+k}\left\{e_{n+k+1-j}\cdot\sum_{i=1}^{j}r^{i-1}\right\}$$

$$= E(S_{n+k})+\sigma\cdot\sum_{j=1}^{n}\left\{e_{n+1-j}\cdot\sum_{i=1}^{j+k}r^{i-1}\right\} + \sum_{j=n+1}^{n+k}\left\{e_j\cdot\sum_{i=1}^{n+k+1-j}r^{i-1}\right\}$$

The logic for calculating the covariance term $Cov(S_n,S_{n+k})$ is similar to that used for the variance above.

$$Cov(S_n,S_{n+k}) = E(S_n\cdot S_{n+k})-E(S_n)\cdot E(S_{n+k})$$

$$Cov(S_n,S_{n+k}) = \sigma^2\cdot\sum_{j=1}^{n}\left\{\left(\sum_{i=1}^{j}r^{i-1}\right)\cdot\left(\sum_{i=1}^{j+k}r^{i-1}\right)\right\} = \frac{\sigma^2}{(1-r)^2}\cdot\sum_{j=1}^{n}\left\{\left(1-r^j\right)\cdot\left(1-r^{j+k}\right)\right\}$$

$$= \frac{\sigma^2}{(1-r)^2}\cdot\left\{n-r\cdot\left(\frac{1-r^n}{1-r}\right)-r^{k+1}\cdot\left(\frac{1-r^n}{1-r}\right)+r^{k+2}\cdot\left(\frac{1-r^{2n}}{1-r^2}\right)\right\}$$

$$= \frac{\sigma^2}{(1-r)^2}\cdot\left\{n-r\cdot(1+r^k)\cdot\left(\frac{1-r^n}{1-r}\right)+r^{k+2}\cdot\left(\frac{1-r^{2n}}{1-r^2}\right)\right\}$$

For the special case in which $r=1$, we can write

$$Cov(S_n,S_{n+k}) = \sigma^2\cdot\sum_{j=1}^{n}\{j\cdot(j+k)\} = \sigma^2\cdot\sum_{j=1}^{n}\{j^2+j\cdot k\}$$

$$= \sigma^2 \cdot \left\{ \frac{n \cdot (n+1) \cdot (2n+1)}{6} + \frac{n \cdot (n+1)}{2} \cdot k \right\}.$$

This completes the derivation of the covariance terms given in section 3.2 of the paper.

As a final observation, we may note that the CV and Covariance expressions are dependent upon σ and r (the reversion parameter), but do not involve the intercept b or the starting point X_0. In other words, we can estimate the variance relative to the mean level of the reserves without having to know the current or long-term inflation rates.

Appendix B: Chain-Ladder ODP Model

The over-dispersed Poisson (ODP) model is useful to illustrate the ideas in this paper since it conveniently balances to the well known chain-ladder reserving method.

We define an incremental loss payment in year y and development period d to be distributed as ODP. The distribution is defined as follows:

Probability Function: $$\text{Prob}(c_{y,d}) = \left(\frac{\mu_{y,d}}{\phi}\right)^{c_{y,d}/\phi} \cdot \frac{e^{-\mu_{y,d}/\phi}}{(c_{y,d}/\phi)!}$$

Mean: $$E(c_{y,d}) = \mu_{y,d}$$

Variance: $$Var(c_{y,d}) = \phi \cdot \mu_{y,d}$$

The parameter ϕ is the "dispersion parameter" and represents a constant variance-to-mean ratio. This parameter will be assumed to be fixed and known, and constant for all accident years and development periods. Mathematically it is just a scaling factor that changes a standard Poisson distribution, defined on the integers {0, 1, 2, 3, ...} to an ODP distribution, defined on evenly spaced values $\{0, \phi, 2\phi, 3\phi, \ldots\}$.

The mean of each cell in the development triangle will then be defined as:

$$\mu_{y,d} = E(c_{y,d}) = \alpha_y \cdot \beta_d$$

In order to calculate the maximum likelihood estimation (MLE) values for these parameters, we need to evaluate the following expression.

$$LogLikelihood = \sum_{y=1}^{n} \sum_{d=1}^{n-y+1} \left\{ \frac{c_{y,d}}{\phi} \cdot \ln(\alpha_y \cdot \beta_d) - \frac{c_{y,d}}{\phi} \cdot \ln(\phi) - \frac{\alpha_y \cdot \beta_d}{\phi} - \ln((c_{y,d}/\phi)!) \right\}$$

However, since we are assuming that the dispersion parameter is fixed, we do not need to include it in our likelihood calculation. Instead, we use a quasi-likelihood (*QLL*) expression including only the portion of the LogLikelihood that is dependent on α_y and β_d.

$$QLL = \sum_{y=1}^{n} \sum_{d=1}^{n-y+1} \left\{ c_{y,d} \cdot \ln(\alpha_y \cdot \beta_d) - \alpha_y \cdot \beta_d \right\}$$

The derivatives with respect to the two parameters are set to zero.

$$\frac{\partial QLL}{\partial \alpha_y} = \sum_{d=1}^{n-y+1} \left\{ \frac{c_{y,d}}{\alpha_y} - \beta_d \right\} = 0$$

and

$$\frac{\partial QLL}{\partial \beta_d} = \sum_{y=1}^{n-d+1} \left\{ \frac{c_{y,d}}{\beta_d} - \alpha_y \right\} = 0.$$

The derivatives imply that the MLE values satisfy two conditions:

$$\sum_{d=1}^{n-y+1} c_{y,d} = \sum_{d=1}^{n-y+1} \alpha_y \cdot \beta_d \quad \text{and} \quad \sum_{y=1}^{n-d+1} c_{y,d} = \sum_{y=1}^{n-d+1} \alpha_y \cdot \beta_d \qquad \forall\, y,d.$$

That is, the row and column totals of the fitted values must equal the row and column totals of the original incremental triangle. Because these conditions do not result in a unique set of parameters, we can add one more constraint $\sum_{d=1}^{n} \beta_d = 1$, which results in $\alpha_1 = \sum_{d=1}^{n} c_{1,d}$. These constraints then mean that the MLE parameters are equivalent to the values in a standard chain-ladder reserve estimate.

This model can then be expanded to include estimates of trend based on the CPI:

$$\mu_{y,d} = E(c_{y,d}) = \alpha_y \cdot \beta_d \cdot CPI(y+d-1)^{\gamma}$$

$$QLL = \sum_{y=1}^{n} \sum_{d=1}^{n-y+1} \left\{ c_{y,d} \cdot \ln(\alpha_y \cdot \beta_d \cdot CPI(y+d-1)^{\gamma}) - \alpha_y \cdot \beta_d \cdot CPI(y+d-1)^{\gamma} \right\}.$$

We find from this expression that the following conditions must again be met:

$$\sum_{d=1}^{n-y+1} c_{y,d} = \sum_{d=1}^{n-y+1} \mu_{y,d} \text{ and } \sum_{y=1}^{n-d+1} c_{y,d} = \sum_{y=1}^{n-d+1} \mu_{y,d} \qquad \forall\, y,d.$$

We must also add the derivative with respect to the CPI curve, γ:

$$\frac{\partial QLL}{\partial \gamma} = \sum_{y=1}^{n}\sum_{d=1}^{n-y+1}\left\{c_{y,d}\cdot\ln(CPI(y+d-1)) - \alpha_y\cdot\beta_d\cdot\ln(CPI(y+d-1))\cdot CPI(y+d-1)^{\gamma}\right\} = 0$$

Which is equivalent to

$$\sum_{y=1}^{n}\sum_{d=1}^{n-y+1}\left\{c_{y,d}\cdot\ln(CPI(y+d-1))\right\} = \sum_{y=1}^{n}\sum_{d=1}^{n-y+1}\left\{\mu_{y,d}\cdot\ln(CPI(y+d-1))\right\}.$$

Unfortunately, there is no longer a convenient closed-form solution for calculating the model parameters, though it can be somewhat simplified using the relation below:

$$\beta_d = \frac{\sum_{y=1}^{n-d+1} c_{y,d}}{\sum_{y=1}^{n-d+1}\left\{\alpha_y\cdot CPI(y+d-1)^{\gamma}\right\}}.$$

The parameters in the model including the external CPI values must be estimated via an iterative calculation. This does not create any great difficulty in our model.

What is more interesting, however, is the relatively little improvement in model fit that is seen when the CPI values are introduced. It makes intuitive sense that loss payments should follow inflation, so why does introducing inflation as an explanatory variable add so little to the goodness of fit?

The answer is that a standard chain-ladder or MLE calculation is already estimating many parameters: one for each accident year α_y and one for each of the first *n*-1 development periods β_d (by constraining these to add to 1.00 we reduce the model by one parameter). This means that in a triangle with n years, we will have $n(n+1)/2$ data points to estimate $2n-1$ parameters; for a 10-year triangle we have 55 incremental payments to estimate 19 parameters. The effects of inflation are "buried" in our otherwise over-parameterized model.

To see this more clearly, we will introduce one more model in which the inflation rate i is assumed to be constant, and is estimated as a parameter of the model.

$$\mu_{y.d} = E(c_{y.d}) = \alpha_y \cdot \beta_d \cdot (1+i)^{y+d}$$

The quasi-likelihood function is given as follows.

$$QLL = \sum_{y=1}^{n} \sum_{d=1}^{n-y+1} \left\{ c_{y,d} \cdot \ln\left(\alpha_y \cdot \beta_d \cdot (1+i)^{y+d}\right) - \alpha_y \cdot \beta_d \cdot (1+i)^{y+d} \right\}$$

Taking the derivative with respect to the inflation rate i, we have

$$\frac{\partial QLL}{\partial i} = \sum_{y=1}^{n} \sum_{d=1}^{n-y+1} \left\{ \frac{c_{y.d} \cdot (y+d)}{(1+i)} - \alpha_y \cdot \beta_d \cdot (y+d) \cdot (1+i)^{y+d-1} \right\} = 0$$

Or equivalently,

$$\frac{\partial QLL}{\partial i} = \sum_{y=1}^{n} \sum_{d=1}^{n-y+1} \left\{ (y+d) \cdot \left(c_{y,d} - \mu_{y,d}\right) \right\} = 0$$

We may note that this condition for the derivative of the loglikelihood with respect to i will automatically be met if we first calculate α_y and β_d via the chain-ladder method (assuming no inflation), and then adjust the numbers as:

$$\alpha_y^* = \alpha_y \cdot (1+i)^{-y} \qquad \beta_d^* = \beta_d \cdot (1+i)^{-d}$$

Such that $\alpha_y^* \cdot \beta_d^* \cdot (1+i)^{y+d} = \alpha_y \cdot (1+i)^{-y} \cdot \beta_d \cdot (1+i)^{-d} \cdot (1+i)^{y+d} = \alpha_y \cdot \beta_d$

The MLE for a model with a <u>constant</u> inflation rate is therefore equal to the chain-ladder model with no inflation.

Appendix C: Variance and Covariance of Products of Random Variables

The general form of the variance of a single random variable X, and its covariance with a second random variable Y, are expressed in the following familiar equations.

$$Var(X) = E(X^2) - E(X)^2$$

$$Cov(X,Y) = E(X \cdot Y) - E(X) \cdot E(Y)$$

The variance of the product of these two random variables has a somewhat more complex expression:

$$Var(X \cdot Y) = E(X^2 \cdot Y^2) - E(X)^2 \cdot E(Y)^2$$

If X and Y are independent, then this can be re-written as follows.

$$Var(X \cdot Y) = Var(X) \cdot Var(Y) + Var(X) \cdot E(Y)^2 + Var(Y) \cdot E(X)^2$$

Proof:

$$Var(X) \cdot Var(Y) = \{E(X^2) - E(X)^2\} \cdot \{E(Y^2) - E(Y)^2\}$$

$$= E(X^2) \cdot E(Y^2) + E(X)^2 \cdot E(Y)^2 - E(X^2) \cdot E(Y)^2 - E(X)^2 \cdot E(Y^2)$$

$$= E(X^2) \cdot E(Y^2) - E(X)^2 \cdot E(Y)^2 - E(X^2) \cdot E(Y)^2 + E(X)^2 \cdot E(Y)^2 - E(X)^2 \cdot E(Y^2) + E(X)^2 \cdot E(Y)^2$$

$$= \{E(X^2 \cdot Y^2) - E(X)^2 \cdot E(Y)^2\} - \{E(X^2) - E(X)^2\} \cdot E(Y)^2 - \{E(Y^2) - E(Y)^2\} \cdot E(X)^2$$

$$Var(X) \cdot Var(Y) = Var(X \cdot Y) - Var(X) \cdot E(Y)^2 - Var(Y) \cdot E(X)^2$$

In a similar fashion, the covariance between two products of random variables can be calculated using the expression below.

$$Cov(X_1 \cdot Y_1, X_2 \cdot Y_2) = E(X_1 \cdot Y_1 \cdot X_2 \cdot Y_2) - E(X_1) \cdot E(Y_1) \cdot E(X_2) \cdot E(Y_2) \backslash$$

Again, if the X's and Y's are independent, the covariance formula can be re-written as follows.

$$\begin{aligned} Cov(X_1 \cdot Y_1, X_2 \cdot Y_2) &= Cov(X_1 \cdot Y_1) \cdot Cov(X_2 \cdot Y_2) \\ &+ Cov(X_1 \cdot X_2) \cdot E(Y_1) \cdot E(Y_2) + Cov(Y_1 \cdot Y_2) \cdot E(X_1) \cdot E(X_2) \end{aligned}$$

The proof follows a similar logic as above for the variance calculation.

Proof: $Cov(X_1 \cdot X_2) \cdot Cov(Y_1 \cdot Y_2)$

$$= \{E(X_1 \cdot X_2) - E(X_1) \cdot E(X_2)\} \cdot \{E(Y_1 \cdot Y_2) - E(Y_1) \cdot E(Y_2)\}$$

$$\begin{aligned} &= E(X_1 \cdot X_2) \cdot E(Y_1 \cdot Y_2) + E(X_1) \cdot E(X_2) \cdot E(Y_1) \cdot E(Y_2) \\ &- E(X_1 \cdot X_2) \cdot E(Y_1) \cdot E(Y_2) \\ &- E(Y_1 \cdot Y_2) \cdot E(X_1) \cdot E(X_2) \end{aligned}$$

$$\begin{aligned} &= \{E(X_1 \cdot X_2) \cdot E(Y_1 \cdot Y_2) - E(X_1) \cdot E(X_2) \cdot E(Y_1) \cdot E(Y_2)\} \\ &- \{E(X_1 \cdot X_2) \cdot E(Y_1) \cdot E(Y_2) - E(X_1) \cdot E(X_2) \cdot E(Y_1) \cdot E(Y_2)\} \\ &- \{E(Y_1 \cdot Y_2) \cdot E(X_1) \cdot E(X_2) - E(X_1) \cdot E(X_2) \cdot E(Y_1) \cdot E(Y_2)\} \end{aligned}$$

$$\begin{aligned} &= Cov(X_1 \cdot Y_2, X_2 \cdot Y_2) \\ &- Cov(X_1 \cdot X_2) \cdot E(Y_1) \cdot E(Y_2) \\ &- Cov(Y_1 \cdot Y_2) \cdot E(X_1) \cdot E(X_2) \end{aligned}$$

Q.E.D.

REFERENCES

[1] Ahlgrim, Kevin C, Stephen P. D'Arcy, Richard W. Gorvett, "Report on Modeling of Economic Series Coordinated with Interest Rate Scenarios," (Section 5) **2001** www.casact.org/research/econ

[2] Brehm, Paul J., "Correlation and the Aggregation of Unpaid Loss Distributions." CAS Forum 2002 (Fall), 1-23.

[3] England, Peter D. and Richard J. Verrall, "A Flexible Framework for Stochastic Claims Reserving," *PCAS* **2001**, Vol. LXXXVIII, 1-38.

[4] Hodes, Douglas M., Sholom Feldblum, Gary Blumsohn, "Workers Compensation Reserve Uncertainty," *PCAS* **1999**, Vol. LXXXVI, 263-392.

[5] Taylor, G.C., "Separation of Inflation and Other Effects from the Distribution of Non-Life Insurance Claim Delays," ASTIN Bulletin, Volume 9, 1977, 219-230.

[6] Working Party on Quantifying Variability in Reserve Estimates, "The Analysis and Estimation of Loss & ALAE Variability: A Summary Report." CAS Forum 2005 (Fall), 29-146.

Abbreviations and notations:
AR(1), autoregressive time-series model dependent on a single prior point
CPI, Consumer Price Index
CV, coefficient of variation = standard deviation / mean
ODP, Over-Dispersed Poisson

Biography of the Author

Dave Clark is Vice President and Actuary with American Re-Insurance. His paper "LDF Curve-Fitting and Stochastic Reserving: A Maximum Likelihood Approach" received the 2003 Reserves Call Paper Prize.

Acknowledgments

The author thanks Micah Woolstenhulme whose comments during the Working Party on Reserve Variability spurred on this research, and Denise Ambrogio, Chuck Thayer and Roger Hayne whose critiques helped the author improve the rigor of the paper in several ways. Any remaining errors in the final paper are, of course, the fault of the author alone.

Supplementary Material

An Excel spreadsheet including an illustrative example of the variance calculation is available upon request from the author.

"Adjusting & Other" Reserves According to the "Loss-Activity Method"

Paul B. Deemer, FCAS, MAAA

Abstract

This paper presents an additional method for calculation "adjusting & other" claim handling expenses. The method is contrasted with other methods present in actuarial literature.

Keywords. Reserving, ULAE, Adjusting and Other Expenses, Claim Handling.

1. INTRODUCTION

Within the scope of the reserving exercise, establishing reserves for Adjusting & Other ("A&O") loss related expenses generally comes last, and in many respects are an after thought. The primary reason for this is the necessity of having established proper reserve levels for losses before attempting to establish reserves for expenses related to managing these losses. An additional factor in the low attention given to this reserve component is the relatively few methods available. This paper presents an additional method for reserving these losses.

The actuarial literature addressing the task of reserving Adjusting & Other loss related expenses includes "aggregate" methods which use loss data at a high level and are, generally speaking, less rigorous. These methods include the classical "paid-to-paid" method and the variation proposed by John Kittel. Other methods, such as those offered by Wendy Johnson, Donald Mango and Craig Allen are more intensive in the usage of data and assumptions. The "loss-activity" approach is properly considered with the former, and thus a detailed comparison is offered. Nonetheless, I will present a discussion contrasting the "loss-activity" method to the Johnson method.

2. APPROACHES TO "ADJUSTING & OTHER" RESERVING

2.1 The Reserving Mindset

As a preface to this paper, it is necessary to frame the discussion with the most general parameters and motivation for loss reserving. The reserving exercise is an effort to reflect ultimate financial reality under all insurance obligations for which the enterprise is liable (losses) or due (premiums). At times the actuarial usage of professional jargon is loose, which can result in a redundant or misleading understanding. What is assumed in the following presentation is:

1. At a certain date, the organization ceases earning new exposures.
2. The organization <u>is not</u> responsible for events occurring beyond that point in time.
3. The organization <u>is</u> responsible for events occurring prior to that point in time, even if made aware of them after that date.
4. "Runoff", as used below, reflects this situation: the organization is settling liabilities previously incurred and not incurring new obligations. This term is not used to imply a "fire-sale" of liabilities, "discounting" or any other term related to an insurance insolvency.

With this backdrop, it can be seen that the reserving mindset is focused on the ultimate answer when all uncertainties and contingencies have emerged (for losses, all claims are closed). As time passes the financial uncertainties of which reserving is concerned will move, to an ever larger degree, from estimate to actual. The reserving exercise is an attempt to determine the ultimate values at point where uncertainty remains.

It is important to establish that a valid methodology for reserving "Adjusting & Other" loss related expenses should explicitly recognize that these reserves are for expenses which are second-order in relation to underlying losses. Stated another way, unless we have a reported claim, a notice of loss or efforts expended in relation to a reported *potential* claim, there can be no claim handling expenses. First we have to have claims. Further, it is fundamentally intuitive that "adjusting & other" costs have a linear relationship to claim activity. The more claims being reported, the larger the claim function will need to be to handle the volume and vice-versa. This is embedded in the most widely used "adjusting & other" reserving method, the classical "paid-to-paid" method.

2.2 Destination: "A&O" Cost Per Unit of "Loss-Activity"

Total "A&O" expenditures in a given year are a known item available from accounting exhibits. What is needed, to make these expenses useful within the context of reserving, is an accurate proxy of what to contrast these costs with. As claim department salary is the vast majority of the "A&O" expense, the crucial task is finding a numerical proxy for the claim department's use of their time. For this method, I have called this proxy "loss-activity". "Loss-Activity" is defined to be the sum of five components:

1) **Current Accident Year Paid + Case Reserve Reported Losses**: These are claims incurred and reported in the most recent accident year. For property claims, they will generally be reported and closed within the year and the reported value is a good representation of their value. For casualty claims, many of these will not be closed at the end of the year and will be subject to future revisions. In the context of financial reporting, an example would be the direct and assumed loss payments plus the direct and assumed case basis unpaid losses (Schedule P, Part 1) for Accident Year 2004 in the 2004 Annual Statement.

2) **Current Accident Year Paid Defense and Cost Containment (DCC)**: This is the DCC (formerly ALAE) component of newly reported claims; as no reserves are established at the case level for this component, paid data suffices. If the claim practice was to establish case reserves for DCC, they should be included here. In the context of financial reporting, an example would be the direct and assumed defense and cost containment payments (Schedule P, Part 1) for Accident Year 2004 in the 2004 Annual Statement.

3) **Prior Accident Years' Reported Losses**: This represents the reporting of lagged IBNR claims or adjustment in value of claims reported in previous years. Note: as with component 2, we would also want to include DCC changes if case reserves were present for that component. In the context of financial reporting, the approach would be similar to that for component (1); only we are looking for the change in paid losses + case reserves for 2003 and prior accident years from the 2003 Annual Statement to the 2004 Annual Statement.

4) **Prior Accident Years' Paid Losses**: This represents the payment in the current calendar year on claims which are from prior accident years which have not yet closed or in some cases had not yet been reported. In the context of financial reporting, the approach would be similar to that for component (3); only we are looking for the change in paid losses for the 2003 and prior accident years from the 2003 Annual Statement to the 2004 Annual Statement.

5) **Prior Accident Years' Paid Defense and Cost Containment**: This represents the payment of DCC on claims which are from prior accident years and have not yet closed. In the context of financial reporting, the approach would be similar to that for component (3);

only we are looking for the change in paid DCC for the 2003 and prior accident years from the 2003 Annual Statement to the 2004 Annual Statement.

While an initial look at these components leaves one wondering as to the reason for what appears to be an artificial summation of data, further consideration using first principles will reveal that these are an excellent proxy for all of the activity of a claim department within a given year. It should be noted that all loss and expense components are included gross of salvage, subrogation and reinsurance recoveries.

2.3 Application

The next step is to relate the total claim handling expense in a year to the total "loss-activity" and thus get a ratio which tells us the "A&O" cost per unit of activity.

$$\text{"}A\ \&\ O\text{"}\ Cost\ Ratio = \frac{Total\ \text{"}A\ \&\ O\text{"}\ Expenses}{Total\ \text{"}Loss - Activity\text{"}}$$

To derive the indicated "A&O" reserve, we multiply the Cost Ratio times the anticipated future "loss-activity". Revisiting the definition of "loss-activity", we see that in a prospective look the first two components fall away.

1) Current Accident Year Reported Losses = 0

2) Current Accident Year Paid Defense and Cost Containment = 0

3) Total Unreported Losses

4) Total Unpaid Losses

5) Total Unpaid Defense and Cost Containment

For purposes of reserving, we no longer have losses occurring; all losses have occurred and what remains is the reporting of IBNR claims and the settling of claims which have and have note been reported. Thus, components (1) and (2) are zero. Component (3) is equal to the calculated ultimate losses less the paid + case reserve losses already reported; component (3) includes pure IBNR claims and development on reported claims. The fourth component is equal to the calculated ultimate losses less paid to date losses, which includes all pure IBNR claims and the settling of claims which have been reported. Component (5) is the calculated ultimate DCC less the paid to date DCC. Collectively, components (3) + (4) + (5) are the anticipated future "Loss-Activity". The A&O reserve is the product of the future "Loss-Activity" and the calculated Cost Ratio:

"A&O" Reserve = Cost Ratio * Anticipated "Loss-Activity"

3. APPLICATION OF THE METHOD

I now present this method applied to the loss experience of a medium sized insurance company writing a mix of property and casualty coverages across both commercial and personal lines. As indicated, all numbers can be easily located within an actuary's reserving work papers or a company's Schedule P data.

Historical Loss Activity ($ Millions)
All Data Gross of Salvage and Subrogation

Calendar Year	Current AY Reported Losses (1)	Current AY DCC Paid (2)	Prior AYs Losses Reported (3)	Prior AYs Losses Paid (4)	Prior AYs DCC Paid (5)	Total "Loss Activity" 1+2+3+4+5
2000	207.3	0.8	16.1	74.9	7.4	306.5
2001	241.6	1.9	32.4	106.2	8.1	390.2
2002	225.1	1.6	36.1	105.3	10.9	379.0
2003	244.4	2.0	53.7	118.8	14.3	433.2
2004	281.6	2.1	33.2	120.1	19.1	456.1

All of the "Loss-Activity" components suggest an organization which is growing, which is indeed the case. To pull these numbers from actuarial reserving work papers, the simplest method is to take the difference of the two most recent diagonals in the loss triangles. For reported losses, the most recent accident year is allocated to column (1), and the remainder of the incremental diagonal to column (3). For paid losses, the most recent accident year is disregarded and the remainder of the incremental diagonal is in column (4). For paid DCC, the most recent accident year is in column (2) and the remainder of the incremental diagonal is in column (5).

Historical "Adjusting & Other" Cost Ratios			
Year	**"Adjusting & Other" Paid ($M)**	**Total "Loss Activity" ($M)**	**Cost Ratio**
2000	17.68	306.5	5.8%
2001	18.69	390.2	4.8
2002	21.32	379.0	5.6
2003	24.93	433.2	5.8
2004	26.16	456.1	5.7

The 2001 cost ratio is clearly an exceptional value. Upon closer inspection, this is largely due to the presence of significant property-catastrophe losses (see the "Current AY Losses Reported" column). These losses are present in the denominator of the Cost Ratio calculation and serve to lower the indicated ratio. This relationship (between "adjusting & other" costs and shock losses in a calendar year) is a distortion to this reserving exercise. Thus, it is preferable to remove the effect of these events from both the losses and the "adjusting & other" expenses. If data is not available to remove the impact of the shock loss event, data points containing significant property catastrophe or other aberrational losses should be given diminished (or even zero) credibility when selecting a final Cost Ratio. For purposes of this demonstration, property-catastrophe losses have been excluded from the calculation of total "loss-activity".

Historical Cost Ratios (Adjusted for CY 'Shock' Loss Activity)			
Year	**"A&O" Paid**	**Adjusted "LOSS ACTIVITY"**	**"A&O" Cost Ratio**
2000	17.22	298.5	5.8%
2001	18.49	344.6	5.3
2002	21.32	379.0	5.6
2003	24.93	433.2	5.8
2004	26.16	456.1	5.7

If the method and proxies are valid, we would expect to find a "per-unit" cost that is not trending upward or downward in any material way, as is the case with this approach. Using a straight average of the five data points, we have a cost of $0.0564 per $1.00 of "loss-activity". With our Cost Ratio in hand, we next need to calculate the amount of anticipated future "loss-activity" in order to produce an indicated "A&O" reserve. The future loss activity is easily attained from the reserving work papers. As noted, the first two components of "loss-activity" are zero since the reserving exercise is not concerned with obligations incurred in the future.

1) Current Accident Year Reported Losses = 0

2) Current Accident Year Paid Defense and Cost Containment = 0

3) Total Unreported Losses = Ultimate Losses - Paid-to-Date Losses - Case Reserve Losses = $100.0 M

4) Total Unpaid Losses = Ultimate Losses – Paid-to-Date Losses = $438.0 M

5) Total Unpaid Defense and Cost Containment = Ultimate DCC – Paid to Date DCC = $69.5 M

As noted above, if the company's practice is to establish case reserves for DCC expenses, then the DCC component would be handled identical to losses. For the company in the example, case reserves for DCC are not established. Adding up the components, there is 100 + 438 + 69.5 = 607.5 M in future "loss-activity". This is larger than the company's carried reserves, due to the inclusion of both unreported and unpaid losses. The product of this anticipated future "loss-activity" and the Cost Ratio is the indicated reserve under this method:

Future "Loss-Activity" X Cost Ratio = 607,500,000 X 5.64% = 34,263,000

3.1 Apparent Difficulties Using the "Loss-Activity" Approach

3.1.1 Inclusion of both *reported* and *paid* losses for the older accident years.

Certainly there may be some losses in both buckets, but for claims which are still open after a year this is arguably very appropriate. These claims may litigate and settle in a future year after the reserve is established. This requires claim staff resources. Thus if we only use reported losses we would be blind to this activity unless the established reserve was to change. Additionally, if we did not use *reported losses* we would not reflect the establishment of IBNR claims, which is very material to the casualty lines of business. For these reasons, the method uses both paid and reported losses, consistent with the reasoning presented by John Kittel in his 1986 paper. Shifting from a "going-concern" mindset to a "runoff" mindset makes it apparent that establishment of the case reserve and the payout on that reserve are both activities which require claim department staff to effect.

3.1.2 Validity of the marginal cost applied prospectively.

It could be argued that the newly reported claims represent a disproportionate cost in the denominator of the Cost Ratio calculation. The argument is an implicit question as to whether the Cost Ratio on "small" claims which open and close in a short time frame is the same as the Cost Ratio on large claims that are open longer. I do not think there is a material distortion - if for no other reason than a significant portion of claim department time in a given year is spent disposing of a high volume of newly reported claims. Thus, it is not unreasonable to suggest that even if the nominal "adjusting & other" dollars per claim is drastically different, the cost relative to claim value (Cost Ratio) is still reasonable for both. It is sensible to argue that bigger claims are going to naturally involve more adjuster time, along with other A & O costs. But the sheer magnitude of the claim value will cause this method to post A & O reserves accordingly. Additionally, it must be pointed out that the logic behind this question breaks apart as you move farther away from an average claim value (in either direction) and actually could be thus interpreted to suggest the opposite.

3.1.3 Properly handling inflationary influences.

We point out that both the numerator and the denominator of the Cost Ratio are subject to inflationary pressures, so there is a degree of "canceling out" which makes gives the ratio a degree of immunity from this type of distortion. Over time, it may be argued, the inflationary pressures on losses are stronger than the inflationary pressures on adjuster salaries and other A&O costs. This may be true, but it disregards efficiency gains due to technology and training. If this was an issue, over time it would serve to be pulling the Cost Ratio downward and with the historical data analyzed in developing this method, this has not been the case.

3.2 Other Methodologies

3.2.2 Paid-to-Paid Method Comparison

The most widely used method in Property-Casualty actuarial practice for establishing A&O reserves is the so-called "paid-to-paid" method. It involves comparing paid A&O expenses to paid losses for the same calendar year. This ratio is then applied to the unpaid losses to determine the needed A&O reserve. Generally there is a significant adjustment necessary to reflect the fact that some of the "A&O" expenses on open claims has already been borne in the establishment of the case reserves. Thus, an assumption of what percentage of the "A&O" cost is incurred at opening of the claim is needed. The full paid-to-paid ratio is applied to true IBNR, and the paid-to-paid ratio times the adjustment factor is applied to the case reserves. This is problematic, and widely known to be so. For reasons of contradistinction and not novelty, I point out:

1. The denominator of the paid-to-paid ratio (paid losses) is a rough proxy for claim department activity. As we isolate scenarios involving operational changes this method of quantifying claim department activity breaks down. For example, in strong growth scenarios, the paid losses increase slowly whereas the paid A&O grows generally in line with the earned exposures, thus increasing the ratio. The artificially high ratio is then applied to the unpaid losses, which are a correct representation of financial reality - including the exposure growth. Without correction, this will lead to an overstated A&O reserve. There are other distortions to which the paid loss denominator is susceptible such as making no recognition for the effort expended on older claims unless a payment is made. This is material for casualty lines of business. The denominator of the Cost Ratio is five components which touch on claim department activity. Paid losses is one of these, but handling IBNR claims, paying ALAE and the case reserves for the current accident year reported losses are also a part. Looking at scenarios involving operational change this method of quantifying claim department activity holds up better. Using the same example, in strong growth scenarios, the current accident year reported losses (a dominant piece of the denominator) increase in line with the exposure growth, consistent with the change in A&O paid.

2. The "percentage paid at open" adjustment factor is not at this time, to my knowledge, prospectively quantifiable with any scientific method. A look at the compliment of this ratio ("percentage *not* paid at open") more quickly leads one to the conclusion that this cannot be quantified with any precision. To do so would involve more than just a collection of motion studies. With these drawbacks in mind, it is seen to be a highly

suspect component of the method, which effectively (and arbitrarily) reduces the paid-to-paid ratio in order to apply it to case reserves where some A&O expense has already been expended. This drawback can result in significantly distorted A&O reserve indications. Traditional usage has gone with the 50/50 rule: 50% "A&O" incurred at open, 50% at closed. This is clearly violated by partial payment lines such as workers compensation and lines involving tremendous litigation such as General Liability. The "loss-activity-method" does not involve speculation about the amount of A&O cost incurred when the claim was opened. This is because the future "loss-activity", to which the Cost Ratio is applied, includes (1) unreported losses, (2) unpaid losses and (3) unpaid Defense and Cost Containment. This acknowledges, and implicitly assumes, that the dominant A&O cost (claim department salary) has a fairly constant marginal cost. Indeed this can be seen from first principles; claim department time has a (relatively) fixed cost since as a functional unit it is a collection of salaried professionals. These three buckets do well quantifying future needs of claim department time to establish reserves on unreported claims as they come in, adjust the claim values as situations warrant and payout all unpaid loss and DCC reserves. Taking a close algebraic look at the "loss-activity" method, the 50/50 proportion can be found present in the handling of the IBNR segment. We note that this is because the Cost Ratio is applied to both the "unreported" and the "unpaid" losses, which for the "pure" IBNR component are identical. But the two methods are working in opposite directions. For the traditional method, the paid-to-paid ratio is multiplied by 50% and applied to the case reserves. This implies many things, but the most obvious is that the current calendar year claim activity was involved with new claims half the time. In a steady state, this may be generally valid, but outside of stable parameters, it is problematic. The "loss-activity" method allows the data to specify the weighting as the data emerges.

3. Beyond the difficulty with quantifying the "percentage paid at open", a change in case reserve adequacy poses further challenges to the classic approach. It is crucial because the method assumes the case reserves are a good representation of the effort already expended to investigate and establish reserves on reported claims. When faced with shifting case reserve adequacy, it is necessary to "lift the hood" on the method. But where? Infusion of an adjustment becomes very arbitrary, both in terms of technique and of calibration. Further, it will compound the difficulties highlighted above for assuming a percentage paid at open. Changes in case reserve adequacy are a

comparatively minor issue for the "loss-activity" method. The method assumes the overall loss reserving exercise has correctly detected the change in adequacy and thus the ultimate losses (and the corresponding unpaid losses) are correct. It is acknowledged that in terms of the denominator of the "A&O" Cost Ratio, the current year reported losses, and reported losses from prior years would be affected by a change in case reserve adequacy. But this change is partially self-correcting as claim department time is involved in re-evaluating case reserves previously established and the future "loss-activity" against which the Cost Ratio is applied is diminished (assuming the ultimate losses are correct). The paid component of "loss-activity" is entirely unaffected. This means the Cost Ratio enjoys a degree of immunity to changes in case reserve adequacy. In other words, if the ultimate losses are correct at the outset, the method will generally roll with the punches successfully.

3.2.3 Johnson Method Comparison

Another technique used in actuarial practice is the method expounded by Wendy Johnson. Her method uses a numerical proxy for claim department activity which is philosophically similar to the approach of the "loss-activity" method. The "weighted number of open claims" is the number of older claims open at the beginning of the year along with the number of claims reported during the year

1. The Johnson method's marginal cost must be trended forward, since it is a cost per *open claim,* which places reliance on a trend factor. We note that the forces appearing as "trend" in the curve fit to the marginal cost for each year are beyond traditional inflationary effects and not necessarily the same in the event of runoff, as discussed above in looking at the proper mindset for reserving. The "loss-activity" method assumes the reserving exercise has properly estimated ultimate losses and thus trending of the Cost Ratio is not warranted.

2. I also note that in application, the number of claims reported during the current year will be the vast majority. Thus, unless working with a severely protracted line of insurance (such as her example with medical malpractice), the year end pending claim counts will receive very little weight. This is further bolstered when looking at the distribution of paid losses. For companies with significant casualty portfolios, the current calendar year paid losses will have a large percentage of payouts from prior years' claims.

3.2.4 Kittel Method Comparison

The last technique I will look at is the method developed by John Kittel. His method is a variant on the paid-to-paid method. In his paper, he acknowledges many theoretical difficulties with the paid-to-paid approach, and concentrates on the flaws with paid losses as the proxy. He proposes replacing this with an average of reported losses and paid losses, particularly in instances where the organization is growing.

As it is a subtle variant of the paid-to-paid approach, my criticisms of the main method apply as well with one limited exception. The Kittel method will be more "accurate" than the paid-to-paid method in certain scenarios such as strong growth or intensive inflation, where the accruing unpaid liabilities exceed the movement in paid losses.

4. CONCLUSIONS

I have a presented a new method for calculating "adjusting & other" loss expense reserves. The method is technically sound and simple to calculate, and it is hoped that it will find a place in the reserving actuary's methods for this reserving task.

Acknowledgment

The author gratefully acknowledges the assistance of Patrick N. Tures, FCAS, Cathy Staats, FCAS, and Brian K. Ciferri, FCAS who have contributed to the paper by providing helpful critique at the various stages of its writing.

5. REFERENCES

[1] Kittel, John, "Unallocated Loss Adjustment Expense Reserves in an Inflationary Environment," *Casualty Actuarial Society Discussion Paper Program*, 1981, 311-331.

[2] W. Johnson, "Determination of Outstanding Liabilities for Unallocated Loss Adjustment Expense," *PCAS* 1989, Vol. LXXVI, 111-125.

[3] Mango, Donald and Craig Allen, "Two Alternative Methods for Calculating the UnallocatedLoss Adjustment Expense Reserve," *Casualty Actuarial Society Forum*, 1999, 233-258.

Abbreviations and notations

A&O, Adjusting & Other Expense (ULAE)
IBNR, Incurred But Not Reported
DCC, Defense and Cost Containment Expense

Biography of the Author

Paul Deemer is the Chief Actuary of North Pointe Holdings, Inc. in Southfield, Michigan. His responsibilities include reserving, financial modeling and actuarial research initiatives. He has a degree in Actuarial Science and Economics from Eastern Michigan University. He is a Fellow of the CAS and a Member of the American Academy of Actuaries. He participates on the CAS examination committee.

Loss Reserving Using Claim-Level Data

James Guszcza, FCAS, MAAA
Jan Lommele, FCAS, MAAA

Abstract

While the actuarial literature devoted to stochastic loss reserving has been developing at an impressive rate, much of this literature has been devoted to the statistical analysis of summarized loss triangles. This restriction limits the benefits that modern statistical techniques can bring to the subject of loss reserving. This paper will sketch one possible framework for estimating future claims payments using claim-level data. The first part of the paper will discuss the use of covariates (or "predictive variables") to improve one's estimates of future payments, especially in cases where the mix of business being analyzed has changed over time. The second part of the paper will describe how the bootstrapping technique can be applied to claim-level data to estimate reserve variability.

Keywords. Reserve Variability; Future Payment Variability; Generalized Linear Model; Over-Dispersed Poisson Model; Bootstrap; Claim-Level Data; Covariates; Predictive Variables; Changing Mix of Business; Chain-Ladder

INTRODUCTION

The recent actuarial literature has enjoyed a growing discussion of statistical methods for performing loss reserve analyses. This discussion has increased the statistical rigor of the subject, and has expanded the set of tools available for estimating reserve variability.

However, much of this recent discussion has been devoted to the statistical analysis of summarized loss triangles. We feel that this limits the potential improvements that predictive modeling can bring to the subject. We will focus on two reasons here. First, summarized loss triangles do not allow the analyst to incorporate predictive variables in his or her reserve analysis. Second, using summarized data limits the accuracy with which an analyst can estimate the variability of his or her loss reserve estimates. It is reasonable to expect that by not "summarizing away" the size-of-loss and loss development information implicit in (un-summarized) claim-level data, potentially better point and variability estimates can result.

Many of the comments in the Discussion of England and Verrall's recent survey paper on stochastic loss reserving [4] expressed this sentiment. Shah's comment is representative:

> The triangulation data that these [Generalized Linear Modeling] techniques have been applied to are just a consequence of history. They come from an era when computing power was expensive. Therefore, I question the value of actually applying such techniques to such limited data. Such sophisticated techniques may be more useful if applied to the underlying claims data, as has been alluded to by several speakers. In view of this, there is a danger that the

> results may be viewed as more scientific than they really are, and may be given more credibility than is truly justified for them.

Tripp's comment also seems to us to be on the mark:

> Why do we throw away information? ... Looking at the life side of our profession, you realise that work like this takes place at policy level detail. If you look within the general insurance part of the actuarial profession, there is a body of thinking that has grown up around premium rating and a body of thinking that has grown up around reserving. Are we getting 'over-siloed'? Could aspects of the methodology and the thinking that has gone into using GLMs for premium rating be brought more into play when it comes to reserving, where, at present, we tend to use aggregated claims data? I wonder whether we are missing out on using information that is available from exposure descriptions and from the circumstances of individual claims.

Motivated by the concerns expressed in these quotes, this paper is an attempt to develop the idea that using un-summarized data will allow one to unleash the full power of modern predictive modeling techniques on the problem of estimating future claim payments. The goals of improving one's reserve point estimates as well as variability estimates will be discussed sequentially in the two parts of this paper.

In Part I we review the well known shortcoming of traditional reserving methods when applied to books of business that have changed over time. A danger of using summarized loss triangles is that they can mask heterogeneous loss development patterns. They also prohibit the use of predictive variables that might be correlated with loss development. We sketch a reserving technique – inspired by the chain-ladder method – that operates on claim-level data. Using simulated data we illustrate how this technique can reflect heterogeneous loss development patterns that the chain ladder misses, resulting in an improved estimate.

We believe that the potential for improved estimates of future loss payments is sufficient motivation to consider the use of claim-level data for reserving. Doing so obviously requires additional effort (not to mention specialist software that goes beyond spreadsheets). But, as Part II of this paper will discuss, it brings a significant side benefit as well.

Namely: once we have claim-level data available for analysis, we can employ the bootstrapping technique (a type of simulation that involves repeatedly sampling with replacement from one's data) to easily compute confidence intervals around our estimates of outstanding losses. Indeed bootstrapping will give us estimates of the entire distribution of our outstanding loss estimator, no matter how complex.

Bootstrapping has been discussed in the recent literature as a promising avenue for estimating reserve variability. But because of the summarized loss triangles that serve as a starting point for most current discussions of reserving, the resampling step of bootstrapping is typically applied to the residuals of various models fit to loss triangles. The idea pursued here is to resample the underlying data points, and then apply one's chosen reserving technique to each of the resulting pseudo-datasets. This is a flexible and perhaps conceptually simpler method of bootstrapping. Also, because its resampling step occurs prior to the building of any model, the pseudo-datasets that it employs are not in themselves dependent on the correctness of the model being fit to the data.

PART I: SUMMARIZED DATA AND THE PROBLEM OF A CHANGING MIX OF BUSINESS

A common criticism of traditional loss reserving techniques is that they can be slow to incorporate changes in the company's mix of business into their estimates of outstanding losses. This is the point of the actuarial road trip joke involving the salesperson with his foot on the gas, the underwriter with his foot on the brake, and the actuary navigating by looking out the rear window.

Bornhuetter and Ferguson state the problem well in "The Actuary and IBNR" [1]:

> The product mix can be an important factor, not so much because two somewhat dissimilar items are combined, but because they may have different rates of growth. For example, a company may have personal and commercial automobile loss development experience combined over the years although, if it were looked at separately, commercial business would require higher loss development factors. As long as the relative exposure between the two categories remains constant there is no problem; however, picture the situation if personal automobile increased at a 5% annual rate while commercial automobile, although relatively small, is growing at a 25% annual rate.

The obvious thing to do in such a situation would be to analyze commercial and personal auto reserves separately. That is, divide the data into two separate loss triangles and proceed as usual. This is helpful as far as it goes, but the approach has its limits. Bornhuetter and Ferguson continue:

> Of course, the volume of data is an important factor in determining what kinds of breakdowns of the data are feasible. If the data are subdivided so finely that most groups have only a small volume of data, the subdivisions may accomplish nothing useful. Or to quote Mr. Longley- Cook's delightful analogy, "We may liken our statistics to a large crumbly loaf cake, which we

> may cut in slices to obtain easily edible helpings. The method of slicing may be chosen in different ways-across the cake, lengthwise, down the cake, or even in horizontal slices, but only one method of slicing may be used at a time. If we try to slice the cake more than one way at a time, we shall be left with a useless collection of crumbs."

For example, it might be nice to set up separate reserve analyses by both coverage and region. But even adding the single additional dimension of "region" might significantly diminish the credibility of the data and thereby threaten the integrity of one's outstanding loss estimate. The goal of the first part of this paper is to suggest a way beyond this impasse.

Our discussion of changing mixes of business is intended only to motivate the method discussed below. Hopefully the method's usefulness is not restricted to this scenario. For example, it might also be useful when, for example, a company moves into a new region or two companies merge.

ENTER PREDICTIVE MODELING

In modern terms, Longley-Cook's image of the crumbly cake is an illustration of the bias-variance tradeoff in predictive modeling. Stated briefly, a complex model (or multiple models fit on sub-segments of the data) will make predictions that are less biased, but at the same time less certain – i.e., more variable – than a simpler model. The tradeoff is that our model should have sufficient complexity to reflect true statistical regularities in the data (thereby reducing bias), yet not have so much complexity that random patterns in the data overwhelm the model and lead to unreliable results (high variance). This is perhaps a special case of Einstein's dictum, "Everything should be made as simple as possible, but not simpler."

An analogy with ratemaking might be helpful. Consider a simple rating plan with the following rating factors:

- Age {<26, 26-50, >50 years}
- Credit {bad, average, good}
- Claim in past 3 years {yes, no}

This rating plan has 3·3·2=18 cells. The most naïve – and over-parameterized – way to proceed would be to simply estimate the loss ratio relativity of each of these cells and base one's rating factors on these parameters. Note that this is equivalent to fitting a regression model with 17 indicator variables. But as Longley-Cook warns, the data in each of these cells is unlikely to have sufficient credibility to produce stable results. Therefore the variance around the resulting rating factor estimates will be large.

For this reason, the modern approach to ratemaking is to employ Generalized Linear Models [GLMs]. Rather than estimate 3·3·2-1=17 parameters, a GLM model in this scenario would estimate 2+2+1=5 parameters. Extending Longley-Cook's analogy, we now get to have our cake and make multivariate estimates with it too. Rather than estimate each of the 17 rating factors each with its own "crumb" of data, we use the loaf to estimate a more modest 5 parameters.

There are three major advantages of deriving one's rating factors from the parameters of a multivariate model, rather than estimating them directly from small "crumbs" of data:

- The resulting rating factors will have less variability (less parameter risk).
- A larger number of rating factors can be used without running into Longley-Cook's "crumbly cake" problem.
- Factors such as Age and Credit can be treated as continuous predictive variables, rather than being arbitrarily divided into discrete bins.

Returning to loss reserving, it is good and accepted practice to perform separate reserve analyses by line of business and by such important subdivisions as Workers Comp Medical vs. Indemnity claims. As we have discussed, this can only be taken so far. But what if (a) claim development patterns vary by a multitude of factors such as Report Lag, Credit Score, Prior Claim, Policy Age… and (b) the mix of business measured by these factors has changed over time? As Bornhuetter and Ferguson point out, it is essential to reflect this shifting mix of business in one's analysis. But as Longley-Cook points out, dividing the data by many of these dimensions will quickly lead to serious credibility problems.

In the light of the ratemaking analogy above, it is perhaps natural to suggest that the way forward is to somehow incorporate a multivariate predictive model into one's reserve analysis. We will sketch one such model below. This model is offered very much in the spirit of taking a first step. We expect that it could be improved or replaced with a better one. Nevertheless, we hope that sketching a sample multivariate loss reserving model that admits covariates will spark further thoughts on the subject.

THE LEVEL OF DATA NEEDED

Multivariate loss reserving requires that one analyze disaggregated data, at the policy or claim level, rather than summarized loss development triangles. The reason for this is clear: predictive variables such as Age, Credit, and Prior Claim pertain to the policy that made the claim. To incorporate policy-level variables such as these, policy-level data must be used in the analysis. There is no way to "attach" such covariates to summarized data. Similarly, if we wish to incorporate variables such as Report Lag or Injury Type into the analysis, claim-level data must be used. Traditional loss triangles do not allow one to use this potentially useful predictive information.

To summarize what has been said so far:

- The traditional approach of separating one's data and performing separate analyses on the resulting loss triangles is an incomplete answer to the problem of a shifting mix of business.
- A plausible approach to this problem is to incorporate covariates into one's reserving technique – that is, build a multivariate reserving model.
- Doing so requires that we use data at the policy or claim (or indeed claimant) level.

For the remainder of this paper phrases such as "reserving using claim-level data" will serve as shorthand for "reserving using policy- or claim- or claimant-level data".

MODEL DESIGN

In this section we propose a claim-level generalization of the simple chain ladder reserving method. As stated above, this is merely one of many possible starting points. For all of its faults, the chain ladder has the virtues of being simple and familiar. Generalizing the chain ladder therefore gives us an intuitive way of illustrating the benefits of using claim-level data to estimate future claim payments.

As discussed above, we assume we have data at the policy, claim, or claimant level. Of course, the finer the level of summarization of one's data, the broader the array of predictive variables one can include in one's model. Deciding on the level of data is a practical decision that does not substantially affect the discussion below. Let us therefore assume that our data is at the claim level.

We therefore assume that we have a database with one record per claim, and multiple variables on each record. These variables can be categorized into three types:

- Predictive variables (Credit Score, Injury Type, Policy Age...)
- Target variables (Loss at 24 months, Loss at 36 months...)
- Informational variables (Accident Year, Zip Code, Agent Number...)

The "informational" variables can sometimes be used to derive further predictive variables (e.g., by using zip code to match such demographic variables as Population Density onto the records). Other times, they are used simply for analytic purposes (e.g., displaying total losses by accident year).

Let us establish some notation. We attempt to be consistent with the notation of England and Verrall. Let C_j denote cumulative losses evaluated as of j months. For example C_{24} denotes the losses (associated with a particular claim) evaluated as of 24 months. $\{C_j\}$ will serve as the target variables in our model design.

Let $\{X_1, X_2, \ldots, X_N\}$ represent the predictive variables. Each value of each predictive variable X_i will appear on each claim-level record. We also assume that the values of each of the predictive variables are measured either at policy inception, or at the claim report date (whichever is appropriate).

Let U_k denote the total ultimate losses for accident year k, summed across all policies: $U_k = \sum C_\infty$. Let R_k denote the outstanding losses (ultimate losses minus losses paid to date) for accident year k. Let U and R denote the sums of U_k and R_k respectively across all accident years. The goal of loss reserving is to calculate an estimate r of R as well as an estimate of variability of, or confidence interval around, R. R is often referred to as a "reserve estimate", but to distinguish it from the quantity that is actually booked in the financial statements, it is probably better to call it the "total outstanding losses" or "total future payments" (see [2]). In the remainder of this paper, the three terms will be used synonymously.

In predictive modeling it is typically the case that we are presented with a single target variable Y (such as pure premium or claim frequency or size of loss) and multiple predictive variables $\{X_1, X_2, \ldots, X_N\}$. We might fit a GLM model of the form:

$$g(\mu) = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_N X_N \equiv \boldsymbol{\beta} \cdot \boldsymbol{X}$$

Where μ denotes E[Y], the expected value of the target variable Y; and g(·) is the link function.

Here, the situation is not so simple. For one thing, we are presented with multiple target variables $\{C_1, C_2, \ldots, C_J\}$ rather than a single target Y. In addition, this (single) target variable is typically the quantity we are ultimately interested in predicting. Here, we are interested in predicting either losses at ultimate or losses as of a certain development period, such as 10 or 20 years. Let us assume that for practical purposes, C_J represents losses at ultimate. That is, $C_J \approx C_\infty$. (That is, let us assume that no tail factors are needed for our analysis.) Then C_J is what we are ultimately interested in predicting; and $\{C_1, C_2, \ldots, C_{J-1}\}$ are intermediate quantities used as stepping stones to estimate C_J.

The reason for this complexity is that C_J is missing on most of the claim-level records in our dataset. Using it as "the" target value analogous to Y in the GLM example above would require us to throw away data points for which Y is unknown. Let us frame our discussion in terms of an example. Suppose we have claim-level records for accident years 1990, 1991,..., 1999. On the 1990 records, we have losses evaluated as of 12, 24, ..., 120 months. On the 1991 records, we have losses evaluated as of 12, 24, ..., 108 months; while losses as of 120 months are unknown ("missing"). On the 1999 records, we have only losses evaluated as of 12 months; $\{C_{24}, C_{36}, \ldots, C_{120}\}$ are all missing.

Of course we have the option of using only the AY 1990 claim records to build a single GLM model; and use this model to predict the ultimate values of the 1991-1999 claims. But in doing so we would throw away the loss development pattern information that traditional reserving methods rely on. This is not a satisfactory option.

Many approaches are possible at this point, but we choose to build – continuing with the same example – 9 successive GLM models, "layered" one on top of the other. Speaking figuratively, we "regress" C_{24} on C_{12}; C_{36} on C_{24}; and so on. Each of these 9 GLM models is analogous to the 9 link ratios in the corresponding chain ladder model that could be run on the summarized 10-by-10 loss triangle. Let us denote these 9 models $M_{24}, M_{36}, \ldots, M_{120}$. The M_{36} model will take as an input either losses evaluated at 24 months (for AY 1990-98); or the predicted value of the M_{24} model (for AY 1999). This is analogous to the way a link ratio is applied in a chain-ladder analysis. Of course in addition to C_{J-1}, the M_J model takes as inputs all of the predictive variables $\{X_1, X_2, \ldots, X_N\}$.

Let us make this abstract discussion more concrete. The motivation for introducing predictive variables is to capture differences in different claims' expected loss development patterns. Given that our basic idea is to "incrementally" model these (potentially heterogeneous) development patterns à la the chain ladder, it makes sense to model each claim's development from period *j*-1 to period *j* as a function of several covariates:

$$\frac{C_j}{C_{j-1}} = \varphi(X_1, X_2, \ldots, X_N)$$

For mathematical convenience, we will further assume that this claim-level "link ratio" is in fact a (pre-specified) monotonic function *f* of a linear combination of the covariates:

$$\frac{C_j}{C_{j-1}} = f(\alpha + \beta_1 X_1 + \beta_N X_2 + \ldots + \beta_N X_N)$$

This is of course the familiar linear modeling trick: we reduce the job of estimating the function φ to estimating the parameters $\{\alpha, \beta_1, \beta_2, \ldots, \beta_N\}$. The monotonic function $f(\cdot)$ might, for example, be the natural exponent function exp(·) or the identity function id(·). The use of linear models (as opposed to, say, generalized additive models or neural networks) is not essential to the basic idea sketched here. But it is fairly flexible and powerful approach that avoids unnecessary complexity.

The above equation implies that the expected development from period *j*-1 to *j* of any given claim is a generalized linear function of the covariates $\{X_1, X_2, \ldots, X_N\}$. We do not need to assume that each claim at period j-1 will have the same expected development to period *j*. Nor do we need to assume that the mix of these (inhomogeneous) claims will stay the same from one accident year to the next.

Suppose, on the contrary, that we did assume perfect claim homogeneity in the sense that all claims have the same expected development. This is tantamount to assuming no variance in claim-level link ratios; and this in turn implies that no covariate X_i could possibly play a statistically significant role in predicting link ratio. Therefore the above equation reduces to a constant:

$$\frac{C_j}{C_{j-1}} = f(\alpha) = Link\_Ratio$$

Thus the chain ladder's link ratio is equivalent to our generalized linear model form with no covariates.

A few more assumptions will let us use the machinery of Generalized Linear Models to estimate the parameters $\{\alpha, \beta_1, \beta_2, \ldots, \beta_N\}$. Let us assume that the function *f* is the exponential function. This is equivalent to assuming the log link function from GLM theory. Let us further assume that the variance of C_{j+1} is proportional to its mean. (This assumption is not essential to the general technique we're tying to develop. This familiar assumption is being made for convenience, and could be altered without substantially affecting the discussion to follow.) In other words, we are assuming the over-dispersed Poisson GLM model form:

$$\log\left(E\left[\frac{C_j}{C_{j-1}}\right]\right) = \alpha + \beta_1 X_1 + \beta_N X_2 + \ldots + \beta_N X_N$$

Equivalently,

$$E\left[\frac{C_j}{C_{j-1}}\right] = \exp\{\alpha + \beta_1 X_1 + \beta_N X_2 + \ldots + \beta_N X_N\}$$

Or,

$$\frac{C_j}{C_{j-1}} = \exp\{\alpha + \beta_1 X_1 + \beta_N X_2 + \ldots + \beta_N X_N\} + \delta$$

where δ is an overdispersed Poisson-distributed error term. Given the quantities $\{C_{j-1}, C_j, X_1, X_2, \ldots, X_N\}$, we can estimate the parameters $\{\alpha, \beta_1, \beta_2, \ldots, \beta_N\}$ of model M_j using any standard GLM package. To be explicit, we would make the following specifications:

- Target: (C_j / C_{j-1})
- Covariates: $\{X_1, X_2, \ldots, X_N\}$
- Weight: C_{j-1}
- Distribution: Poisson
- Link: Log

Recasting the above equation as follows will allow us an alternate way of conceptualizing the above model form. Let us multiply both sides of the equation by C_{j-1}:

$$C_j = C_{j-1} \cdot \exp\{\alpha + \beta_1 X_1 + \beta_N X_2 + \ldots + \beta_N X_N\} + \varepsilon$$

which is equivalent to:

$$C_j = \exp\{\log(C_{j-1}) + \alpha + \beta_1 X_1 + \beta_N X_2 + \ldots + \beta_N X_N\} + \varepsilon$$

This is perhaps a more useful conceptualization of our model. The target variable is C_j, there is no weight variable, and $\log(C_j\text{-}1)$ serves as the "offset term". Explicitly:

- Target: C_j
- Offset: $\log(C_{j-1})$
- Covariates: $\{X_1, X_2, \ldots, X_N\}$
- Weight: none
- Distribution: Poisson
- Link: Log

(Note that all standard GLM packages allow one to specify an offset term.) The offset term essentially functions as a regressor whose corresponding "beta" parameter is constrained to be 1. This conceptualization illustrates the chain ladder-esque idea that we are building a model that estimates the expected value of C_j as a "generalized linear link function" $\exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_N X_N)$ applied to C_{j-1}, the (known or estimated) losses as of j-1.

HOW TO HANDLE IBNR

Note: this short section, and the appendix it refers to, outlines a method for extending the model design to handle IBNR claims. The authors suggest skipping it on the first reading. Indeed, this section can be skipped altogether if the reader takes the attitude that the model outlined can be used for losses on reported claims only; with IBNR claims being estimated in a separate analysis.

This model design also allows us a way of incorporating incurred but not reported (IBNR) losses into our model. For simplicity, let us assume that all claims that are unreported at 12 months are reported by 24 months. Therefore there will be records in our data with $C_{12}=0$ and $C_{24}>0$. In the M_{24} model, we add to the database one record for each in-force 1999 policy that had no claim as of 12 months from its effective date. On this record, we would force the offset term $\log(C_{12})$ to be zero. We would also include on all records an indicator variable X_0 as a covariate in M_{24} that takes on the value 1 if $C_{12}=0$, and 0 otherwise. Finally, we would neutralize all predictive variables that measure claim-level information.

("Neutralize" typically means that we recode missing variables to the median value.) As with all of the other AY 1999 records in the database the values of $\{C_{24}, C_{36}, \ldots, C_{120}\}$ are all missing.

Doing this will "allocate" a portion of the 12→24 IBNR (estimated from the AY 1990-1998 data) to each 1999 in-force policy that has no claim reported as of 12 months. The γ_0 parameter of the X_0 indicator functions in place of the offset term, which was forced to be zero on each of the 1999 zero-claim records. In other words, $\exp(\gamma_0)$ is the average expected 12→24 IBNR for each AY 1999 policy. The expected IBNR for an individual policy is $e^{\gamma_0}\exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_N X_N) = \exp(\alpha + \gamma_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_N X_N)$. The successive models M_{36}, M_{48},... will "develop" this allocated IBNR loss along with the other losses.

An example might clarify this discussion. Suppose that the total IBNR (as of 24 months) from AY 1990-98 was $400,000 and that during this time period, there were 4000 policies without claims as of 12 months. This is an average of $100 per claim-free policy. The value of γ_0 would therefore be $\log(100)\approx 4.6$.

Note that this method of treating IBNR assumes that the covariates $\{X_1, X_2, \ldots, X_N\}$ affect the allocation and development of IBNR in the same way that they affect the development of other losses. We could refine the model by including the interactions $\{X_0{*}X_1, X_0{*}X_2, \ldots\}$ as further model covariates. These covariates would be non-zero only for the records corresponding to policies with no losses as of 12 months. This idea is more fully explicated in the Appendix.

SIMULATION APPROACH

We will now apply the above model to a (very rudimentary) simulated dataset. The advantage of using simulated data is twofold. First, by construction we know which covariates are truly related to the various claims' differential development over time. Because of this, we can illustrate the operation of the model without the distraction of having to convince ourselves that a set of covariates is reasonably complete or significantly correlated with the claims' differential loss development.

Second, we can simulate our data "to ultimate", and set aside the (otherwise unknown) losses at ultimate as a standard against which we can compare our model's predictions with the predictions of the traditional chain-ladder model.

Of course a major disadvantage of using simulated data is that our sample results will give little indication of the degree to which out proposed model will produce improved predictions on real-world data.

However, it is our hope that the potential of this approach will be intuitive to many readers. The authors' experience in building predictive models for ratemaking and underwriting applications suggests that it is nearly always possible to find traditional and non-traditional predictive variables that are significantly correlated with size-of-loss. Given that larger claims are known to develop more slowly, one expects that that many of these same predictive variables will be correlated with loss development patterns.

SIMULATION ASSUMPTIONS

We illustrate our model with a simulated dataset that is very simple, yet with sufficient structure to illustrate the potential advantage of this model over the traditional chain ladder.

By construction, our claim-level dataset has the following characteristics:

- **Near-homogeneity of data:** the claims in our book of business all have identical expected loss development patterns except for one characteristic: whether the policyholder that made the claim had "good" credit or "bad" credit.
- **Differential development:** The claims of bad credit policies are expected to develop more slowly than the claims of good credit policies.
- **Changing mix of business:** A greater proportion of bad credit policies have been written in recent years.

As Bornhuetter and Ferguson point out, the differential loss development of bad/good credit policies' claims would present no special problem to the traditional methods were it not for the changing mix of business. However, the greater proportion of bad credit policies written in more recent years implies that the overall development patterns will shift from year to year. In particular, the expected development pattern for the most recent accident year will not be adequately represented by an average development pattern derived from the prior accident years' claims in a loss triangle.

The simulation incorporates the idea that a measurable quantity – here, credit – is correlated with loss development. Therefore by including credit in our reserving model, we are reflecting the shifting mix of business in our analysis. Put another way, the shifting proportion of bad credit policies is a "leading indicator" of a slow-down in the book's loss development. Using credit as a covariate in our reserving model allows us to quantify this slow-down, rather than judgmentally adjust for it after a traditional reserving exercise.

We simulate 5000 data points, each representing one claim. By design there are 500 claims for each of the accident years 1990, 1991, ..., 1999. Each of the 5000 records has 10 loss fields C_{12}, C_{24}, ..., C_{120}. We will describe how the values of $\{C_{12}, C_{24}, \ldots, C_{120}\}$ are assigned to each claim.

Finally, two simplifying assumptions are made. First, we assume that there is no IBNR: all claims are reported by 12 months from the beginning of the accident year. (See the discussion above and the Appendix for a discussion estimating IBNR in the current model framework.) Second, we assume that losses are fully developed as of 120 months: for each accident year k, $U_k=\sum C_{120}$.

Next we describe our simulation of the loss fields $\{C_{12}, C_{24}, \ldots, C_{120}\}$. We draw the losses at 12 months (C_{12}) from a lognormal distribution; and then successively apply 9 randomly generated "link" factors to these losses. The means and standard deviations of the distributions used to generate the losses and link factors were selected by judgment.

In more detail, the 5000 values of C_{12} were drawn from a lognormal distribution with parameters $\mu=8$ and $\sigma=1.3$:

$$\log(C_{12}) \sim n(8, 1.3)$$

For good credit claims, the values of $\{C_{24}, \ldots, C_{120}\}$ were determined by the following algorithm:

$$C_{j+1} = C_j * (\text{link}_j^{\text{good}} * \varepsilon_j)$$

The similar algorithm for bad claims is:

$$C_{j+1} = C_j * (\text{link}_j^{\text{bad}} * \varepsilon_j)$$

where

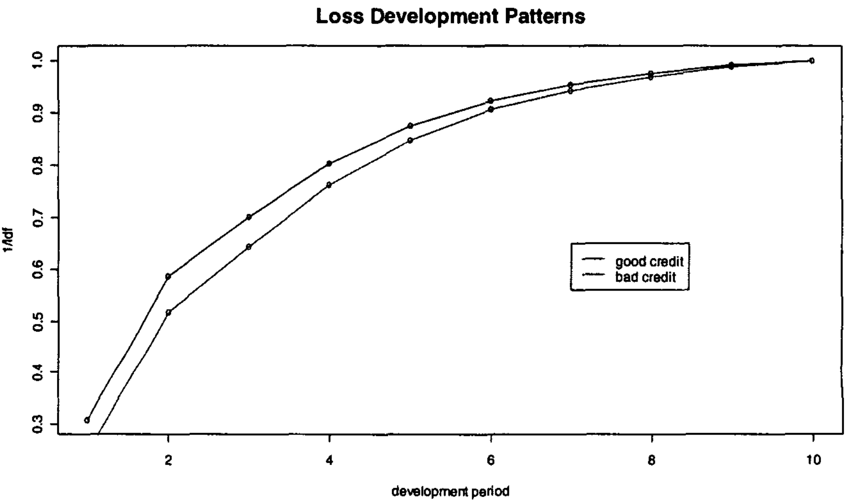
$link^{good} = \{1.8, 1.17, 1.13, 1.08, 1.05, 1.03, 1.02, 1.015, 1.008\}$

and $link^{bad} = (link^{good}-1)*1.25 + 1$:

$link^{bad} = \{2, 1.2125, 1.1625, 1.1, 1.0625, 1.0375, 1.025, 1.01875, 1.01\}$.

Finally, ε_j is a normally distributed "shock" term with mean 1 and a standard deviation that is a function of the value of the link ratio.

The development patterns (1/LDF) implied by the above expected link ratios are graphed below. This graph illustrates that by construction, bad credit claims develop more slowly than good credit claims.



In summary, each claim at each time period is assigned its own randomly generated link ratio; but the *expected* link ratios for bad/good credit claims are the ones stated above. (A word about motivation: the number of claims, size-of-loss distribution, and the general magnitude of the link ratios were judgmentally chosen to result in a summarized loss triangle similar to an actual Workers Comp loss triangle studied by one of the authors. The differing $link^{bad}$ and $link^{good}$ development patterns were selected purely judgmentally.)

So far, we have discussed the "homogeneity" and "differential development" assumptions. Regarding the "changing mix of business", we randomly apply the "bad" and "good" link ratios in the following proportions across the accident years:

shifting exposure base

year	%bad credit	%good credit
1990	30%	70%
1991	35%	65%
1992	40%	60%
1993	45%	55%
1994	50%	50%
1995	55%	45%
1996	60%	40%
1997	65%	35%
1998	70%	30%
1999	75%	25%

Note that the simulation approach we have laid out allows us to assign values of $\{C_{12}, C_{24}, \ldots, C_{120}\}$ to each claim, *regardless of accident year*. We will apply both our model and the traditional chain ladder to the data elements that would be available in an actual reserving exercise – namely those that form the upper half of the loss triangle. At the same time, we can use the data elements that would be unknown in an actual reserving exercise – the lower half of the triangle – as the "truth" against we can judge the success of both our method and the chain ladder.

The simulated data, summarized to the accident year level, is displayed below:

	Losses in $1000's											
	@12	@24	@36	@48	@60	@72	@84	@96	@108	@120	ultimate	o/s
1990	3,522	6,562	7,766	8,850	9,627	10,144	10,473	10,700	10,875	10,970	10,970	**0**
1991	3,527	6,623	7,876	9,011	9,817	10,361	10,705	10,942	11,123	11,223	11,223	**99**
1992	3,681	6,939	8,235	9,428	10,274	10,833	11,194	11,444	11,635	11,739	11,739	**295**
1993	3,780	7,152	8,539	9,791	10,666	11,262	11,642	11,902	12,100	12,210	12,210	**567**
1994	2,912	5,563	6,644	7,629	8,329	8,808	9,112	9,321	9,484	9,571	9,571	**763**
1995	3,724	7,167	8,573	9,850	10,763	11,393	11,796	12,070	12,282	12,397	12,397	**1,634**
1996	3,213	6,202	7,423	8,540	9,337	9,885	10,232	10,473	10,656	10,757	10,757	**2,217**
1997	3,335	6,445	7,727	8,887	9,721	10,281	10,643	10,890	11,083	11,187	11,187	**3,460**
1998	3,596	6,975	8,387	9,662	10,589	11,207	11,604	11,876	12,090	12,204	12,204	**5,229**
1999	3,327	6,481	7,817	9,018	9,889	10,483	10,860	11,123	11,323	11,432	11,432	**8,105**
												22,369
implied												
Link ratios	1.964	1.209	1.149	1.094	1.060	1.036	1.022	1.017	1.009	1.000		
LDFs	3.436	1.750	1.448	1.260	1.152	1.087	1.049	1.026	1.009	1.000		

The "unknown" data elements (those that would be known as of 12/31/2000 or after) are shaded, and will not be used to fit models. Note that the "ultimate" column is the same as the "at 120 months" column, and represents the "true", though unknown ultimate losses (u_k). Similarly, the "o/s" column represents the "true" outstanding losses as of 12/31/1999 (r_k). Thus the "true" value that we wish to estimate is $\varrho=\sum \varrho_k$ =\$22.369M.

Note that the link ratios computed from this summarized data are essentially weighted averages of the link^{bad} and link^{good} ratios stated above. This is representative of the way important patterns can be "summarized away" when the data is summarized to the triangle level.

MODEL RESULTS

We applied our sequence of 9 Poisson GLM models to the 5000 simulated data points. The exact steps of this process are sketched below:

Step 1: Regress the 4500 data points with non-missing values of C_{24} (i.e. the claims from AY 1990-98) on credit score, using log(C_{12}) as the offset term. This model is then applied to the 500 claims with unknown values of L_{24} (i.e. the AY 1999 claims) to produce *predicted* values of C_{24}.

Step 2: Regress the 4000 data points with non-missing values of C_{36} (i.e. the claims from AY 1990-97) on credit score, using log(C_{24}) as the offset term. This model is then applied to the 1000 claims with unknown values of L_{36} (i.e. the AY 1998-99 claims) to produce *predicted* values of C_{36}. Note that the AY 1998 values of C_{36} are based on *actual* values of C_{24}; whereas the AY 1999 values of C_{36} are based on *predicted* values of C_{24}.

...

Step 9: Regress the 500 data points with non-missing values of C_{120} (i.e. the claims from AY 1990) on credit score, using log(C_{108}) as the offset term. This model is then applied to the 4500 claims with unknown values of C_{120} (i.e. the AY 1991-99 claims) to produce *predicted* values of C_{120}. Note that the AY 1990 values of C_{120} are based on *actual* values of C_{108}; whereas the AY 1991-99 values of C_{120} are based on *predicted* values of C_{108}.

Step 10: The ultimate loss estimate is the sum of C_{120} across all claims and across all accident years: $u=\sum\sum C_{120}$. The estimate of total outstanding losses r equals u minus the total claims paid as of 12/31/1999.

The way in which the model M_j is applied to the predicted values of model M_{j-1} is analogous to the way the chain ladder's link ratios are multiplied together to produce loss development factors.

The results of these 10 steps, summarized to the accident year level, are displayed below. They can be compared to the display of the "truth" above:

GLM predictions (shaded) Losses in $1000's	@12	@24	@36	@48	@60	@72	@84	@96	@108	@120	ultimate	o/s
1990	3,522	6,562	7,766	8,850	9,627	10,144	10,473	10,700	10,875	10,970	10,970	-
1991	3,527	6,623	7,876	9,011	9,817	10,361	10,705	10,942	11,123	11,222	11,222	99
1992	3,681	6,939	8,235	9,428	10,274	10,833	11,194	11,444	11,635	11,738	11,738	294
1993	3,780	7,152	8,539	9,791	10,666	11,262	11,642	11,904	12,106	12,214	12,214	572
1994	2,912	5,563	6,644	7,629	8,329	8,808	9,112	9,323	9,485	9,573	9,573	765
1995	3,724	7,167	8,573	9,850	10,763	11,387	11,786	12,063	12,277	12,392	12,392	1,629
1996	3,213	6,202	7,423	8,540	9,337	9,877	10,222	10,461	10,646	10,745	10,745	2,205
1997	3,335	6,445	7,727	8,896	9,729	10,293	10,654	10,904	11,097	11,202	11,202	3,475
1998	3,596	6,975	8,378	9,665	10,584	11,207	11,605	11,882	12,096	12,212	12,212	5,237
1999	3,327	6,484	7,795	8,999	9,859	10,442	10,816	11,075	11,276	11,384	11,384	8,057
												22,333
implied												
link	1.954	1.208	1.152	1.093	1.059	1.036	1.023	1.017	1.009	1.000		
LDF	3.422	1.751	1.450	1.258	1.151	1.087	1.049	1.026	1.009	1.000		

Note that the implied LDFs at the bottom of this display were calculated by dividing the predicted ultimate values by the losses for that accident year as of 12/31/99. The implied link ratios were then derived from the implied LDFs.

Finally, the results of a chain ladder exercise are displayed in the following table:

Chain Ladder predictions (shaded) Losses in $1000's	@12	@24	@36	@48	@60	@72	@84	@96	@108	@120	ultimate	o/s
1990	3,522	6,562	7,766	8,850	9,627	10,144	10,473	10,700	10,875	10,970	10,970	-
1991	3,527	6,623	7,876	9,011	9,817	10,361	10,705	10,942	11,123	11,223	11,220	97
1992	3,681	6,939	8,235	9,428	10,274	10,833	11,194	11,444	11,635	11,739	11,733	289
1993	3,780	7,152	8,539	9,791	10,666	11,262	11,642	11,902	12,100	12,210	12,200	558
1994	2,912	5,563	6,644	7,629	8,329	8,808	9,112	9,321	9,484	9,571	9,536	728
1995	3,724	7,167	8,573	9,850	10,763	11,393	11,796	12,070	12,282	12,397	12,298	1,535
1996	3,213	6,202	7,423	8,540	9,337	9,885	10,232	10,473	10,656	10,757	10,637	2,097
1997	3,335	6,445	7,727	8,887	9,721	10,281	10,643	10,890	11,083	11,187	11,031	3,304
1998	3,596	6,975	8,387	9,662	10,589	11,207	11,604	11,876	12,090	12,204	11,873	4,898
1999	3,327	6,481	7,817	9,018	9,889	10,483	10,860	11,123	11,323	11,432	10,793	7,466
												20,972
implied												
link	1.906	1.192	1.146	1.090	1.055	1.033	1.022	1.016	1.009	1.000		
LDF	3.244	1.702	1.428	1.246	1.143	1.083	1.048	1.025	1.009	1.000		

(Note that this calculation can be verified by the reader in a spreadsheet. The spreadsheet-based results will differ from the above o/s loss estimate by $2000 (0.01%). This is due to rounding errors: the above table was generated by a computer program using un-rounded losses in the upper triangle.)

For convenience, the results of both methods – together with the simulated "truth" – are displayed below:

	Losses in $1000's													
	@12	@24	@36	@48	@60	@72	@84	@96	@108	@120		C-L	truth	proposed
1990	3,522	6,562	7,766	8,850	9,627	10,144	10,473	10,700	10,875	10,970	10,970	0	**0**	0
1991	3,527	6,623	7,876	9,011	9,817	10,361	10,705	10,942	11,123		11,220	97	**99**	99
1992	3,681	6,939	8,235	9,428	10,274	10,833	11,194	11,444			11,734	289	**295**	294
1993	3,780	7,152	8,539	9,791	10,666	11,262	11,642				12,200	558	**567**	572
1994	2,912	5,563	6,644	7,629	8,329	8,808					9,537	728	**763**	765
1995	3,724	7,167	8,573	9,850	10,763						12,298	1,535	**1,634**	1,629
1996	3,213	6,202	7,423	8,540							10,637	2,097	**2,217**	2,205
1997	3,335	6,445	7,727								11,031	3,304	**3,460**	3,475
1998	3,596	6,975									11,873	4,898	**5,229**	5,237
1999	3,327										10,792	7,466	**8,105**	8,057
												20,972	**22,369**	22,333
C-L	1.906	1.192	1.146	1.090	1.055	1.033	1.022	1.016	1.009	1.000		**-6.25%**		**-0.16%**
	3.244	1.702	1.428	1.246	1.143	1.083	1.048	1.025	1.009	1.000				
truth	1.964	1.209	1.149	1.094	1.060	1.036	1.022	1.017	1.009	1.000				
	3.436	1.750	1.448	1.260	1.152	1.087	1.049	1.026	1.009	1.000				
proposed	1.954	1.208	1.152	1.093	1.059	1.036	1.023	1.017	1.009	1.000				
	3.422	1.751	1.450	1.258	1.151	1.087	1.049	1.026	1.009	1.000				

Because the chain ladder is slow to pick up the changing mix of business (i.e., increasing proportion of bad credit policies that produce slower-developing claims), its estimates are too low for each accident year. This effect is most pronounced for the later accident years (shaded). In this example, the chain ladder's total outstanding loss estimate is approximately 6% too low.

By comparison, the proposed method's total outstanding loss estimate is almost exactly correct. It goes without saying that this is because our losses were simulated to develop in the multiplicative fashion assumed by the chain ladder; and because by construction only one covariate – credit – has a statistically significant relationship with loss development. Of course real-world data present no such conveniences. The above results are therefore suggestive at best. Still, the point remains that the proposed method is able to reflect changes in the mix of business (assuming that these changes can be measured by covariates capable of being collected and put into a model) that the chain ladder misses.

THE PROPOSED METHOD IS A PROPER GENERALIZATION OF THE CHAIN LADDER

By now it should be clear that the proposed loss reserving framework is intended to function as a GLM/micro-data-based analog of the chain ladder. One can go further and state that it is a true generalization of the chain ladder, in the sense that it produces the same results as the chain ladder when no covariates are present.

We verified this with the simulated data analyzed above. That is, we simply fit the above sequence of 9 GLM models, replacing the credit variable with a constant. The proposed method results in *exactly* the same results as the chain ladder. These results are summarized below.

acc. year	losses @ 12/99	true ultimate	true o/s	our method	chain ladder
1990	10,970	10,970	-	-	-
1991	11,123	11,223	**99**	**97**	**97**
1992	11,444	11,739	**295**	**289**	**289**
1993	11,642	12,210	**567**	**558**	**558**
1994	8,808	9,571	**763**	**728**	**728**
1995	10,763	12,397	**1,634**	**1,535**	**1,535**
1996	8,540	10,757	**2,217**	**2,097**	**2,097**
1997	7,727	11,187	**3,460**	**3,304**	**3,304**
1998	6,975	12,204	**5,229**	**4,898**	**4,898**
1999	3,327	11,432	**8,105**	**7,466**	**7,466**
			22,369	**20,972**	**20,972**
				-6.25%	-6.25%

It is generally a bad idea to exclude a statistically significant covariate from the GLM models. Here we see that doing so reproduces the chain ladder's (understated) reserve estimate. This lends a statistical perspective to where the chain ladder goes wrong when applied to a book of business whose development patterns have changed over time.

PART II: THE PROBLEM OF ESTIMATING RESERVE VARIABILITY

From a statistical perspective, R is an *estimator* of outstanding losses. It is a function of the values of the random variables $\{C_{12}, X_1, X_2, \ldots, X_N\}$ for each data point. In other words, it is a complicated function of several random variables. Like any such estimator, it has a probability distribution that is a complicated function of the distributions of the underlying random variables.

As we have demonstrated above, it is fairly straightforward to calculate the expected value of R. This is our outstanding loss *estimate*. It summarizes what the data (and our model) tells us to expect about the amount of future claim payments. But we would also like a measure of how strongly we should believe this estimate. To do this, we need further information – other than the expected value – about the distribution of our estimator of outstanding losses. For example, what are the cutoffs of a 95% confidence interval around the estimate?

This problem – sometimes referred to as the problem of *reserve variability* – has received a lot of attention in the recent loss reserving literature. The recent report of the CAS Working Party on Quantifying Variability in Reserve Estimates [2] puts the matter this way:

> A risk bearing entity wishes to know its financial position on a particular date. In order to do this, among other items it must understand the future payments it will be liable to make for obligations existing at the date of the valuation. For an insurance situation, these future payments are not known with certainty at the time of the valuation.
>
> The fundamental question that the risk bearing entity asks itself is:
> *Given any value (estimate of future payments) and our current state of knowledge, what is the probability that the final payments will be no larger than the given value?*

A full answer to this question would involve the assessment of model risk, and is beyond the scope of this paper. But even a limited answer would go beyond supplying a mere confidence interval or variability estimate. Ideally, we would like an estimate of the entire *probability distribution* of the outstanding loss estimator.

This seems like a lot to ask. After all, both the loss distribution underlying our claims data as well as our estimators of outstanding losses are fairly complex. Surprisingly, modern statistics supplies us with a simulation-based technique – called *bootstrapping* – that allows us to estimate this distribution with fairly little effort.

ENTER THE BOOTSTRAP

The Bootstrap was introduced by Bradley Efron in the late 1970s. Since then, it has become a commonly used technique in any number of problems in applied statistics. The classic text is Efron and Tibshirani [3]. Put briefly, bootstrapping is a simulation-based technique for estimating potentially "difficult" distributional properties – such as the standard deviation or the 90th percentile – of potentially complex estimators. We typically do not know the "true" distribution of such estimators. The basic idea of the Bootstrap is therefore to use the *actual*, empirical distribution (i.e., the data) as a proxy for the true, unknown distribution. Once this conceptual leap is made, many otherwise intractable problems become fairly straightforward exercises in statistical computing.

An analogy lies at the heart of bootstrapping. Just as our *actual* distribution is one of an infinite number of *possible* draws from the "true" theoretical distribution; we can take a large number of *resamples* of our actual distribution to form an arbitrarily large number of "pseudo-datasets".

Actual distribution : "true" distribution :: resampled datasets : actual distribution

Just as we would know everything we need to know about the "true" distribution if we could draw a large number of samples from it, we can *estimate* much of what we would like to know about the "true" distribution by treating the actual distribution as a proxy, and drawing multiple resamples from it.

We can illustrate this idea by applying it to a very simple problem for which we know the answer in advance. Suppose we draw 500 observations $\boldsymbol{X}=\{X_1,\ldots, X_{500}\}$ from a normal distribution with $\mu=5000$ and $\sigma=100$: $n(5000,100)$. Let m denote the sample average of this data:

$$m = \frac{1}{n}\sum_{i=1}^{5000} X_i$$

m is an estimate of the true value μ, just as we derived an estimate of the "true" outstanding losses in the previous sections. m therefore tells us "what we think" about the true value of μ based on the data. We would also like a measure of "how sure we are". In this simple example, the obvious thing to do is construct a confidence interval by appealing to the elementary fact that:

$$s.d.(m) = \frac{\sigma}{\sqrt{n}} = \frac{100}{\sqrt{5000}} \approx 4.47$$

Let us apply bootstrapping to this problem to see how close we can come to the answer (4.47) that we know in advance.

The following table records some facts about our data:

- # obs: 500
- Mean: 4995.79
- Stdev: 98.78
- 2.5^{th} %^{ile}: 4812.30
- 97.5^{th} %^{ile}: 5195.58

We can resample from this dataset a large number of times to create multiple "pseudo-datasets". "Resampling" means sampling with replacement as many times as there are points in your initial dataset (here, 500). Explicitly: pull a point at random from $\{X_1,\ldots, X_{500}\}$; record it; throw it back in; repeat this until we have our first pseudo-dataset containing 500 observations. Let us denote this pseudo-dataset $\boldsymbol{X}^*_1$.

We now repeat this process as many times as we would like, say 999 additional times. We therefore have 1000 pseudo-datasets $\boldsymbol{X}^*_1 \ldots \boldsymbol{X}^*_{1000}$. We can compute the sample average m on each one of these datasets. Denote these $\{m^*_1 \ldots m^*_{1000}\}$. These 1000 estimates constitute an estimate of the ***distribution*** of our estimator m. With this distribution $\{m^*_1 \ldots m^*_{1000}\}$ in hand, we can very easily estimate nearly any distributional property of m that we would like. In particular: the sample standard deviation of m based on our 1000 resamples is 4.43:

$$s.d.(m) \approx \frac{1}{999}\sum_{i=1}^{1000}\left(m^*_i - \frac{1}{1000}\sum_{k=1}^{1000} m^*_k\right)^2 \approx 4.43$$

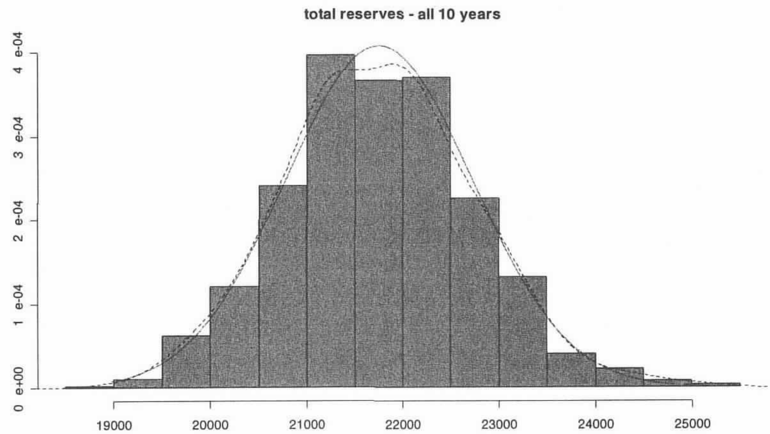
This differs from the true value (4.47) by less than a percentage point.

Bootstrapping in this toy example is therefore a complete success. The key point to note is that the unlike our analytic formula for s.d.(m), the bootstrapping technique does not assume any knowledge of the underlying distribution of $\boldsymbol{X}$. All that was required was computing power. Because of this, it is possible to execute essentially the same process on the loss data analyzed in the previous sections.

BOOTSTRAPPING RESERVE ESTIMATES

Having introduced the concept and run through a simple example, there is little to say in this section, other than to report the results. Let ***S*** denote our database of 5000 claims. We resampled ***S*** 500 times to get the 500 pseudo-datasets $S^*_1, \ldots S^*_{500}$. We then ran the above 9 GLM models on each of these 500 pseudo-datasets and computed outstanding losses on each pseudo-dataset: $\{R^*_1, \ldots, R^*_{500}\}$. Although it might seem excessive to fit 4500 GLM models to estimate the distribution of outstanding losses, doing so took less than 15 minutes on a standard laptop equipped with the shareware statistical software package R.

The estimated distribution of the outstanding loss estimator R is plotted below:



The bars are simply a histogram of the 500 estimates of outstanding losses. The solid curve is a superimposed normal distribution. The dotted curve is a kernel density estimate of the distribution underlying the histogram. Some basic statistics of this distribution are reported below:

- Mean: $21.751M
- Median: $21.746M
- Stdev: $0.982M

- C.V.: 4.5%

This kernel density estimate in the graph suggests that the distribution of our outstanding loss estimator is normal, to a reasonable degree of approximation. The fact that the mean is nearly equal to the median reinforces this judgment. Therefore a 95% confidence interval around our reserve estimate can be calculated in one of two ways:

- Record the 2.5 and 97.5 percentiles of the bootstrap distribution.
- Calculate 21.751M $\pm$ (1.96)*(0.982M).

Both of these methods produce the same answer, to within the nearest $100K:

($19.8M, $23.7M)

In short (ignoring model risk), we have 95% confidence that the true outstanding loss is within $\pm$ 9% of our estimated value. We remind the reader that this result is based on a rudimentary simulation, and is only intended to be suggestive.

DISCUSSION

Before concluding this paper, we would like to make four points about the bootstrapping technique illustrated above. First, bootstrapping is uncommonly generous to the practitioner in that it gives one an estimate of the entire *distribution* of an arbitrarily complex estimator without asking for *any* knowledge of the distributions underlying the data. Nearly any question we would typically ask about the outstanding loss distribution (standard deviation, skewness, percentiles, probability of ruin...) can be addressed with mere computation.

Second, the bootstrap method illustrated above is not specific to our GLM-based reserving technique. Indeed, if the claim-level data is available, one can also use this technique to bootstrap chain-ladder, Bornhuetter-Ferguson, or any other reserve estimates. To do this, we would summarize each of our pseudo-datasets to the triangle level; and apply our favorite technique to each of the resulting triangles. The 1000 outstanding loss estimates (assuming 1000 pseudo-datasets, as in the above illustration) resulting from each of the 1000 pseudo-triangles will constitute the distribution of our outstanding loss estimate.

Third, bootstrapping has been the subject of some discussion in the recent loss reserving literature. But there is an important difference between these discussions and the technique illustrated here. To the best of our knowledge, these discussions have been offered in the context of analyses of summarized loss triangles, not claim-level data.

The excellent survey paper by England and Verrall [4] is an example. England and Verrall apply a GLM model to a summarized loss triangle, and resample the standardized *residuals* of this model. They resample the distribution of residuals (there will be 55 data points for a 10-by-10 loss triangle) a large number of times. Each time they add the pseudo-dataset of residuals to the original loss triangle to form a pseudo-history to which they can again apply their GLM. Doing so allows them to estimate the prediction error of their estimate.

The difference between England and Verrall's approach and the approach illustrated here is generic, and found in most textbook discussions of bootstrapping. When bootstrapping model predictions, it is possible either to bootstrap *cases* (our approach) or *residuals* (England-Verrall). When dealing with small loss triangles it is not meaningful to bootstrap cases. However bootstrapping cases *is* meaningful when claim-level data is available.

As noted in the final paragraph of the introduction, our approach of resampling cases occurs prior to any reserving model being fit to the data. In other words, the very validity of our pseudo-datasets does not depend on the adequacy of the model being fit. In this sense, the cases-based resampling strategy is less sensitive to the correctness of ones model than the residual-based resampling strategy.

One final comment: bootstrapping is not the last word on the topic of reserve variability. In particular, nothing we have said addresses the problem of *model* risk. Suppose, for example, that we bootstrapped the traditional chain ladder applied to our simulated data. The bootstrapped confidence interval would not reflect the bias due to excluding the credit covariate in our reserving model. What is perhaps the biggest challenge in reserve variability is therefore left untouched by this discussion. Still, by giving us a practical way of estimating the predictive distribution of outstanding losses, bootstrapping potentially allows one to devote more attention to model risk.

CONCLUSION

The traditional summarized loss triangle is in general not a "sufficient statistic" for estimating outstanding losses. There will be times when we can do better by basing reserve and reserve variability estimates on un-summarized claim-level data.

As the first half of our paper illustrates, loss triangles can suppress heterogeneous loss development patterns that could be used to improve our predictions of outstanding losses. At the same time, summarized data does not allow us to use predictive variables that might be correlated with different loss development patterns.

Furthermore, as noted in the second half of our paper, loss triangles potentially summarize away variability information that could be used to make improved estimates of reserve variability. Using claim-level data allows us to bootstrap *cases*, not merely residuals from models applied to loss triangles with small numbers of observations.

In short, the use of claim-level data, together with relevant predictive variables, has the potential to improve actuaries' estimates of outstanding losses. In addition, it makes available a powerful and conceptually simple method for estimating reserve variability.

Acknowledgment

The authors would like to thanks Gerry Kirschner, Keith Curley, Peter Wu, and Frank Zizzamia for helpful conversations.

APPENDIX: ADDING IBNR TO THE MODEL

This appendix outlines a method by which one can enhance the model to predict INBR losses. Alternately, one can simply use the model outlined in the body of this paper to model the development of reported claims (as is done in the simulation example to follow); and build a separate model to estimate IBNR.

The 12→24 model (M_{24}) not modified to reflect IBNR takes the form:

$$C_{24} = \exp\{\log(C_{12}) + \alpha + \beta_1 X_1 + \beta_N X_2 + ... + \beta_N X_N\} + \varepsilon$$

The idea is to introduce a record for each *policy* with no losses as of 12 months (C_{12}=0) from its effective date. (Note that the other records in our database are at the claim level.) We set the offset term $\log(C_{12})$ on these records to be zero. We also include on all records an indicator variable X_0 that takes on the value 1 if C_{12}=0, and 0 otherwise. Finally, on the (claim-free) policy-level records we would neutralize all predictive variables that measure claim-level information. ("Neutralize" might mean that we recode missing values of a variable to the median value of that variable.)

For the 1990-98 policy-level records, we let $\{C_{24}, C_{36}, \ldots, C_{120}\}$ equal the total IBNR evaluated at these various evaluation points. As with all of the other AY 1999 records in the database the values of $\{C_{24}, C_{36}, \ldots, C_{120}\}$ are all missing. We add the indicator variable X_0 in the model. At this point our model takes the form:

$$C_{24} = \exp\{\log(C_{12}) + \alpha + \gamma_0 X_0 + \beta_1 X_1 + \beta_N X_2 + ... + \beta_N X_N\} + \varepsilon$$

Note that in this model form, the offset term only "applies" to the claim-level records with a non-zero value of C_{12}; similarly, the term $\gamma_0 X_0$ "applies" only to the policy-level records with C_{12}=0. The remaining terms apply to both types of records. In other words, each of the β parameters simultaneously models development of losses reported as of 12 months, as well as allocates IBNR losses at 24 months.

If this dual functioning of the β parameters is unsatisfactory, it is possible to let the β parameters *only* model the development of *reported* claims (as in the original model with no IBNR component) by introducing interaction terms. Suppose that $X_1 \ldots X_{N-p}$ are the policy-level covariates (such as policy age and credit score) in the model. (Claim-level variables such as report lag or injury type do not apply to policy-level records.) We add the interaction terms $X_0 * X_1 \ldots X_0 * X_{N-p}$ into the model:

$$C_{24} = \exp\{\log(C_{12}) + \alpha + \gamma_0 X_0 + \beta_1 X_1 + \ldots + \beta_N X_N + \gamma_1 X_0 * X_1 + \ldots + \gamma_{N-p} X_0 * X_{N-p}\} + \varepsilon$$

If this seems somewhat complex, it is because we have really designed "two models in one". The 12→24 development of a claim C^*_{12} is given by the following equation:

$$C^*_{24} = \exp\{\log(C^*_{12}) + \alpha + \beta_1 X_1 + \ldots + \beta_N X_N\}$$

All of the terms with X_0 drop out because X_0 is assumed to be 0 on (claim-level) records with non-zero C_{12}. In other words, we are back to the model form given at the beginning of this appendix.

On the other hand, the allocated IBNR at 24 months for a policy with no loss at 12 months is given by the following equation:

$$C_{24} = \exp\{\alpha + \gamma_0 + (\beta_1 + \gamma_1) X_1 + \ldots + (\beta_{N-p} + \gamma_{N-p}) X_{N-p} + \kappa\}$$

Here κ denotes the terms $\{\beta_{N-p+1} X_{N-p+1} + \ldots + \beta_N X_N\}$. These terms reduce to a constant κ because the claim-level variables $\{X_{N-p+1} \ldots X_N\}$ were neutralized on the policy-level records. In addition, note that the offset term was forced to be zero on these policy-level records.

It might be helpful to note that $\exp\{\alpha + \gamma_0 + \kappa\}$ is the *average* IBNR allocated to each of the policies that were claim-free as of 12 months. The multiplier $\exp\{(\beta_1 + \gamma_1) X_1 + \ldots + (\beta_{N-p} + \gamma_{N-p}) X_{N-p}\}$ adjusts each policy's allocated IBNR based on the values of the policy-level covariates $X_1 \ldots X_{N-p}$. As with expected claim development, the fact that the allocation of IBNR is "tailored" to the individual policy according to that policy's characteristics allows the model to reflect changes in the mix of business being analyzed.

Models M_{36}, ..., M_{120} can similarly be modified to handle the further emergence and development of IBNR.

REFERENCES

[1] Bornhuetter, R. and R. Ferguson. 1972. The Actuary and IBNR. *PCAS* LIX:181-95.

[2] CAS Working Party on Quantifying Variability in Reserve Estimates. 2005. The Analysis and Estimation of Loss & ALAE Variability: A Summary Report. Casualty Actuarial Society *Forum*, Fall 2005.

[3] Efron, B. and R. J. Tibshirani. 1993. *An Introduction to the Bootstrap*. San Francisco: Chapman & Hall.

[4] England, P. D. and R. J. Verrall. 2002. Stochastic Claims Reserving in General Insurance. *British Actuarial Journal* 8:443-544.

Parameter Estimation for Bornhuetter/Ferguson

Thomas Mack, Munich Re

Abstract: The Bornhuetter/Ferguson loss reserving method consists of selecting a development pattern and, for each accident year, an initial ultimate loss ratio. From these, the reserve estimate is derived. In this paper, the usual way to obtain the development pattern from the chain ladder link ratios is criticized because it assumes a multiplicative connection between past and future loss amounts whereas the Bornhuetter/Ferguson method establishes an additive connection (i.e. an independence). Therefore, an alternative approach to derive and select a development pattern is proposed.

Furthermore, the raw data usually contain some implicit information about the underwriting cycle. This paper shows how this information can be extracted from the data and used in the selection of the initial ultimate loss ratios.

Altogether the proposed approach is believed to align with the concepts of Bornhuetter and Ferguson better than the conventional approach does. The result is a standalone reserving method which does not rely upon the use of chain ladder elements.

Keywords. Loss reserving, Bornhuetter/Ferguson, Development pattern, Initial ultimate loss ratio

1. Introduction

Let $C_{i,k}$ denote the cumulative loss amount (either paid or incurred) of accident year i after k years of development, $1 \leq i, k \leq n$, and v_i be the premium volume of accident year i. Then $C_{i,n+1-i}$ denotes the current loss amount of accident year i. Let further $S_{i,k} = C_{i,k} - C_{i,k-1}$ denote the incremental loss amount (with $C_{i,0} = 0$) and U_i the (unknown) ultimate loss amount of accident year i. Then $R_i = U_i - C_{i,n+1-i}$ is the (unknown true) loss reserve for accident year i. For an easier exposition of the ideas, we assume in the beginning that n is large enough such that there is no significant loss development beyond development year n. We will eliminate this assumption at the end of section 3.

Bornhuetter/Ferguson (BF) introduced their method to estimate R_i in 1972 in order to cope with a major weakness of the chain ladder (CL) method. Therefore, we will first examine this weakness: The CL uses link ratios $\hat{f}_k$ in order to project the current loss amount $C_{i,n+1-i}$ to ultimate, i.e. it estimates $\hat{U}_i^{CL} = C_{i,n+1-i}\hat{f}_{n+2-i} \cdot ... \cdot \hat{f}_{n-1}\hat{f}_n$. Therefore, the CL reserve is

$$\hat{R}_i^{CL} = \hat{U}_i^{CL} - C_{i,n+1-i} = C_{i,n+1-i}\left(\hat{f}_{n+2-i} \cdot ... \cdot \hat{f}_n - 1\right).$$

This means that the reserve is heavily dependent upon the current loss amount $C_{i,n+1-i}$. This can lead to a nonsensical reserve $\hat{R}_i^{CL}$ = 0 for accident years where currently no claims are paid or reported which is not unusual in excess-of-loss reinsurance for the most recent accident year(s).

The BF method avoids this dependency upon the current loss amount $C_{i,n+1-i}$. The indicated BF reserve is defined as

$$\hat{R}_i^{BF} := \left(1-\hat{b}_{n+1-i}\right)\hat{U}_i$$

where

$\hat{U}_i = v_i\hat{q}_i$ with an *a priori* estimate $\hat{q}_i$ of the ultimate loss ratio (ULR) $q_i := U_i/v_i$ for accident year i,

$b_k \in [0, 1]$ is the percentage of ultimate losses expected to be known after development year k.

Note that $\hat{q}_i$ is called the *a priori* (or *initial*) estimate of the ULR, in contrast to the posterior estimate $(C_{i,n+1-i} + \hat{R}_i^{BF})/v_i$ of the ULR. This a priori estimate is different from the posterior estimate if and only if $C_{i,n+1-i} \neq \hat{b}_{n+1-i}v_i\hat{q}_i$. The percentages $(b_1, b_2, ..., b_n)$ constitute the expected cumulative development pattern (with $b_n = 1$ due to our preliminary assumption regarding n) and $1-\hat{b}_{n+1-i}$ is therefore the expected outstanding loss percentage of accident year i.

Thus, in order to apply the BF method, the actuary has to estimate the parameters q_i and b_k for all i and k. In practice, the b_k are derived from the CL link ratios in the following way:

$$b_n = 1,\ \hat{b}_{n-1} = \hat{f}_n^{-1},\ \hat{b}_{n-2} = \left(\hat{f}_{n-1}\hat{f}_n\right)^{-1},\ldots,\ \hat{b}_1 = \left(\hat{f}_2\cdot\ldots\cdot\hat{f}_n\right)^{-1}.$$

The method itself does not provide an objective approach for the determination of the a priori estimate $\hat{q}_i$. In practice, the q_i are estimated in a variety of ways, often based upon last year's estimate and/or pricing and market information. At worst, this practice can make the estimate $\hat{q}_i$ appear manipulated in order to achieve a reserve of a desired size. At best, the use of the CL pattern makes it difficult to view the BF method as a standalone reserving method.

Moreover, the use of the CL link ratios assumes that the unknown losses are a direct multiple of the already known losses at each point of the development. This contradicts the basic idea of the independence between $C_{i,n+1-i}$ and $\hat{R}_i^{BF}$ which was fundamental to the origin of the BF method.

Therefore, this paper develops an alternative approach to estimating the BF parameters q_i and b_k without the use of CL concepts along with rather clear guidance on how to arrive at an a priori estimate for the ultimate loss ratio q_i. Through this approach, the BF method becomes a true alternative to the CL method.

2. Estimation of the Development Pattern

If we already have an a priori estimate for U_i (e.g. from the traditional approach as outlined above), we are able to estimate the appropriate development pattern. From the BF reserve formula $\hat{R}_i^{BF} = \left(1-\hat{b}_{n+1-i}\right)\hat{U}_i$ we deduce

$$\hat{b}_{n+1-i} = 1-\frac{\hat{R}_i}{\hat{U}_i} = \frac{\hat{U}_i-\hat{R}_i}{\hat{U}_i} \approx \frac{C_{i,n+1-i}}{\hat{U}_i}.$$

As previously stated, the ≈-sign is a strict equality only if the a priori estimate $\hat{U}_i$ equals the posterior $C_{i,n+1-i} + \hat{R}_i$, i.e. if $C_{i,n+1-i} = \hat{b}_{n+1-i}\hat{U}_i$. This will not be the case for every i but should be true on average, at least approximately, otherwise the pattern $\hat{b}_1, \hat{b}_2, \ldots$ would not fit to the data. Therefore, the previous approximate equation suggests the estimator

$$\hat{b}_k := \sum_{i=1}^{n+1-k} C_{ik} \Big/ \sum_{i=1}^{n+1-k} \hat{U}_i$$

as weighted average of the ratios $C_{i,k}/\hat{U}_i$. This direct way of estimating the cumulative pattern b_1, b_2, ... may lead to inversions, i.e. $\hat{b}_k > \hat{b}_{k+1}$, because each $\hat{b}_k$ is based on a different number of accident years. In order to avoid such inversions, we use the corresponding increments

$$\hat{\beta}_k := \sum_{i=1}^{n+1-k} S_{ik} \Big/ \sum_{i=1}^{n+1-k} \hat{U}_i$$

and obtain $\hat{b}_k$ by adding up the $\hat{\beta}_k$, i.e. take

$$\hat{b}_k = \hat{\beta}_1 + \ldots + \hat{\beta}_k$$

and supplement it with $\hat{b}_{n+1} = 1$.

This is the development pattern as suggested by the BF reserve formula itself. This pattern is different from the CL pattern as can be seen e.g. from the numerical example below. Of course, the $\hat{\beta}_k$ should be smoothed and decreasing towards 0. This can be achieved by smoothing selections much as one would do when selecting CL link ratios. We will apply such a procedure together with the estimation of the ultimate loss ratio in the next section. But the actuary who wants to stay with the traditional BF way to arrive at an estimate for U_i can stop reading here and just use the specific BF pattern derived above.

3. Estimation of the Initial Ultimate Loss Ratios

As said in the introduction, the BF method aims at developing an estimate for q_i which does not directly depend on the losses $C_{i,n+1-i}$ known to-date and can be similarly obtained by another actuary. The procedure proposed here employs a three-steps approach. The first step considers the average incremental loss ratio (ILR)

$$\hat{m}_k := \sum_{i=1}^{n+1-k} S_{ik} \Big/ \sum_{i=1}^{n+1-k} v_i$$

of development year k observed to-date. The sum $\hat{m}_1 + \ldots + \hat{m}_n$ of all average ILRs is an a priori estimate of the ultimate loss ratio of an average accident year (if the development is assumed to be finished after n years). Note that in determining this a priori estimate, the known loss experience $C_{i,n+1-i}$ of any fixed accident year i is taken into account only marginally (as opposed to the CL estimate for U_i).

In the second step, we leverage the fact that the ultimate loss ratio q_i of accident year i is highly influenced by the level of the rate adequacy of that particular year. The rate adequacy is determined by two factors: the rate level and the loss level, which together yield the level of the loss ratio. But whereas in rate making we have to determine a sufficient absolute rate level – sufficient to pay all costs of the business -, for reserving purposes it is sufficient to judge the relative level of rate adequacy of an accident year as compared to the other accident years. With this information we can translate the (almost) known loss ratio of the oldest accident year(s) into predictions for the more recent accident years. Thus, we have to estimate the rate level change and the loss cost trend only. This is much easier because, at the time of reserving, we know the degree to which any rate changes have been realized and we know already some part of the losses of each accident year. This information should therefore be used for the assessment of the rate adequacy in addition to the information from the time of rate making.

Thus, we analyze what the run-off data tell us about the rate adequacy. If an accident year i has a below average rate adequacy (as compared to the other accident years considered), then the premium volume v_i is smaller than it should be for an average accident year. Therefore, most of its observed individual incremental loss ratios

$$\frac{S_{i,1}}{v_i}, \frac{S_{i,2}}{v_i}, \dots, \frac{S_{i,n+1-i}}{v_i}$$

will be higher than the corresponding averages

$$\hat{m}_1, \hat{m}_2, \dots, \hat{m}_{n+1-i},$$

at least after we have eliminated any unusually large individual losses as is normally done with any loss reserving method. In order to arrive at a single figure indicating the emerged relative rate adequacy level of accident year i (as compared to the average level of all accident years considered) we use the weighted average

$$r_i := \sum_{k=1}^{n+1-i} \frac{\hat{m}_k}{\sum \hat{m}_j} \cdot \frac{S_{ik}/v_i}{\hat{m}_k} = \sum_{k=1}^{n+1-i} S_{ik} \Big/ \sum_{k=1}^{n+1-i} (v_i \hat{m}_k) = \frac{C_{i,n+1-i}/v_i}{\sum_{k=1}^{n+1-i} \hat{m}_k}$$

of the ratios of $S_{i,k}/v_i$ and $\hat{m}_k$. Thus, r_i is the ratio of the current individual loss ratio $C_{i,n+1-i}/v_i$ of accident year i divided by the corresponding a priori average loss ratio. Therefore, r_i can be called a *loss ratio index*.

As seen from the premium perspective, r_i indicates the factor by which the premium v_i has to be multiplied in order to adjust it to the average rate adequacy level of the accident years i = 1, …, n considered. From this perspective, r_i can be called an *on-level premium factor*. Again, the factor r_i does not necessarily bring the premium v_i to the sufficient absolute size; it only achieves that – in relation to $v_i r_i$ instead of v_i - all accident years have approximately the same ultimate loss ratio $U_i/(v_i r_i) \approx \hat{m}_1 + \dots + \hat{m}_n$, may the latter be profitable or not. At this stage we can already state that, if the r_i's and the $\hat{m}_k$'s are plausible, then

$$(\hat{m}_1 + \dots + \hat{m}_n) r_i$$

is a reasonable a priori estimate of the ultimate loss ratio $q_i = U_i/v_i$ (if the development is assumed to be finished after n years).

As a third step, we have to check the plausibility of r_i. Initially we realize that the paid data and the incurred data will yield different values for r_i. But of course, these should be identical because they relate to the same premium v_i and losses U_i for either set of data. Without additional knowledge, we would therefore use the straight average $\left(r_i^{paid} + r_i^{inc}\right)/2$ or – as we deal with factors - rather the geometric mean

$$\overline{r}_i = \sqrt{r_i^{paid} \cdot r_i^{inc}} \; .$$

The calculation of the r_i's should be based on the data of a rather large portfolio in order to have the factors r_i be as reliable as possible. This large portfolio could be comprised of several run-off triangles for which the reserving is done separately, but which are assumed to have undergone similar changes in rate adequacy level.

Normally, we also have some information from pricing available, i.e. the rate changes effected and an estimate of the loss trend. The ratio r_i/r_{i-1} of any two consecutive years should be checked against the ratio of the loss trend and the effective rate change imbedded in v_i (in combination these represent the indicated change of the rate adequacy level). For instance, if from year i-1 to year i a loss increase of +10% is expected but a rate change of only +5% has been achieved, the ratio r_i / r_{i-1} should be close to 1.10/1.05 indicating a deterioration of the loss ratio by 4.8% (= 1.10/1.05 - 1). If not, we have to make a decision between these two ratios, e.g. form a credibility-weighted average of both values.

For the most recent accident years $i=n$ and $i=n$-1 we probably will trust the pricing information more than the r_i-estimate from the data, as the latter only relies on one or two entries in the triangle. At an extreme, r_i could be 0, which would be nonsensical and must obviously be adjusted. The size of r_1 for the first accident year can in principle be chosen arbitrarily, because its rate adequacy level (loss ratio level) will be taken into account in a subsequent adjustment of $\hat{m}_k$, see below. Therefore it can be left as it comes out of the formula in order to keep the $\hat{m}_k$ at the intuitive incremental loss ratio level.

What really matters are the relativities r_i/r_{i-1}. Therefore, we first select the values for these relativities based on all information available and then, starting with a selection for r_1^*, derive from these the resulting selections r_i^* for each accident year i. With these selected r_i^*, all adjusted premium volume figures $v_i r_i^*$, $1 \leq i \leq n$, should ultimately lead to (approximately) the same rate adequacy level, i.e. yield similar values of $U_i/(v_i r_i^*)$.

At next year's reserve calculation, the data triangle will contain an additional diagonal which will result in changes to all r_i. But the ratios r_i/r_{i-1} have the same interpretation as before. Therefore, due to the arbitrariness of r_1^*, we can keep the "old" r_1^* and – as long as no changes in the ratios r_i^*/r_{i-1}^* are indicated – also keep the other r_i^* and just add a new r_{n+1}^* based on a plausible ratio r_{n+1}^* / r_n^*.

Before using r_i^* for the estimation of q_i we have to adjust the average incremental loss ratios $\hat{m}_k$ because these were based on the unadjusted premium volume figures v_i. Therefore we replace $\hat{m}_k$ with

$$\hat{\hat{m}}_k := \sum_{i=1}^{n+1-k} S_{ik} \Big/ \sum_{i=1}^{n+1-k} (v_i r_i^*).$$

Often this will result in minor changes only. Major changes may happen for the last two or three development years or generally with data where the sizes of v_i or r_i^* vary significantly.

The adjusted ILRs $\hat{\hat{m}}_k$ of the last few development years could still produce unintuitive results, again due to the limited number of data points. Of course, these incremental values should be smooth and decreasing towards 0. Therefore, a smoothing approach is reasonable, and we denote the ILRs finally selected with $\hat{m}_k^*$.

At this point we will abandon the unrealistic assumption of not having any development beyond development year n. This is simply achieved by selecting an average tail ratio $\hat{m}_{n+1}^*$ (which may be 0 or even negative, like any other $\hat{m}_k^*$), to supplement the ILRs $\hat{m}_k^*$, $1 \leq k \leq n$, already selected.

Using these selected ILRs, we now have

$$\hat{m}^* := \hat{m}_1^* + ... + \hat{m}_n^* + \hat{m}_{n+1}^*$$

as an adjusted estimate for the ULR at average rate adequacy level. Of course, the paid data should have the same estimated ULR $\hat{m}^*$ as the incurred data. If that is not the case, we must adjust some $\hat{m}_k^*$, especially $\hat{m}_{n+1}^*$, to achieve the equality $\hat{m}_{paid}^* = \hat{m}_{inc}^*$. This finally yields the a priori estimate $\hat{q}_i := r_i^* \hat{m}^*$ for the ULR of accident year i and the corresponding amount $\hat{U}_i := v_i r_i^* \hat{m}^*$.

In contrast to the traditional BF procedure, this procedure gives the actuary the possibility to consolidate the general pricing and market information available with the trends and relativities contained in the paid and incurred data triangle. Moreover, this procedure uses a detailed decomposition of the initial ultimate loss ratio $\hat{q}_i = r_i^* (\hat{m}_1^* + ... + \hat{m}_{n+1}^*)$ into its components rate adequacy and development pattern. This makes the procedure easier to be followed or peer-reviewed by any other actuary.

4. Estimation of the Development Pattern (continued)

Now, we insert the result $\hat{U}_i = v_i r_i^* \hat{m}^*$ of the previous section into the formula derived for $\hat{\beta}_k$ in section 2 and obtain

$$\hat{\beta}_k = \frac{\sum_{i=1}^{n+1-k} S_{ik}}{\sum_{i=1}^{n+1-k} \hat{U}_i} = \frac{\sum_{i=1}^{n+1-k} S_{ik}}{\sum_{i=1}^{n+1-k} v_i r_i^* \hat{m}^*} = \frac{\hat{\hat{m}}_k}{\hat{m}^*}.$$

Here we see that the numerator $\hat{\hat{m}}_k$ may differ from the finally selected $\hat{m}_k^*$, as the denominator reflects the selected ILRs. Therefore it is logical to select

$$\hat{\beta}_k^* := \frac{\hat{m}_k^*}{\hat{m}^*}.$$

This finally leads to

$$\hat{b}_k^* := \hat{\beta}_1^* + \ldots + \hat{\beta}_k^* = \frac{\hat{m}_1^* + \ldots + \hat{m}_k^*}{\hat{m}_1^* + \ldots + \hat{m}_{n+1}^*}.$$

This is the genuine BF development pattern which is different from the CL pattern (see the numerical example below).

5. Putting it all Together

Altogether, we have the following steps of calculation:

$\hat{m}_k = \sum_{i=1}^{n+1-k} S_{ik} \Big/ \sum_{i=1}^{n+1-k} v_i$ raw incremental loss ratio (ILR) at development year k

$r_i = \sum_{k=1}^{n+1-i} S_{ik} \Big/ \sum_{k=1}^{n+1-i} (v_i \hat{m}_k)$ raw on-level premium factor for accident year i

r_i^* = selected on-level premium factor for accident year i (same for paid and incurred)

$\hat{m}_k^*$ = selected average ILR at development year k

(smoothed version of $\hat{\hat{m}}_k = \sum_{i=1}^{n+1-k} S_{ik} \Big/ \sum_{i=1}^{n+1-k} (v_i r_i^*)$)

$\hat{q}_i = r_i^* \left(\hat{m}_1^* + \ldots + \hat{m}_n^* + \hat{m}_{n+1}^*\right)$ *a priori* ULR for accident year i, including tail ratio $\hat{m}_{n+1}^*$

$\hat{U}_i$	$= v_i\hat{q}_i = v_i r_i^*\left(\hat{m}_1^* + \ldots + \hat{m}_{n+1}^*\right)$	*a priori* estimate of ultimate losses for accident year *i*
$\hat{b}_k^*$	$= \dfrac{\hat{m}_1^* + \ldots + \hat{m}_k^*}{\hat{m}_1^* + \ldots + \hat{m}_{n+1}^*}$	avg. cumulative percentage paid (incurred) at development year *k*
$\hat{R}_i$	$= \left(1 - \hat{b}_{n+1-i}^*\right)\hat{U}_i = v_i r_i^*\left(\hat{m}_{n+2-i}^* + \ldots + \hat{m}_{n+1}^*\right)$	loss reserve for accident year *i*

With this way of estimating its parameters q_i and b_k, the BF method is truly a standalone reserving method which is completely independent of the CL method. As shown in section 2, this way of calculating the pattern b_1, b_2, ... can also be used if the a priori estimates $\hat{q}_i$ and $\hat{U}_i = v_i\hat{q}_i$ are arrived at in a different (e.g. traditional) way. Thus, even if one does not like to work with m_k and r_i, one should at least adopt the estimation of the pattern as outlined above and avoid using the CL pattern.

6. Numerical Example

Data from General Liability Excess business are used to demonstrate the method. Exhibit A contains the premiums v_i and the incremental amounts $S_{i,k}$ of the incurred and the paid losses for the accident years 1992 – 2004 and development years 1 to 13. Some negative amounts have been kept in order to demonstrate that this does not lead to distortions. Exhibits B and C show the detailed results of the calculations for the incurred and the paid data respectively. These two exhibits are subdivided into three column blocks and two row blocks indicating the order of calculation: Columns (A) through (C) and rows (1) through (2) are the given data in aggregated form. From these the various components are calculated in the following order:

Rows (3) through (4),

Columns (D) through (G),

Rows (5) through (9),

Columns (H) through (M).

In the headings of column (H) and row (9), (8#) stands for the last number in row (8), i.e. $\hat{m}^*$. The suffix $_{+k}$ in rows (2), (3) and (5) stands for summation over *i*, i.e. $\sum_{i=1}^{n+1-k}$, The term "post." in columns (L) and (M) stands for "posterior". The bold headings **r_i***, **m_k*** and **Tail-ILR** indicate those positions where selections were required. These selections have been made in the following way:

Before selecting r_i^* we looked at Exhibit D where the raw r_i from column (E) are plotted for both paid and incurred data. The graph shows that the two sets of data are reasonably consistent, except for accident year 2004. Therefore, for i = 1992, ..., 2002, we selected r_i^* as the geometric mean between the paid r_i and the incurred r_i. For i = 2003 and 2004, we have set r_i^* = 0.50 for both, incurred and paid. The latter choice is not based on any further information. It is just an example. As mentioned earlier, information from pricing should also be used when making the selection. But even without this, the resulting r_i^* seem to give a realistic picture of the rather extreme rate adequacy level changes over the years considered. These r_i^* correspond to the following adequacy changes:

i-1→i	92→93	93→94	94→95	95→96	96→97	97→98	98→99	99→00	00→01	01→02	02→03	03→04
r_i^*/r_{i-1}^*	0.89	0.95	0.94	1.52	1.49	1.26	1.54	0.72	0.66	0.79	0.67	1.00

If we interpret r_i a loss ratio index, the above figures imply that we assume a decrease of the loss ratio index r_i from 1992 to 1993 of 11% (= 0.89 - 1) and an increase of 52% from 1995 to 1996.

m_k^* has been taken from row (6) ($\tilde{m}_k$) for development years k = 1, ..., 7. All the other m_k^* have been selected in order to make the development smoothly decreasing. Of course, other selections would have been possible. The **Tail-ILR** for incurred has been selected to be 0 and the **Tail-ILR** for paid has been selected such that the sum $\hat{m}^*$ of all paid ILRs equals that of the incurred-ILRs which is 137.9%. Note that the traditional way to apply BF will yield exactly the same reserve R_i as obtained in column (K) if we use $1.379 \cdot r_i^*$ as initial loss ratio and the pattern from row (9).

Finally, Exhibit E shows a comparison between the raw development pattern as proposed here and the pattern derived from the raw CL factors. More precisely, the BF pattern is a plot of $\hat{b}_k^{raw} = \frac{\hat{m}_1 + \ldots + \hat{m}_k}{\hat{m}_1 + \ldots + \hat{m}_n}$ using the raw ILR's m_k of row (4), whereas the CL pattern is a plot of $\hat{b}_k^{CL} = (\hat{f}_{k+1} \cdot \ldots \cdot \hat{f}_n)^{-1}$ with $\hat{f}_k = \sum_{i=1}^{n-k} C_{i,k+1} \Big/ \sum_{i=1}^{n-k} C_{i,k}$. We see that the raw BF pattern is clearly different from the raw CL pattern for either data set.

7. Final Remarks

As with any reserving method, this approach to estimating the parameters (i.e. the reserve) relies on implicit assumptions. One main assumption has already been addressed in the beginning: the data observed to-date and the amounts still outstanding are independent. This assumption is a cornerstone of the BF method. As the assumption should hold at any point in time, it essentially means that all incremental amounts $S_{i,1}$, ..., $S_{i,n}$ of each accident year are assumed to be independent. This would be violated if claim payments or bookings of case reserves were not done in the same way each year, especially if high payments in one calendar year would be followed by rather delayed payments in the following year(s). Similarly, the independence of the accident years is implicitly assumed in the estimation of m_k. This independence assumption is normally less problematic but could also be violated by calendar year effects. A more critical assumption is that the development pattern is consistent across all accident years. Of course, this assumption is not unique to this approach, as it is also implicit in the traditional BF method, as well as in the CL. This assumption should be especially borne in mind when selecting the accident years upon which the parameter estimates are to be based.

The way in which the parameters r_i and m_k are estimated consists of starting with an estimate for m_k which then is used to estimate r_i. The latter is adjusted and then used to arrive at an improved estimate for m_k. Thus, it may be tempting to again use this improved estimate of m_k to improve the estimate for r_i. But one must be cautious here. External judgment has already been applied in developing these parameters, and therefore any further changes based on the run-off data would only serve to dilute the (presumably desired) impacts of those judgments. Similarly, a purist might be tempted to iterate the estimations without any adjustments in between, i.e. to start with $\hat{m}_k$ and r_i as given in section 4, and with $\hat{\hat{m}}_k$ as in section 3, but then to use the latter for calculating $\tilde{r}_i = \sum_{k=1}^{n+1-i} S_{ik} \Big/ \sum_{k=1}^{n+1-i} \left(v_i \hat{\hat{m}}_k\right)$. This would then be iterated by calculating new estimates, first for m_k then for r_i by using the corresponding estimates obtained immediately before. Indeed, this procedure will quickly converge upon and yield exactly the same reserves as the CL does (for a full triangle only). This is not surprising, since proceeding in this way implies that we fully believe all the information contained in the data, without any input of external information. Thus we see that the input of external information is vital for the BF method.

For the CL, a methodology of assessing the variability of the reserves has been established in recent years. See e.g. the papers by Murphy or Mack in the 1994 CAS Spring Forum. Therefore, one would like to have this for BF, as well. For this purpose, we refer to the fact that our way of modeling the BF method can be seen as a cross-classified model, as in automobile rating, based upon the assumption $E(S_{i,k}/v_i) = r_i m_k$. Thus it can be treated using Generalized Linear Models.

However, this would use the "wrong" volume v_i instead of $v_i r_i$. Moreover, an appropriate assumption for the variance is necessary, too. Therefore, it may seem easier to use the alternative approach of embedding this BF model into the classical credibility IBNR model (see the author's paper "Improved Estimation of IBNR Claims by Credibility Theory" in the journal *Insurance: Mathematics & Economics* of 1990). In this way, the rate level r_i would be treated as a random variable. In any case, the issue of reserve variability deserves a separate paper.

8. References

Bornhuetter, Ronald L. and Ronald E. Ferguson, "The Actuary and IBNR", 1972 *PCAS*.

Mack, Thomas, "Improved Estimation of IBNR Claims by Credibility Theory", *Insurance: Mathematics & Economics* 1990.

Mack, Thomas, "Measuring the Variability of Chain Ladder Reserve Estimates", 1994 *CAS Spring Forum, Vol. 1*.

Murphy, Daniel M., "Unbiased Loss Development Factors", 1994 *CAS Spring Forum, Vol. 1*, and 1994 *PCAS*.

Abbreviations

BF: Bornhuetter/Ferguson

CL: chain ladder

ILR: incremental loss ratio

ULR: ultimate loss ratio

About the author

The author is retired (since 2006) chief actuary non-life of Munich Re. He studied Mathematics at the universities of Munich and Mannheim (Germany) and then worked more than 30 years for Munich Re. He is a member of the German Actuarial Association and a honorary member of the Institute of Actuaries (UK) and of the SAV (Switzerland). He has written several papers, mainly published in the ASTIN Bulletin. He has been awarded the 2nd-3rd prize at the CAS prize paper competition on variability of loss reserves in 1992 and the Charles A. Hachemeister prize in 1994.

Exhibit A

Incremental Incurred Loss Amounts

Acc.Year	Premium	Dev.Yr.	1	2	3	4	5	6	7	8	9	10	11	12	13
1992	41020		7362	3981	4881	5080	3806	2523	792	731	-1	241	-347	3	-115
1993	57547		5400	7208	7252	4946	4394	3198	3039	-771	988	-495	-182	1251	
1994	60940		2215	12914	6494	5585	2211	3363	2126	445	421	118	849		
1995	63034		1109	6581	5833	4827	5672	8638	12	146	4054	-625			
1996	61256		6220	10065	10343	11259	9032	1207	26	4221	378				
1997	57231		1324	6579	16428	17453	20457	3209	7103	101					
1998	91137		5772	12714	22918	33920	20709	33941	28483						
1999	96925		8563	47206	59695	60043	50458	5129							
2000	167021		11771	48696	84750	77361	39404								
2001	148494		11259	27000	38648	51890									
2002	165410		11855	27183	25927										
2003	228239		6236	18214											
2004	226454		7818												

Incremental Paid Loss Amounts

Acc.Year	Premium	Dev.Yr.	1	2	3	4	5	6	7	8	9	10	11	12	13
1992	41020		234	4643	6249	3530	6539	2737	2546	1815	335	110	18	26	-1
1993	57547		1994	4936	4825	6180	7659	1951	5110	611	776	409	48	1327	
1994	60940		-75	3208	7853	7127	5360	3876	3426	1440	1283	67	1616		
1995	63034		236	2202	4125	5003	4189	9064	2202	2064	3244	1179			
1996	61256		976	4719	9397	13253	6106	4975	3049	4719	2715				
1997	57231		-730	3353	12904	10642	16491	8886	7228	8512					
1998	91137		539	5238	14901	24865	20274	17769	32934						
1999	96925		725	14900	34676	43595	52621	27480							
2000	167021		312	6442	43596	88702	38812								
2001	148494		2988	9921	20357	34585									
2002	165410		260	7181	22202										
2003	228239		994	3049											
2004	226454		2411												

Exhibit B

Reserve Calculation for Incurred Data

(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)
Acc. Year i	v_i	$C_{i,n+1-i}$	Σm_k	r_i	r_i^*	$v_i r_i^*$	q_i	U_i	$1-b_{n+1-i}$	R_i	post. U_i	post. ULR
			from (4)	(C)/(B)/(D)	selected	(B)*(F)	(F)*(8#)	(B)*(H)	from (9)	(I)*(J)	(C)+(K)	(L)/(B)
1992	41,020	28,937	132.6%	0.53	0.57	23,421.1	78.7%	32,299.9	0.0%	0.0	28,937.0	70.5%
1993	57,547	36,228	132.9%	0.47	0.51	29,206.9	70.0%	40,279.1	0.1%	29.2	36,257.2	63.0%
1994	60,940	36,741	131.6%	0.46	0.48	29,464.7	66.7%	40,634.6	0.2%	88.4	36,829.4	60.4%
1995	63,034	36,247	131.4%	0.44	0.46	28,717.6	62.8%	39,604.3	0.6%	229.7	36,476.7	57.9%
1996	61,256	52,751	131.8%	0.65	0.69	42,376.1	95.4%	58,440.6	1.3%	762.8	53,513.8	87.4%
1997	57,231	72,654	129.7%	0.98	1.03	58,985.8	142.1%	81,346.9	2.8%	2,241.5	74,895.5	130.9%
1998	91,137	158,457	128.3%	1.36	1.30	118,381.1	179.1%	163,258.7	6.4%	10,417.5	168,874.5	185.3%
1999	96,925	231,094	118.7%	2.01	2.01	194,439.7	276.7%	268,150.6	15.5%	41,569.7	272,663.7	281.3%
2000	167,021	261,982	107.1%	1.46	1.44	240,660.2	198.7%	331,893.1	24.0%	79,509.4	341,491.4	204.5%
2001	148,494	128,797	84.7%	1.02	0.94	140,323.8	130.3%	193,519.8	38.7%	74,977.1	203,774.1	137.2%
2002	165,410	64,965	52.4%	0.75	0.74	122,950.0	102.5%	169,559.7	60.5%	102,656.5	167,621.5	101.3%
2003	228,239	24,450	24.4%	0.44	0.50	114,119.5	69.0%	157,381.6	80.5%	126,690.1	151,140.1	66.2%
2004	226,454	7,818	5.9%	0.58	0.50	113,227.0	69.0%	156,150.7	95.0%	148,318.1	156,136.1	68.9%

(1) Dev.Yr. k		1	2	3	4	5	6	7	8	9	10	11	12	13	
(2) S_{+k}		86,904	228,341	283,169	272,364	156,143	61,208	41,581	4,873	5,840	-761	320	1,254	-115	
(3) v_{+k}	from (B)	1,464,708	1,238,254	1,010,015	844,605	696,111	529,090	432,165	341,028	283,797	222,541	159,507	98,567	41,020	
(4) m_k	(2)/(3)	5.9%	18.4%	28.0%	32.2%	22.4%	11.6%	9.6%	1.4%	2.1%	-0.3%	0.2%	1.3%	-0.3%	
(5) $(vr^*)_{+k}$	from (G)	1,256,273.4	1,143,046.4	1,028,926.9	905,976.9	765,653.0	524,992.9	330,553.2	212,172.2	153,186.4	110,810.3	82,092.7	52,628.0	23,421.1	
(6) $\tilde{m}_k$	(2)/(5)	6.9%	20.0%	27.5%	30.1%	20.4%	11.7%	12.6%	2.3%	3.8%	-0.7%	0.4%	2.4%	-0.5%	**Tail-ILR**
(7) m_k^*	selected	6.9%	20.0%	27.5%	30.1%	20.4%	11.7%	12.6%	5.0%	2.0%	1.0%	0.5%	0.2%	0.1%	0.0%
(8) $\Sigma(7)$		6.9%	26.9%	54.4%	84.5%	104.9%	116.5%	129.1%	134.1%	136.1%	137.1%	137.6%	137.8%	137.9%	137.9%
(9) b_k	(8)/(8#)	5.0%	19.5%	39.5%	61.3%	76.0%	84.5%	93.6%	97.2%	98.7%	99.4%	99.8%	99.9%	100.0%	100.0%

Exhibit C

Reserve Calculation for Paid Data

(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)
Acc. Year i	v_i	$C_{i,n+1-i}$	Σm_k	r_i	r_i^*	$v_i r_i^*$	q_i	U_i	$1-b_{n+1-i}$	R_i	post. U_i	post. ULR
			from (4)	(C)/(B)/(D)	selected	(B)*(F)	(F)*(8#)	(B)*(H)	from (9)	(I)*(J)	(C)+(K)	(L)/(B)
1992	41,020	28,781	114.5%	0.61	0.57	23,421.1	78.7%	32,299.9	3.5%	1,118.6	29,899.6	72.9%
1993	57,547	35,826	114.5%	0.54	0.51	29,206.9	70.0%	40,279.1	4.9%	1,979.1	37,805.1	65.7%
1994	60,940	35,181	113.1%	0.51	0.48	29,464.7	66.7%	40,634.6	6.4%	2,585.8	37,766.8	62.0%
1995	63,034	33,508	112.1%	0.47	0.46	28,717.6	62.8%	39,604.3	8.5%	3,381.8	36,889.8	58.5%
1996	61,256	49,909	111.3%	0.73	0.69	42,376.1	95.4%	58,440.6	12.2%	7,109.0	57,018.0	93.1%
1997	57,231	67,286	108.3%	1.09	1.03	58,985.8	142.1%	81,346.9	17.2%	14,024.5	81,310.5	142.1%
1998	91,137	116,520	102.7%	1.24	1.30	118,381.1	179.1%	163,258.7	25.2%	41,168.2	157,688.2	173.0%
1999	96,925	173,997	89.6%	2.00	2.01	194,439.7	276.7%	268,150.6	37.6%	100,850.1	274,847.1	283.6%
2000	167,021	177,864	75.1%	1.42	1.44	240,660.2	198.7%	331,893.1	48.2%	160,000.6	337,864.6	202.3%
2001	148,494	67,851	52.4%	0.87	0.94	140,323.8	130.3%	193,519.8	63.2%	122,259.5	190,110.5	128.0%
2002	165,410	29,643	24.3%	0.74	0.74	122,950.0	102.5%	169,559.7	82.2%	139,350.9	168,993.9	102.2%
2003	228,239	4,043	6.4%	0.28	0.50	114,119.5	69.0%	157,381.6	94.9%	149,426.8	153,469.8	67.2%
2004	226,454	2,411	0.7%	1.44	0.50	113,227.0	69.0%	156,150.7	99.4%	155,171.6	157,582.6	69.6%

(1) Dev.Yr. k		1	2	3	4	5	6	7	8	9	10	11	12	13
(2) S_{+k}		10,864	69,792	181,085	237,482	158,051	76,738	56,495	19,161	8,353	1,765	1,682	1,353	-1
(3) v_{+k}	from (B)	1,464,708	1,238,254	1,010,015	844,605	696,111	529,090	432,165	341,028	283,797	222,541	159,507	98,567	41,020
(4) m_k	(2)/(3)	0.7%	5.6%	17.9%	28.1%	22.7%	14.5%	13.1%	5.6%	2.9%	0.8%	1.1%	1.4%	0.0%

(5) $(vr^*)_{+k}$	from (G)	1,256,273.4	1,143,046.4	1,028,926.9	905,976.9	765,653.0	524,992.9	330,553.2	212,172.2	153,186.4	110,810.3	82,092.7	52,628.0	23,421.1	
(6) $\tilde{m}_k$	(2)/(5)	0.9%	6.1%	17.6%	26.2%	20.6%	14.6%	17.1%	9.0%	5.5%	1.6%	2.0%	2.6%	0.0%	**Tail-ILR**
(7) m_k^*	selected	0.9%	6.1%	17.6%	26.2%	20.6%	14.6%	17.1%	11.0%	7.0%	5.0%	3.0%	2.0%	2.0%	4.8%
(8) Σ(7)		0.9%	7.0%	24.6%	50.8%	71.4%	86.0%	103.1%	114.1%	121.1%	126.1%	129.1%	131.1%	133.1%	137.9%
(9) b_k	(8)/(8#)	0.6%	5.1%	17.8%	36.8%	51.8%	62.4%	74.8%	82.8%	87.8%	91.5%	93.6%	95.1%	96.5%	100.0%

Exhibit D

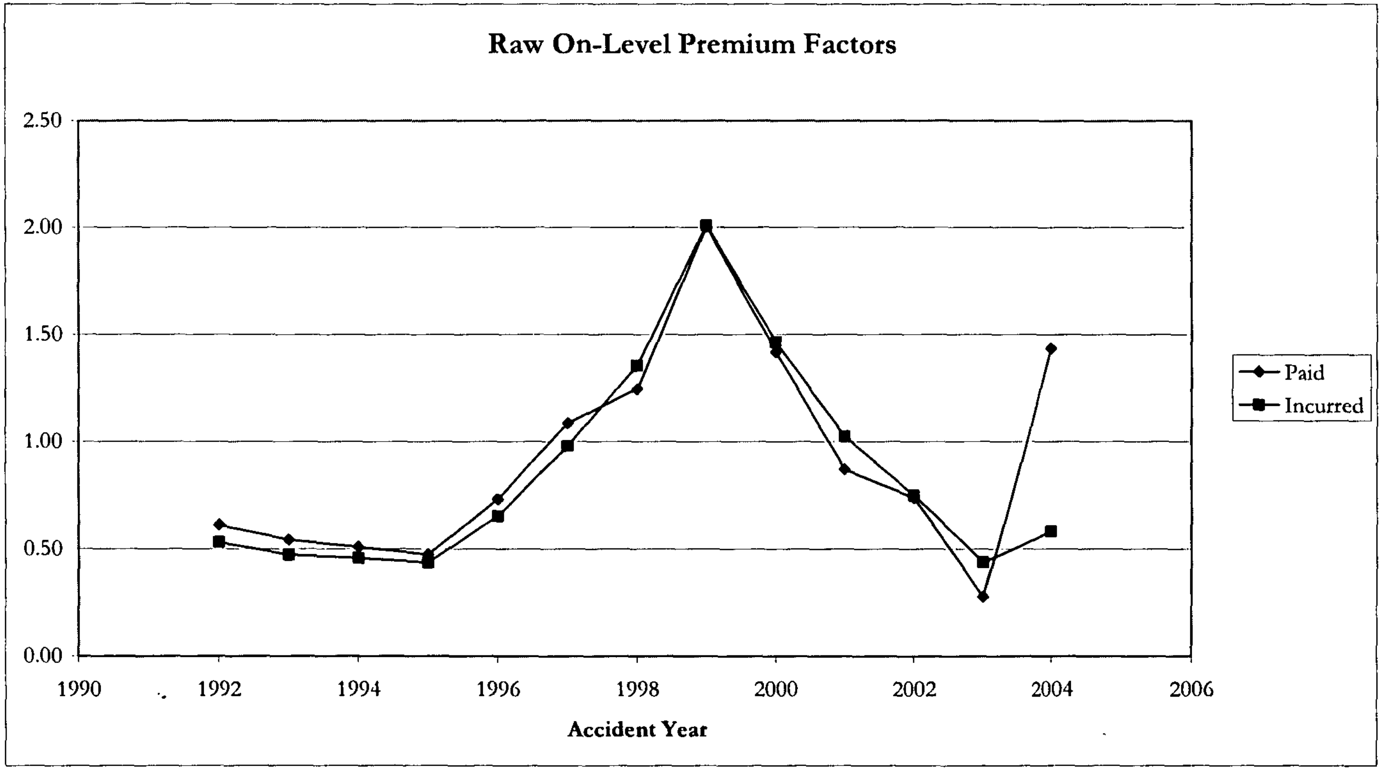
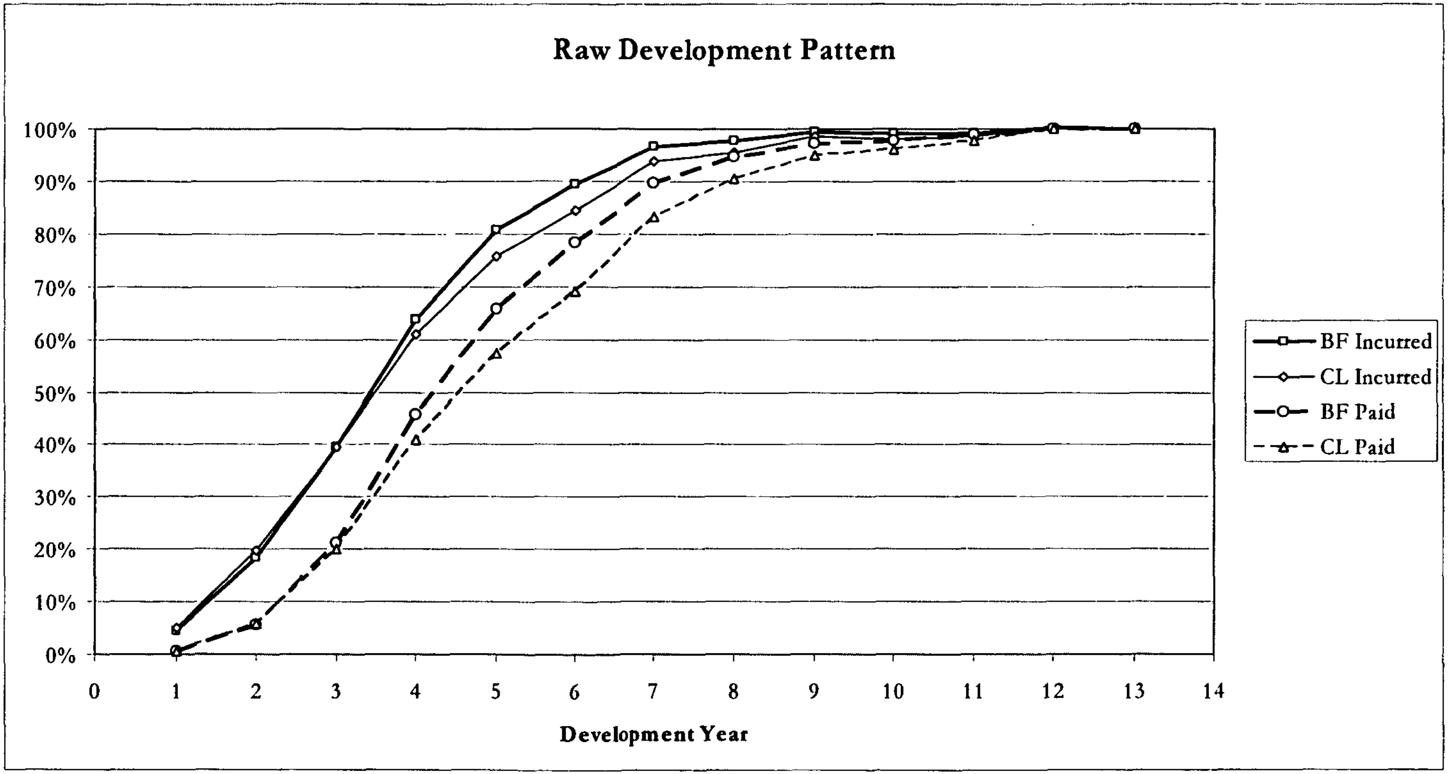

Raw On-Level Premium Factors
2.50
2.00
1.50
1.00
0.50
0.00
1990
1992
1994
1996
1998
2000
2002
2004
2006
Accident Year
Paid
Incurred

Exhibit E


Raw Development Pattern
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Development Year
BF Incurred
CL Incurred
BF Paid
CL Paid

Estimating Predictive Distributions for Loss Reserve Models

By

Glenn Meyers, FCAS, MAA, Ph.D.

Abstract
This paper demonstrates a Bayesian method for estimating the distribution of future loss payments of individual insurers. The main features of this method are: (1) the stochastic loss reserving model is based on the collective risk model; (2) predicted loss payments are derived from a Bayesian methodology that uses the results of large, and presumably stable, insurers as its prior information; and (3) this paper tests its model on large number of insurers and finds that its predictions are well within the statistical bounds expected for a sample of this size. The paper concludes with an analysis of reported reserves and their subsequent development in terms of the predictive distribution calculated by this Bayesian methodology.

Key Words
Reserving Methods, Reserve Variability, Uncertainty and Ranges, Schedule P, Suitability Testing, Collective Risk Model, Fourier Methods, Bayesian Estimation, Hypothesis Testing

1. Introduction

Over the years, there have been a number of stochastic loss reserving models that provide the means to statistically estimate confidence intervals for loss reserves. In discussing these models with other actuaries, I find that many feel that the confidence intervals estimated by these methods are too wide. The reason most give for this opinion is that experienced actuaries have access to information that is not captured by the particular formulas they use. These sources of information can include intimate knowledge of claims at hand. A second source of information that many actuarial consultants have is the experience gained by setting loss reserves for other insurers.

As one digs into the technical details of the stochastic loss reserving models, one finds many assumptions that are debatable. For example Mack, [1993], Barnett and Zehnwirth [2000], and Clark [2003] all make a number of simplifying assumptions on the distribution of an observed loss about its expected value. Now it is the essence of predictive modeling to make simplifying assumptions. Which set of simplifying assumptions should we use? Arguments based on the "reasonability" of the assumptions can (at least in my experience) only go so far. One should also test the validity of these assumptions by comparing the predictions of such a model with observations that were not used in fitting the model.

Given the inherent volatility of loss reserve estimates, testing a single estimate is unlikely to be conclusive. How conclusive is the following statement?

"Yes, the prediction falls somewhere within a wide range."

A more comprehensive test of a loss reserve model should involve testing its predictions on many insurers.

The purpose of this paper is to address at least some of the issues raised above.

- The methodologies developed in this paper will be applied to the Schedule P data submitted on the 1995 NAIC Annual Statement for each of 250 insurers.

- The stochastic loss model underlying the methods of this paper will be the collective risk model. This model combines the underlying frequency and severity distributions to get the distribution of aggregate losses. This approach to stochastic loss reserving is not entirely new. Hayne [2003] uses the collective risk model to develop confidence regions for the loss reserve, but they assume that the expected value of the loss reserve is known. This paper makes explicit use of the collective risk model to first derive the expected value of the loss reserve.

- Next, this paper will illustrate how to use Bayes' Theorem to estimate the predictive distribution of future paid losses for an individual insurer. The prior distributions used in this method will be "derived" by an analysis of loss triangles for other insurers. This method will provide some of the "experience gained by setting loss reserves for other insurers" that is missing from existing statistical models for calculating loss reserves. An advantage of such an approach is that all assumptions (i.e., prior distributions) and data will be clearly specified.

- Next, this paper will test the predictions of the Bayesian methodology on data from the corresponding Schedule P data in the corresponding 2001 NAIC Annual Statements. The essence of the test is to use the predictive distribution derived from the 1995 data to estimate the predicted percentile of losses posted in the 2001 Annual Statement for each insurer. While the circumstances of each individual insurer may be different, the predicted percentiles of the observed losses should be uniformly distributed. This will be tested by standard statistical methods.

- Finally, this paper analyzes the reported reserves and their subsequent development in terms of the predictive distributions calculated by this Bayesian methodology.

The main body of the paper is written to address a general actuarial audience. My intention is to make it clear "what" I am doing in the main body. I will discuss additional details needed to implement the methods described in some of the sections in the Appendix.

2. Exploratory Data Analysis.

The basic data used in this analysis was the earned premium and the incremental paid losses for accident years 1986 to 1995. The incremental paid losses were those reported as paid in each calendar year through 1995.

The data used in this analysis was taken from Schedule P of the 1995 NAIC Annual Statement, as compiled by the A.M. Best Company. I chose the Commercial Auto line of business because the payout period was long enough to be interesting but short enough so that ignoring the tail did not present a significant problem. The estimation of the tail is beyond the scope of this paper.

I selected 250 individual insurance groups from the hundreds that were reported by A.M. Best, based on the following criteria. First, there had to be at least some exposure in each of the years 1986 to 1995. Second, the payment pattern had to, in my judgment, "look reasonable."

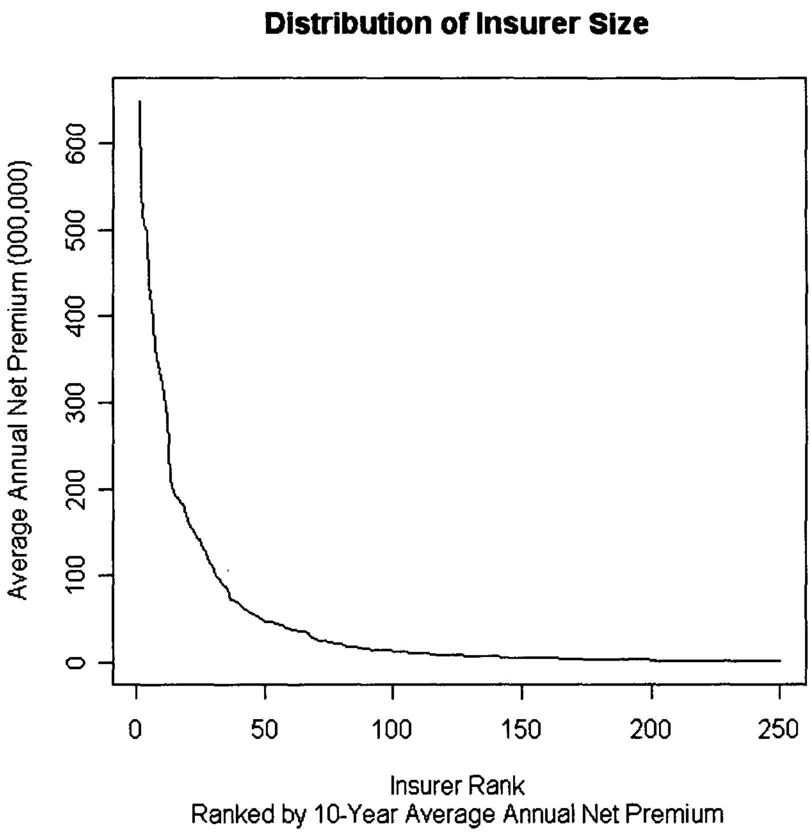
Occasionally, the reported incremental paid losses were negative. In this case, I treated the losses as if they were zero. I believe this had minimal effect on the total loss reserve.

Let's look at some graphic summaries of the data. Figure 1, below, shows the distribution of insurer sizes, ranked by 10-year average earned premium. It is worth noting that 16 of the insurers accounted for more than half of the total premium of the 250 insurers.

Figure 2, below, shows the variability of payment paths (i.e., proportion of total reported paid loss segregated by settlement lag) for the accident year 1986. This figure makes it clear that payment paths do vary by insurer. How much these differences can be attributed to systematic differences between insurers, versus how much can be attributed to random processes, is unclear at this point.

Figure 3, below, shows the aggregate payment patterns for four groups, each accounting for approximately one quarter of the total premium volume.

Figure 1



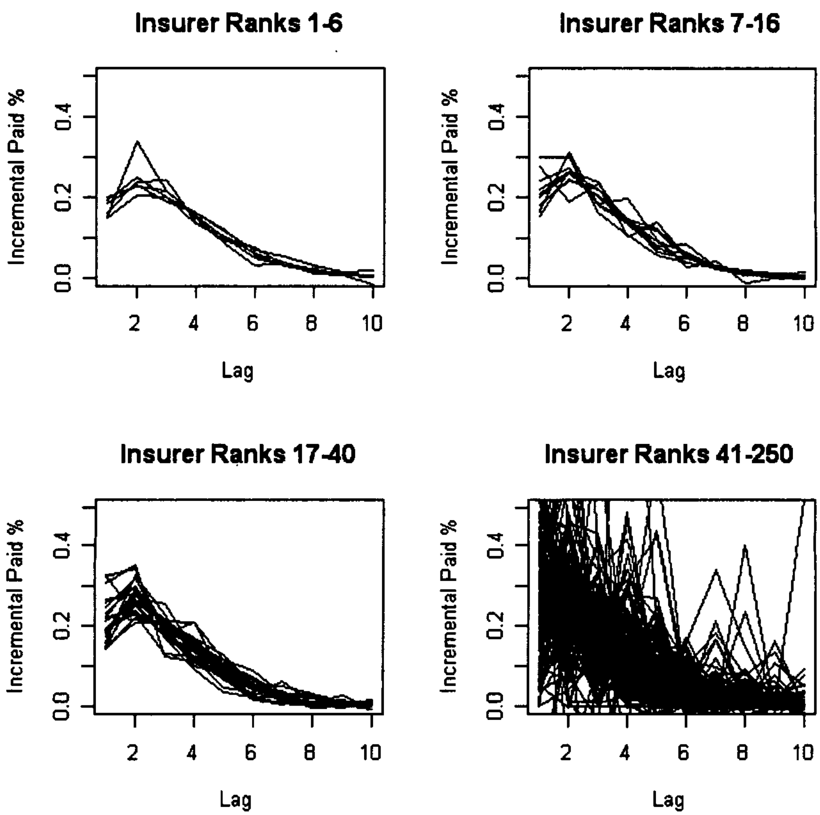
1

- Insurers ranked 1-6, 7-16, 17-40 and 41-250 each accounted for about one quarter of the total premium

Figure 2

Empirical Payment Paths for Accident Year 1986

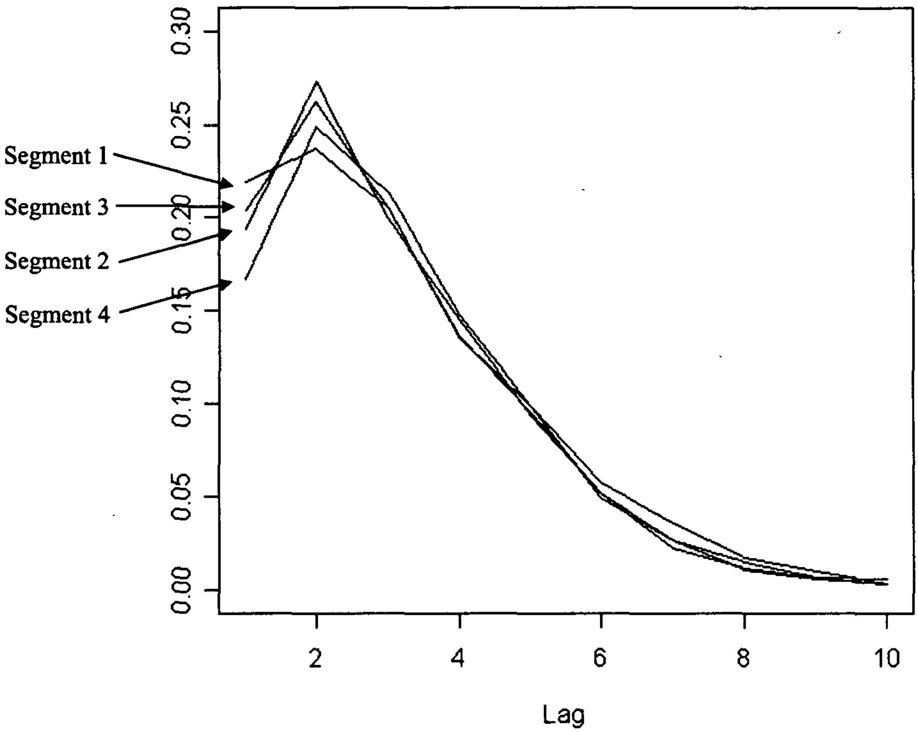
Incremental Paid Losses as a Proportion of 10-Year Total



- Each plot represents approximately one quarter of the total premium volume.
- The variability of the incremental paid loss factors increases as the size of the insurer decreases.

Figure 3

Empirical Payment Paths for the Four Industry Segments



- Segment 1 – Insurers ranked 1-6, Segment 2 – Insurers ranked 6-16, Segment 3 – Insurers ranked 17-40, Segment 4 – Insurers ranked 41-250.
- There is no indication of any systematic differences in payout patterns by size of insurer.

3. A Stochastic Loss Reserve Model

The goal of this paper is to develop a loss reserving model that makes testable predictions. And then actually perform the tests. Let's start with a more detailed outline of how I intend to reach this goal.

1. The model for the expected payouts will be fairly conventional. It will be similar to the "Cape Cod" approach first published by Stanard [1985]. This approach assumes a constant expected loss ratio across the 10-year span of the data.

2. Given the expected loss, the distribution of actual losses around the expected will be modeled by the collective risk model – a compound frequency and severity model. As mentioned above, this approach has precedents with Hayne [2003]. This will conclude Section 3.

3. In Section 4, I will turn to estimating the parameters for the above models. The initial estimation method will be that of maximum likelihood.

4. I will then discuss testing the predictions of the model in Section 5. Initially, the tests will be on the same data that was used for fitting the models. (The tests on data in the 2001 Annual Statements will come later.) As mentioned above, the test will consist of calculating the percentiles of each of the observed loss payments and testing to see that those predictions are uniformly distributed.

As we proceed, I will focus on the 40 largest insurers. I do this because, in my judgment, the models are responding mainly to random losses for the smaller insurers. As we shall see, the results of the fitted models for the 40 largest insurers will form the basis for the Bayesian analysis that will be applied to each insurer, large and small. Implicit in this approach is the assumption that main systematic differences in the loss payment paths are somehow captured by the largest 40 insurers.

Let's proceed.

Assume that the expected losses are given by the following model.

$$E\left[Paid\ Loss_{AY,Lag}\right] = Premium_{AY} \times ELR \times Dev_{Lag}, \tag{1}$$

where:

- AY (1986 = 1, 1987 = 2,…) is an index for accident year.
- Lag = 1, 2, ..., 10 is the settlement lag reported after the beginning of the accident year.
- *Paid Loss* is the incremental paid loss for the given accident year and settlement lag.
- *Premium* is the earned premium for the accident year.
- ELR is an unknown parameter that represents the expected loss ratio.
- Dev_{Lag} is an unknown parameter that depends on the settlement lag.

As with Stanard's "Cape Cod" method, the ELR parameter will be estimated from the data.

The "Cape Cod" formula that I used to estimate the expected loss is by no means a necessary feature of this method. Other formulas, like the chain ladder model, can be used.

A common adjustment that one might make to Equation 1 is to multiply the ELR by a premium index to adjust for the "underwriting cycle." I tried this, but it did not appreciably increase the accuracy of the predictions *for this data and time period.* Thus I chose to use the simpler model in this paper. But one should consider using a premium index in other circumstances.

Let $X_{AY,Lag}$ be a random variable for an insurer's incremental paid loss in the specified accident year and settlement lag. Assume that $X_{AY,Lag}$ has a compound negative binomial (*CNB*) distribution, which I will now describe.

- Let Z_{Lag} be a random variable representing the claim severity. Allow each claim severity distribution to differ by settlement lag.
- Given $E[Paid\ Loss]_{AY,\ Lag}$, define the expected claim count, $\lambda_{AY,Lag}$ by

$$\lambda_{AY,Lag} \equiv E\left[Paid\ Loss_{AY,Lag}\right] / E\left[Z_{Lag}\right]. \tag{2}$$

- Let $N_{AY,Lag}$ be a random variable representing the claim count. Assume that the distribution of $N_{AY,Lag}$ is given by the negative binomial distribution with mean $\lambda_{AY,Lag}$ and variance $\lambda_{AY,Lag} + c \cdot \lambda_{AY,Lag}^2$.
- Then the random variable $X_{AY,Lag}$ is defined by

$$X_{AY,Lag} = Z_{Lag,1} + Z_{Lag,2} + \ldots + Z_{Lag,N_{AY,Lag}}$$

While the above defines how to express the random variable, $X_{AY,Lag}$, in terms of other random variables $N_{AY,Lag}$ and Z_{Lag}, later on we will need to calculate the likelihood of observing $x_{AY,Lag}$ for various accident years and settlement lags. The details of how to do this are in the technical appendix. Here I will give a high-level overview of what will be done below.

1. The distributions of Z_{Lag} were derived from data reported to ISO as part of its regular increased limits ratemaking activities. Like the substantial majority of insurers that report their data to ISO, the policy limit will be set to $1,000,000. The distributions varied by settlement lag with lags 5-10 being the most severe. See Figure 4 below. For this application I discretized the distributions at intervals h, which depended on the size of the insurer. h was chosen so the 2^{14} (16,384) values spanned the probable range of losses for the insurer.

2. I selected the value of 0.01 for the negative binomial distribution parameter, *c*. My paper, Meyers [2006], analyzes Schedule P data for Commercial Auto and provides justification for this selection.

3. Using the Fast Fourier Transform (FFT), I then calculated the entire distribution of a discretized $X_{Ay,Lag}$, rounded to the nearest multiple of *h*. The use of Fourier Transforms for such calculations is not new. References for this in CAS literature include my joint paper, Heckman and Meyers[1983], along with Wang [1998].

4. Whenever the probability density of a given observation $x_{AY,Lag}$ given $E\left[\textit{Paid Loss}_{AY,Lag}\right]$, was needed I rounded the $x_{AY,Lag}$ to the nearest multiple of *h* and did the above calculation. The resulting distribution function is denoted by:

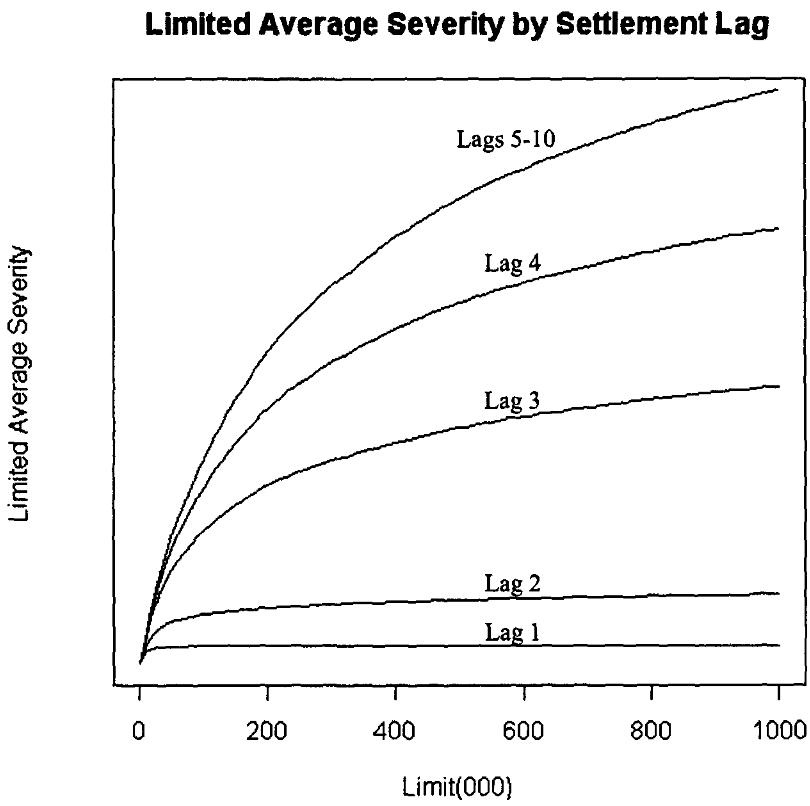
$$CNB\left(x_{AY,Lag} \mid E\left[\textit{Paid Loss}_{AY,Lag}\right]\right). \tag{3}$$

This specifies the stochastic loss reserving model used it this paper. The parameters that depend on the particular insurer are the *ELR* and the 10 Dev_{Lag} parameters. I will now turn to showing how to estimate these parameters, given the earned premiums and the Schedule P loss triangle.

Clark [2003] has taken a similar approach to loss reserve estimation. Indeed, I credit Clark for the inspiration that led to the approach taken in this section and the next. Clark used the Weibull and loglogistic parametric models where I used Equation 1 above. In place of the *CNB* distribution described above, Clark used what he calls the "overdispersed Poisson" (*ODP*) distribution[1]. He then estimated the parameters of his model by maximum likelihood. This is where I am going next.

[1] A random variable has an overdispersed Poisson distribution if it is an ordinary Poisson random variable times a constant scaling factor.

Figure 4


Limited Average Severity by Settlement Lag
Lags 5-10
Lag 4
Lag 3
Lag 2
Lag 1
Limited Average Severity
0
200
400
600
800
1000
Limit(000)

4. Maximum Likelihood Estimation of Model Parameters

The data for a given insurer consists of earned premium by accident year, indexed by $AY = 1, 2, \ldots, 10$, and a Schedule P loss triangle with losses $\{x_{AY,Lag}\}$ and $Lag = 1,\ldots,(11 - AY)$. With this data, one can calculate the probability, conditional on the parameters *ELR* and $Dev_{AY,Lag}$, of obtaining the data by the following equation.

$$L\left(\{x_{AY,Lag}\}\right) = \prod_{AY=1}^{10} \prod_{Lag=1}^{11-AY} CNB\left(x_{AY,Lag} \mid E\left[Paid\ Loss_{AY,Lag}\right]\right). \tag{4}$$

Generally one calls $L(\cdot)$ the likelihood function of the data.

For this model, maximum likelihood estimation refers to finding the parameters *ELR* and Dev_{Lag} that maximize Equation 4 (indirectly through Equation 1). There are a number of mathematical tools that one can use to do this maximization. The particular method I used is described in the Appendix.

After examining the empirical paths plotted in Figures 2 and 3, I decided to put the following constraints in the Dev_{Lag} parameters.

1. $Dev_1 \leq Dev_2$.
2. $Dev_j \geq Dev_{j+1}$ for $j = 2, 3, \ldots, 9$.
3. $Dev_7/Dev_8 = Dev_8/Dev_9 = Dev_9/Dev_{10}$.
4. $\sum_{i=1}^{10} Dev_i = 1$.

The third set of constraints was included to add stability to the tail estimates. They also reduce the number of free parameters that need to be estimated from eleven to nine. The last constraint eliminated an overlap with the *ELR* parameter and maintained a conventional interpretation of that parameter.

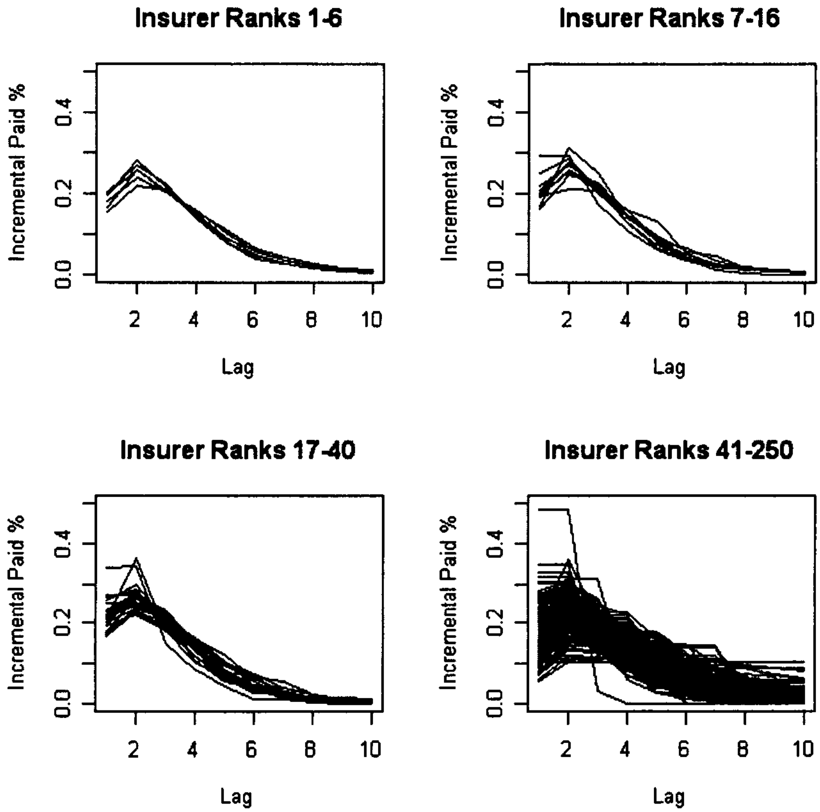
Figure 5 plots the fitted payment paths for each of the 250 insurers. You might want to compare these payment paths with the empirical payment paths in Figure 2.

Figure 6 gives histograms of the 250 *ELR* estimates.

Figure 5

Maximum Likelihood Estimates of

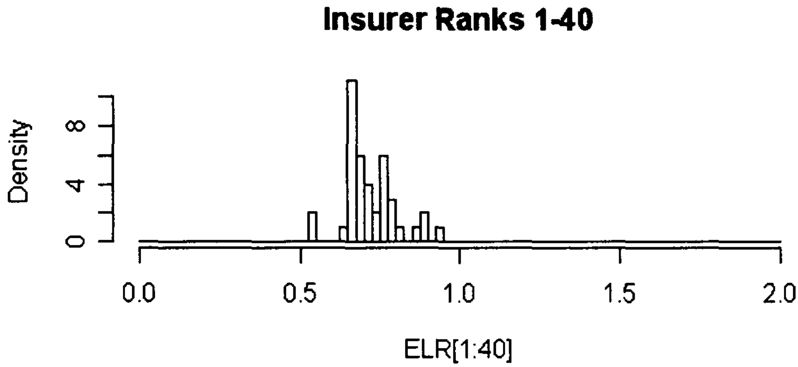
Incremental Paid Development Factors

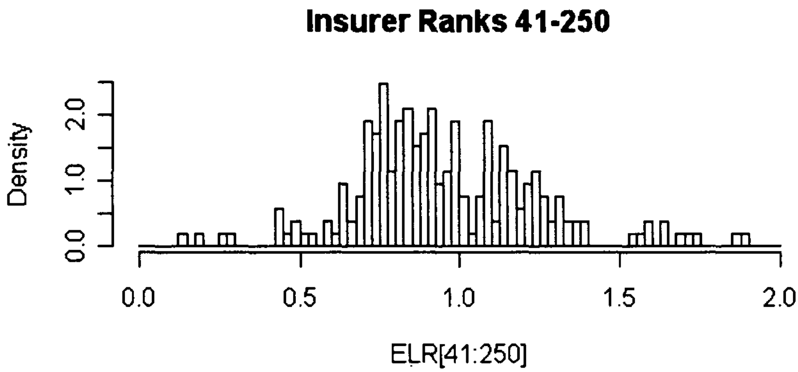


Note the wide variability of the fitted payment paths for the smallest insurers.

Figure 6

Maximum Likelihood Estimates of the ELR Parameters





Note the high variability of the *ELR* estimates for the smallest insurers.

5. Testing the Model

Given parameter estimates, *ELR* and $Dev_{AY,Lag}$, one can use the model specified by Equations 1-3 above to calculate the percentile of any observation $x_{AY,Lag}$ by first calculating the expected loss, then the expected claim count, and finally the distribution of losses about the expected loss by the *CNB* distribution. Whatever the expected losses, accident year or settlement lag, the percentiles should be uniformly distributed. One can also include the calculated percentiles of several insurers to give a more conclusive test of the model.

The hypothesis that any given set of numbers has a uniform distribution can be tested by the Kolmogorov-Smirnov test. See (for example) Klugman, Panjer and Willmot (KPW) [2004, p.428] for a reference on this test. The test is applied in our case as follows. Suppose you have a sample of numbers, $F_1, F_2, \ldots, F_n$, between 0 and 1, sorted in increasing order. One then calculates the test statistic:

$$D = \max_i \left| F_i - \frac{i}{n+1} \right|. \tag{5}$$

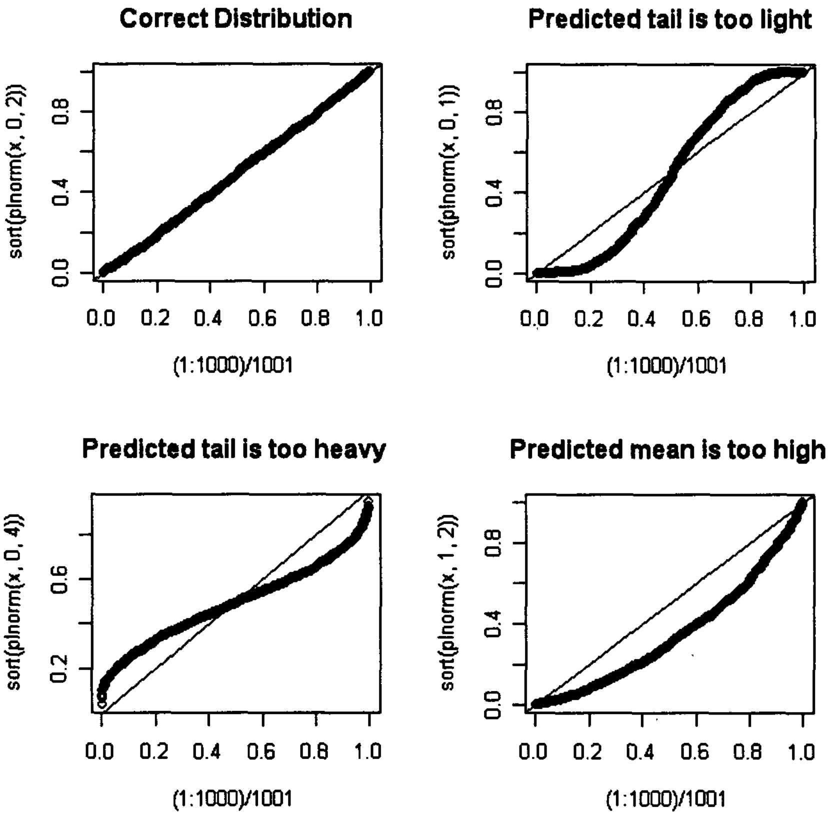
If D is greater than the critical value for a selected level, α, we reject the hypothesis that the F_i's are uniformly distributed. The critical values depend upon the sample size. Commonly used critical values are $1.22/\sqrt{n}$ for $\alpha = 0.10$, $1.36/\sqrt{n}$ for $\alpha = 0.05$, and $1.63/\sqrt{n}$ for $\alpha = 0.01$.

A graphical way to test for uniformity is a p-p plot, which is sometimes called a probability plot. A good reference for this is KPW [2004, p.424]. The plot is created by arranging the observations $F_1, F_2, \ldots, F_n$, in increasing order and plotting the points $(i/(n+1), F_i)$ on a graph. If the model is "plausible" for the data, the points will be near the 45° line running from (0,0) to (1,1). Let d_α be a critical value for a Kolmogorov-Smirnov test. Then the p-p plot for a plausible model should lie within $\pm\, d_\alpha$ of the 45° line.

A nice feature of p-p plots is that they provide, to the trained eye, a diagnosis of problems that may arise from an ill-fitting model. Let's look at some examples. Let x be a random sample of 1,000 numbers from a lognormal distribution with parameters $\mu = 0$ and $\sigma = 2$. Let's look at some p-p plots when we mistakenly choose a lognormal distribution with different μ's and σ's. On Figure 7a, sort(plnorm(x, μ, σ)) on the vertical axis will denote the sorted F_i's predicted by a lognormal distribution with parameters μ and σ.

Figure 7a

Sample p-p plots



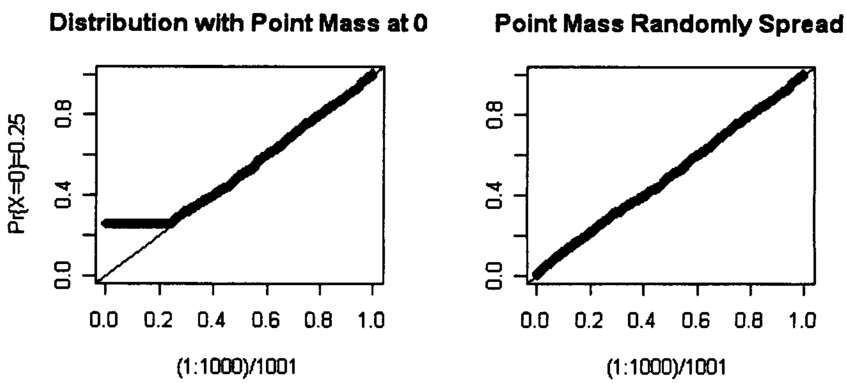
- On the first graph, μ and σ are the correct parameters, and the p-p plot lies on a 45° line as expected.
- On the second graph with $\sigma = 1$, the low predicted percentiles are lower than expected, while the high predicted percentiles are higher than expected. This indicates that the tails are too light.

- On the third graph with $\sigma = 4$, the low predicted percentiles are higher than expected, while the high predicted percentiles are lower than expected. This indicates that the tails are too heavy.
- On the fourth graph, with $\mu = 1$, almost all the predicted percentiles are lower than expected. This indicates that the predicted mean is too high.

If a random variable X has a continuous cumulative distribution function $F(x)$, the F_i's associated with a sample $\{x_i\}$ will have a uniform distribution. There are times when we want to use a p-p plot with a random variable X, which we expect to have a positive probability at $x = 0$. The left side of Figure 7b shows a p-p plot for a distribution with $\Pr\{X=0\} = 0.25$. The Kolmogorov-Smirnov test is not applicable in this case. However we can "transform" the F_i's into a uniform distribution by multiplying the $F_i = F(x_i)$ by a random number that is uniformly distributed whenever $x_i = 0$. We can then use the Kolmogorov-Smirnov test of uniformity. The right side of Figure 7b illustrates the effect of such an adjustment. All of the p-p plots below will have this adjustment.

Figure 7b

Sample p-p plots



Now let's try this for real.

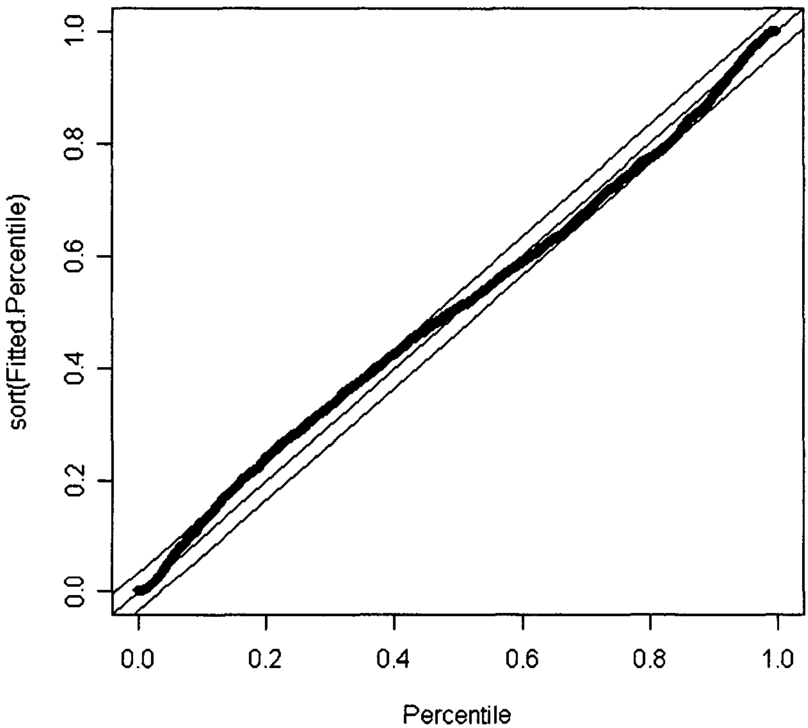
Figure 8 gives a p-p plot for the percentiles predicted for the data that was used to fit each model for the top 40 insurers. Overall there were 2,200 (= 40 × 55) calculated percentiles. The Kolmogorov-Smirnov *D* statistic for this sample was 0.042. This is higher than the critical values of 0.035 at the $\alpha = 1\%$ level and 0.029 at 5% level. So we must reject the hypothesis that our model gives a good fit to the data. By examining Figure 8, we see that the fitted model has tails that are a bit too heavy.

Let me make a personal remark here. In my many years of fitting models to data, it is a rare occasion when a model passes such a test with data consisting of thousands of observations. I was delighted with the goodness of fit. Nevertheless, I investigated further to see what "went wrong." Figure 9 shows p-p plots for the same data segregated by settlement lag. These plots appear to indicate that the main source of the problem is in the distributions predicted for the lower settlement lags.

Figure 10 shows p-p plots for the percentiles predicted for the data used in fitting the smallest 210 insurers. Suffice it to say that these plots reveal serious problems with using this estimation procedure with the smaller insurers. I think the problem lies in fitting a model with nine parameters to noisy data consisting of 55 observations. On the other hand, the procedure appears to work fairly well for large insurers with relatively stable loss payment patterns. See Figures 2, 3, 5 and 6. I suspect the same problem with small insurers occurs with other many-parameter models such as the chain ladder method.

Figure 8

P-P Plot for the Top 40 Insurers



- The Kolmogorov-Smirnov *D* statistic for this sample of 2,200 observations was 0.042. Compare this with the critical value of 0.035 at the 1% level and 0.029 at the 5% level. The sample consisted of 2,200 individual observations.
- The lines that are 0.035 above and below the 45° lines enclose the confidence band for the p-p plot at the 1% level.

Figure 9

P-P Plots for the Top 40 Insurers by Settlement Lag

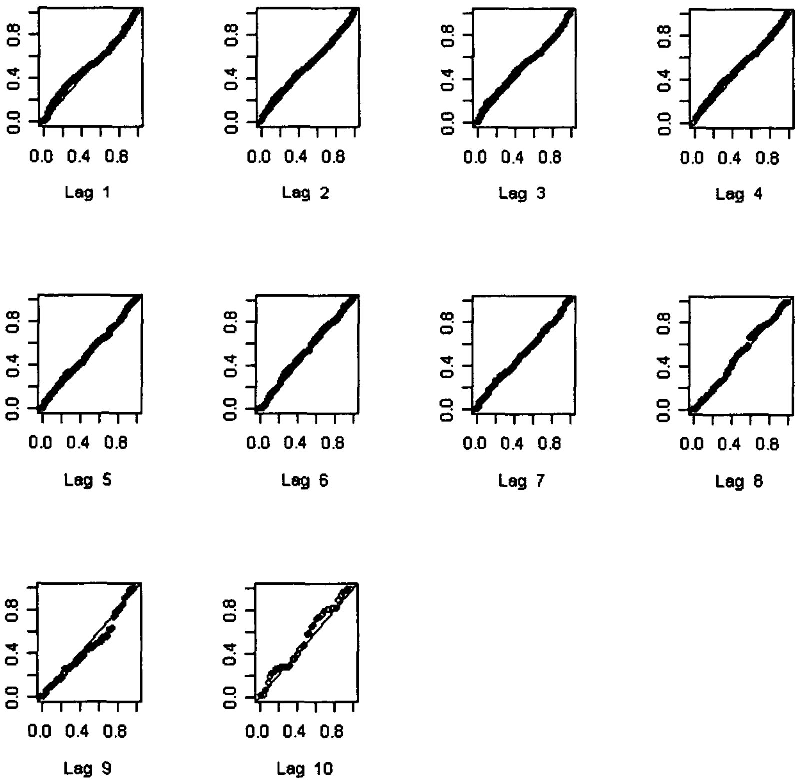
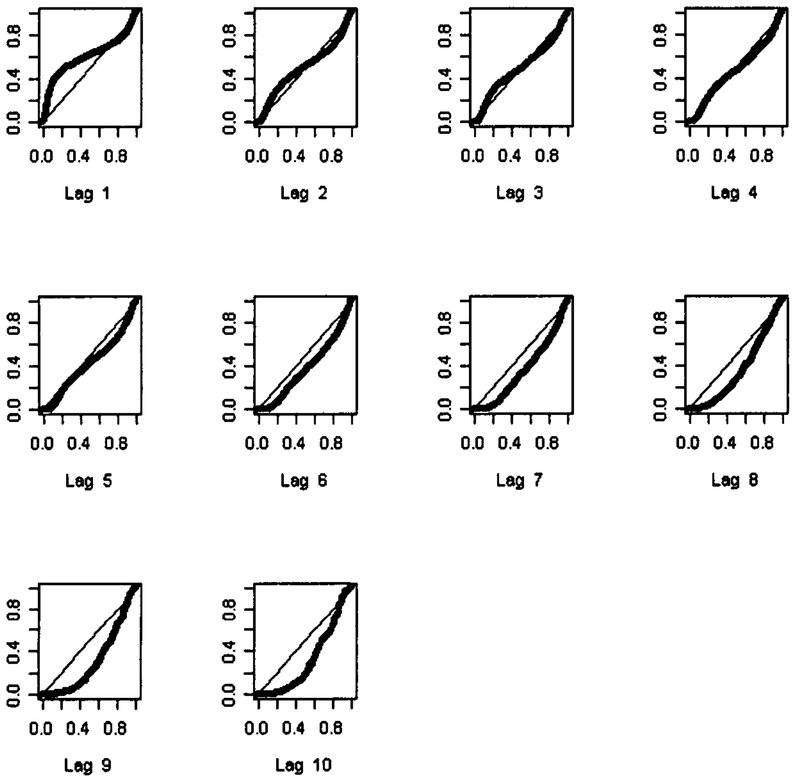


Figure 10

P-P Plots by Settlement Lag for Insurers Ranked 41-250



- These p-p plots reveal serious problems with fitting the model to smaller insurers.

6. Predicting Future Loss Payments Using Bayes' Theorem

The failure of the model to predict the distribution of losses for the smaller insurers and the comparatively successful predictions of the model on larger insurers leads one to ask the following. Is there any information that can be gained from the larger insurers that would be helpful in predicting the loss payments of the smaller insurers? That is the topic of this section.

Let $\Omega = \{ELR, Dev_{Lag}, Lag = 1, 2, \ldots, 10\}$ be a set of models, indexed by ω, that determine the expected losses in accordance with Equation 1. These models are distinguished only by the values of their parameters, and not by the assumptions or methods that were used to generate the parameters. Using Equation 4, one can combine each expected loss model $\omega \in \Omega$ with the parameters as assumptions underlying Equations 2 and 3 to calculate the likelihood of the a given loss triangle $\{x_{AY,Lag}\}$. Each likelihood can be interpreted as:

$$L = \text{Probability}\{\text{data} \mid \text{model}\} \equiv \Pr\left\{\left\{x_{AY,Lag}\right\}\middle|\omega\right\}. \tag{6}$$

Then using Bayes' Theorem one can then calculate:

$$\text{Probability}\{\text{model} \mid \text{data}\} \propto \text{Probability}\{\text{data} \mid \text{model}\} \times \text{Prior}\{\text{model}\}.$$

Stated more mathematically:

$$\Pr\left\{\omega\middle|\left\{x_{AY,Lag}\right\}\right\} \propto \Pr\left\{\left\{x_{AY,Lag}\right\}\middle|\omega\right\} \times \Pr\{\omega\}. \tag{7}$$

Each $\omega \in \Omega$ will consist of forty $\{Dev_{Lag}\}$ combinations taken from maximum likelihood estimates of the top 40 insurers above. I judgmentally selected equal probabilities for each $\omega \in \Omega$. Each of the forty $\{Dev_{Lag}\}$ combinations will be independently crossed with nine potential *ELRs* starting with 0.600 and increasing by steps of 0.025 to 0.800. Thus Ω has 360 parameter sets. I judgmentally selected the prior probability of the *ELRs* after an inspection of the distribution of maximum likelihood estimates. See Figure 11 and Table 1 below.

Figure 11

Comparing the Selected Prior Distribution of *ELR* with the Maximum Likelihood Estimates of ELR for the Top 40 Insurers

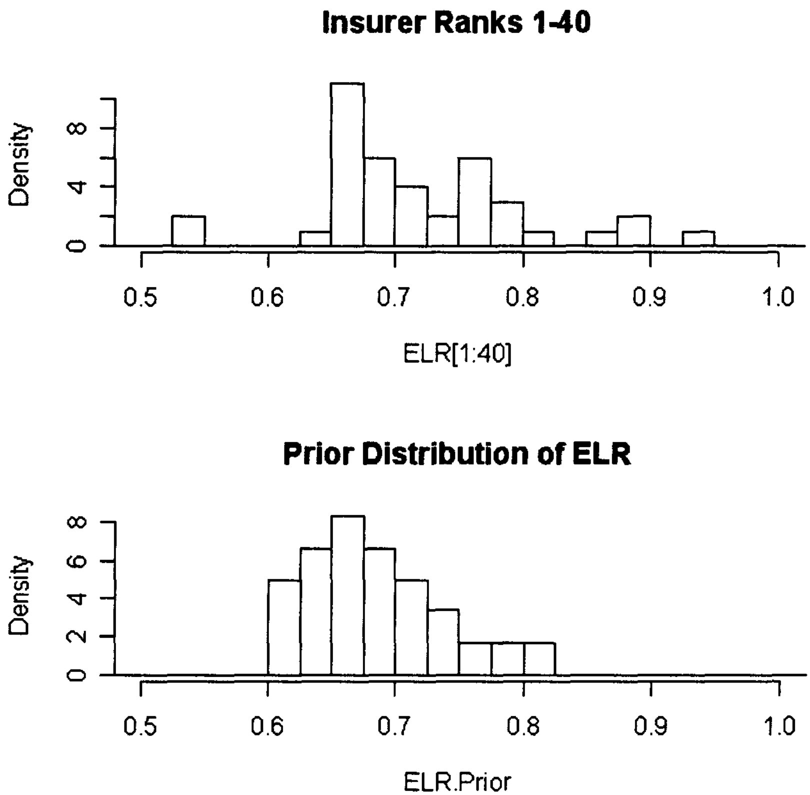


Table 1

Prior Probabilities for *ELR*

ELR	Prior	*ELR*	Prior	*ELR*	Prior
0.600	3/24	0.675	4/24	0.750	1/24
0.625	4/24	0.700	3/24	0.775	1/24
0.650	5/24	0.725	2/24	0.800	1/24

So we are given a loss triangle $\{x_{AY,Lag}\}$, and we want to find a stochastic loss model for our data. Here are the steps we would take to do this.

1. Using Equation 4, calculate $\Pr\left\{\left\{x_{AY,Lag}\right\}\middle|\omega\right\}$ for each $\omega \in \Omega$.

2. The posterior probability of each $\omega \in \Omega$ is given by

$$\Pr\left\{\omega\middle|\left\{x_{AY,Lag}\right\}\right\} = \frac{\Pr\left\{\left\{x_{AY,Lag}\right\}\middle|\omega\right\} \times \Pr\{\omega\}}{\sum_{\omega\in\Omega} \Pr\left\{\left\{x_{AY,Lag}\right\}\middle|\omega\right\} \times \Pr\{\omega\}}. \tag{8}$$

In words, the final stochastic model for a loss triangle is a mixture of all the models $\omega \in \Omega$, where the mixing weights are proportional to the posterior probabilities.

Here are some technical notes.

- In doing these calculations for the 250 insurers, it happens that almost all the weight is concentrated on at most a few dozen models. So, instead of including all models ω in the original Ω, I sorted the models in decreasing order of posterior probability and dropped those after the cumulative posterior probability summed to 99.9%.
- When calculating the final model for any of the top 40 insurers, I excluded that insurer's parameters $\{Dev_{Lag}\}$ from Ω and added the parameters for 41st largest insurer in its place. I did this to reduce the chance of overfitting.

The stochastic model of Equation 8 is not the end product. Quite often, insurers are interested in statistics such as the mean, variance, or a given percentile of the total reserve. I will now show how to use the stochastic model to calculate these "statistics of interest."

At a high level, the steps for calculating the "statistics of interest" are as follows.

1. Calculate the statistic conditional on ω for each accident year and settlement lag of interest.
2. Aggregate the statistic over the desired accident years and settlement lags for each ω.
3. Calculate the unconditional statistic by mixing (or weighting) the conditional statistics of Step 2, above, with the posterior probabilities of each ω.

These steps should become clearer as we look at specific statistics. Let's start with the expected value.

1. For each accident year and settlement lag, calculate the expected value for each ω using Equation 1.

$$E\left[Paid\ Loss_{AY,Lag}\middle|\omega\right] = Premium_{AY} \times ELR(\omega) \times Dev_{Lag}(\omega).$$

2. To get the total expected loss for each ω, sum the expected values over the desired accident years and settlement lags.

$$E\left[Paid\ Loss \mid \omega\right] = \sum_{AY,Lag} E\left[Paid\ Loss_{AY,Lag}\middle|\omega\right].$$

3. The unconditional total expected loss is the posterior probability weighted average of the conditional total expected losses, with the posterior probabilities given by Equation 8.

$$E\left[Paid\ Loss\right] = \sum_{\omega\in\Omega} E\left[Paid\ Loss\middle|\omega\right] \times \Pr\left\{\omega\middle|\left\{x_{AY,Lag}\right\}\right\}.$$

Note that for each ω, the conditional expected loss will differ. Our next "statistic of interest" will be the standard deviation of these expected loss estimates. This should be of interest to those who want a "range of reasonable estimates."

The first two steps are the same as those for finding the expected loss above. In the third step we calculate $E[Paid\ Loss]$ as above but, in addition, we calculate the second moment:

3. $$SM\left[\hat{E}\left[Paid\ Loss\right]\right] = \sum_{\omega\in\Omega} E\left[Paid\ Loss\middle|\omega\right]^2 \times \Pr\left\{\omega\middle|\left\{x_{AY,Lag}\right\}\right\}.$$ Then:

$$\text{Standard Deviation}\left[\hat{E}\left[Paid\ Loss\right]\right] = \sqrt{SM\left[\hat{E}\left[Paid\ Loss\right]\right] - E\left[Paid\ Loss\right]^2}.$$

As the second example begins to illustrate, the three steps to calculating the "statistic of interest" can get complex.

Our third statistic of interest is the standard deviation of the actual loss. Before we begin, it will help to go over the formulas involved in finding the standard deviation of sums of losses.

First, recall from Equation 2 that our model imputes an expected claim count, $\lambda_{AY,Lag}$ by dividing the expected loss by the expected claim severity for the settlement lag.

Next recall the following bullet from the description of the *CNB* distribution above.

- Let $N_{AY,Lag}$ be a random variable representing the claim count. Assume that the distribution of $N_{AY,Lag}$ is given by the negative binomial distribution with mean $\lambda_{AY,Lag}$ and variance $\lambda_{AY,Lag} + c \cdot \lambda_{AY,Lag}^2$.

The negative binomial distribution can be thought of as the following process.

1. Select the random number, χ, from a gamma distribution with mean 1 and variance c.
2. Select $N_{AY,Lag}$ from a Poisson distribution with mean $\chi \cdot \lambda_{AY,Lag}$

Consider two alternatives for applying this to the claim count for each settlement lag in a given accident year.

1. Select χ independently for each settlement lag.
2. Select a single χ and apply it to each settlement lag.

If one selects the second alternative, the multivariate distribution of $\{N_{AY,Lag}\}$ is called the negative multinomial distribution. This does not change the distribution of losses of an individual settlement lag. It does generate the correlation between the claim counts by settlement lag.

I will assume that the multivariate claim count for settlement lags within a given accident year has a negative multinomial distribution. The thinking behind this is that the χ is the result of an economic process that affects how many claims occur in a given year.

Clark [2006] provides an alternative method for dealing with correlation between settlement lags.

Let $F_{Z_{Lag}}$ be the cumulative distribution for Z_{Lag}. Mildenhall [2006] shows that (stated in the notation of this paper) the distribution of $\sum_{Lag} X_{AY,Lag}$ has a *CNB* distribution with expected claim count $\lambda_{AY,Tot} = \sum_{Lag} \lambda_{AY,Lag}$ and claim severity distribution

$$F_{Z_{AY,Tot}} = \sum_{Lag} \lambda_{AY,Lag} \cdot F_{Z_{Lag}} / \sum_{Lag} \lambda_{AY,Lag}.$$

Now let's describe the three steps to calculate the standard deviation of the actual loss.

1. For each accident year and settlement lag, calculate the expected claim count, $\lambda_{AY,Lag}(\omega)$ using Equation 2.

2. The aggregation for each ω takes place in two steps.

   a. Calculate the first and second moments of each accident year's actual loss.

$$E\left[Paid\ Loss_{AY} \middle| \omega\right] = \lambda_{AY,Tot}(\omega) \cdot E\left[Z_{Tot}\right].$$

$$SM\left[Paid\ Loss_{AY} \middle| \omega\right] = \lambda_{AY,Tot}(\omega) \cdot SM\left[Z_{AY,Tot}\right] + (1+c) \cdot \lambda_{AY,Tot}(\omega)^2 \cdot E\left[Z_{AY,Tot}\right]^2.$$

   b. Sum the first and second moments over the accident years.

$$E\left[Paid\ Loss \middle| \omega\right] = \sum_{AY} E\left[Paid\ Loss_{AY} \middle| \omega\right].$$

$$SM\left[Paid\ Loss \middle| \omega\right] = \sum_{AY} SM\left[Paid\ Loss_{AY} \middle| \omega\right].$$

3. $$E\left[Paid\ Loss\right] = \sum_{\omega \in \Omega} E\left[Paid\ Loss \middle| \omega\right] \times \Pr\left\{\omega \middle| \left\{x_{AY,Lag}\right\}\right\}.$$

$$SM\left[Paid\ Loss\right] = \sum_{\omega \in \Omega} SM\left[Paid\ Loss \middle| \omega\right] \times \Pr\left\{\omega \middle| \left\{x_{AY,Lag}\right\}\right\}.$$

$$\text{Standard Deviation}\left[Paid\ Loss\right] = \sqrt{SM\left[Paid\ Loss\right] - E\left[Paid\ Loss\right]^2}.$$

The final "statistic of interest" is the distribution of actual losses. We are fortunate that the *CNB* distribution of each individual $X_{AY,Lag}$ is already defined in terms of its Fast Fourier Transform (FFT). To get the FFT of the sum of losses, we can simply multiply the FFTs of the summands. Other than that, the three steps are similar to those of calculating the standard deviation of the actual losses. To shorten the notation, let X denote *Paid Loss*.

1. For each accident year and settlement lag, calculate the expected claim count, $\lambda_{AY,Lag}(\omega)$ using Equation 2.

2. The aggregation for each ω takes place in two steps.

   a. Calculate the FFT,

   $$\Phi_{X_{AY,Tot}|\omega}(\vec{\mathbf{q}}) = \sum_{Lag=1}^{11-AY} \lambda_{AY,Lag} \cdot \Phi_{X_{AY,Lag}|\omega}(\vec{\mathbf{q}}_{Lag}) \Big/ \sum_{Lag=1}^{11-AY} \lambda_{AY,Lag},$$

   for each accident year.

   b. The FFT for the sum of all accident years is given by:

   $$\Phi_{X|\omega}(\vec{\mathbf{q}}) = \prod_{AY} \Phi_{X_{AY,Tot}|\omega}(\vec{\mathbf{q}})$$

3. The distribution of actual losses is obtained by inverting the FFT:

$$\Phi_X(\vec{\mathbf{q}}) = \prod_{\omega \in \Omega} \Phi_{X|\omega}(\vec{\mathbf{q}}) \times \Pr\left\{\omega \middle| \left\{x_{AY,Lag}\right\}\right\}.$$

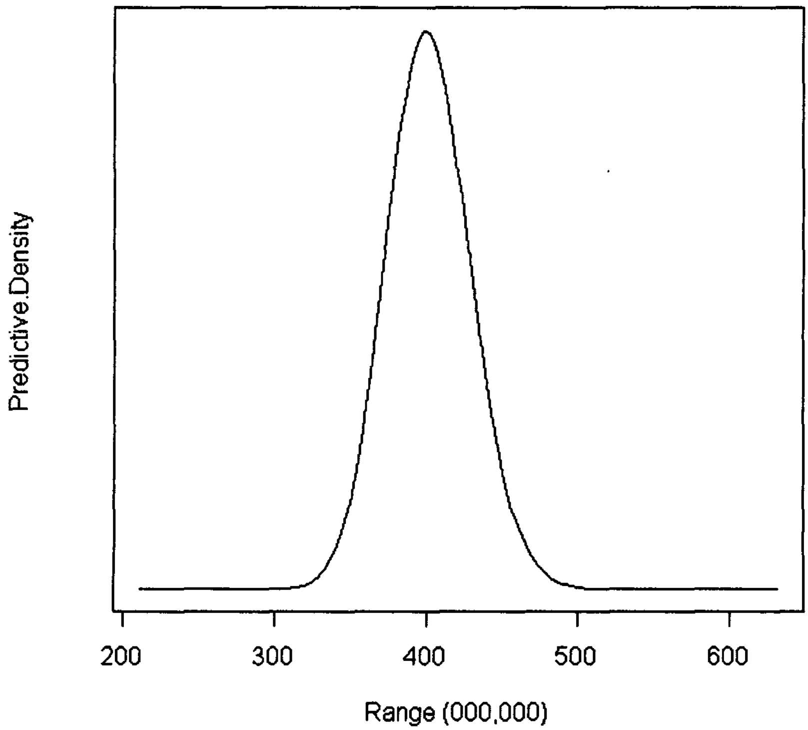
See the Appendix for additional mathematical details of working with FFTs.

Figures 12 and 13 below show each of the three statistics for two insurers for the outstanding losses for accident years 2,...,10 up to settlement lag 10. The insurer in Figure 12 has ten times the predictive mean reserve as the insurer in Figure 13. Figure 14 plots the predictive coefficient of variation against the predictive mean reserve. The decreased variability that comes with size should not come as a surprise. The absolute levels of variability will be interesting only if I can demonstrate that this methodology can predict the distribution of future results. That is where I am going next.

Figure 12

Predictive Distribution of Actual Losses for Total Reserve

Insurer Rank 7

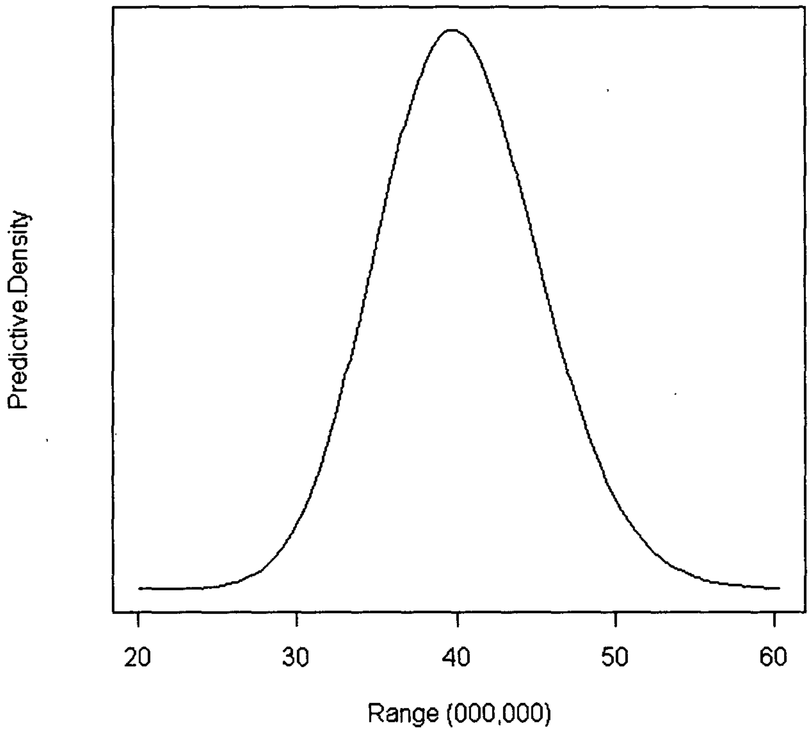


- Predictive Mean = 401,951,000 (roughly ten times that in Figure 13).
- Coefficient of Variation for the Actual Loss = 6.9%.

Figure 13

Predictive Distribution of Actual Losses of Total Reserve

Insurer Rank 97

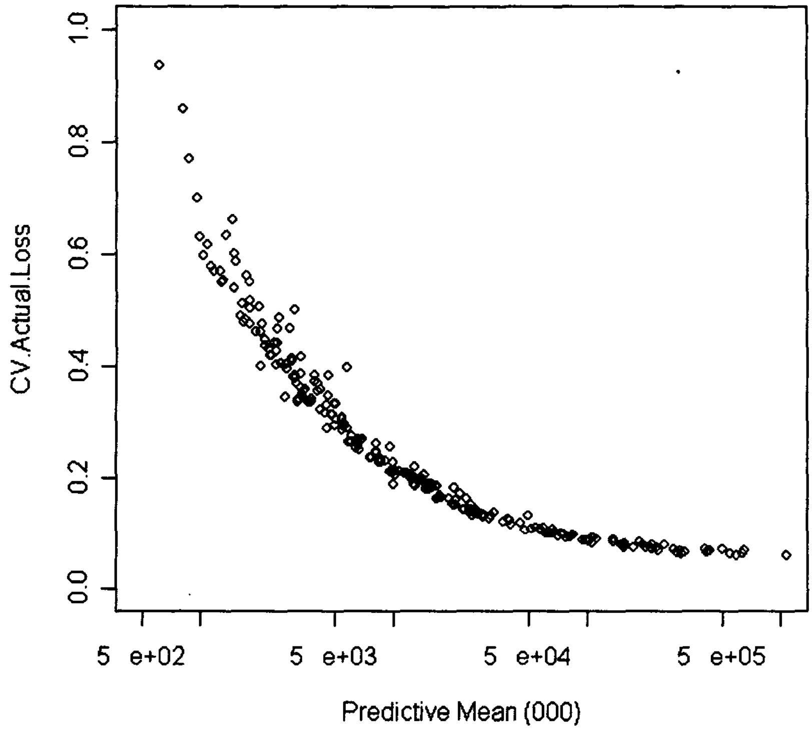


- Predictive Mean = 40,277,000 (roughly one tenth of that in Figure 12).
- Coefficient of Variation for the Actual Loss = 12.6%.

Figure 14

Predictive Coefficient of Variation Plotted

With the Predictive Mean for 250 Insurers



7. Testing the Predictions

The ultimate test of a stochastic loss reserving model is its ability to correctly predict the distribution of future payments. While the distribution of future payments will differ by insurer, when one calculates the predicted percentile of the actual payment, the distribution of these predicted percentiles should be uniform.

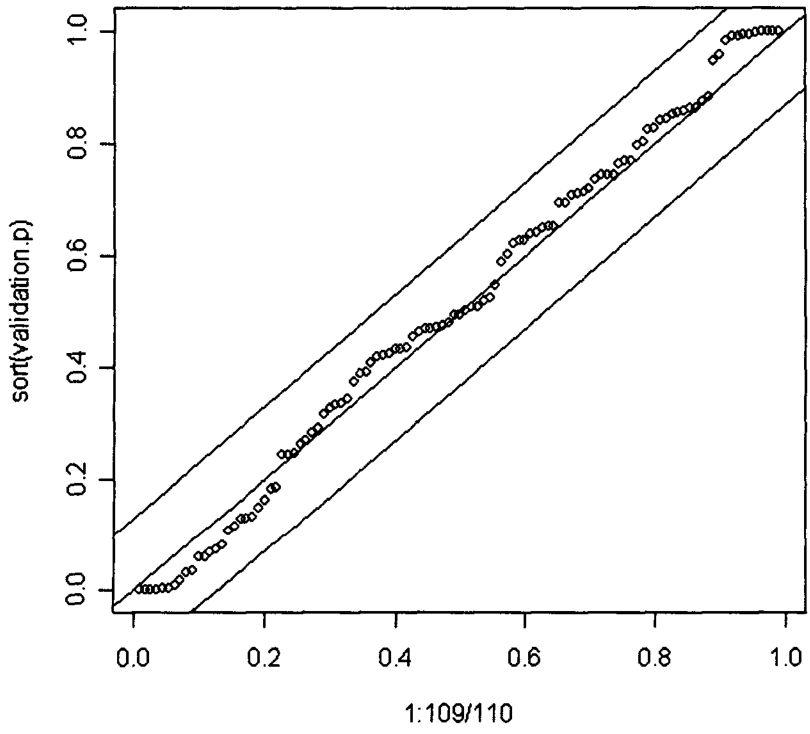
To test the model, we examined Schedule P from the 2001 NAIC Annual Statement. The losses reported in these statements contain six subsequent diagonals on the four overlapping years from 1992 through 1995. Earned premiums and losses in the overlapping diagonals for the 1995 and 2001 Annual Statements agreed in 109 of the 250 insurers, so I used these 109 insurers for the test.

Using the predictive distribution described in the last section, I calculated the predicted percentile of the total amount paid for the four accident years in the subsequent six settlement lags. These 109 percentiles should be uniformly distributed. Figure 15 shows the corresponding p-p plot and the confidence bands at the 5% level as determined by the Kolmogorov-Smirnov test. The plot lies well within that band. While one can never "prove" a model is correct with statistics, one gains confidence in a model as we fail to reject the model with such statistical tests. I believe this test shows that the Bayesian *CNB* model deserves serious consideration as a tool for setting loss reserves.

Figure 15

P-P Plot of Predicted Percentiles for

Paid Losses from 1996 to 2001



- The critical values for a Kolmogorov-Smirnov test at the 5% level are ±13.03%.

8. Comparing the Predictive Reserves with Reported Reserves

This section provides an illustration of the kind of analysis that can be done externally with the Bayesian methodology described in this paper. Readers should exercise caution in generalizing the conclusions of this section beyond this particular line of business in this particular time period.

This paper makes no attempt to pin down the methods used in setting the reported reserves. However there are many actuaries that expect reported reserves to be more accurate than a formula derived purely from the paid data reported on Schedule P. As stated in the introduction to this paper, those who set those reserves have access to more information that is relevant to estimating future loss payments.

The comparisons below will be performed to two sets of insurers – the entire set of 250 insurers and the subset of 109 insurers for which the overlapping accident years 1992-95 agree. Testing the latter will enable us to compare the predictions based on information available in 1995 with the incurred losses reported in 2001.

The first test looks at aggregates summed over all insurers in each set. Table 2 compares the predictions of this model with the actual reserves reported on the 1995 annual statement. The "actual reserve" is the difference between the total reported incurred loss, as of 1995 for the "initial" reserve, and 2001 for the "retrospective" reserve, minus the total reported paid loss, as of 1995.

Table 2

Predicted and Reported Loss Reserves

		Reported 1995 Reserve (000)	
	Predictive Mean (000)	Initial @ 1995	Retrospective @ 2001
250 Insurers AY 1986-1995	14,873,303	16,221,998 - 9.1%	---
109 Insurers AY 1992-1995	1,798,794	1,976,299 – 9.9%	1,842,104 – 2.4%

For the 250 insurers, the reported initial reserve was 9.1% higher than the predictive mean. For the 109 insurers the corresponding percentage was 9.9%. The lowering of the percentage reserves from 1995 to 2001 to 2.4% suggests that for the industry, reserves were redundant for Commercial Auto in 1995[2].

For the remainder of this section let's suppose that the expected value of the Bayesian *CNB* model described above is the "best estimate" of future loss payments. From the above, there are two arguments supporting that proposition.

1. Figure 15 in Section 7 above shows that the Bayesian *CNB* model successfully predicted the distribution of payments for the six years after 1995 well within the normal statistical bounds of error.
2. The final row of Table 2 shows that the expected value predicted by the Bayesian *CNB* model, in aggregate, comes closer to the 2001 reserve than did the reported reserves for 1995.

Now let's examine some of the implications of this proposition for reported reserves.

There are many actuaries who argue that reported reserves should be somewhat higher than the mean. See, for example, Paragraph 2.17 on page 5 of Report of the Insurer Solvency Working Party of the International Actuarial Association [2004]. Related to this, I recently saw a working paper by Grace and Leverty [2006] that tests various hypotheses on insurer incentives.

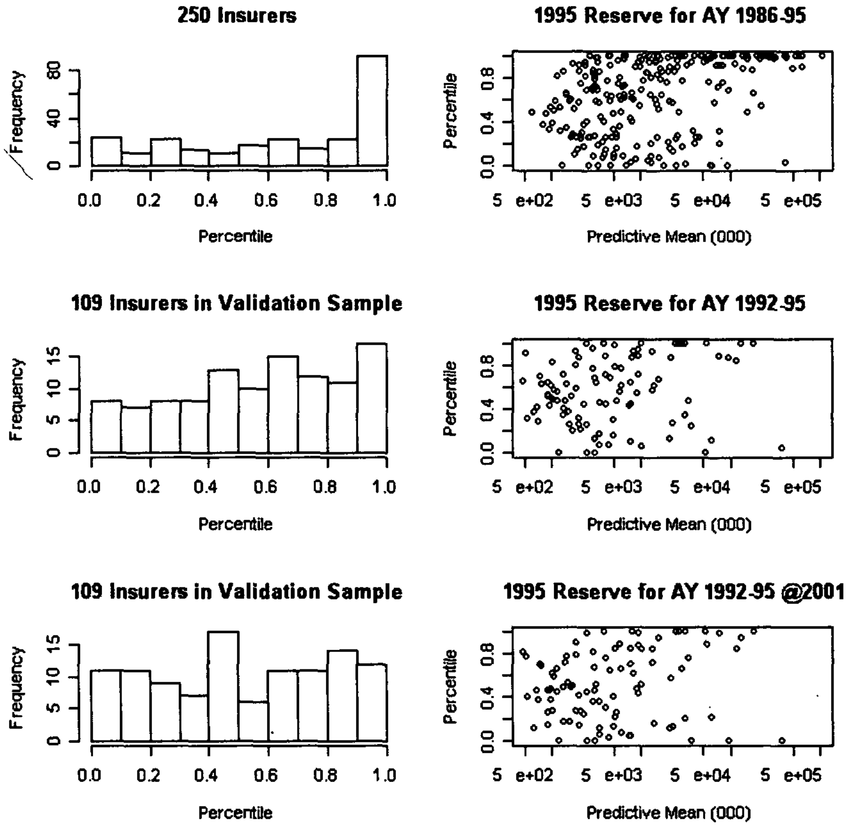
If insurers were deliberately setting their reserves at some conservative level, we would expect to see that the reported reserves are at some moderately high percentile of the predictive distribution. Figure 16 shows that some insurers appear to be reserving conservatively. But there are also many insurers for which the predictive percentile of the reported reserve is below 50%. But by 2001, the percentiles of the retrospective reserve for 1995 were close to being uniformly distributed.

---

[2] There are some potential biases in these figures. First, the predictive means may be somewhat understated since they ignore development after ten years. Second, the downward development from 1995 to 2001may continue in future years.

Figure 16

Predictive Percentiles of Reported Reserves



- The greater number of insurers reserved above the 50th percentile indicates that some insurers have conservative estimates of their loss reserves posted in 1995.
- The right side of this figure shows that the spread of the reserve percentiles spans all insurer sizes.

If there is a bias in the posted reserves, we would see corrections in subsequent years. The 109 insurers for which we have subsequent development provide data to test potential bias. To perform such a test, I divided the 109 insurers into two groups. The first group consisted of all insurers that posted reserves in 1995 that was lower than their predictive mean. The second group consisted of all insurers that posted reserves higher than their predictive mean.

As Figure 17 and Table 3 show, the first group shows an upward adjustment and the second group shows a more pronounced downward adjustment. The plots show that we cannot attribute these adjustments to only a few insurers. However, there are some insurers in the first group that show a downward adjustment, and other insurers in the second group that show an upward adjustment.

The fact that the total adjustments only go part way to the predictive mean suggests that some insurers may be able to make more accurate estimates with access to information that is not provided on Schedule P.

Figure 17

Analysis of Subsequent Reserve Changes for 109 Insurers

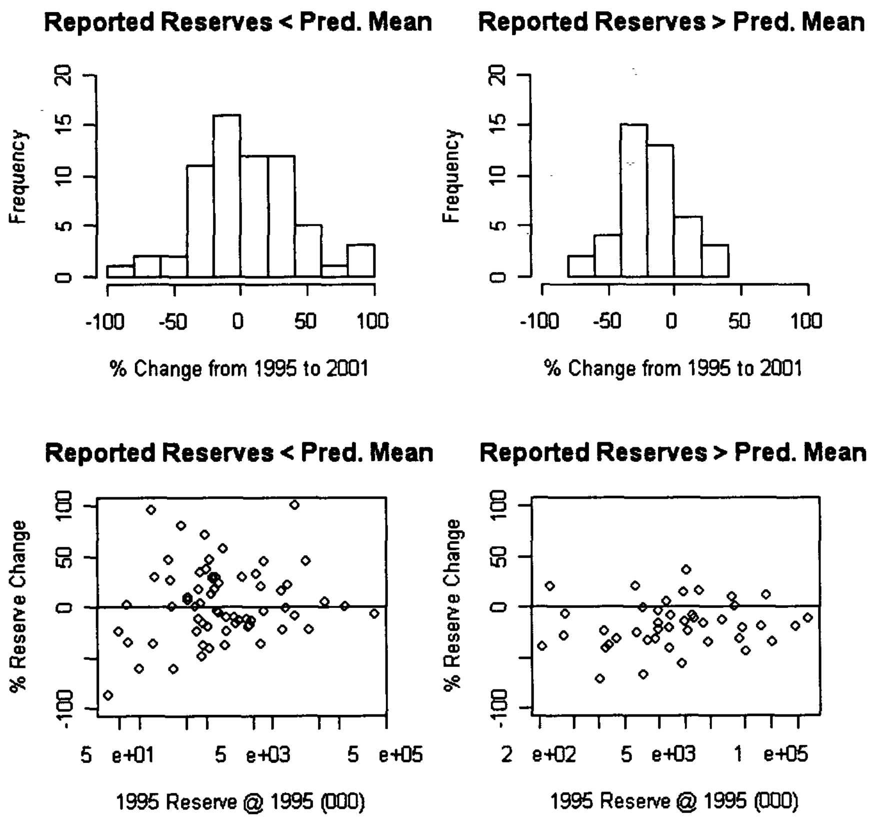


Table 3

Summary Statistics for the Plots Above

	Reported Reserve @ 1995	
	< Predictive Mean (000)	> Predictive Mean (000)
Number of Insurers	66	43
Total Predictive Mean	926,134	872,660
1995 Reserve @ 1995	803,175	1,173,124
1995 Reserve @ 2001	856,393	985,711

9. Summary and Conclusions

This paper demonstrates a method, which I call the Bayesian *CNB* model, for estimating the distribution of future loss payments of individual insurers. The main features of this method are as follows.

- The stochastic loss reserving model is based on the collective risk model. While other stochastic loss reserving approaches make use of the collective risk model, this approach uses it as an integral part of estimating the parameters of the model.
- Predicted loss payments are derived from a Bayesian methodology that uses the results of large, and presumably stable, insurers as its "prior information." While insurers do indeed differ in their claim payment practices, the underlying assumption of this methodology is that these differences are reflected in this collection of large insurers.
- Loss reserving models should be subject to testing their predictions on future payments. Tests on a single insurer are often inconclusive because of the volatile nature of the loss reserving process. But it is possible to test a stochastic loss reserving method on several insurers simultaneously by comparing its predicted percentiles of subsequent losses to a uniform distribution. This paper tests its model on 109 insurers and finds that its predictions are well within the statistical bounds expected for a sample of this size.
- By making the assumption that the Bayesian *CNB* model provides the "best estimate" of future loss payments, the analysis in this paper suggested that there are some insurers that post reserves conservatively, while others post reserves with a downward bias. Readers should exercise caution in generalizing these conclusions beyond this particular line of business in this time period.

I view this paper as an initial attempt at a new method for stochastic loss reserving. To gain general acceptance, this approach should be tested on other lines of insurance and by other researchers. This method requires considerable statistical and actuarial expertise to implement. It also takes a lot of work. In this paper, I have tried to make the case that we should expect that such efforts could yield fruitful results.

References

1. Glen Barnett and Ben Zehnwirth, "Best Estimate for Reserves," *Proceedings of the Casualty Actuarial Society*, LXXXVII, 2000, pp. 245-321.
   http://www.casact.org/pubs/proceed/proceed00/00245.pdf

2. David R. Clark, "LDF Curve Fitting and Stochastic Loss Reserving: A Maximum Likelihood Approach," *CAS Forum*, Fall 2003, pp 41-91.
   http://www.casact.org/pubs/forum/03fforum/03ff041.pdf

3. David R Clark, "Variance and Covariance Due to Inflation," *CAS Forum*, Fall 2006 (to appear)

4. Martin F. Grace and J. Tyler Leverty, "Property Liability Insurer Reserve Error-Motive, Manipulation, or Mistake," This is a working paper. Contact the authors at mgrace@gsu.edu or ty-leverty@uiowa.edu for the most recent version.

5. Roger M. Hayne, "Measurement of Reserve Variability," *CAS Forum*, Fall 2003, pp 141-172.
   http://www.casact.org/pubs/forum/03fforum/03ff141.pdf

6. Philip E. Heckman and Glenn G. Meyers, "The Calculation of Aggregate Loss Distributions for Claim Severity and Claim Count Distributions," *Proceedings of the Casualty Actuarial Society*, LXX, 1983, pp. 41-61, addendum in LXXI, 1984, pp. 49-56.
   http://www.casact.org/pubs/proceed/proceed83/83022.pdf
   http://www.casact.org/pubs/proceed/proceed84/84049.pdf

7. Stuart A. Klugman, Harry H. Panjer and Gordon E. Willmot, *Loss Models: From Data to Decisions*, 2004, Wiley Series in Probability and Statistics.
   http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471215775.html

8. Insurer Solvency Working Party, *A Global Framework for Insurer Solvency Assessment*, 2004, International Actuarial Association.
   http://www.actuaries.org/LIBRARY/Papers/Global_Framework_Insurer_Solvency_Assessment-members.pdf

9. Thomas Mack, "Distribution-Free Calculation of Chain Ladder Reserve Estimates," ASTIN Bulletin, Vol 23, No. 2, November 1993, pp. 213-226.
   http://www.casact.org/library/astin/vol23no2/213.pdf

10. Glenn G. Meyers, "The Common Shock Model for Correlated Insurance Losses," *CAS Forum*, Winter 2006, pp. 231-254.
    http://www.casact.org/pubs/forum/06wforum/06w235.pdf

11. Stephen J. Mildenhall, "Correlation and Aggregate Loss Distributions with an Emphasis on the Iman-Conover Method," *CAS Forum*, Winter 2006, pp. 103-205
    http://www.casact.org/pubs/forum/06wforum/06w107.pdf

12. James N. Stanard, "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques," *Proceedings of the Casualty Actuarial Society*, LXXII, 1985, pp. 124-148.
    http://www.casact.org/pubs/forum/06wforum/06w107.pdf

13. Shaun Wang, "Aggregation of Correlated Risk Portfolios: Models and Algorithms," *Proceedings of the Casualty Actuarial Society*, LXXXV, 1998, pp. 848-939.
    http://www.casact.org/pubs/proceed/proceed98/980848.pdf

Appendix

This appendix gives the mathematical details that implement the methodologies described in Sections 3 and 4.

A.3.1 Discretizing the Claim Severity Distributions

The first step is to determine the discretization interval length h. h, which depended on the size of the insurer, was chosen so the 2^{14} (16,384) values spanned the probable range of annual losses for the insurer. Specifically, let h_1 be the sum of the insurer's ten-year premium divided by 2^{14}. The h was set equal to 1,000 times the smallest number from the set {5, 10, 20, 25, 40, 50, 100, 125, 200, 250, 500, 1000} that was greater than $h_1/1000$. This last step guarantees that a multiple, m, of h would be equal to the policy limit of 1,000,000.

The next step is to use the mean-preserving method (described in KPW, p. 656) to discretize the claim severity distribution for each settlement lag. Let $p_{i,Lag}$ represent the probability of a claim with severity $h \cdot i$ for each settlement lag. Using the limited average severity (LAS_{Lag}) function determined from claim severity distributions provided by ISO, the method proceeds in the following steps.

1. $p_{0,Lag} = 1 - LAS_{Lag}(h)/h$.
2. $p_{i,Lag} = (2 \cdot LAS_{Lag}(h \cdot i) - LAS_{Lag}(h \cdot (i-1)) - LAS_{Lag}(h \cdot (i+1)))/h$ for i = 1, 2, ..., m-1.
3. $p_{m,Lag} = 1 - \sum_{i=0}^{m-1} p_{i,Lag}$.
4. $p_{ik} = 0$ for $i = m + 1, ..., 2^{14} - 1$.

A.3.2 Calculating the Conditional Density of the *CNB* Distribution

The purpose of this section is to show how to calculate $CNB\left(x_{AY,Lag} \mid E\left[Paid\ Loss_{AY,Lag}\right]\right)$.

The calculation proceeds in the following steps.

1. Set $\vec{\mathbf{p}}_{Lag} = \left\{p_{0,Lag}, \ldots p_{2^{14}-1,Lag}\right\}$.
2. Calculate the Fast Fourier Transform (FFT) of $\vec{\mathbf{p}}_{Lag}, \Phi_{Z_{Lag}}\left(\vec{\mathbf{p}}_{Lag}\right)$.
3. Calculate the expected claim count, $\lambda_{AY,Lag}$, for each accident year and settlement lag using Equation 2, $\lambda_{AY,Lag} \equiv E\left[Paid\ Loss_{AY,Lag}\right] / E\left[Z_{Lag}\right]$.

4. Calculate the FFT of each aggregate loss random variable, $X_{AY,Lag}$, using the formula

$$\Phi_{X_{AY,Lag}}\left(\vec{\mathbf{q}}_{Lag}\right)=\left(1-c\cdot\lambda_{AY,Lag}\cdot\left(\Phi_{Z_{Lag}}\left(\vec{\mathbf{p}}_{Lag}\right)-1\right)\right)^{-1/c}.$$

   This formula is derived in KPW [2004, Equation 6.28]. Note the different but equivalent parameterization. The probability generating function for the negative binomial distribution is given in Appendix B of KPW. It is written as $P_N(z)=\left(1-\beta(z-1)\right)^{-r}$. In this paper's notation $\lambda=\beta\cdot r$ and $c = 1/r$.

5. Calculate $\vec{\mathbf{q}}_{AY,Lag}=\Phi^{-1}\left(\Phi_{AY,Lag}\left(\vec{\mathbf{q}}_{AY,Lag}\right)\right)$, the inverse FFT of the expression in Step 4 above.

6. Set i equal to the multiple of h that is nearest to $x_{AY,Lag}$. Then

$$CNB\left(x_{AY,Lag}\mid E\left[\textit{Paid Loss}_{AY,Lag}\right]\right)=\text{ the } i^{th} \text{ component of } \vec{\mathbf{q}}_{AY,Lag}.$$

Note that calculating this probability requires one to first calculate a vector of length 16,384 by inverting an FFT and reading off a single component. (To increase efficiency, one should calculate $\Phi_{Z_{Lag}}\left(\vec{\mathbf{p}}_{Lag}\right)$ for each settlement lag in advance.) Using the R computing language (www.r-project.org) on my 3GHz personal computer with 1GB Ram, I estimate it takes about 1/20th of a second to evaluate a single *CNB* probability. Evaluating a likelihood for a loss triangle with 55 $x_{AY,Lag}$s 1,000 times (typical for what follows below) takes about 45 minutes. Implementing this methodology requires the patience that I was fortunate to develop in the early days of actuarial computing.

A.4 Maximizing the Likelihood for the *CNB* Model

The purpose of this section is to show how to find the ELR and $\{Dev_i\}$ parameters that maximize the likelihood

$$L\left(\left\{x_{AY,Lag}\right\}\right)=\prod_{AY=1}^{10}\prod_{Lag=1}^{11-AY} CNB\left(x_{AY,Lag} \mid E\left[Paid\ Loss_{AY,Lag}\right]\right), \tag{4}$$

subject to the following constraints in the Dev_{Lag} parameters.

1. $Dev_1 \leq Dev_2$.
2. $Dev_j \geq Dev_{j+1}$ for $j = 2, 3, \ldots, 7$.
3. $Dev_7/Dev_8 = Dev_8/Dev_9 = Dev_9/Dev_{10}$.
4. $\sum_{i=1}^{10} Dev_i = 1$.

The maximization was done using the R programming language *optim* function using the Nelder-Mead parameter search method. This method is described in KPW [2004, p.664] and is considered to be robust but slow. At this stage of the research, I value "robust" over "fast."

Primarily because of habits I developed using Excel Solver, I elected not to use standard constraints provided by the function. Instead I coded a "*tdev to Dev*" function that mapped all of $\mathbb{R}^9$ into a subset of $\mathbb{R}^{11}$ that satisfied those constraints. Here is a description of *tdev to Dev.*

1. $Dev_1' = e^{-tdev_1^2}/2$.
2. $Dev_2' = Dev_1' \cdot \left(1+e^{-tdev_2^2}\right)$.
3. $Dev_i' = Min\left[\left(1-\sum_{j=1}^{i-1} Dev_j'\right), Dev_{i-1}'\right] \cdot e^{-tdev_i^2}$ for i=2,...,7.
4. $Dev_i' = Min\left[\left(1-\sum_{j=1}^{i-1} Dev_j'\right), Dev_{i-1}'\right] \cdot e^{-tdev_8^2}$ for i=8, 9, 10.
5. $Dev_i = Dev_i' / \sum_{j=1}^{10} Dev_j'$.
6. $ELR = tdev_9^2 \cdot \sum_{j=1}^{10} Dev_j'$.

As noted in the previous section, the *CNB* model requires a lot of time to calculate. This time can be significantly reduced if one has a good set of starting values for the *optim* function. To get these starting values, I replaced the *CNB* distribution with the "overdispersed Poisson" (*ODP*) distribution given in Clark [2003] to find the ELR and $\{Dev_i\}$ parameters that maximize the logarithm of following expression.

$$L\left(\left\{x_{AY,Lag}\right\}\right)=\prod_{AY=1}^{10}\prod_{Lag=1}^{11-AY} E\left[Paid\ Loss_{AY,Lag}\right]\cdot e^{x_{AY,Lag}\cdot E\left[Paid\ Loss_{AY,Lag}\right]-E\left[Paid\ Loss_{AY,Lag}\right]}.$$

The maximization proceeds in the following steps.

1. Pick a starting vector in $\vec{s}\in\mathbb{R}^9$ e.g. (1,1,1,1,1,1,1,1,1).
2. Set $\vec{t}$ = *tdev to Dev*($\vec{s}$) and use it to calculate $E\left[Paid\ Loss_{AY,Lag}\right]$.
3. Use $E\left[Paid\ Loss_{AY,Lag}\right]$ to calculate the *ODP* likelihood above.
4. Use the Nelder-Mead algorithm to calculate an updated vector $\vec{s}$.
5. Return to Step 2 and repeat until convergence.
6. After convergence is obtained with the *ODP* likelihood, use the current $\vec{s}$ as a starting value for the *CNB* likelihood in Equation 4.
7. Set $\vec{t}$ = *tdev to Dev*($\vec{s}$) and use it to calculate $E\left[Paid\ Loss_{AY,Lag}\right]$.
8. Use $E\left[Paid\ Loss_{AY,Lag}\right]$ to calculate the *CNB* likelihood above.
9. Use the Nelder-Mead algorithm to calculate an updated vector $\vec{s}$.
10. Return to Step 7 and repeat until convergence.
11. Set $\vec{t}$ = *tdev to Dev*($\vec{s}$) to obtain the maximum likelihood estimate of ELR and $\{Dev_i\}$.

Run time was short for the *ODP*. For the *CNB*, I found that it generally took, on average, 1,000 iterations of Steps 7-10 to achieve R's *optim* function default convergence criteria. With the warning that individual results may vary, I felt comfortable in limiting the number of iterations to 300.

I am providing code to calculate the above maximum likelihood estimates on sample data to be placed on the CAS website with the publication of this paper.

Acknowledgements

The author would like to thank Irene Bass, David Clark, Daniel Crifo, Eugene Gaydos, Roger Hayne, Stan Khury, John Kollar, Stephen Mildenhall and Vladimir Shander who reviewed early drafts of this paper and offered several helpful comments.

Biography of the Author

Glenn Meyers is the Chief Actuary for ISO Innovative Analytics, a division of ISO devoted to predictive modeling. He holds a bachelor's degree in mathematics and physics from Alma College, a master's degree in mathematics from Oakland University, and a Ph.D. in mathematics from the State University of New York at Albany. Glenn is a Fellow of the Casualty Actuarial Society and a member of the American Academy of Actuaries. Before joining ISO in 1988, Glenn worked at CNA Insurance Companies and the University of Iowa.

Glenn's current responsibilities at ISO include the development of scoring products. Prior responsibilities have included working on ISO Capital Management products, increased limits and catastrophe ratemaking, ISO's, and Property Size-of-Loss Database (PSOLD), ISO's model for commercial property size of loss distributions.

Glenn's work has been published in *Proceedings of the Casualty Actuarial Society (CAS).* He is a three-time winner of the Woodward-Fondiller Prize, a two-time winner of the Dorweiller Prize and a winner of the Dynamic Financial Analysis Prize. He is a frequent speaker at CAS meetings and seminars.

His service to the CAS has included membership on various education and research committees. He currently serves on the International Actuarial Association Solvency Committee and the CAS Board of Directors.

A Method For Projecting Individual Large Claims

Karl Murphy FIA and Andrew McLennan FIAA, FIA

Abstract

Motivation. The paper will address the issue of estimating the uncertainty in the run off of individual large claims in insurance portfolios, which is often the primary source of uncertainty in the reserving risk component of insurance risk.

Method. The paper begins by reviewing current methodologies for estimating the uncertainty in loss reserves. Methods until now have focused on aggregate modeling of gross or net of reinsurance loss reserves, and no direct connection between the distribution of gross and net reserves.

The paper develops a non-parametric framework to simulate the distribution of ultimate position of large claims, both reported and large IBNR claims. The method samples the development of individual claims based on the historic development of large claims, incorporating information at an aggregate level surrounding reserving strength. The model also predicts when claims will settle, and the timing of claim payments.

Results. The method developed is not intended to replace existing aggregate modeling, but is an improvement to traditional methods which estimate the variability of gross of reinsurance loss reserves, and is a useful tool to allow for reinsurance recoveries more accurately.

By individually projecting the ultimate position of large claims, we can explicitly allow for policy or contract limits. Further, we can apply any reinsurance program structure to the gross results, including allowance for aggregate deductibles, incomplete placements, retrocessions to captive reinsurers, indexation clauses, and different treaty attachment rules (ie Losses Occurring During vs Risks Attaching).

The paper then shows how the variability of attritional claims can be estimated using traditional stochastic methods, and the attritional and large results can be combined to estimate the variability of the aggregate portfolio of loss reserves.

Keywords. Reserving, Large Claims, Reinsurance, Stochastic Modeling, Simulation, Capital Modeling, IBNR.

1. INTRODUCTION

With an increased focus on understanding variability in claims reserves, a series of papers have been published which develop and add to existing literature on stochastic reserving, in particular England and Verrall[1]. However, almost universally, these papers consider aggregate claims triangles, and do not consider the range of possible outcomes of individual claims. We believe that for many classes of business, the primary source of uncertainty in reserve run-off stems from the uncertainty in large claims, and so a natural extension to the developments in stochastic claims reserving methods would be to produce stochastic outcomes of individual claims.

The paper develops a practical framework to simulate the distribution of ultimate position of large claims, both reported and large IBNR claims. The method samples the development of individual claims based on the historic development of large claims, and applies this development to the current position of claims. The model also predicts when claims will settle, and the timing of claim payments.

A practical by-product of having individually projected the ultimate position of large claims is that we can apply any policy contract limits to any claims, and any reinsurance program structure to the gross results in order to derive stochastic net results that are consistent with the gross without having to make simplifying approximations. For example, by having individual large claims information, excess of loss reinsurance can be properly allowed for. Other more complicated arrangements can also be considered, including allowance for aggregate deductibles, incomplete placements, retrocessions to captive reinsurers, indexation clauses, and different treaty attachment rules (i.e. "losses occurring during" treaties compared to "risks attaching" treaties). Reinsurance recoveries can then be allocated to specific contracts, enabling easier commutation and reinsurance bad debt calculations.

The paper then shows how the variability of attritional claims can be estimated using aggregate stochastic methods, and the attritional and large results can be combined to estimate the variability of the aggregate portfolio of loss reserves. By separating large and attritional claims in the estimation of the uncertainty in loss reserves, changes to the mix (by size and numbers) of large claims can be directly allowed for and modeled.

The structure of the paper will be as follows: first we are going to briefly discuss the main existing stochastic methods for estimating reserving risk. We will then look at a new method which we believe better identifies the main source of uncertainty in reserving risk. We will then show how the method can make exact explicit allowance for any historic reinsurance programs that protect the portfolio. By doing this, we show how to provide a very explicit link between gross and net reserving risk.

2. A BRIEF OUTLINE OF STOCHASTIC MODELLING TECHNIQUES

This section of the paper is intended to be a general review of existing techniques; hence we have kept existing theory to a minimum, quoting other papers or literature where a more theoretical explanation is required. In particular, readers are directed to the recent paper by England and Verrall [1] which sets out most techniques in theoretical detail.

Many stochastic techniques to date are based on some form of chain ladder technique.

Mack's model [2] was one of the first models used in practice to understand the variability in future claim amounts. Mack provided the first two moments of the future cumulative claim amounts, and assumed the model to be "distribution free". Ultimately, however, we are interested in the full predictive distribution of claims, rather than the first two moments. England and Verrall [1] provide a solution to this assuming the cumulative claims are normally distributed.

Renshaw and Verrall [3] introduced a statistical model assuming the incremental claim amount in each accident period and development period are independent random variables with an over-dispersed Poisson distribution.

Verrall [4] developed on the over-dispersed Poisson chain ladder model with the over-dispersed negative binomial model. A key difference between this and the over-dispersed Poisson model is the assumption that incremental claim payments are dependent on the cumulative claim amount at the previous period, similar to Mack's model.

In general, techniques to date have been designed for use on aggregate, portfolio level triangles of claim payment or incurred triangles. Making adequate, explicit allowance for reinsurance in practice has been, at best, an after-thought, often made using a deterministic gross to net ratio for each accident period, selected using information from aggregate modeling of the central estimate using traditional actuarial techniques. Techniques described above assume that all claims develop, on average, in a similar way, or that the mix of claims with different development patterns is constant throughout history. Due to the highly volatile occurrence and size of large claims, this may not be appropriate.

3. A METHOD FOR PROJECTING INDIVIDUAL LARGE CLAIMS

3.1 Introduction

One of the key assumptions in the aggregate stochastic methods described above is that the mix of claims with different development patterns over origin periods is stable. No allowance is made, for example, for increased variability for an accident year with "known" poor large claims experience. Also, no allowance is made for the status (i.e. open/settled) of large claims.

Perhaps more importantly, aggregate stochastic methods do not provide a process for linking the variability of gross and net of reinsurance reserves, where non-trivial treaties (such as quota shares) are in place.

We propose a model designed to cope with the problems described above, by separating the major source of uncertainty, large claims, from the remaining attritional losses, with a separate projection of individual large claims.

The remainder of this section will detail the specifics of our proposed method: uncertainty in known (reported) large claims, uncertainty in the numbers and amounts of unknown (un-reported, and reported, but not yet large) large claims, attritional claims and the aggregation of results.

3.2 Known Large Claims

We must first define by what we mean as "large". There are a number of practical considerations in choosing the threshold of large claims. The main concern is if we are going to use the results for calculating reinsurance recoveries under an Excess of Loss (XoL) program, we must choose a threshold below any historic excess of loss programs. Secondly, as we shall see, we need a significant pool of claims to sample from. To balance the above points, in the limit, we could apply this method to all claims in the portfolio, however computational and time limitations necessitate a cap on the size of the pool. It is important to frame question of choosing a threshold within context of the portfolio, for example, by considering the size of claims which are managed by the complex or large claims unit. In general, we have found this method produces reasonable results with as few as 200 individual large claims with the oldest years having had up to ten years of development.

We include all claims which were "ever" large in our method, that is to say, we include claims which could ultimately be small (or nil) but which were once estimated to be large.

We propose to adopt a stochastic chain ladder projection on individual large claims, where the simulated chain ladder factors are sampled from the observed chain ladder factors in historic large claims. Further, when simulating the development factor of the claim, we also sample the subsequent status of the claim. We therefore simulate chain ladder factors for open claims from historic claims which were open at the same point in development. Closed claims can be simulated at subsequent development periods from similar closed claims to allow for the possibility of re-opening; to the extent that they are present in the historic data.

Consider the following claims. For simplicity, assume all claims are settled by development year 3. To finalize the projection of large claims, we need to project claim D and E for one year and claim F for two years.

Table 1

Incurred Amounts			
	Development Year		
Claim	**1**	**2**	**3**
A	400,000	800,000	800,000
B	500,000	1,600,000	850,000
C	1,000,000	1,000,000	1,500,000
D	200,000	500,000	
E	300,000	200,000	
F	150,000		

Table 2

Development Factors		
Claim	Year 1 to Year 2	Year 2 to Year 3
A	2.00	1.00
B	3.20	0.53
C	1.00	1.50
D	2.50	
E	0.67	
F		

Table 3

Claim Status			
	Development Year		
Claim	1	2	3
A	Open	Closed	Closed
B	Open	Open	Closed
C	Open	Open	Closed
D	Open	Open	
E	Open	Closed	
F	Open		

To develop claim D to ultimate, we pick a claim that was open at development year 2. In this case, B and C were open at development year 2, and so we can either develop claim D by a chain ladder factor of 0.53 or 1.5.

To develop claim E to ultimate, we pick a claim that was closed at development year 2. In this simple example, only claim A was closed at the same point. Therefore, to simulate the ultimate position of claim E, we pick the chain ladder factor from claim A, that is 1.0.

To develop claim F, we must first project the position to development year 2 from open claims. Therefore, it can simulate chain ladder factors from any of claims A to E, with equal probability. If the claim follows the experience of claim B, C or D to development year 2, the claim remains open, and develops by a chain ladder factor of 3.2, 1.0 or 2.5 respectively. If the claim follows the experience of either claim A or claim E, then the claim closes and develops by a chain ladder factor of 2.0 or 0.67. Developing the position from year 2 to year 3 depends on whether the simulated claim closed in year 2 or remained open. If it remained open (i.e. was simulated from either B, C or D), then the development from years 2 to 3 is simulated from claims B or C (with equal probability) in a similar manner to claim D; if it closed (i.e. was simulated from A or E), then the development is simulated from claim A only (in a similar manner to claim E).

Based on this set of data, the possible range of outcomes for claim D is $265,625 to $750,000, for claim E is $200,000, and for claim F is $100,000 to $720,000 (the lower end of the range is attained if the simulation chooses claim E and then A, the upper end of it chooses claim B and then C). Note that the implied total ultimate chain ladder factor for the maximum simulated value of claim F is 4.8. This is more extreme than any ultimate chain ladder factors seen to date.

By explicitly identifying open and closed claims, we are adding extra information to the basic chain ladder model. The model will then capture the increased volatility of origin years which have a larger number or amount of large claims than average, and the reduced volatility of origin years with fewer large claims.

3.3 IBNR Large Claims

The above section deals with the uncertainty around claims which are already large. This is clearly only part of the picture. We must also deal with claims which become large at some point in the future. These claims can arise from genuinely new claims which have been incurred but not reported, and claims which have been reported, but which are not yet (or have never been) large.

Both the number and size of these claims need quantifying. The following sections detail how the method deals with these.

3.3.1 IBNR Large Claim Numbers

In dealing with the known large claims, we allow for the possibility that a currently large claim will ultimately settle below the large threshold. In our large number projection, we need a definition of large claim numbers that can cope with these outcomes. We deal with this by projecting a triangle of claim numbers, where a claim is counted once in the development year it became large. Claims which subsequently fall below the threshold are included in this triangle. We therefore are not making any assumption about how many of these claims will ultimately settle for less than the threshold in this step of the projection.

Standard stochastic chain ladder techniques can be applied to this data if desired, however we believe this may not be appropriate in this particular case. In particular, due to the generally small number of claims which are reported as large in development years one and two, the projected number of large claims for the most recent origin periods may be artificially unstable.

Further, we must ask ourselves if it is intuitive to suggest that if the most recent origin period has twice as many large claims per unit of exposure reported in development year one as the historical average, then it will have twice the number of large claims per unit of exposure ultimately. This does not seem to make sense in practice. Given that most aggregate stochastic methods are based on chain-ladder projections, in this instance the mean number of large claims may tend to be over-stated.

We suggest a more appropriate model for large claims numbers would be to assume the claim frequency per unit of exposure in each development period is independent of previous or subsequent development periods. The definition of exposure could include earned policy count, vehicle years, rate-adjusted earned premium or ultimate number of attritional claims.

Assuming the number of claims in a unique origin and development period follows a Poisson (or negative binomial) distribution, a number of claims that become large in each future time period can be simulated.

3.3.2 IBNR Large Claim Severity

A number of options are available to simulate the ultimate size of individual IBNR large claims.

The method we suggest is to sample from the (simulated) known large claims, where the claims are selected from the claims which became large in that development period. It may be necessary to group older development periods together to gain a significant pool of claims to sample from. By adopting this approach, we are allowing for any potential differences in average claim size by reporting development period, including the propensity for a claim to be ultimately small, and avoid the need to specify the claim size distribution. Appropriate adjustments for inflation are also required; a further refinement would allow the inflation factor selected to be stochastic.

A simplification to this method could be to sample from all simulated known claims, however if we are interested in the finalization date of claims, for example to calculate reinsurance recoveries under an excess of loss program with an indexation clause, we can run the risk of claims being finalized before they were reported as being large.

Instead of sampling from the simulated known claims, it is possible to parameterize the probability of a reported large claim finalizing as large, the finalization period of a large claim and the severity of ultimately large claims. These can be calculated from historic data, usually using a Bernoulli distribution for the probability of a reported large claim finalizing as large, a discrete distribution for the finalization period, and an appropriate distribution (perhaps Pareto or generalized Pareto) for the severity. These various distributions can then be reviewed against other market or portfolio benchmarks if available.

3.4 Combining Known and IBNR Large Claims

Now that we have separately generated the simulated ultimate position of known large claims and IBNR large claims, combining these results gives us the full picture of large claims in the run-off of reserves.

It is possible to apply a dependency structure to allow for correlations between the run-off of the known claims and the number and severity of large IBNR claims. Applying a positive correlation has an intuitive appeal; however it is very difficult to estimate the strength or shape of this relationship. We recommend at the very least scenario testing the results using various correlation strength and dependency shapes.

3.5 Non-Large Claims

To understand the variability of the aggregate reserve distribution, we need to allow for the variability of the non-large claims.

To do this we recommend using an aggregate triangle where each claim is "capped" at a certain value. For example, if a capping level of $100,000 is chosen, then the capped triangle contains all development up to the point where it reaches $100,000, and any amount in excess of this is omitted from the triangle. A claim which is reserved at $50,000 in year 1, $99,000 in year 2 and $150,000 in year 3 is included as $50,000, $99,000 and $100,000 for each respective development year. We prefer the use of a capped triangle as opposed to a triangle where large claims have been completely removed for a number of reasons, as we find it produces more stable results, and the historic triangle does not change when new diagonals of data are added (as large claims drop below the threshold and new large claims develop). In the example above, if large claims are removed from the triangle, then the development from the example claim is $50,000, $99,000, $0.

Once a capped triangle has been calculated, one of the traditional aggregate stochastic reserving methods described in Section 2 can be used to determine a range of outcomes for the "capped" reserve. This aggregate distribution can then be calculated as the sum of the capped claims and the excess of cap large claim amounts.

When selecting the capping level for the attritional claims, we recommend using a level above the "large" claim threshold. By selecting a cap above the large claim threshold, we are using information about the claims which are currently just below the cap and have a good chance of increasing above the cap at some stage in their development.

Again, it may be appropriate to introduce a dependency between the run-off of the capped and excess of cap claims.

4. REFINEMENTS TO THE METHOD AND KEY ASSUMPTIONS

4.1 Model Refinements

There are a number of refinements to the basic method that are worth outlining for completeness.

When simulating the known large claims, consideration should be given to measuring the development period as the time since the claim became large rather than as the time since accident (such that it is on a reporting period basis). This may be more appropriate for large claims due to the claim management and legal processes these claims are subject to, and generally these progress in a similar manner from the time a claim becomes large rather than from the time the accident occurs. Alternatively, a further split can be made by considering those reported "early" and "late", although this tends to reduce the sample from which to simulate from further.

We suggest splitting the large claims into at least two layers, to allow for different development patterns in the extremely large claims. For example, whereas a claim movement from $500,000 to $5 million is possible, it is perhaps less likely for a claim of $5million to increase to $50 million. Including the development factors from smaller large claims in the pool to project the extremely large claims may overstate the variability of possible outcomes for these claims. In determining which claims are in the upper layers (and indeed in the original large definition), it is important to standardize the historical claims for inflationary effects so as to not bias the claims towards more recent origin periods. It is also important to recognize claims can be in different layers at different development periods.

Selecting the very large threshold(s) is a difficult choice, and there is no one single correct method. We have found a threshold that varies by development period, such that between 10% and 20% of claims are in the top layer produces enough claims to sample from, and produces reasonably reliable results.

4.2 Key Assumptions

There are a series of assumptions underlying the model, which are worth pointing out so that their appropriateness or otherwise can be assessed.

We are assuming the historic observed chain ladder, and settlement patterns, contain the entire population of possible values. Clearly, over 1 period, this is not appropriate. However, as we are interested in the ultimate position of claims, often over a significant time period, the possible number of ultimate development factors (i.e. the product of the 1 period factors) even for a small number of possible factors (e.g. 50 at each period) becomes very large, and this assumption is not unreasonable.

We assume that chain ladder factors from one period to the next are independent, other than for changes in layer and claim status. This assumption is consistent with most other stochastic reserving methods. Further, we have assumed that individual claims develop independently within each period. This is potentially optimistic as there may be changes to internal case estimation procedures which affect all open claims, and there are external factors which also affect all open claims such as legal changes and economic factors. These global external effects can be allowed for within the model by overlaying these effects on the underlying process. By projecting claim status into the future, the effects can be applied only to open claims, as would happen in practice. If these effects are overlaid on the claims, it is important to remove any historic effects from the data to avoid double counting these shocks. Applying future inflation effects on top of the underlying projection is useful if this modeling is carried out as part of a wider capital modeling project, as it links in the reserving risk with the global economic scenarios.

As seen with the above simple example, for very new claims, the method can produce very wide ranging results. If the resulting range is thought to be too unstable, for example when considering the implied reinsurance recoveries at high layers, it may be appropriate to either adjust the range of possible results, or use a method similar to that developed for the IBNR large claims described above.

5. ALLOWANCE FOR HISTORIC REINSURANCE STRUCTURES

As we have now projected the ultimate position of all large claims, we can calculate any reinsurance recoveries exactly. For known large claims, we know all the reinsurance details which attach to the claim, and any quota share arrangements can be applied to the aggregate results.

It may be necessary to introduce a further refinement to the model if, say, the excess of loss treaty is placed on a risks attaching basis. For known large claims, we will know the underwriting year of the policy. For IBNR large claims, the underwriting year to which the claim attaches can be simulated. Typically the probability would be in proportion to the exposure that each underwriting year contributes to the accident year.

6. RECONCILIATION OF RESULTS WITH AGGREGATE MODELLING

Invariably, this work will form part of a larger piece of work; usually an outstanding claims review or part of a capital modeling project. The actuary may form a view of the reserves based on aggregate deterministic methods. This will not correspond with the results of the above method, or indeed any of the methods described in Section 2. This is less than ideal, as the practitioner would like to understand the variability around their central estimate, rather than some other result.

One way of ensuring consistency is to scale results by origin year so that the mean simulated result equals the actuary's best estimate of reserves, or try a different method. This can be done by either applying a multiplicative scaling factor for each accident year, or alternatively by adding on a fixed loading for each accident year. This can lead to undesirable results, either with negative reserves in some instances of additive scaling, or extreme results if the multiplicative scaling factor is large.

If the outstanding claims review uses consistent data (in terms of separately modeling capped and excess claims, and considers ultimate counts as well as amounts), then there are additional diagnostics available to the actuary, such as the following:

- Large Claim Frequency;
- Large Ultimate Claim Frequency;
- Large Excess Ultimate Claim Size;
- Large Excess Ultimate Burning Cost / Loss Ratio;
- Capped Claim Burning Cost / Loss Ratio;
- Large Excess Cost as a Percentage of the Total Claims Cost.

With these, the actuary has the ability to understand which piece of the projection is producing results inconsistent with the aggregate modeling.

7. CASE STUDY

7.1 Introduction

The concepts described above are more readily visualized as a case study. The data modeled is from a UK auto account, and contain 16 years of historic data. For individual large claims above £100,000, the data included the accident date, report date, and the year-end paid and incurred positions, as well as a history of the claim status.

The layers were chosen such that 80% of the claims in each development period were in the lower layer, and 20% in the upper layer. The actual layer limits can be seen in Appendix 1.

7.2 Analysis of the Gross Results

Figure 1 shows the simulated development of a claim which has just been reported as being large, with a current incurred position of £125,000. The lighter shades of gray represents the more extreme percentiles, with the dotted lines representing the 90^{th}, 75^{th}, 50^{th}, 25^{th} and 10^{th} percentiles. The mean development is represented by the solid line. As can be seen, we expect the case reserve to be ultimately inadequate, with the expected ultimate amount being just above £300,000. However, using the method described in this paper, can see that 90% of the time, the claim will settle for £700,000 or less. Occasionally, however, the claim develops much more significantly. Figure 2 shows an individual simulation where the claim grows to more than £1,000,000.

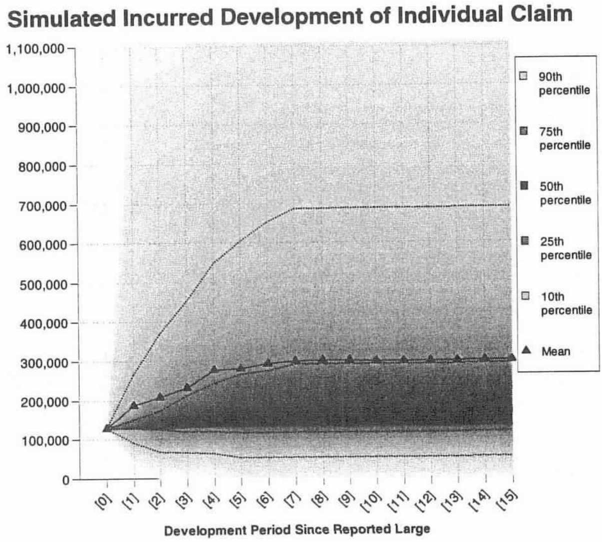


Figure 1

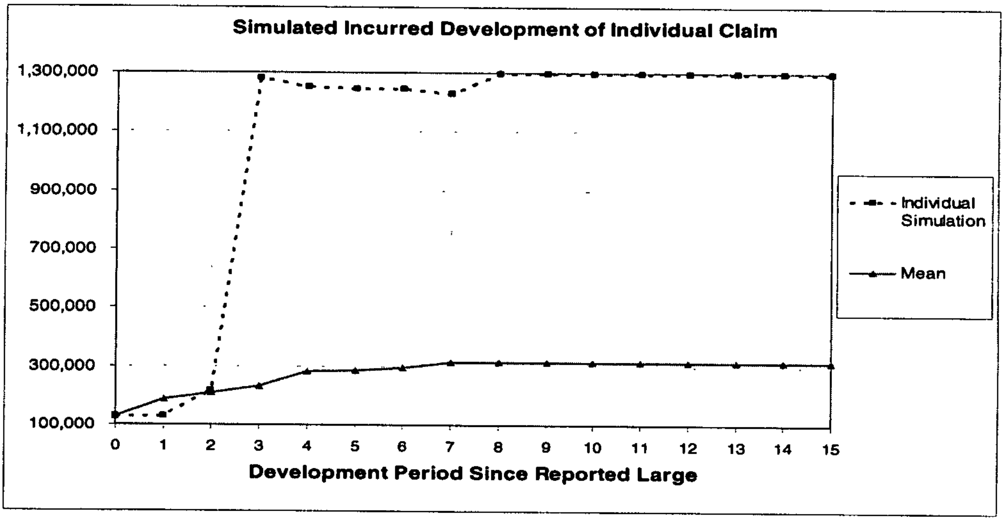


Figure 2

Even for claims that have been reported as large for several years, there is uncertainty over the development. Figure 3 shows the simulated development for a claim that has been reported large for four years, using the same percentile descriptions as for Figure 1. On average, the claim is expected to run off at an increase to the current incurred. Note that the variability around this is still quite significant.

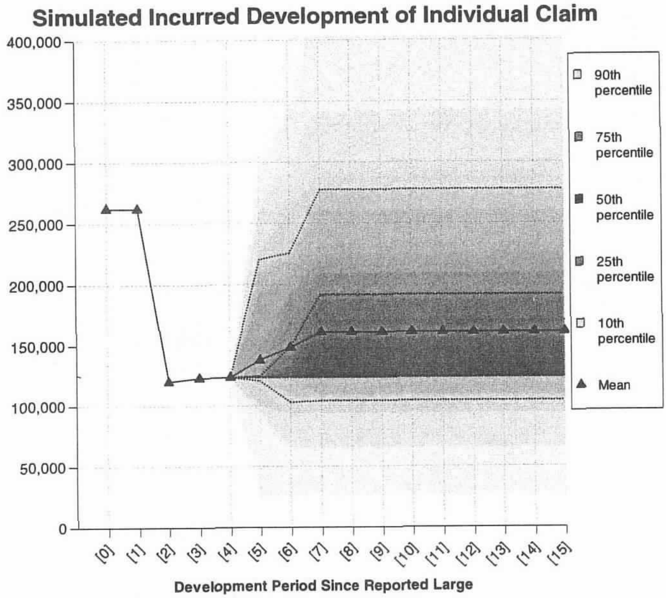


Figure 3

We mentioned in Section 4 that we would typically expect to see different loss development factors for individual "small" large claims than for "large" large claims. This is illustrated in Figure 4.

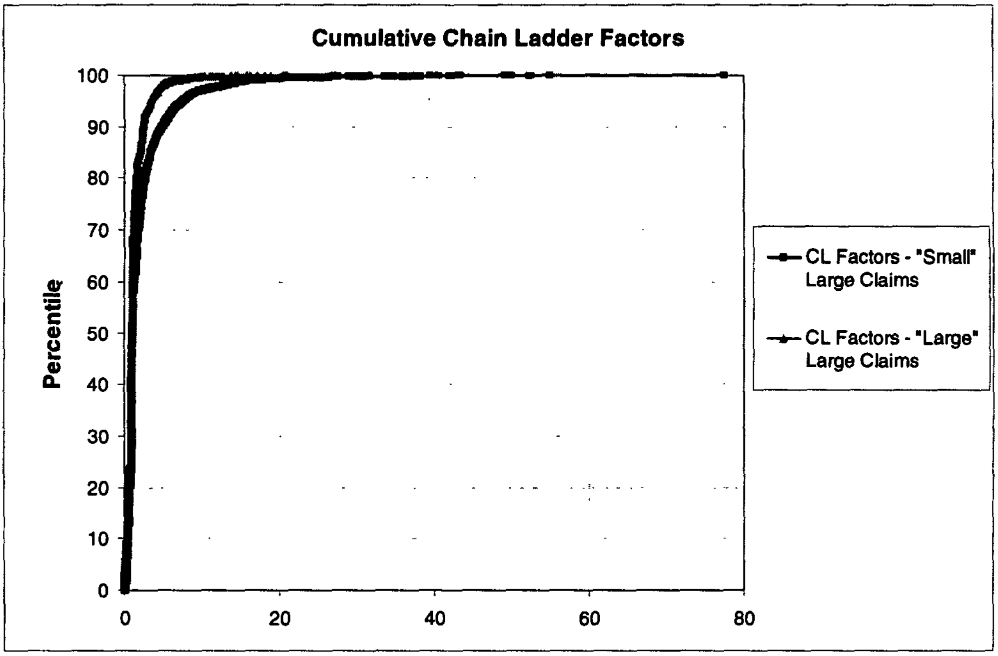


Figure 4

The darker line represents the distribution of cumulative loss development factors for "small" large claims in the first development period, the lighter line the distribution for "large" large claims. As expected, it is much more unlikely to have a large development factor for the "large" large claims, although it is quite possible.

To analyze full accident year results, we have estimated the uncertainty surrounding the attritional claims using Mack's method on a triangle based on a combination of incurred and paid data. Figure 5 shows the percentile plot of the total unpaid liabilities of the capped claims.

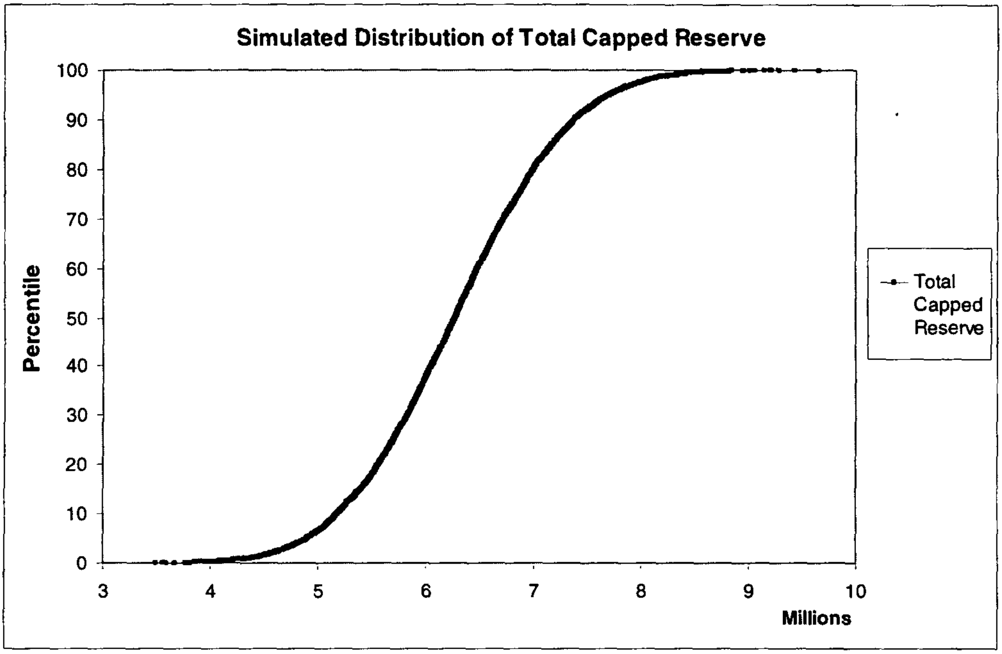


Figure 5

The table below shows the results of our projections and compares the results with those obtained by modeling the aggregate triangle using a Mack bootstrap. The 75th and 95th percentiles are given as percentages of the mean reserve. The coefficient of variation (C.o.V.) indicates the variability in the results.

Table 4

		Individual Claim Projection Method			Mack Bootstrap		
Accident Year	Mean Reserve	C.o.V.	75th percentile	95th percentile	C.o.V.	75th percentile	95th percentile
1998 (and Prior)	499,653	60.28%	120.51%	212.55%	15.60%	110.21%	126.61%
1999	2,836,912	16.52%	107.26%	135.45%	8.65%	105.73%	114.29%
2000	4,525,560	21.08%	109.64%	137.95%	25.06%	116.66%	141.34%
2001	6,582,895	24.46%	112.33%	144.18%	23.08%	115.10%	138.96%
2002	7,073,142	28.99%	114.39%	153.25%	27.41%	117.91%	146.37%
2003	12,608,970	32.81%	113.21%	161.58%	18.46%	112.19%	130.81%
2004	12,265,893	25.46%	113.75%	147.09%	29.15%	118.97%	149.76%
2005	15,134,996	30.66%	114.44%	154.19%	30.97%	120.17%	153.40%
Total	61,528,020	12.53%	107.29%	121.70%	13.56%	109.04%	123.04%

It can be seen that similar estimates are produced by the two methods for the C.o.V. of the total gross reserve. However the results for individual accident years can be significantly different. Figures 6, 7 and 8 show the gross reserve distributions for 2003, 2004 and 2005 respectively. In all three graphs, the individual claim projection method produces a distribution which is heavier in the upper tail than the aggregate modeling.

On investigation, the cohort of claims in 2003 contain a higher proportion of open large claims than average, including one claim of £6m, which results in the greater uncertainty than implied by the aggregate projection. The extra information provided by the individual claim projections arguably enables a more realistic projection of the true underlying uncertainty in the liabilities.

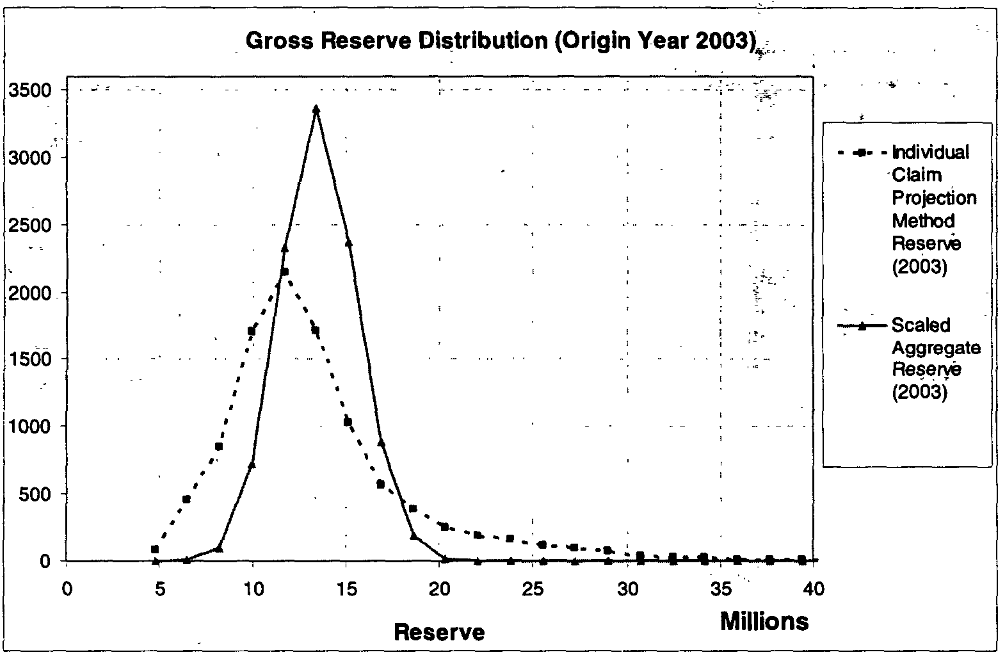

Gross Reserve Distribution (Origin Year 2003)
3500
3000
2500
2000
1500
1000
500
0
0
5
10
15
20
25
30
35
40
Reserve
Millions
Individual Claim Projection Method Reserve (2003)
Scaled Aggregate Reserve (2003)

Figure 6

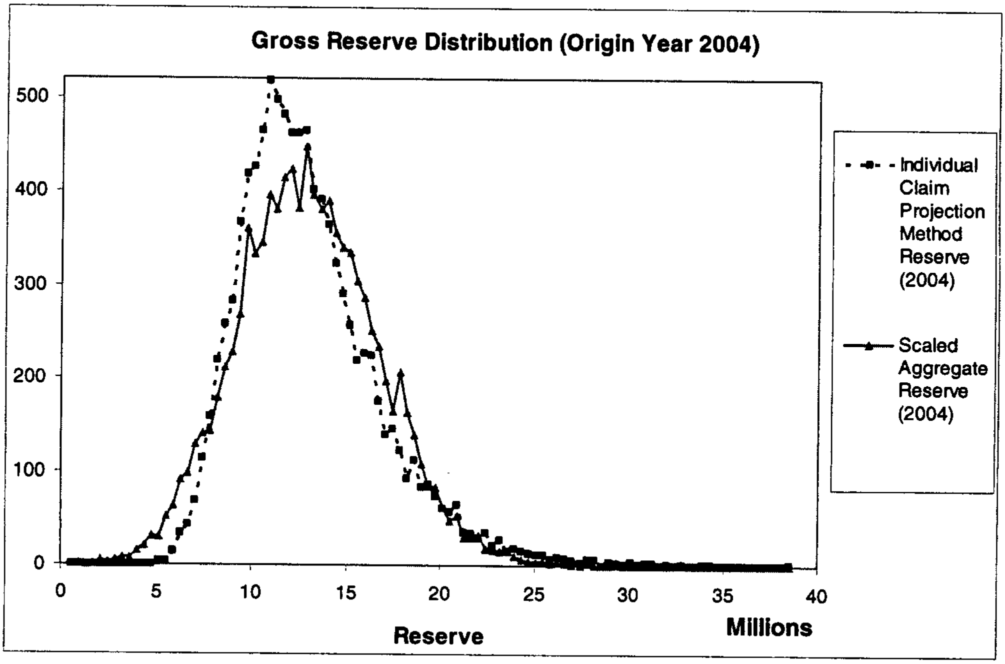

Gross Reserve Distribution (Origin Year 2004)
500
400
300
200
100
0
0
5
10
15
20
25
30
35
40
Reserve
Millions
Individual Claim Projection Method Reserve (2004)
Scaled Aggregate Reserve (2004)

Figure 7

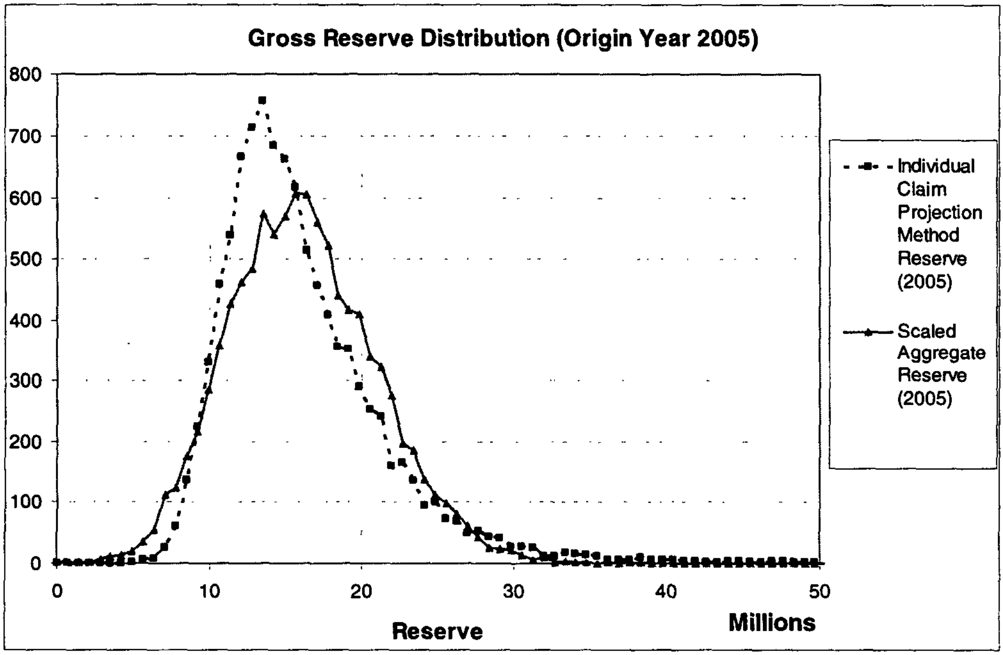


Figure 8

Figure 9 shows the cumulative distribution of the aggregate unpaid liabilities across all accident years based on the two methods. It can be seen that the two methods produce very similar results for the total gross reserve although the individual claim projection method produces a slightly heavier upper tail. This is highlighted in Figure 10, which compares the two distributions in the upper tail.

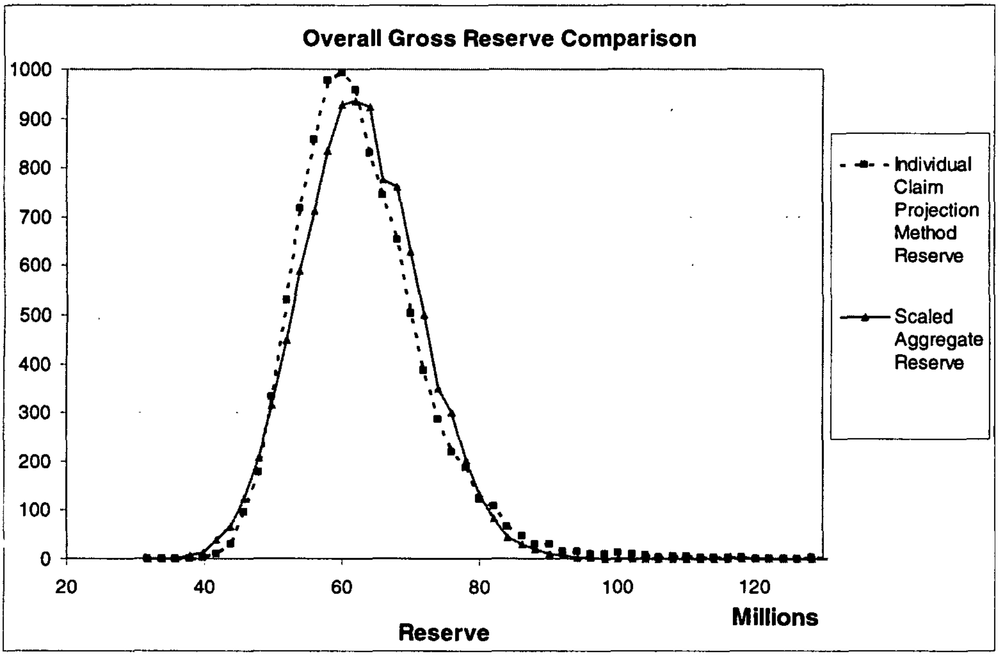

Overall Gross Reserve Comparison
1000
900
800
700
600
500
400
300
200
100
0
20
40
60
80
100
120
Millions
Reserve
Individual Claim Projection Method Reserve
Scaled Aggregate Reserve

Figure 9

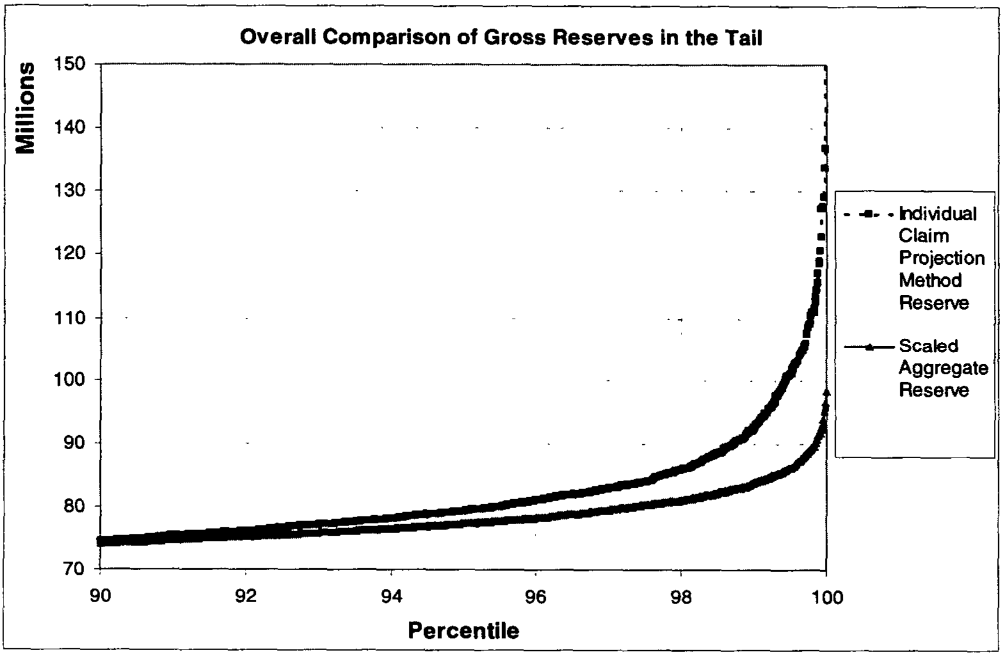

Overall Comparison of Gross Reserves in the Tail
Millions
150
140
130
120
110
100
90
80
70
90
92
94
96
98
100
Percentile
Individual Claim Projection Method Reserve
Scaled Aggregate Reserve

Figure 10

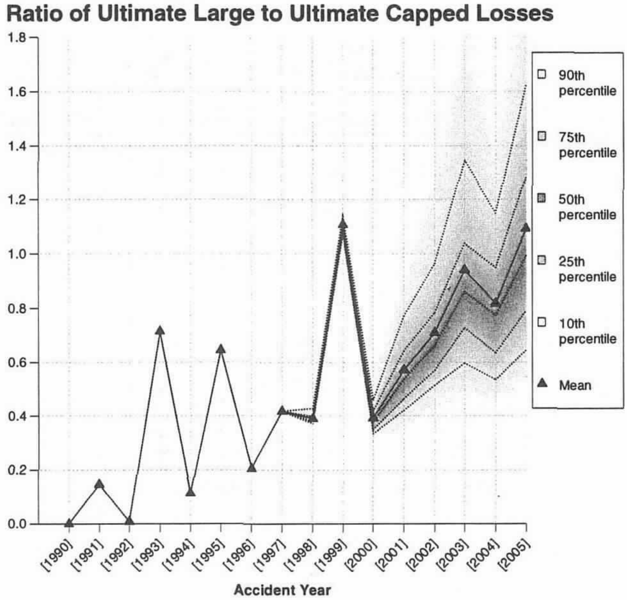


Figure 11

One reason for the heavier upper tail produced by projecting the individual large claims can be seen in Figure 11 (using the same percentile description as in Figure 1). The graph implies that the ultimate large claim proportion is increasing in recent years, the appropriateness of which can be tested in the aggregate modeling. This trend, if true, will not be allowed for adequately in the aggregate stochastic methods.

7.3 Analysis of the Net Results

Once we are comfortable with the gross results, we can calculate reinsurance recoveries on individual claims using the appropriate reinsurance terms. Table 5 shows the net results for the individual claim projection method.

Table 5

		Individual Claim Projection Method		
Accident Year	Mean Reserve	C.o.V.	75th percentile	95th percentile
1998 (and Prior)	465,519	46.96%	119.36%	200.79%
1999	977,976	26.68%	108.93%	147.16%
2000	2,983,380	13.90%	109.34%	124.26%
2001	3,496,184	18.09%	112.00%	131.59%
2002	3,457,150	23.91%	116.35%	141.11%
2003	5,755,890	17.85%	111.07%	131.46%
2004	8,518,805	16.93%	110.45%	129.51%
2005	9,999,956	16.62%	110.57%	128.90%
Total	35,654,860	8.19%	105.26%	114.12%

Figure 12 shows the net and gross reserves for the 2005 accident year. The gross reserves have been scaled to have the same mean as the net reserves. As would be expected, the netting down has resulted in a large reduction in variability.

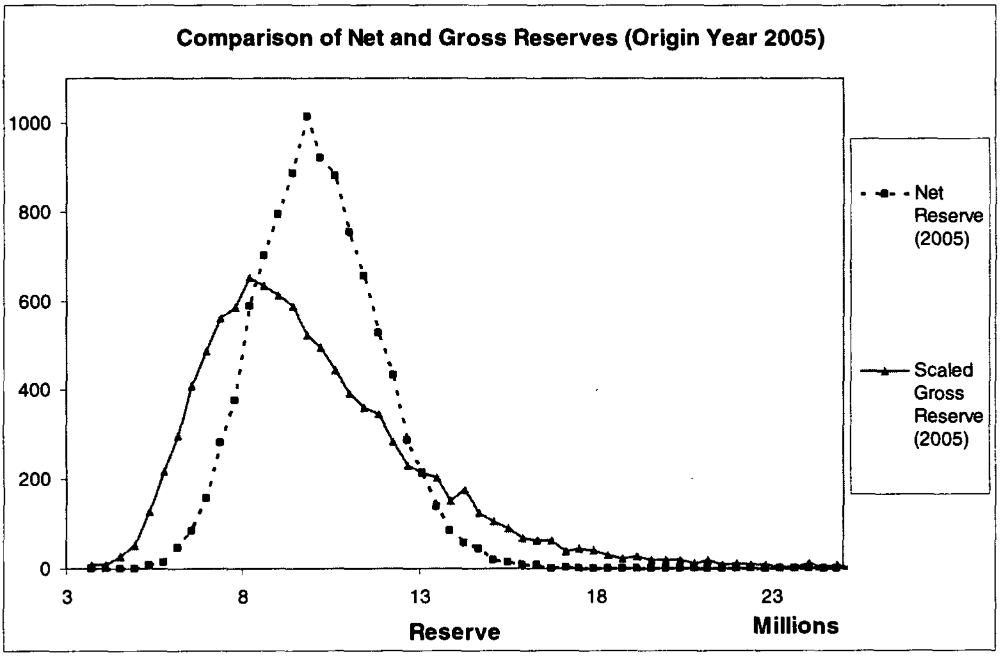


Figure 12

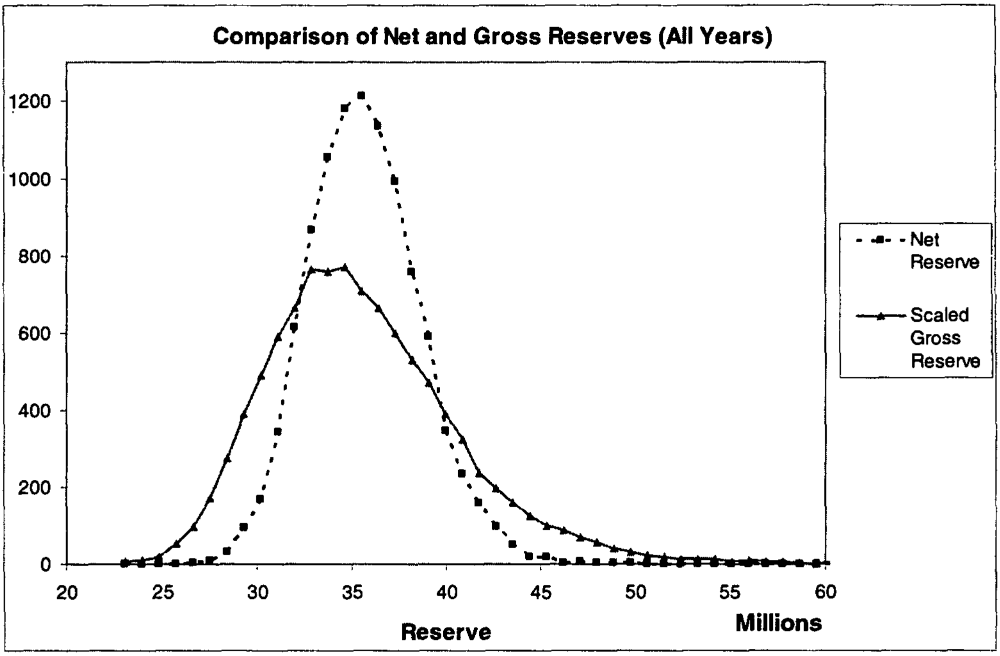


Figure 13

Figure 13 shows the overall net and gross reserves. The net reserve again shows a substantial reduction in variability.

It has already been noted that one of the additional benefits of the method described in this paper is the ability to accurately examine the performance of reinsurance cover. Figure 14 shows the distribution of recoveries associated with an aggregate deductible of £2.25m attaching to a layer of £400k in excess of £600k for the 2002 accident year. As can be seen, approximately 10% of the time the deductible is fully blown and losses pass through to the reinsurer.

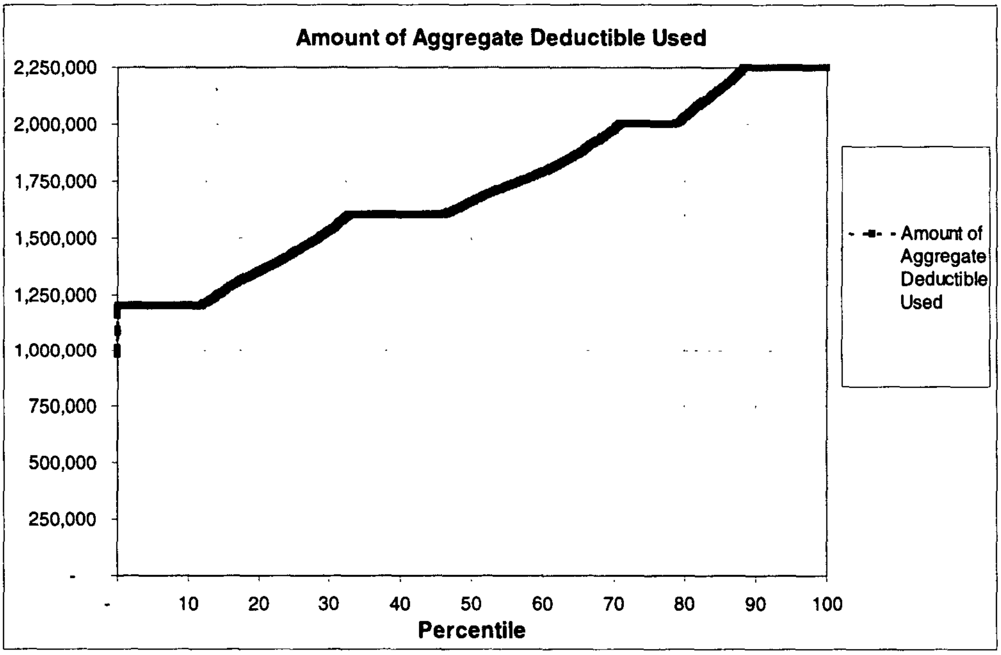


Figure 14

This allows us to consider the value of this contract and whether it represents value for money.

The method described in this paper also provides the complete predictive distributions of the gross ultimate position and ultimate reinsurance recoveries of individual large claims. Therefore the mean net ultimate position for each simulated claim can be correctly calculated. Some netting down methodologies we have seen used in practice implicitly assume that the mean of the reinsurance recoveries equals the mean of the gross claim less the retention. The one-sided nature of reinsurance means that this is flawed. The error associated with this assumption can be seen in Table 5, which shows the gross, reinsurance (RI) and net ultimate incurred position for two claims, with an excess of £1,000,000. The ultimate figures have been calculated on the our stochastic basis and also on a deterministic basis. The final two columns correspond to the stochastic calculations, where the mean net position takes into account the variability of the ultimate gross position.

Table 6

Netting Down - Comparison of Methodologies						
	Current Incurred	Mean Gross Ultimate	*Deterministic Ultimate Reinsurance Recoveries*	*Deterministic Net Ultimate*	Mean Ultimate Reinsurance Recoveries	Mean Net Ultimate
Claim 1	500,000	829,180	0	829,180	226,604	506,962
Claim 2	1,000,000	1,337,416	337,416	1,000,000	464,964	872,452

In this case, the deterministic basis is likely to lead to an overestimation of the net position, and is therefore a conservative basis. While in itself this is not a cause for concern, a desirable property of any reserving exercise would be to ensure a consistent basis for gross and net reserves.

8. INTEGRATION AND APPLICATION WITHIN CAPITAL MODELS

In recent years, there has been considerable time invested in the development of capital models to understand and quantify the risks faced by an insurance business. A significant piece of this work has been an analysis of reserving risk, which forms part of the wider insurance risk. In our experience of the UK market, there are two main methods used by practitioners to estimate net of reinsurance reserving risk. Both methods project gross aggregate triangles, with a different approach to netting down for reinsurance recoveries.

The first method arrives at net results using a deterministic net to gross ratio applied to the stochastic gross results. This method has the advantage of simplicity and transparency, however it in effect gives no credit for the expected reduction in volatility that non-proportional reinsurance should provide.

The second method projects both gross and net triangles, with some link between the projections in an attempt to ensure consistency and nonsensical simulations are avoided (for example, simulations where net reserves are higher than gross). While this should allow for the reduction in variability not captured by the first method, it is likely that the reinsurance has changed over the years (for example, reinsurance excess points have changed), and the observed historical figures may not be appropriate to apply to the newer accident years.

In forecasting the ultimate large claim severity, it is important to allow for parameter uncertainty. We would further recommend including development uncertainty. Currently, most ultimate loss generators are parameterized from some projected ultimate claim figures, allowing for IBNER, which is assumed to be known and fixed. Parameter error, using various techniques is included in the forecasting of the large claims. However the ultimate position of the claims used to parameterize the distribution are not known or fixed. To not make allowance for this will understate the true uncertainty of the underlying distribution.

9. CONCLUSIONS

Existing methods available to help gain understanding of the variability of insurance liabilities have focused on aggregate gross data, with no explicit allowance for changing mix of claims, and with no obvious adjustment to allow for non-trivial reinsurance. We have developed a method based on a small number of key assumptions to explicitly project the development of individual large claims. We show how various refinements can be made to the standard method and implement this method via a case study using actual data from a UK motor injury portfolio.

By explicitly projecting individual claims we show how to make appropriate allowance for policy limits and the reduction in variability arising from non-proportional reinsurance. By separately considering attritional and large claims, we can directly allow for changes in the mix of claims in our portfolios.

A range of diagnostics is available to the practitioner to aid understanding of the results, and to ensure it is not applied in a mechanical fashion.

Appendix 1 – Example Data

The following table shows the layer limits used in the Case Study. The lower layer lower bound is the threshold above which claims are individually simulated. The upper layer lower bound defines the boundary between 'small' large claims and 'large' large claims, in order to partition the development factors. Development Periods 11 and above have been grouped due to scarcity of data.

Table 7

Development Period	Lower Layer Lower Bound	Upper Layer Lower Bound
1	100,000	360,000
2	100,000	500,000
3	100,000	520,000
4	100,000	400,000
5	100,000	680,000
6	100,000	500,000
7	100,000	630,000
8	100,000	320,000
9	100,000	310,000
10	100,000	260,000
11+	100,000	120,000

ACKNOWLEDGEMENTS

The authors are very grateful for the work put in by John Berry to help complete this paper.

REFERENCES

[1] England, P.D. and Verrall, R.J., "Stochastic Claims Reserving in General Insurance," *B.A.J.*, **2002**, Vol 8, Part III, No. 37

[2] Mack, T., "Distribution-free calculation of the standard error of chain-ladder reserve estimates," *ASTIN Bulletin*, **1993**, Vol 30, 213-225.

[3] Renshaw, A.E. and Verrall, R.J., "A stochastic model underlying the chain-ladder technique," *B.A.J.*, **1998**, Vol 4, 903-923.

[4] Verrall, R.J., "An investigation into stochastic claims reserving models and the chain-ladder technique," *Insurance: Mathematics and Economics*, **2000**, Vol 26, 91-99

Biography(ies) of the Author(s)

Karl Murphy is Partner at EMB Consultancy LLP in the UK. Karl graduated from Trinity College, Dublin in 1990, with a first class honors degree in Mathematics and Economics. His main areas of study were in statistics and econometrics, and received the Gold Medal for his final exam results.

On graduation, Karl joined the non-life department of consulting actuaries Bacon & Woodrow, before joining EMB in 1993. Karl has considerable experience in many areas of non-life insurance, and clients include insurance companies, Lloyd's syndicates, captives and brokers. His main expertise is in personal lines insurance, and is considered to be one of the UK's leading personal lines pricing experts. Karl has played an important role in the development of EMBLEM, EMB's pricing tool that is now used by 100 insurers around the world. More recently Karl has being introducing the new capital modeling techniques developed by EMB to personal lines clients.

Karl has spoken at conferences throughout the world, including at several American Casualty Actuarial Society seminars, the South African Actuarial Society's General Insurance Convention, the UK's GIRO conference, and for the Actuarial Society of India. He co-authored a paper entitled "Using Generalized Linear Models to Build Dynamic Pricing Systems" that appeared in the CAS Forum, and has written several articles that have appeared in the insurance press on various topics.

He has served on the Institute of Actuaries' GIRO Committee and Research Steering Committee, and is currently on the Institute of Actuaries' GRIP (Pricing) committee.

Andrew McLennan is Senior Consultant at EMB Consultancy LLP in the UK. Andrew joined EMB in 2003, having previously worked in the non-life department of Trowbridge Consulting in Australia. Since moving to the UK, Andrew has been involved in a number of capital modeling projects for both personal lines clients and London Market clients. Andrew has also been involved in developing new actuarial techniques for application in the personal lending market. Andrew had considerable experience in the Australian general insurance market, including significant work with major government accident compensation schemes, including the design of the funding policy of one authority, and was involved in the due diligence review of the world's largest private equity floatation in 2002. Andrew qualified as an actuary in 2003.

The authors can be contacted by email at karl.murphy@emb.co.uk

Measuring Loss Reserve Uncertainty

William H. Panning

EVP, Willis Re
88 Pine Street, 4th Floor
Wall Street Plaza
New York, NY 10005

Phone: 212-820-7680 Cell: 502-387-5411 Email: bill.panning@willis.com

---

Motivation. For property casualty insurers, loss reserves are by far their largest liability. These are actuarial estimates of future loss payments resulting from accidents that have already occurred. In fact, the actual future loss payments may deviate – sometimes substantially – from the amount that was estimated. Senior managers, shareholders, rating agencies, and regulators all have an interest in knowing the magnitude of these potential deviations – deviations whose distributions we here call *loss reserve uncertainty* -- since firms with large potential deviations need more capital or reinsurance than other firms with smaller potential deviations. Actuarial journals provide several proposed procedures for measuring loss reserve uncertainty. But in practice they are rarely used, since they typically require specialized software and use statistically complex procedures that are unfamiliar to most actuaries. Moreover, in at least some cases, these procedures provide estimates of loss reserve uncertainty that depend on very strong assumptions that virtually assume the conclusions obtained.
Method. In this report I provide a simple method for measuring loss reserve uncertainty that is easily implemented with a spreadsheet model, that relies on data available for all US insurers and all lines of business, and that makes relatively few easily accepted assumptions.
Results. The method for estimating loss reserve uncertainty explained and demonstrated here has five important advantages. First, it is simple, and easy to implement. This report even provides the relevant Excel formulas for implementing crucial steps in the method. Second, it avoids severe statistical problems that affect numerous rival methods, as explained in detail. Third, the method is validated (rather than merely illustrated) by applying it to simulated data in which answers are known, and demonstrating that its estimates agree closely with these known answers. Fourth, the measure of loss reserve uncertainty used here – the standard deviation of loss reserves as a percentage of the estimated reserve -- is scalable, so that it can be applied to reserves estimated by other methods. Fifth, the resulting measure of loss reserve uncertainty can be directly compared across different lines of business in a single firm, or for the same line of business across different firms.
Conclusion. The method presented here appears to be the first instance of a method for estimating loss reserves and loss reserve uncertainty that is thoroughly validated by comparing its estimates to those of a simulation with known parameters. Its results can assist CEO's, CFO's, Chief Risk Officers, actuaries, rating agencies, regulators, and stock analysts in estimating the variability of loss reserves, in estimating a firm's capital adequacy, in forecasting the distribution of possible loss reserve payments during the next calendar year, and in determining whether current or past calendar year deviations from expected loss payments are sufficiently large to deserve special attention.
Availability. To obtain the model presented here, email Bill.Panning@Willis.com.

Keywords. Loss reserve uncertainty, regression, reserving, Enterprise Risk Management

---

1. WHAT LOSS RESERVE UNCERTAINTY IS AND WHY IT MATTERS

1.1 Defining Loss Reserve Uncertainty

Property-casualty loss reserves are estimates of the total future payments that will be required to settle claims on accidents that have already occurred. Because such estimates are inherently imprecise, for reasons discussed later on, insurers may ultimately pay out more or less to claimants than is forecast in the firm's current reserve. **Loss reserve uncertainty (LRU) is a measure of the magnitude of this potential difference between forecast and actual loss payments.**

In this paper I propose, explain, and justify a particular method of estimating loss reserve uncertainty. This method has several important virtues. First, it is simple, and so can be implemented on a spreadsheet and applied to universally available data. Second, the method is accurate, since it addresses and avoids a number of pitfalls in statistical estimation that would otherwise produce biased and misleading results. Third, the resulting estimates are comparable across different lines of business and different firms. Fourth, the measure of LRU is scalable, so that it is applicable to reserves that have been estimated in different ways. Finally – and this is particularly significant – the method has been thoroughly validated by demonstrating that its estimates of reserves and loss reserve uncertainty closely match the known parameters underlying 10,000 simulated loss reserve triangles.

1.2. Why Loss Reserve Uncertainty Matters

A method for estimating LRU that has these characteristics is likely to be extremely useful to insurers, investors, regulators, and rating agencies, for estimating surplus adequacy, for pricing and capital allocation, and for determining the potential significance of reserve developments.

1.2.1. Estimating Surplus Adequacy.

The uncertainty of an insurer's loss reserve has direct implications for its required surplus or reinsurance. The greater an insurer's LRU, the greater the surplus or reinsurance it needs to cope with potential scenarios in which ultimate losses exceed forecast losses. In the absence of an accepted measure of LRU, these various audiences have relied on indirect measures of surplus adequacy such as premium-to-surplus or reserve-to-surplus ratios relative to peer companies or to industry averages. Such relative evaluations can be quite misleading in an industry that exhibits profound swings in pricing and reserve adequacy.

The problem of estimating surplus adequacy is a fundamental issue in Enterprise Risk Management, which attempts to estimate the total capital needed by an insurer to withstand potential losses from all sources of risk. Before total enterprise risk can be managed, it must first be measured. For most property casualty firms, the principal sources of risk are loss reserve uncertainty, asset risk (principally due to equities), pricing risk (potential differences between forecast losses and the incurred losses initially booked), and credit risk on receivables and recoverables. Of these, LRU is often one of the largest and one of the most difficult to estimate.

1.2.2. Pricing and Capital Allocation.

Many insurers allocate capital to different lines of business and evaluate pricing adequacy by the return on capital achieved in each line. Although firms may employ different methods for allocating capital among different lines of business, there is consensus that the capital allocated to a particular line should reflect the degree to which estimated losses are uncertain. Consequently, the capital allocated to a line of business should reflect the rapidity with which its reserve runs off and the magnitude of uncertainty involved. Measuring loss reserve uncertainty can therefore inform and improve capital allocation and pricing.

1.2.3. Managerial Feedback.

An insurer's loss reserve is a forecast of all future loss payments, including those anticipated during the next calendar year, from accidents that have already occurred. The measure I propose can be adapted to estimate the uncertainty of this calendar year estimate. What makes this important is that this estimated uncertainty provides a useful benchmark against which any difference between actual and forecast loss payments can be evaluated. For example, if calendar year paid losses are 20% higher than forecast, this is of little concern when the standard deviation of those forecast losses is 15%. But if, instead, the standard deviation is 6%, then the 20% deviation should trigger significant managerial concern. Since managerial attention is a scarce and valuable resource, the ability of this method to distinguish significant deviations from those that are not should prove to be quite useful.

2. PRIOR STUDIES OF LOSS RESERVE UNCERTAINTY

Given the potential importance of measuring LRU, it is not surprising that the number of papers on the subject has grown significantly during the past decade. Relevant papers include Ashe (1986), Barnett and Zehnwirth (2000), Braun (2004), Brehm (2002), England and Verrall (1999, 2001, 2002), Halliwell (1996), Hayne (2003), Hodes, Feldblum, and Blumsohn (1996), Holmberg (1994), Kloek (1998), Mack (1993, 1994, 1995, 1999), Murphy (1994), Taylor (1987, 2004), Taylor and Ashe (1983), and Verrall (1994). Rather than describing each paper individually, I shall comment on this body of work taken as a whole.

2.1 Chain Ladder Focus

First, a central assumption of much of this literature is that the chain ladder method for estimating reserves is the obligatory starting point for estimating reserve uncertainty. For example, in their excellent review of a variety of models and techniques for estimating reserves and reserve uncertainty, England and Verrall (2002) note that a principal objective of the models they review is "to give the same reserve estimates as the chain-ladder technique" (p. 448). By contrast, there are relatively few studies like Stanard (1985), Narayan and Warthen (1997), Barnett and Zehnwirth (2000), and Taylor (2003) that focus on the key assumptions and comparative adequacy of the chain ladder method. Here I make no attempt to ensure that my proposed method agrees with the chain ladder method in estimating the unknown parameters of some paid loss triangle. The point is to obtain estimates that are correct, whether or not they agree with a widely-used method.

To validate the method I shall use known parameters to simulate thousands of paid loss triangles and determine whether my proposed method is able to accurately estimate these parameters and the corresponding simulated reserves and simulated reserve uncertainty. Agreement with the chain ladder method is simply irrelevant to this validation procedure, especially since the chain ladder method has itself not been definitively validated in a comparable manner.

2.2 Absence of Estimation Criteria

Second, apart from the special place accorded to the chain ladder method, much of the literature seems to assume a kind of algorithmic democracy, in which one technique for estimating reserves or LRU is considered as good as any other. (This assumption reaches its inevitable conclusion when the results obtained from different methods are averaged.) With few exceptions, there is no discussion of criteria that must be met in order for estimates of reserves or LRU to be accurate. The notable exceptions here are Ashe (1986), Barnett and Zehnwirth (2000), Halliwell (1996), Taylor (1987), and Taylor and Ashe (1983), but even here the relevant issues are typically either assumed or discussed very briefly. Here I explain at some length conditions that are crucial to accurate estimation, and show specifically what must be done to meet those conditions.

It is important here to recognize the significant differences between estimating reserves on the one hand and estimating LRU on the other. Some methods for estimating reserves are totally incapable of being extended to estimating LRU. Moreover, there is an enormous difference, at least in my view, between methods that principally focus on estimating reserves but only incidentally focus on LRU, and methods that principally aim to estimate LRU. The former are especially prevalent, and invite strong assumptions with little guidance on ways to test their validity or to estimate the sensitivity of LRU estimates to slight changes in these assumptions. The latter are rare, and include the method presented here.

2.3 Complexity

Third, the procedures proposed to estimate LRU are typically quite complex. Moreover, some recommended procedures, such as generalized least squares (GLS) and generalized linear models (GLM) in fact typically require very strong a priori assumptions about variances and covariances. Checking and, when necessary, appropriately modifying these assumptions is indeed feasible, but at the expense of making a complex procedure even more vulnerable to the temptation to over-fit the model, thereby "finding" what one has really assumed. Here I utilize a much simpler procedure that is less elegant but, in this respect, more robust.

3. A MEASURE OF LOSS RESERVE UNCERTAINTY AND ITS MERITS

No single method of estimating loss reserve uncertainty is appropriate under all circumstances. Much depends upon the type and extent of data available for such an analysis. For example, actuaries within an insurance firm may have access to data that is far more extensive and detailed than the data available to external analysts. Given these differences in available data, internal and external analysts may appropriately utilize different methods to estimate loss reserve uncertainty. Nonetheless, I believe that the results obtained from the method presented here can be applied directly to reserve estimates obtained using other methods and more extensive data.

The procedure I propose has two steps. The first is estimating the loss reserve itself, in dollars. The second is estimating the standard deviation of that loss reserve, again in dollars. Because both of these estimates are in dollars, comparisons across lines of business or across different firms are essentially meaningless, since differences in these numbers will principally be affected by differences in the volume of business in each line or each firm. But if we instead express reserve uncertainty as a **coefficient of variation** (the standard deviation of the estimated reserve as a percentage of the estimated reserve), we arrive at a measure that has three important properties.

First, it can be compared across different lines of business within a particular firm. A line of business in which the coefficient of variation is 6% is clearly less risky (in this respect, at least) than one in which the coefficient of variation is, say, 15%.

Second, this measure of LRU can be compared across different firms for the same line of business. If the coefficient of variation for workers' compensation is smaller for one firm than for another, it is pretty clear that this line of business is less risky for the first firm than for the second. This fact has enormous implications for the measurement of capital adequacy.

The results of both of these comparisons must be interpreted carefully, since they depend on the volume of business written as well as on supposedly intrinsic differences between different lines of business. As the central limit theory implies, the coefficient of variation for a line of business will tend to decrease with the volume of business written. This principle is confirmed by the fact that, for a particular line of business, the coefficient of variation for the industry as a whole is typically smaller than that same measure for any particular firm.

Third, I believe that this measure of LRU can be applied to reserves that have been estimated by methods other than the one recommended here. My argument here is very simple. Suppose that I utilize the method and data proposed here to forecast future loss payments (i.e., the reserve) for some insurer and obtain a value R. Suppose also that the firm's own actuaries, utilizing a different method and far more extensive data, obtain an estimated reserve value of R*, where R* = aR (i.e., some positive constant times the value R obtained using the method and data recommended here). Under rather broad conditions it is the case that if R* = aR, then the standard deviation S* = aS, where S* is the standard deviation of the R* and S is the standard deviation of R. If this is so, then it is necessarily true that S*/R* = S/R. In other words, the coefficient of variation S/R will be (approximately) the same regardless of the method used to estimate reserves.

4. DATA NEEDED TO MEASURE LOSS RESERVE UNCERTAINTY

Comparing LRU across different lines of business and, in particular, across different firms, requires that data that is commonly available and consistently defined. The data utilized here consists of the paid loss triangles reported in Schedule P, Part 3, of the Annual Statement required by the National Association of Insurance Commissioners. This data is publicly available for all insurance companies licensed in the United States.

Table 1 is an example of such data. The rows of this table are **accident years**: the calendar years in which accidents occurred. The columns are **development years**: calendar years in which claims payments for those accidents were actually made. A single accident can trigger multiple claim payments occurring in different development years. For example, an auto accident in November 1995 could trigger a payment for physical damage to the insured's vehicle in December of that same year, and an additional claim payment, for bodily injury medical costs, in 1996. Litigation, if it occurs, may delay claim payments into later years. Table 1 shows that, for all accidents occurring in 1994, $ 624 million in claims were paid that same year, a cumulative total of $ 2.1 billion had been paid by the year-end 1996, and $ 2.9 billion by year-end 2003.

Table 1: Cumulative Paid Losses (millions)

Year Losses Were Incurred	Development Year 1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
1994	624	1,595	2,066	2,366	2,559	2,685	2,765	2,818	2,860	2,895
1995		695	1,503	1,975	2,295	2,496	2,631	2,727	2,784	2,831
1996			668	1,477	1,968	2,263	2,447	2,562	2,645	2,707
1997				696	1,540	2,055	2,357	2,551	2,699	2,806
1998					770	1,670	2,225	2,583	2,822	2,985
1999						690	1,515	2,051	2,436	2,666
2000							544	1,321	1,859	2,191
2001								563	1,355	1,852
2002									593	1,416
2003										621

Table 2: Accident Year x Development Year Cumulative Paid Losses (millions)

t	Development Year 0	1	2	3	4	5	6	7	8	9
0	624	1,595	2,066	2,366	2,559	2,685	2,765	2,818	2,860	2,895
1	695	1,503	1,975	2,295	2,496	2,631	2,727	2,784	2,831	
2	668	1,477	1,968	2,263	2,447	2,562	2,645	2,707		
3	696	1,540	2,055	2,357	2,551	2,699	2,806			
4	770	1,670	2,225	2,583	2,822	2,985				
5	690	1,515	2,051	2,436	2,666					
6	544	1,321	1,859	2,191						
7	563	1,355	1,852							
8	593	1,416								
9	621									

In Table 2, which is a reformatted version of Table 1, each row after the first has been shifted to the left, and development years have been renumbered, from zero to nine, to represent the number of years that have elapsed since the year in which the accident occurred. (I shall refer to these development years as DY0, DY1, and so on, and to accident years, also renumbered, as AY0, AY1, and so on.) The rearranged data in Table 2 more clearly shows how the claim payments for an accident year develop over time, represented by the number of development years subsequent to the year of the accident. As before, these are cumulative claim payments. Table 2 is typical of the data commonly used to estimate loss reserves for a single line of business. To estimate reserves, one must estimate, for all accident years, the difference between the amounts already paid and the ultimate amounts that will have been paid when all claims are finally settled. (This may occur well after DY9; if so, then years prior to AY0 will also have to be analyzed, using separate data. Here I ignore all prior years.)

5. THE SOURCES OF LOSS RESERVE UNCERTAINTY

Property-casualty loss reserves are estimates – forecasts -- of the total future payments that will be required to settle claims on accidents that have already occurred. The actual future payments may deviate from the forecast amount for several reasons, each of which reflects a different risk.

5.1 Types of Risk

To distinguish between the different types of risk that are practically important in estimating LRU, it may be helpful to consider a simple phenomenon with which we are all familiar: flipping a coin. Let us postulate that we receive payoffs that correspond to the proportion of heads that are flipped. In the first place, even if we know for certain that the probability of flipping heads is p, the fact remains that the proportion of heads actually flipped can deviate substantially from p. This is **process risk**. By contrast, **parameter risk** reflects the fact at the true probability of flipping heads is unknown to us, and must either be assumed or inferred from the outcomes we observe. We may, for example, infer that a coin with five heads in ten flips is fair and another with eight or nine flips in ten is biased. Parameter risk reflects the possibility that in both instances we may be wrong. In most practical situations we are exposed to both process risk and parameter risk, and find it difficult to distinguish between the two.

Note, by the way, that it is process risk that gives rise to parameter risk. If a coin with a true probability of ½ of flipping heads always produced five heads in ten flips, then parameter risk would not exist.

Finally, the inferential process -- inferring whether a coin is biased by observing the outcome of multiple flips -- itself relies on a crucial assumption that is seldom made explicit: that the probability of the coin flipping heads is constant and independent of prior and subsequent flips. If, by contrast (as many gamblers assume), the probability of flipping heads is mean-reverting, so that flipping tails is more likely after a long series of heads, or if (by contrast) outcomes are positively serially correlated, so that flipping tails becomes even more likely after a preceding series of tails, then the previously described inferences from observed outcomes are too simple to reflect reality. The point here is that making inferences from the outcomes we observe depends upon an explicit or implicit mental model of how those outcomes are generated. If our mental model is wrong, and assumes serial correlation where it is absent or assumes the absence of serial correlation where it is present, we may draw the wrong conclusions from what we observe. This possibility that our mental model is wrong is called **model risk**. All three types of risk – process risk, parameter risk, and model risk – are important in estimating loss reserve uncertainty.

5.2 Process Risk

Some degree of uncertainty is inherent in the process of settling claims payments. The amount actually paid in a given development year is a complex result of numerous factors – among them the uncertain outcomes and costs of medical diagnoses and treatments, and of court proceedings or settlement negotiations. None of these factors can be easily forecast. Consequently, even at an aggregate level, attempts to predict future claim payments are inescapably imprecise. Process risk explains why our models fit past paid losses only imperfectly, and why they require an error term in the prediction equation discussed below.

5.3 Parameter Risk

Actuarial methods necessarily use past experience to forecast future patterns. But past experience can be misleading. The culprit here is the relatively short period of time – ten years – covered by a typical paid loss triangle, so that parameter estimates are derived from a relatively small number of observations. The paid losses in the triangle are all affected by process risk, but the small number of observations creates substantial **sampling error**. The result is that past data may, simply by chance, reflect unusually favorable or unfavorable claims experience, and thereby affect the model parameters we are trying to estimate.

As an example, consider the step in the chain ladder method in which one of several weighting methods is used to produce a ratio of cumulative losses in DY5 to cumulative losses in DY4. This and other similar ratios are key parameters in the chain ladder model. But note that in DY5 there are only five cumulative AY losses from which ratios can be formed. If one or more of these five cumulative loss numbers is affected by an unusually large, or unusually small, claim payment in DY5 or in any preceding DY, then the resulting ratio will be atypically large or small. As this example suggests, this problem of sampling error is more acute for firms and lines of business that have few claims involving large payments than for firms with many small claims. Sampling error is likely to be less relevant to private passenger auto than to, say, product liability or D&O.

Especially in low-frequency high-severity lines of business, then, sampling error can lead to distorted estimates of key loss reserve parameters. This important consequence of sampling error can be called parameter risk, since it pertains to the accuracy with which we can use past data to estimate key parameters in our model of reserves or reserve uncertainty. Unlike process risk, which is inherent in the claims settlement process, parameter risk reflects our ignorance of the true parameters that characterize that process and the consequent need for us to use imperfect data to estimate them.

5.4 Model Risk

All reserve estimates require extrapolation from the past to the future. We use data from the past to create a model of the evolution of claims payments, and we then use this model to forecast future payments. Implicit in this process are two crucial assumptions. One is that we have **correctly** modeled the past: that we have included all the relevant variables and specified the correct functional form of the model. A second implicit assumption is that the pattern of future claims payments will continue to conform to this model. That is, the way claims are settled in the future will closely resemble the way they have been settled in the past. This implicit assumption may become misleading if there are fundamental changes – known as **regime changes** – in the claims settlement process at a particular firm (perhaps as a consequence of regulatory or judicial decisions), or in the types of claims being settled (which may change over time due to changes in business mix). These two components of reserve uncertainty can be called model risk, since they pertain to the capability of a model to correctly extrapolate from the parameters of past experience to estimates of future payments.

Regime changes that have occurred in the past can often, although not invariably, be identified and corrected by means of a thorough analysis of the differences (residuals) between the fitted values of past paid losses obtained from a model and the actual paid losses that have been observed. (The paper by Barnett and Zehnwirth (2000) provides an excellent example of the analysis of residuals.) If these residuals exhibit a trend or a sudden temporal shift, then there is good reason to suspect that a regime change has occurred. This possibility can be confirmed by testing a more advanced model that incorporates temporal changes in the value of key parameters. Unfortunately, there are tradeoffs in introducing additional variables, since doing so is likely to increase our estimate of LRU. Introducing additional variables may better fit past loss payments, but at the expense of greater uncertainty in forecasting future loss payments. The brutal fact is that a simplified but imperfect model of past losses may be superior to a more complex model in its ability to precisely forecast future loss payments.

It is important to note, however, that regime changes occurring in the future can invalidate the results of our analysis, which consist of forecasts and estimates concerning that future. The risk measures proposed here implicitly assume a stable future environment, and do not incorporate the risk of future regime changes. If such changes do occur, then the results obtained by the method presented here may become totally irrelevant to the changed circumstances.

5.5 Summary

Process risk essentially reflects the fact that some aspects of the claims settlement process are inherently unpredictable. **Parameter** risk reflects the fact that, even if we have a correct model of the evolution of paid losses, our estimates of the parameters of this correct model will necessarily be somewhat imprecise. **Model** risk reflects the possibility that the model we are using may itself be incorrect, so that our ability to predict future loss payments from past paid losses may be impaired. A satisfactory approach to estimating LRU should address all three of its sources. In particular, it should provide systematic ways to avoid, minimize, or detect model risk in the past, and it should quantify both process and parameter risk.

6. CRITERIA FOR ACCURATELY ESTIMATING RESERVES AND RESERVE UNCERTAINTY

The method presented here uses linear regression to fit past loss payments, forecast future loss payments, and estimate LRU. But the use of linear regression – or any other method, for that matter -- will not produce accurate estimates of reserves and LRU unless certain crucial problems are avoided or corrected. Despite their huge potential impact on estimates of loss reserves and LRU, and the enormous attention devoted to them by even elementary econometrics texts, these problems are typically assumed away if they are discussed at all. Here I describe these problems, their relevance, and what can be done about them.[1]

6.1 Linear Regression

In linear regression we initially assume a simple relationship between some dependent variable Y (here specified as paid losses) and one or more independent variables X. The relationship between the two is represented by the equation $Y = \beta X + \varepsilon$, where β is one or more parameters to be estimated, and ε is an error term that represents random disturbances or deviations from the

[1] In this section I rely heavily on Kennedy (2003), a superb elementary presentation of the essentials of econometrics, and on Greene (2000), one of the most widely used advanced texts.

predicted relationship between Y and X. In the simplest possible model, X consists of a single independent variable. I shall refer to this as model 1.

In this simple model, process risk is represented by ε, which consists of disturbances that are assumed to have an expected value of zero and a standard deviation that is constant across all observations. Parameter risk is reflected in the fact that the resulting estimated value of β, represented by b, is assumed to be correct, so that $b = \beta$, which may not be true. Finally, model risk is represented in several ways. For example, model 1 directly assumes that the relationship between Y and X is indeed linear, that all variables pertinent to Y are included in X, and that β is constant. All of these may in fact be false, but can typically be checked by thoroughly examining the residuals from the model – the deviations between actual and fitted paid losses.

6.2 Bias

If important variables affecting Y are omitted from model 1, the error term is likely to have a nonzero mean, the fitted and forecast values from the model will be biased – their estimated values will systematically deviate from their true values. In the absence of specific data concerning the omitted variables, we can take their influence into account by adding an intercept term to the original model, which now becomes $Y = \alpha + \beta X + \varepsilon$. I shall refer to this model as model 2. Since the unnecessary use of an intercept term affects our estimate of LRU, we should use model 2 only when there is convincing evidence that the error terms from model 1 have a mean that significantly differs from zero.

6.3 Varying Parameters

Another source of model risk is change over time in the value of β. This may occur due to changes in (a) the firm's claims settlement process, (b) judicial decisions or regulatory requirements, (c) the composition of the firm's policyholders in that line of business, or (d) the structure of a firm's reinsurance program (since paid losses are reported net of reinsurance recoverable). These and other possible changes may produce sudden or gradual changes over time in the true value β, but these changes that will not be reflected in its estimated value b. Fortunately, situations of this sort exhibit a characteristic pattern of residuals, and can be corrected by using a slightly more complex model in which β is assumed to change linearly over time, so that $\beta = \beta_0 + \beta_1 t$, where $t = 0, 1, \ldots n$ is a time index. When this is substituted into the original model we have a new model, which I shall refer to as model 3: $Y = \beta_0 X + \beta_1 tX + \varepsilon$. If $\beta_1 = 0$, then this model collapses into the original, simpler model 1.

6.4 Correlated Disturbances

Linear regression models assume that the disturbance terms for each observation are uncorrelated with one another. For the data in Table 2, it sensible to assume – as many others have -- that the disturbances in different accident years are uncorrelated. The important question is whether the disturbance terms within the same accident year are correlated across development years. I will show that they are in cumulative data.

Suppose that, for a given line of business and a given accident year, the expected paid losses are \$40, \$30, \$20, and \$10 in development years zero through three. However, in any given development year the actual paid losses will deviate from these expected paid losses due to a variety of random factors whose net effect in those development years is ε_0, ε_1, ε_2, and ε_3, respectively. There are good reasons for assuming that these four random terms are independent of one another and of all other similar random terms affecting other accident years and development years. However, if we create a table of cumulative paid losses, as in Table 2, we will create correlations among these random terms, since the new disturbance term for AY1, for example, is now $\varepsilon_0 + \varepsilon_1$, which is clearly positively correlated with ε_0, the disturbance term for AY0. In cumulative data, then, an unusually large disturbance in any development year will be reflected in all subsequent cumulative paid losses for that accident year.

The typical consequence of correlated disturbances, explained in both Kennedy and Greene, is that a given model will appear to fit the historical data better than it actually does, so that process error will be underestimated. This, in turn will result in LRU being underestimated as well.

Fortunately, the remedy for the correlated disturbances in cumulative paid loss triangles is simple: we should use incremental paid losses rather than cumulative ones. Consequently, the data we will utilize to estimate reserves and reserve uncertainty will be incremental, like that shown in Table 3, which is derived from Table 2. (The boxes in Table 3 are explained later.) Hallowell's (1996) alternative solution is discussed below.

Table 3: Accident Year x Development Year Incremental Paid Losses (millions)

	Development Year									
t	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	**624**	971	**471**	300	193	126	80	53	42	35
1	**695**	808	**473**	319	201	135	96	57	47	
2	**668**	809	**491**	295	184	115	83	63		
3	**696**	844	**515**	302	194	148	107			
4	**770**	900	**555**	358	239	162				
5	**690**	825	**536**	384	231					
6	**544**	777	**537**	332						
7	**563**	792	**497**							
8	**593**	823								
9	**621**									

6.5 Heteroskedasticity

Linear regression assumes that the disturbance terms for past observations are homoskedastic – i.e., have a constant variance or standard deviation as measured here in dollars (and not in percentage terms). This assumption is clearly violated in paid loss triangles like Table 3, for the variability of disturbances typically decreases from one development year to the next. Heteroskedastic (non-constant) disturbances reduce the precision of reserve estimates and especially of estimates of LRU.

There are two remedies for heteroskedasticity that are relevant to the problem at hand. One is to use a procedure known as Generalized Least Squares (GLS), which is a variation of linear regression that incorporates the use of an assumed or estimated variance-covariance matrix of disturbances (Halliwell, 1996; Taylor and Ashe, 1983). One typical assumption, for example, is that the standard deviation of disturbances is proportional to the observed losses themselves. Besides its complexity, there is a fundamental problem with the use of GLS for estimating reserves and LRU. Whether the variance-covariance matrix is assumed or estimated, the use of GLS introduces additional parameter risk that is not taken into account in the estimate of LRU. Moreover, however useful GLS may be in increasing the accuracy of reserves estimates, when it is applied to the problem of estimating LRU it comes dangerously close to assuming precisely what we are trying to estimate.

A second and far simpler remedy is to assume – quite plausibly – that the standard deviation of disturbance terms is constant within the same development year. What this implies, in practice, is the need to perform separate regressions on each development year. While this procedure may be less elegant than performing a single comprehensive regression for the whole paid loss triangle, it avoids the need to make problematic assumptions about variances and covariances.

6.6 Zero Correlation Between Disturbances and Independent Variables

Halliwell (1996) correctly points out that the classical linear regression model required that the independent variables be non-stochastic. However, both Greene (2000) and Kmenta (1977, pp. 297ff.) demonstrate that this stringent and seldom-met condition can be replaced by one that is far less demanding, namely, that the disturbance terms be independent of the values of the independent variables. However, even if the correlation is slightly positive rather than zero, the effect on the resulting estimates of reserves and LRU is imperceptible, as I shall demonstrate later on.

6.7 Implications and Summary

In using linear regression to estimate reserves and LRU, it is essential to avoid the various pitfalls just described. If one or more of these problems do occur, then estimates of reserves and LRU may be seriously affected. It should be noted that this conclusion applies not only to the use of linear regression, but to the use of other estimation procedures as well.

The immediate implications for modeling reserves and LRU can be summarized as follows: (a) if bias appears to be a problem, use model 2 rather than model 1; (b) if the model parameters appear to change over time, use model 3; (c) to avoid correlated errors, use incremental paid loss triangles; (d) to avoid heteroskedasticity, analyze different development years separately; (e) the use of non-stochastic independent variables, as advocated by Halliwell (1996) is unnecessary provided that there is no correlation between disturbances and the independent variables.

7. ESTIMATING LOSS RESERVES

I will present the full procedure for estimating reserves in this section, and for estimating LRU in the next one. In both, the presentation will focus initially on DY0 through DY7 and subsequently on DY8 and DY9, where data is minimal and extrapolation from the results for preceding DY's becomes necessary.

7.1 Fitting and Forecasting Losses for DY1 to DY7

The procedure I use here is linear regression. As explained in the previous section, we will analyze each DY separately. The independent variable X used to fit each DY is the column of paid losses in DY0, shown in the left box in Table 3. We will illustrate the procedure by fitting DY2, the right box in Table 3, as the dependent variable. In model 1, these are the only two variables.[2]

In the absence of specialized statistical software, one would typically perform the linear regression in Excel to obtain the regression coefficients and fitted values. Then one would obtain forecast values and, finally, calculate their forecast standard deviations. This final step can be especially complex. Here I introduce a method first suggested by Salkever (1976), later recommended by Kennedy, and described briefly but clearly by Greene (pp. 308-310), that makes it possible to do all three steps simultaneously.

Table 4: Fitting and Forecasting DY2

	Y	X		
	DY2	DY0	D9	D8
	471	624	0	0
	473	695	0	0
	491	668	0	0
	515	696	0	0
	555	770	0	0
	536	690	0	0
	537	544	0	0
	497	563	0	0
	0	593	0	-1
	0	621	-1	0
b		0.77	477	456
se_b		0.03	62	62
R^2, se_{est}		0.99	59.0	
t-statistic		24.3	7.7	7.4

Table 4 shows the key steps in the Salkever algorithm. First, augment the dependent variable Y with zeros so that it is the same length as DY0. Second, for each of the zeros added to Y, create

[2] Some readers have asked why I don't propose using DY1 and DY2 in estimating DY3, and so on. The answer is simple: very soon one has more variables than observations. This point is actually a specific case of an important and more general principle: increasing the number of independent variables can actually increase loss reserve uncertainty by reducing degrees of freedom.

additional columns in the independent variable X, in each of which there is a single entry, -1, corresponding to one of the zeros in Y. These additional variables are known as "dummy" variables, and so I have labeled them as D9 and D8, since their nonzero entries correspond to AY9 and AY8. In this example X now consists of three variables. Third, perform the linear regression (LINEST in Excel, with no intercept). The results are shown at the bottom of Table 4, with one slight difference from those obtained in Excel: I have reversed the left-to-right order of the first two rows of regression results, so that they now appear in the same order as the three variables in X, thereby doing what Microsoft should have done.

7.2 Results

The first column of results is identical to what one would have obtained by simply regressing Y against X. The estimated regression coefficient b is 0.77, which indicates that the losses in DY2 are about 77% of those in DY0. The standard error of b, in the second row, tells us that b has an estimated standard deviation of 0.03. The t-statistic, in the fourth row, is the ratio of b to its standard error. As a general rule of thumb, t-statistics with absolute values greater than 2.0 are considered significantly different from zero. The two numbers in the third are R^2 and the standard error of the estimate, which is the estimated standard deviation of the error terms, the differences between fitted and actual values of Y. In the absence of an intercept R^2 is typically high, so the standard error of the estimate is a better measure of goodness of fit.

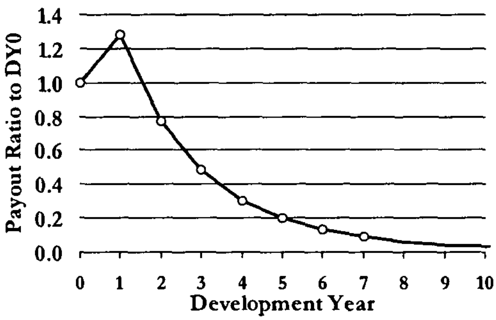
Table 5: Fitted Values & Regression Coefficients For Each Development Year

	Development Year							
t	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0		800	480	303	187	124	85	54
1		890	534	337	208	138	95	61
2		855	513	324	200	133	91	58
3		891	535	338	209	139	95	
4		986	592	374	231	153		
5		884	530	335	207			
6		697	418	264				
7		722	433					
8		760						
9								
b		1.28	0.77	0.49	0.30	0.20	0.14	0.09
$\mathbf{se_b}$		0.05	0.03	0.02	0.01	0.01	0.01	0.00
$\mathbf{se_{est}}$		91	59	40	15	12	9	4

The real value of the Salkever procedure lies in the remaining two columns of regression results. The regression coefficients in the first row are the **forecast paid losses** for AY9 and AY8, respectively, and the values in the second row are the corresponding **forecast standard errors.**[3]

The results from applying this procedure to DY1 through DY7 are summarized in Table 5. The top part shows the fitted values of past loss payments; the lower part shows the regression coefficients and other summary measures of goodness of fit. (In all cases R^2 was 0.99.) For DY1 through DY7, all estimated coefficients b were relatively precise, as indicated by their small standard errors. In each DY goodness of fit, as measured by the standard error of the estimate, is likewise small relative to the average paid loss. Particularly noteworthy, though, is the fact that the standard error of the estimate varies dramatically across development years. This validates our concern about heteroskedasticity, described in section 6.5.

Figure 1: Model 1 Payout Pattern



The overall pattern of the regression coefficients is shown in Figure 1. These regression coefficients can be used in a fashion similar to chain ladder link ratios. The regression coefficient for any given DY is the estimated incremental dollars paid in that DY relative to the dollars paid in DY0. In DY1, for example, one can anticipate paying, on average, about 28% more than was paid out in DY0. **For a given AY, then, the remaining payments can be estimated by adding up the coefficients for the remaining DY's and then multiplying by the amount paid in DY0.** In this example, the sum of the coefficients is 3.43 when one includes the tail. For AY9 the estimated

[3] Readers who attempt to replicate these results may obtain slightly different parameter estimates, since the data in Tables 3 and 4 are rounded values. The actual data is available from the author on request.

remaining paid losses are 3.43 times the $621 million loss in DY0, or $2.130 billion, so that the estimated ultimate AY total is $2,751 billion.

7.3 Analysis of Residuals

Table 6: Residuals from Fitted Values

	Development Year							
t	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0		171	-9	-3	5	2	-5	-1
1		-82	-62	-18	-7	-4	1	-3
2		-47	-22	-29	-16	-18	-8	4
3		-47	-19	-36	-15	10	12	
4		-86	-37	-16	8	9		
5		-59	5	49	24			
6		80	119	68				
7		70	64					
8		63						

Table 6 shows the residuals -- the difference between actual and fitted values – for the data analyzed here. Two questions are central to the analysis of these residuals. First, do they exhibit patterns that may alert us to variables or unusual conditions not reflected in Model 1? Second, are the magnitudes of any particular residuals significant or noteworthy?

Table 7: Standardized Residuals

	Development Year							
t	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0		1.9	-0.2	-0.1	0.3	0.1	-0.6	-0.3
1		-0.9	-1.0	-0.5	-0.5	-0.3	0.1	-0.8
2		-0.5	-0.4	-0.7	-1.1	-1.6	-0.9	1.1
3		-0.5	-0.3	-0.9	-1.0	0.9	1.3	
4		-0.9	-0.6	-0.4	0.5	0.8		
5		-0.6	0.1	1.2	1.5			
6		0.9	2.0	1.7				
7		0.8	1.1					
8		0.7						

Table 6 provides a basis for answering the first question, and Table 7, which shows standardized residuals (residuals divided by the DY standard error), facilitates answering the second. Table 7 shows, for example, that only one standardized residual has an absolute value greater than 2, which can be expected to occur about five percent of the time, or in about two instances of the 42 values shown in the table. Although the signs of the residuals show a suspicious pattern across accident years, the magnitude of the deviations is not sufficiently great to add additional variables to the analysis. (It is also the case, as numerous studies have shown, that we psychological anticipate that truly random variables will be even more "random" than is in fact the case. The signs and magnitudes of the residuals in Table 7 are quite consistent with an assumption of random residuals.)

The fact that an extensive discussion of the art of residual analysis is beyond the scope of this paper should by no means obscure its fundamental importance. The estimation of loss reserves and LRU should not be a mechanical application of a standard algorithm to standard data. As experienced actuaries and analysts know, the scientific model-building that lies at the core of actuarial science must necessarily be accompanied by skillful judgment in determining how those models are applied and interpreted for particular firms and lines of business.

7.4 Forecasting the Tails

Table 8 shows the forecast future paid losses obtained from applying model 1 to the data in Table 3 as well as the estimated payments for the tails, DY8 and beyond. The procedures used to obtain these tail estimates makes two important assumptions. The first is that the regression coefficients from DY4 and beyond decrease exponentially. Figure 1 already demonstrated that this assumption does not hold for earlier DYs. Focusing on DY4 and beyond makes it possible to apply this procedure to lines of business with long tails. The second assumption is that the rate of exponential decay can be estimated from the coefficients already obtained for DY4 through DY7. I now describe the two steps needed to derive forecasts from these assumptions.

In step one we extrapolate the regression coefficients already obtained to DYs beyond DY7. Because we have assumed that the coefficients decrease exponentially, it is appropriate to use logarithmic regression. We create a variable W that consists of the regression coefficients for DY4 through DY7, shown previously in Table 5. We also create a variable V consisting simply of the numbers 4, 5, 6, and 7. Then we obtain estimates a and b of the coefficients α and β in the logarithmic regression $\ln W = \alpha + \beta V + \varepsilon$. From this we obtain b, the estimated value of β, which is the logarithm of the rate of exponential decay. In this analysis I use d = exp(b) = 0.66 as the estimated rate at which the coefficients decrease from one year to the next in the tail.[4]

Table 8: Forecast Future Paid Losses

	Development Year										
t	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Tail**
0											48
1										27	53
2									35	24	47
3								61	43	29	57
4							105	67	47	31	62
5						137	94	60	42	28	56
6					163	108	74	47	32	21	42
7				274	169	112	77	49	33	22	44
8			456	288	178	118	81	52	35	23	46
9		796	477	302	186	124	85	54	37	24	48
Development Year Total		796	933	863	696	600	517	390	305	230	504

In step two we create robust forecasts of the paid losses in subsequent development years by using an average of three separate forecasts. For each accident year, the paid loss for DY8 is forecast as $P_8 = (P_4d^4 + P_5d^3 + P_6d^2)/3$. The three terms in parentheses are three different forecasts of P8 created from the actual or forecast paid losses in DY4, DY5, and DY6. The forecasts for P_9 are made in the same way, except that the exponents of d are each increased by one. Finally, the forecast value for the tail, consisting of paid losses for all development years after nine, is calculated as the forecast for P_{10} multiplied by $1/(1\text{-}d)$, the formula for the sum of the infinite exponentially decreasing series $(1 + d + d^2 + \ldots)$. The results of this procedure have already been shown in Table 8. The estimated reserve for these data, based on Model 1, is $5,835 million.

---

[4] Technically, unless the logarithmic regression perfectly fits the data, one should include a slight adjustment for the error term in order to obtain the mean estimated value of b. By deliberately failing to include this adjustment I instead obtain the median value of b, which is presumably more robust. In most cases the difference is miniscule and difficult to explain to a non-technical audience.

8. ESTIMATING LOSS RESERVE UNCERTAINTY

As in estimating loss reserves, here we deal first with DY1 through DY7, and then tackle DY8 and beyond. We will first estimate the uncertainty of the total forecast payments for each DY. Then we will estimate total LRU by appropriately aggregating the uncertainties obtained for each DY.

8.1 Estimating the Uncertainty of DY Forecast Totals

For DY1 there is only one future payment to be forecast, and we can obtain that forecast and its forecast standard deviation directly from the Salkever method.

Table 9: Calculating the Standard Deviation of Forecast Paid Losses for DY2

Step 1: Assemble the Input Data: X, X_0, s, I

	DY0
X=	624
	695
	668
	696
	770
	690
	544
	563
X0=	593
	621

The standard error of the estimate, se_{est}, shown in Tables 4 and 5:

s= 59

The Identity Matrix I (for DYn it is n x n)

I=

1	0
0	1

Step 2: Calculate the Variance-Covariance Matrix VCV

$$VCV = s^2[I + X_0(X'X)^{-1}X_0']$$

=

3,838	369
369	3,872

Step 3: Calculate the square root of the sum of the entries in the VCV matrix

$$(\Sigma\, VCV)^{1/2} = 8{,}447^{1/2} = 92$$

This is the standard deviation of the sum of forecast paid losses for DY2

For subsequent DYs the problem is more complex, since forecast future payments within a DY share the same parameters and are therefore correlated since they share common parameter risk. What this means, in concrete terms, is that if the regression coefficient b is too high relative to its true value β, then all forecasts will be too high, and so will be correlated with one another, although not perfectly. When we estimate LRU we must take into account not only the forecast standard errors for each entry in the lower right portion of the loss reserve triangle, but also the estimated covariances among these forecasts. Fortunately, the Salkever procedure provides a relatively simply way to do this.

The estimation procedure for DY2 and subsequent DYs is shown in Table 9. The input data are shown at the top of the table. One is the column of paid losses in DY0. Recall that in Table 4 the first eight entries of DY0 were used to fit the eight paid losses already observed in DY2, and the remaining two entries were used to forecast future paid losses. Here we need to split DY0 into two separate parts, which we label X and X_0, to correspond to the notation used by Greene (2000, p. 309). Another input, s, is the standard error of the estimate for DY2, already reported in Tables 4 and 5. Finally, we need an identity matrix I, a square matrix of size n, where n is the number of the DY, with one's on its main diagonal and zero's elsewhere. From these inputs we obtain VCV= $s^2[I + X_0(X'X)^{-1}X_0']$, the variance-covariance matrix for forecast errors. We then sum these entries and take the square root of that result to obtain the standard deviation of the DY2 sum of forecast paid losses, which here is 92.

The standard deviations of the sum of forecast paid losses for DY3 to DY7 are calculated in the same way.[5] Note that as we move from one DY to the next we must increase the number of entries in X_0 by one, correspondingly decrease the number in X by one, and increase the dimension of I by one. The results are reported in Table 11, to which we shall return after we first obtain standard deviations for paid losses in DY8 and beyond.

8.2 Estimating the Standard Deviations of Forecast Tail Paid Losses

Salkever's method, applied to DY1 through DY7, provided forecasts of future paid losses (shown in Table 8) as well as standard errors (standard deviations) for these forecast values, are shown in Table 10. The table also shows the estimated standard errors of forecast losses for DY8 and beyond, which we calculated as follows.

As with regression coefficients, the assumption is that the standard errors decrease exponentially in the tail. As before, we use logarithmic regression to estimate the rate of decrease. Here, however, the dependent variable U consists of the average standard error for each DY from DY1 to DY7, and the independent variable T consists of the numbers from 1 to 7. For the regression $\ln U = \alpha + \beta T + \varepsilon$, we obtain an estimate b such that the rate of decrease $g = \exp(b) = 0.61$. In a manner identical to the one used for paid losses, we forecast the standard deviations for DY8 in each AY as $E8 = (E_4g^4 + E_5g^3 + E_6g^2)/3$, an average of three forecasts. Here each E within parentheses is the standard

[5] The Excel array formula for VCV, where range names are shown in boldface type, is this: VCV= S*(**I**.2+ MMULT(MMULT(**XZERO**,MINVERSE(MMULT(TRANSPOSE(**X**),**X**))),TRANSPOSE(**XZERO**))).

error of the forecast (for cells with forecast values) or the standard error of the estimate (for cells with observed values).

Table 10: Standard Errors of Forecast Paid Losses

	Development Year										
t	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Tail**
0											5
1										2	5
2									3	2	5
3								5	3	2	5
4							10	5	4	2	5
5						13	10	5	4	3	5
6					16	12	10	4	4	2	5
7				42	16	12	10	4	4	2	5
8			62	42	16	12	10	5	4	2	5
9		96	62	43	16	12	10	5	4	3	5

The last step requires that we obtain the standard errors of the sum of forecast paid losses for development years eight, nine, and the tail. To do this requires that we estimate what the variance-covariance matrices for those years might look like. We can in fact do this by examining the matrices already calculated for earlier development years.

The function of the variance-covariance matrix is to reflect interrelationships among the forecast errors. These interrelationships exist because the various forecast values all depend upon a common underlying parameter, β, whose estimate, b, may incorporate error. Any error in b will simultaneously affect all the forecast values. Moreover, as the number of observations on which estimates of β is based decreases, the interrelationships among forecast errors increase.

Were it not for these interrelationships among forecast errors, we could very easily calculate the standard deviation of total forecast paid losses by assuming that these forecasts and their errors were independent. In this case, the standard deviation of total forecast paid losses for development year n, which has n forecast values, would be $\sigma^* = (n\sigma_i^2)^{1/2} = \sigma_i n^{1/2}$, where σ^* is the standard deviation of total forecast paid losses assuming independence, and σ_i is the standard error of individual forecasts, here assumed to be equal (which is approximately true). In fact, however, we need to take into account the fact that the off-diagonal elements in the variance-covariance matrix are non-zero. Here we assume that these elements are identical in value (again, approximately true) and equal to $k\sigma_i^2$, where k is some constant to be estimated. In this case, the correct standard deviation of total forecast paid losses, σ, is $\sigma_i(n+kn(n-1))^{1/2}$. If we now calculate the ratio of σ to σ^* we obtain the quantity $(1+k(n-1))^{1/2}$, which is a multiplier: it is the amount by which σ^*, which assumes independence, must be multiplied to obtain σ, which does not. This approximation, when applied development years one through seven, produces results that are nearly an exact match to those obtained by having the actual variance-covariance matrix.

The key to applying this method is having a value for k, without which the multiplier cannot be calculated. The procedure used here was, first, to obtain the value of k from the variance-covariance matrices calculated for development years one through seven, second, to use linear regression to fit these values to an independent variable consisting of the number one through seven, and third, to forecast values of k for development years eight and nine and for the tail, for which the independent variable was nine plus the tail's weighted average length in years.[6]

Table 11: Standard Errors of Forecast Paid Losses
By Development Year, Total Reserve, and Calendar Year

Development Year	1	2	3	4	5	6	7	8	9	tail	Total
A. Sum of Forecast Paid Losses	796	933	863	696	600	517	390	305	**230**	**504**	**5,835**
B. Standard Deviation of Forecast	96	92	81	37	34	33	17	18	**15**	**45**	**175**
C. Coefficient of Variation (=B/A)	12%	10%	9%	5%	6%	6%	4%	6%	6%	9%	**3.0%**
D. Calendar Year 2004 Forecast Paid Losses											**2,070**
E. Standard Deviation of CY Forecast											**124**
F. CY Coefficient of Variation (=E/D)											**6.0%**

8.3 Results

Table 11 shows the combined results of applying these-procedures. Line A shows the sum of the forecast paid losses for each development year, as previously reported in Table 8. Line B shows the

[6] Recall that d is the estimated ratio, in the tail, of the paid loss in one development year to the paid loss in the prior development year, so that d < 1. The average length of the tail, L, is calculated as a ratio in which the numerator is the infinite series $1+2d+3d^2+4d^3+\ldots$, and the denominator is the infinite series $1+d+d^2+d^3+\ldots$. The numbers 1, 2, and so on are the number of years subsequent to development year 9 in which payments occur, and each year is weighted by the percentage of total tail payments occurring in that year. The denominator is total tail payments. The value of the numerator is $1/(1-d)^2$, and the value of the denominator is $1/(1-d)$, so that the value of their ratio, L, is $1/(1-d)$. Consequently, for purposes of estimating k to calculate the multiplier for the tail, the number of the tail development year is $9 + 1/(1-d)$, which in this case is 12.

standard deviations of the values in line A. These differ from the values shown in Table 10, which are the standard deviations of the individual components of the sums in line A. Both, however, reflect parameter risk as well as process risk. The Total in line B is obtained by taking the square root of the summed squares of the values in that row. This assumes independence, which is appropriate since by using incremental paid losses we have eliminated correlations across development years.

Line C shows the coefficients of variation, the standard deviations divided by the forecast paid losses. For the total estimated reserve of $5.8 billion, the standard deviation of $175 million is approximately 3.0% of the reserve. The fact that a consistent methodology was here used to estimate both the reserve and its standard deviation underscores a point made earlier: **even if other methods or information are used to obtain a different estimated reserve, this estimate of loss reserve uncertainty, the coefficient of variation, should nonetheless remain valid.**

Line C also validates the concern about heteroskedasticity discussed in section 6.5. In Table 5 we shows that the standard deviations of the disturbance terms in our regression results varied considerably, in dollar terms, across different development years. Line C shows that heteroskedasticity remains even when the standard deviation of the disturbance terms are expressed as a percentage of forecast paid losses (i.e., as coefficients of variation). What this means is that the convenient assumptions often utilized in generalized linear models (GLM) or generalized least squares (GLS) may not be valid. In practice, these estimation procedures focus principally on estimating loss reserves, so that estimates of LRU are purely secondary. By contrast, the model presented here focuses principally on estimating LRU, and estimates of loss reserves are of secondary importance.[7]

Table 11 also shows, in line D, the sum of the forecast paid losses for calendar year 2004, which consists of the sum of the forecast losses in Table 8. The standard deviation of this value, shown in line E, is $124 million, or about 6% of the estimated calendar year total forecast payments of $2.07 billion. This calendar year measure of LRU can be especially important for helping managers to determine whether actual calendar year paid losses (for AY1 to AY9) deviate significantly from their forecast total.

I hasten to observe that the Coefficient of Variation (Table 11, Line C) and the Calendar Year Coefficient of Variation (Table 11, Line F) are both atypically low. Although I have deliberately not identified the firm nor the line of business analyzed here, I will point out that this firm has a high volume of business in this line and deliberately targets its exposures to the less risky end of the risk

[7] This distinction is not trivial. Estimates of loss reserves may in fact be improved by using estimates of LRU that are relatively correct but nonetheless absolutely wrong by orders of magnitude.

spectrum. In a subsequent report I will describe the typical parameters and risk measures for the principal firms in each line of business.

Finally, I simply note that the Calendar Year Coefficient of Variation (CYCV), 6%, shown in Table 11 Line F, is greater than the Total Reserve Coefficient of Variation (CV) shown at the far right of line C. This result is consistent with what one would anticipate, since the Total Reserve CV includes all forecast future loss payments, which are imperfectly correlated, whereas the entry on line F includes only the forecast future loss payments occurring in the next calendar year. The Total Reserve CV is therefore considerably more diversified than the CYCV, and consequently is smaller.

9. VALIDATING THE RESULTS

Here I validate the results just obtained by demonstrating that the same methods accurately estimate the future paid losses and LRU's of 10,000 simulated paid loss triangles with known parameters and outcomes.

To create simulated paid loss triangles I begin with an underlying deterministic payout pattern in which paid losses decrease exponentially from an initial value in DY0 that is identical for all accident years. (In this particular simulation, paid losses in each DY are half those in the preceding one.) I then add to each of these expected payments a random deviation drawn from a normal distribution, with a mean of zero and a standard deviation that increases linearly from 10% of the expected paid loss in DY0 to 100% in DY9 and 110% in DY10 and beyond. The simulations in fact generate the entire path of paid losses to the point where they become miniscule. Consequently, the ultimate paid losses can, in principle and in fact, depart considerably from the expected values established by the underlying pattern, and the standard deviations of these simulated variations can be calculated.

The results of the simulation are shown in Table 12. The first half of the table reports the accuracy of the method used here in forecasting DY sums of future loss payments. Line A shows the DY sums of expected future loss payments **before** random disturbances are added. Line B shows the average, over the 10,000 scenarios, of the simulated DY sums of future loss payments. Section C reports the results of using the procedure used in this paper to estimate DY sums of forecast future loss payments.

When the independent variable is stochastic, and consists of the simulated loss payments in development year zero, the results are only trivially different from those obtained by using as the independent variable the expected (i.e., deterministic) loss payments in DY0 as if they were in fact known. **This confirms the assertion in section 6 that the use of a stochastic independent variable is not a problem if its disturbances are independent of those that affect the**

dependent variable. The admittedly ad hoc procedure used here to calculate the tail values overestimates them somewhat. This is undoubtedly due to the fact that **using the exponential decay function to project tail payments rules out negative payments, while the simulation does not.** Nonetheless, forecasts of paid losses in the next calendar year, shown in the last column in Table 12, are remarkably accurate.

Table 12: Monte Carlo Results for Estimating Reserves and Reserve Uncertainty

DY:	1	2	3	4	5	6	7	8	9	10+	Total	CY
<u>DY Sum of Future Loss Payments</u>												
A. Underlying mean values	400	400	300	200	125	75	44	25	14	16	1,598	800
B. Average simulated values	399	401	301	200	124	75	44	25	14	16	1,599	799
C. Average forecast values												
-- using stochastic X	397	397	298	198	124	75	43	34	22	33	1,622	796
-- using fixed X	400	400	300	200	125	75	44	34	22	33	1,633	802
<u>Standard Deviation (SD) of DY Sum of Future Loss Payments</u>												
D. True SD from parameters	80	85	69	50	34	21	13	8	5	5	151	112
E. SD of simulated payments	80	85	71	50	34	21	13	8	5	5	152	112
F. Parameter risk multipliers		1.1	1.1	1.2	1.3	1.4	1.6					
G. True SD plus parameter risk		**89**	**77**	**60**	**43**	**30**	**21**					
H. Estimated SD												
-- using stochastic X	91	**97**	**82**	**62**	**45**	**32**	**19**	18	13	31	189	127
-- using fixed X	82	**92**	**79**	**61**	**44**	**31**	**19**	17	13	31	180	118

The second part of Table 12 verifies the accuracy of the procedure for estimating the standard deviations of forecast future loss payments. Line D shows the actual standard deviations used in the simulation, and line E shows the standard deviation of the simulated losses. As one would hope from a properly conducted simulation, the two are virtually identical. Line F shows the multipliers for parameter risk obtained from the modeled variance-covariance matrices, and line G shows the true standard deviations in line D multiplies by the corresponding values in line F. These values in line G are the values one would hope to obtain in estimating LRU. The actual estimates obtain, both with a stochastic X and a fixed X, are shown in section H. The two sets of estimates in this section agree closely with each other and with the target values in line G. However, it appears that using a fixed X, as recommended by Halliwell (1996) <u>slightly</u> improves the estimates for DY1 and DY2. For the total reserve, both stochastic and fixed X's produce a similar result, and substituting one for the other would have an imperceptible effect on the coefficient of variation (CV). For stochastic X the CV is 11.65%, and for fixed X it is 11.02%.

10. SUMMARY AND CONCLUSIONS

The method I have presented here for estimating loss reserve uncertainty – the coefficient of variation of estimated future loss payments -- has a number of merits. First, it can be used to address significant issues in surplus management, in pricing and capital allocation, and in the management of uncertainty. Second, it uses a measure of loss reserve uncertainty that facilitates comparison across different lines of business and can be applied to reserve estimates obtained through alternative methods. Third, it uses a publicly-available source of data that facilitates comparison across different firms. Fourth, the method avoids a number of serious pitfalls that can distort estimates of reserves or LRU. Fifth, the method is simple, at least as compared to some of the alternative methods advocated in the relevant literature. In particular, its use of Salkever's method provides an extremely useful shortcut for obtaining results. And sixth, the method accurately captures the key parameters of simulated paid loss trajectories. The reserve estimates are extremely accurate, and the estimates of reserve uncertainty, which include parameter risk, agree closely with benchmark calculations.

At the same time, the method proposed here has important limitations. First, I have used linear regression as a model for forecasting future loss payments. Linear regression is often advocated as a maximum likelihood procedure for estimating model coefficients. This is indeed the case when residuals are assumed to have a normal distribution. Here I make no such assumption, and so rely on linear regression as a procedure that estimates parameters so as to minimize squared error between fitted and actual values of the dependent variable. This is quite legitimate, but potentially disturbing to statistical perfectionists. Second, I make no assumption concerning the nature of the distribution of disturbances. The inferences from the model I present concern only the mean and standard deviation of loss reserves. The information needed to derive, say, an 80th percentile of the distribution of ultimate loss payments cannot be obtained from the method presented here.

I hope that I have convinced readers that the method presented here for estimating loss reserve uncertainty that is both accurate and reasonably simple to implement. I also hope that my presentation of it is accessible to a large number of professional colleagues, who are invited to apply it in their own work and to extend it to novel uses.

Acknowledgements

Although it has its origin in earlier work (Panning 1986, 1999), the method presented here was developed to assist the Risk Management Group at Progressive, Inc., in their ongoing Enterprise Risk Management project. I am indebted to Progressive's Risk Management and Actuarial staff for providing extensive data and feedback on the method described here.

I thank my colleagues at Willis Re for valuable discussions, comments and suggestions. I am especially indebted to Henry Mikrut for providing critical and easily accessible Schedule P data, to Ken Vincent for providing market share data, and to Yves Provencher for his patient and thorough evaluation of two versions of this paper. I likewise thank Richard Goldfarb, of Ernst and Young, for his insightful comments and careful verification of key results, Chris Gross at Christopher Gross Consulting for his detailed and thorough comments and discussions, Yisheng Bu of Liberty Mutual for his expert observations and questions, and Richard Derrig of Opal Consulting for his counsel and encouragement as well as for many years of stimulating conversations on this and many other topics. I similarly appreciate insightful questions and comments from Helmut Grundl of Humboldt University in Berlin, from Aaron Halpert of KPMG, and from a number of clients who have seen and commented on this method as applied to their own reserves and those of competitors. Given such expert advice, I can hardly avoid taking responsibility for any errors that remain.

The opinions and recommendations expressed in this paper are those of the author, and do not necessarily represent the views of Willis Re.

11. REFERENCES

[1] Ashe, F. 1986. An Essay at Measuring the Variance of Estimates of Outstanding Claim Payments. ASTIN Bulletin 16, S99-S113.

[2] Barnett, G. and B. Zehnwirth B. 2000. Best Estimates for Reserves. Proceedings of the Casualty Actuarial Society 87, 245-321.

[3] Braun, C. 2004. The Prediction Error of the Chain Ladder Method Applied to Correlated Run-off Triangles. ASTIN Colloquium.

[4] Brehm, P. 2002. Correlation and the Aggregation of Unpaid Loss Distributions. CAS Loss Reserve Seminar.

[5] England, P. D. and R. J. Verrall. 1999. Analytic and bootstrap estimates of prediction errors in claims. Insurance: Mathematics and Economics 25, 281-293.

[6] England, P. D. and R. J. Verrall. 2001. A Flexible Framework for Stochastic Claims Reserving. Proceedings of the Casualty Actuarial Society 88, 1-38.

[7] England, P. D. and R. J. Verrall. 2002. Stochastic Claims Reserving in General Insurance. British Actuarial Journal 8, 443-544.

[8] Greene, W. H. 2000. Econometric Analysis, Prentice Hall, Upper Saddle River, New Jersey. Fourth Edition.

[9] Halliwell, L. J. 1996. Loss Prediction by Generalized Least Squares. Proceedings of the Casualty Actuarial Society 88, 436-489.

[10] Hayne, R. M. 2003. Measurement of Reserve Variability. Casualty Actuarial Society Forum, 141-171.

[11] Hodes, D. M., S. Feldblum, and G. Blumsohn. 1996. Workers Compensation Reserve Uncertainty. Casualty Actuarial Society Forum, 61-150.

[12] Holmberg, R. D. 1994. Correlation and the Measurement of Loss Reserve Variability. Casualty Actuarial Society Forum 1, 247-277.

[13] Kennedy, P. 2003. A Guide to Econometrics. MIT Press, Cambridge, Massachusetts. 5th edition.

[14] Kloek,T. 1998. Loss development forecasting models: an econometrician's view. Insurance: Mathematics and Economics 23, 251-261.

[15] Kmenta, J. 1971. Elements of Econometrics. Macmillan, New York, New York.

[16] Mack, T. 1993. Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates. ASTIN Bulletin 23, 213-225.

[17] Mack, T. 1994. Measuring the Variability of Chain Ladder Reserve Estimates. Casualty Actuarial Society Forum, 101-182.

[18] Mack, T. 1995. Which Stochastic Model is Underlying the Chain Ladder Method? Casualty Actuarial Society Forum, 229-240.

[19] Mack, T.. 1999. The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor. ASTIN Bulletin 29, 361-366.

[20] Maddala, G. S. 1977. Econometrics. McGraw-Hill, New York, New York.

[21] Murphy, D. M. 1994. Unbiased Loss Development Factors. Proceedings of the Casualty Actuarial Society 81, 154-222.
[22] Narayan, P. and T. Warthen. 1997. A Comparative Study of the Performance of Loss Reserving Methods Through Simulation. Casualty Actuarial Society Forum, 1, 175-196.
[23] Panning, W. 1986. Estimated Loss Reserves and their Confidence Intervals. Casualty Loss Reserve Seminar, 838-855.
[24] Panning, W. 1999. The Strategic Uses of Value at Risk. North American Actuarial Journal 3:2, 82-105.
[25] Salkever, D. 1976. The Use of Dummy Variables to Compute Predictions, Prediction Errors, and Confidence Intervals. Journal of Econometrics 4, 393-7.
[26] Stanard, J. N. 1985. A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques. Proceedings of the Casualty Actuarial Society 72, 124-148.
[27] Taylor, G. 2003. Chain Ladder Bias. ASTIN Bulletin 33, 313-330.
[28] Taylor, G. 2004. Loss Reserving with GLM's: A case study. Centre for Actuarial Studies. University of Melbourne.
[29] Taylor, G. . 1987. Regression Models in Claims Analysis I: Theory. Proceedings of the Casualty Actuarial Society 87, 354-383.
[30] Taylor, G. and F. R. Ashe. 1983. Second Moments of Estimates of Outstanding Claims. Journal of Econometrics 23, 37-61.
[31] Verrall, R. J. 1994. Statistical Methods for the Chain Ladder Technique. Casualty Actuarial Society Forum, 393-446.

Biography of the Author

Bill advises insurers and other financial firms on issues pertaining to Enterprise Risk Management (ERM). Current and past Enterprise Risk Management clients include Progressive Corporation and United Grain Growers (now Agricore United), where his consulting team helped to put in place an insurance-based hedge that CFO Magazine called the "deal of the decade." This project is the subject of case studies at the Harvard Business School and the University of South Carolina. In April 2006 Bill became the first recipient of the ERM Research Excellence Award, established by the Actuarial Foundation to recognize excellence in contributions to the growing body of knowledge and research in enterprise risk management.

Early in his thirty-year career Bill taught at the Wharton School and other universities. Subsequently he held positions as the principal property-casualty investment strategist at Aetna, as Director of Asset-Liability Management and Director of Investment Risk Management at The Hartford, as a senior portfolio manager (of $4 billion) and director of a portfolio analytics group at an investment firm, as Chief Operating Officer of Willis' Advanced Risk Management Services division, and as Chief Investment Officer and co-President at Integrity Life. After the sale of that firm he joined Willis Re in March 2000 to help build its analytical capabilities.

A frequent speaker at industry events (with some forty presentations) and a regular columnist for ***Best's Review***, Bill has published more than thirty articles. Two have been selected by the Casualty Actuarial Society for their examination syllabus, one was selected by the Society of Actuaries, and yet another won a Graham & Dodd Award of Excellence for the best articles published in the ***Financial Analysts Journal***. Bill holds a Ph.D. from the University of Pennsylvania and a B.A. from the University of Kansas.

Phone: 212-820-7680 Cell: 502-387-5411 Email: bill.panning@willis.com

Methods and Models of Loss Reserving Based on Run-Off Triangles: A Unifying Survey

Klaus D. Schmidt
Lehrstuhl für Versicherungsmathematik
Technische Universität Dresden

Abstract.
The present paper provides a unifying survey of some of the most important methods and models of loss reserving which are based on run-off triangles. The starting point is the thesis that the use of run-off triangles in loss reserving can be justified only under the assumption that the development of the losses of every accident year follows a development pattern which is common to all accident years. This assumption can be viewed as a primitive stochastic model of loss reserving.
The notion of a development pattern turns out to be a unifying force in the comparison of methods which to a large extent can be summarized under a general version of the Bornhuetter-Ferguson method. It is shown that the loss-development method and the chain-ladder method as well as the Cape-Cod method and the additive method can be viewed as special cases of the general Bornhuetter-Ferguson method.
Some of these methods can be justified by general principles of statistical inference applied to suitable and more sophisticated stochastic models. It is shown that credibility prediction and Gauss-Markov prediction as well as maximum-likelihood estimation can contribute in a substantial way to the understanding of various methods of loss reserving.

Keywords. Bornhuetter-Ferguson principle; credibility prediction; development pattern; Gauss-Markov prediction; loss reserving; maximum-likelihood estimation.

1. INTRODUCTION

We start with the general modelling of loss-development data by a family of random variables representing incremental or cumulative losses and with the run-off triangles representing the observable incremental or cumulative losses (Section 2).

We then introduce the central notion of a development pattern which can be expressed in three different but equivalent ways and turns out to be a powerful and unifying concept for the interpretation and comparison of several methods and models of loss reserving (Section 3).

The subsequent three sections are devoted to methods, least-squares prediction, and maximum-likelihood estimation.

In the section on methods (Section 4), we start with a general version of the Bornhuetter-Ferguson method which provides a general framework into which several other methods, like the loss-development method, the chain-ladder method, the Cape-Cod method and the additive method, can be embedded as special cases. We also consider two variants of the chain-ladder method which have no practical interest but are needed as a link between the chain-ladder method and certain stochastic models.

In the section on least-squares prediction (Section 5), we study credibility prediction and Gauss-Markov prediction. It is shown that, under certain model assumptions, these methods of prediction yield predictors of the Bornhuetter-Ferguson type.

In the section on maximum-likelihood estimation (Section 6), we study maximum-likelihood estimation for a large class of stochastic models for claim counts. It is shown that in many cases, but not always, the maximum-likelihood estimators of the expected ultimate cumulative losses are identical with the chain-ladder predictors of the ultimate cumulative losses.

In the final section (Section 7) we collect some conclusions.

Throughout this paper, let $(\Omega, \mathcal{F}, P)$ be a probability space on which all random variables are defined. We also assume that all random variables are square integrable. Moreover, all equalities and inequalities involving random variables are understood to hold almost surely with respect to the probability measure P.

2. LOSS DEVELOPMENT DATA

We consider a portfolio of risks and we assume that each claim of the portfolio is settled either in the accident year or in the following n development years. The portfolio may be modelled either by incremental losses or by cumulative losses.

2.1 Incremental Losses

To model a portfolio by incremental losses, we consider a family of random variables $\{Z_{i,k}\}_{i,k\in\{0,1,\dots,n\}}$ and we interpret the random variable $Z_{i,k}$ as the loss of *accident year* i which is settled with a delay of k years and hence in *development year* k and in *calendar year* $i+k$. We refer to $Z_{i,k}$ as the *incremental loss* of accident year i and development year k.

We assume that the incremental losses $Z_{i,k}$ are *observable* for *calendar years* $i+k \le n$ and that they are *non-observable* for calendar years $i+k \ge n+1$. The observable incremental losses are represented by the following *run-off triangle*:

Accident Year	Development Year								
	0	1	…	k	…	$n-i$	…	$n-1$	n
0	$Z_{0,0}$	$Z_{0,1}$	…	$Z_{0,k}$	…	$Z_{0,n-i}$	…	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1,0}$	$Z_{1,1}$	…	$Z_{1,k}$	…	$Z_{1,n-i}$	…	$Z_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$Z_{i,0}$	$Z_{i,1}$	…	$Z_{i,k}$	…	$Z_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$Z_{n-k,0}$	$Z_{n-k,1}$	…	$Z_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$Z_{n-1,0}$	$Z_{n-1,1}$							
⋮	⋮								
n	$Z_{n,0}$								

The problem is to *predict* the non-observable incremental losses.

2.2 Cumulative Losses

To model a portfolio by cumulative losses, we consider a family of random variables $\{S_{i,k}\}_{i,k\in\{0,1,\ldots,n\}}$ and we interpret the random variable $S_{i,k}$ as the loss of *accident year* i which is settled with a delay of *at most* k years and hence *not later than* in *development year* k. We refer to $S_{i,k}$ as the *cumulative loss* of accident year i and development year k, to $S_{i,n-i}$ as a *cumulative loss of the present calendar year* n, and to $S_{i,n}$ as an *ultimate cumulative loss.*

We assume that the cumulative losses $S_{i,k}$ are *observable* for *calendar years* $i+k\le n$ and that they are *non-observable* for calendar years $i+k\ge n+1$. The observable cumulative losses are represented by the following *run-off triangle*:

Accident Year	Development Year								
	0	1	…	k	…	$n-i$	…	$n-1$	n
0	$S_{0,0}$	$S_{0,1}$	…	$S_{0,k}$	…	$S_{0,n-i}$	…	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$	…	$S_{1,k}$	…	$S_{1,n-i}$	…	$S_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$S_{i,0}$	$S_{i,1}$	…	$S_{i,k}$	…	$S_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$S_{n-k,0}$	$S_{n-k,1}$	…	$S_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$S_{n-1,0}$	$S_{n-1,1}$							
⋮	⋮								
n	$S_{n,0}$								

The problem is to *predict* the non-observable cumulative losses.

2.3 Remarks

Of course, modelling a portfolio by incremental losses is equivalent to modelling a portfolio by cumulative losses:

- The cumulative losses are obtained from the incremental losses by letting

$$S_{i,k} := \sum_{l=0}^{k} Z_{i,l}.$$

- The incremental losses are obtained from the cumulative losses by letting

$$Z_{i,k} := \begin{cases} S_{i,k} & \text{if } k=0 \\ S_{i,k} - S_{i,k-1} & \text{else.} \end{cases}$$

In the sequel we shall switch between incremental and cumulative losses as necessary.

Correspondingly, prediction of non-observable incremental losses is *essentially* equivalent to prediction of non-observable cumulative losses:

- If $\{\hat{Z}_{i,k}\}_{i,k\in\{0,1,\ldots n\},\ i+k\geq n+1}$ is a family of predictors of the non-observable incremental losses, then a family of predictors of the non-observable cumulative losses is obtained by letting

$$\hat{S}_{i,k} := S_{i,n-i} + \sum_{l=n-i+1}^{k} \hat{Z}_{i,l}.$$

- If $\{\hat{S}_{i,k}\}_{i,k\in\{0,1,\ldots n\},\ i+k\geq n+1}$ is a family of predictors of the non-observable cumulative losses, then a family of predictors of the non-observable incremental losses is obtained by letting

$$\hat{Z}_{i,k} := \begin{cases} \hat{S}_{i,n-i+1} - S_{i,n-i} & \text{if } k = n-i+1 \\ \hat{S}_{i,k} - \hat{S}_{i,k-1} & \text{else.} \end{cases}$$

For the ease of notation and to avoid the distinction of cases as in the previous definition, we shall also refer to $Z_{i,n-i}$ and $S_{i,n-i}$ as predictors of $Z_{i,n-i}$ and $S_{i,n-i}$, although these random variables are, of course, observable.

Warning: Whenever prediction is subject to an optimality criterion, it cannot be guaranteed in general that the previous formulas lead from optimal predictors of incremental losses to optimal predictors of cumulative losses or vice versa.

The enumeration of accident years and development years starting with 0 instead of 1 is widely but not yet generally accepted; see Taylor [2000] as well as Radtke and Schmidt

[2004]. It is useful for several reasons:

- For losses which are settled within the accident year, the delay of settlement is 0. It is therefore natural to start the enumeration of development years with 0.
- Using the enumeration of development years also for accident years implies that the incremental or cumulative loss of accident year i and development year k is observable if and only if $i+k \leq n$. In particular, the cumulative losses $S_{i,n-i}$ are those of the present calendar year n and are crucial in most methods of loss reserving.

After all, the notation used here simplifies mathematical formulas.

3. DEVELOPMENT PATTERNS

The use of run-off triangles in loss reserving can be justified only if it is assumed that the development of the losses of every accident year follows a development pattern which is common to all accident years. This vague idea of a development pattern can be formalized in various ways.

In the present section we consider three types of development patterns which are formally distinct but can easily be converted into each other. These development patterns and their equivalence provide a key to the comparison of several methods of loss reserving.

The assumption of an underlying development pattern can be viewed as a primitive stochastic model of loss reserving.

3.1 Incremental Quotas

The development pattern for incremental quotas compares the expected incremental losses with the expected ultimate cumulative losses:

Development Pattern for Incremental Quotas: *There exist parameters* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ *with* $\sum_{l=0}^{n} \vartheta_l = 1$ *such that the identity*

$$\vartheta_k = \frac{E[Z_{i,k}]}{E[S_{i,n}]}$$

holds for all $k \in \{0,1,\ldots,n\}$ *and for all* $i \in \{0,1,\ldots,n\}$.

The assumption means that, for every development year $k \in \{0,1,\ldots,n\}$, the *incremental quotas*

$$\vartheta_{i,k} = \frac{E[Z_{i,k}]}{E[S_{i,n}]}$$

are identical for all accident years.

In the case of a run-off triangle for *paid losses* or *claim counts*, it is usually reasonable to assume in addition that $\vartheta_k > 0$ holds for all $k \in \{0, 1, \ldots, n\}$. In the case of *incurred losses*, however, this additional assumption may be inappropriate since, due to conservative reserving, the expected incremental losses of development years $k \in \{1, \ldots, n\}$ may be negative.

3.2 Cumulative Quotas

The development pattern for cumulative quotas compares the expected cumulative losses with the expected ultimate cumulative losses:

Development Pattern for Cumulative Quotas: *There exist parameters* $\gamma_0, \gamma_1, \ldots, \gamma_n$ *with* $\gamma_n = 1$ *such that the identity*

$$\gamma_k = \frac{E[S_{i,k}]}{E[S_{i,n}]}$$

holds for all $k \in \{0, 1, \ldots, n\}$ *and for all* $i \in \{0, 1, \ldots, n\}$.

The assumption means that, for every development year $k \in \{0, 1, \ldots, n\}$, the *cumulative quotas*

$$\gamma_{i,k} = \frac{E[S_{i,k}]}{E[S_{i,n}]}$$

are identical for all accident years.

In the case of a run-off triangle for paid losses or claim counts, it is usually reasonable to assume in addition that $0 < \gamma_0 < \gamma_1 < \ldots < \gamma_n$. In the case of incurred losses, however, this additional assumption may be inappropriate since, due to conservative reserving, the sequence of the expected cumulative losses may be decreasing.

The development patterns for incremental and cumulative quotas can be converted into each other:

- If $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ is a development pattern for incremental losses, then a development pattern for cumulative losses is obtained by letting

$$\gamma_k := \sum_{l=0}^{k} \vartheta_l .$$

- If $\gamma_0, \gamma_1, \ldots, \gamma_n$ is a development pattern for cumulative losses, then a development pattern for incremental losses is obtained by letting

$$\vartheta_k := \begin{cases} \gamma_0 & \text{if } k=0 \\ \gamma_k - \gamma_{k-1} & \text{else.} \end{cases}$$

Furthermore, the condition $\vartheta_k > 0$ is fulfilled for all $k \in \{0, 1, \ldots, n\}$ if and only if $0 < \gamma_0 < \gamma_1 < \ldots < \gamma_n$.

3.3 Factors

The development pattern for factors compares subsequent expected cumulative losses:

Development Pattern for Factors: *There exist parameters* $\varphi_1, \ldots, \varphi_n$ *such that the identity*

$$\varphi_k = \frac{E[S_{i,k}]}{E[S_{i,k-1}]}$$

holds for all $k \in \{1, \ldots, n\}$ *and for all* $i \in \{0, 1, \ldots, n\}$.

The assumption means that, for every development year $k \in \{1, \ldots, n\}$, the *factors*

$$\varphi_{i,k} = \frac{E[S_{i,k}]}{E[S_{i,k-1}]}$$

are identical for all accident years.

In the case of a run-off triangle for paid losses or claim counts, it is usually reasonable to assume in addition that $\varphi_k > 1$ holds for all $k \in \{1, \ldots, n\}$. In the case of incurred losses, however, this additional assumption may be inappropriate since, due to conservative reserving, the sequence of the expected cumulative losses may be decreasing.

The development patterns for cumulative quotas and for factors can be converted into each other:

- If $\gamma_0, \gamma_1, \ldots, \gamma_n$ is a development pattern for cumulative losses, then a development pattern for factors is obtained by letting

$$\varphi_k := \frac{\gamma_k}{\gamma_{k-1}}.$$

- If $\varphi_1, \ldots, \varphi_n$ is a development pattern for factors, then a development pattern for cumulative losses is obtained by letting

$$\gamma_k := \prod_{l=k+1}^{n} \frac{1}{\varphi_l}$$

(such that $\gamma_n = 1$).

Furthermore, the condition $\gamma_0 < \gamma_1 < \ldots < \gamma_n$ is fulfilled if and only if $\varphi_k > 1$ holds for all $k \in \{1, \ldots, n\}$.

Combining this result and that of the previous subsection, it is evident that also the development patterns for incremental quotas and for factors can be converted into each other. We omit the corresponding formulas since they will not be needed in the sequel.

3.4 Estimation

At the first glance, there is little hope to estimate the parameters of the development patterns for incremental or cumulative quotas since the only obvious estimators of ϑ_k and γ_k are the observable quotients $Z_{0,k} / S_{0,n}$ and $S_{0,k} / S_{0,n}$, respectively.

Fortunately, the situation is quite different for the development pattern for factors: For every development year $k \in \{1, \ldots, n\}$, each of the *individual development factors*

$$\hat{\varphi}_{i,k} := \frac{S_{i,k}}{S_{i,k-1}}$$

with $i \in \{0, 1, \ldots, n-k\}$ is a reasonable estimator of φ_k, and this is also true for every weighted mean

$$\hat{\varphi}_k := \sum_{j=0}^{n-k} W_{j,k} \hat{\varphi}_{j,k}$$

with random variables (or constants) satisfying $\sum_{j=0}^{n-k} W_{j,k} = 1$. The most prominent estimator of this large family is the *chain-ladder factor*

$$\hat{\varphi}_k^{\text{CL}} := \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

which can also be written as

$$\hat{\varphi}_k^{\mathrm{CL}} = \sum_{j=0}^{n-k} \frac{S_{j,k-1}}{\sum_{h=0}^{n-k} S_{h,k-1}} \hat{\varphi}_{j,k}$$

and is used in the chain-ladder method.

Due to the correspondence between the three development patterns, it is then clear that in the same way estimators of factors can be converted into estimators of cumulative quotas and hence into estimators of incremental quotas.

3.5 Remarks

In the case of a run-off triangle for paid claims or claim counts, the intuitive cumulative interpretation of the development patterns of incremental or aggregate quotas would be their interpretation as incremental or cumulative probabilities. This interpretation is helpful, but it is not quite correct since the parameters of the development pattern are defined as *quotients of expectations* instead of *expectations of quotients* and since these quantities are in general distinct.

One may thus argue that the definitions of development patterns are inconvenient since they do not exactly correspond to intuition. In the following two sections, however, it will be shown that the definitions given here are nevertheless reasonable since they provide a powerful and unifying concept for the interpretation and the comparison of many methods and models of loss reserving.

4. METHODS

The present section provides a unifying presentation of the most important methods of loss reserving. The starting point is a general version of the Bornhuetter-Ferguson method which is closely related to the notion of a development pattern for cumulative quotas and turns out to be a unifying principle under which various other methods of loss reserving can be subsumed.

4.1 Bornhuetter-Ferguson Method

The Bornhuetter-Ferguson method is based on the assumption that there exist parameters $\alpha_0, \alpha_1, \ldots, \alpha_n$ and $\gamma_0, \gamma_1, \ldots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$E[S_{i,k}] = \gamma_k \alpha_i$$

holds for all $i, k \in \{0, 1, \ldots, n\}$. Then we have

$$E[S_{i,n}] = \alpha_i$$

and hence

$$E[S_{i,k}] = \gamma_k E[S_{i,n}]$$

such that the parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ form a development pattern for cumulative quotas.

The Bornhuetter-Ferguson method is also based on the additional assumption that *prior estimators*

$$\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_n$$

of the expected ultimate cumulative losses $E[S_{i,n}]$ and *prior estimators*

$$\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_n$$

of the development pattern are given and that $\hat{\gamma}_n = 1$.

Comment: Prior estimators may be obtained from information provided by various sources:

- *Internal information*: This is any information which is *contained* in the run-off triangle of the portfolio under consideration. Internal information could be used, e. g., by estimating the development pattern from the given run-off triangle.
- *External information*: This is any information which is *not contained* in the run-off triangle of the portfolio under consideration. External information could be obtained, e. g., from market statistics, from other portfolios which are judged to be similar to the given one, or from premiums or other volume measures of the portfolio under consideration; see Section 4.6.

Of course, prior estimators may also be obtained by combining internal and external information. In any case, the choice of prior estimators is an important decision to be made by the actuary.

The *Bornhuetter-Ferguson predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\hat{S}_{i,k}^{\mathrm{BF}} := S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\hat{\alpha}_i.$$

The definition of the Bornhuetter-Ferguson predictors reminds of the identity

$$E[S_{i,k}] = E[S_{i,n-i}] + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\hat{\alpha}_i$$

which is a consequence of the model assumption.

The definition of the Bornhuetter-Ferguson predictors shows that the prior estimators are dominant for young accident years whereas they are less important for old development years. Also, in the extreme case where the prior estimators are completely determined by external information, the major part of the run-off triangle is ignored and only the cumulative losses of the present calendar year are used. This is reasonable when the quality of the data from older calendar years is poor.

Example A. We consider the following reduced run-off triangle for cumulative losses which contains the cumulative losses of the present calendar year and is complemented by the prior estimators of the expected ultimate cumulative losses and of the development pattern:

Accident		**Development Year** k					
Year i	$\hat{\alpha}_i$	0	1	2	3	4	5
0	3517						3483
1	3981					3844	
2	4598				3977		
3	5658			3880			
4	6214		3261				
5	6325	1889					
$\hat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

Computing now the Bornhuetter-Ferguson predictors, the run-off triangle is completed as follows:

Accident		**Development Year** k					
Year i	$\hat{\alpha}_i$	0	1	2	3	4	5
0	3517						3483
1	3981					3844	*4043*
2	4598				3977	*4391*	*4621*
3	5658			3880	*4785*	*4389*	*5577*
4	6214		3261	*4442*	*5436*	*5995*	*6306*
5	6325	1889	*3344*	*4546*	*5558*	*6127*	*6443*
$\hat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

When the cumulative losses of the present calendar year are judged to be reliable, it may be desirable to modify the Bornhuetter-Ferguson predictors in order to strengthen the weight of the cumulative losses of the present calendar year and to reduce that of the prior estimators of the expected ultimate cumulative losses. This goal can be achieved by iteration.

For example, if on the right hand side of the previous formula the prior estimators $\hat{\alpha}_i$ are replaced by the Bornhuetter-Ferguson predictors $\hat{S}_{i,n}^{\text{BF}}$, then the resulting predictors are the *Benktander Hovinen predictors*

$$\hat{S}_{i,k}^{\text{BH}} := S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\,\hat{S}_{i,n}^{\text{BF}}$$

which in the case $\hat{\gamma}_{n-1} < \hat{\gamma}_k$ increase the weight of the cumulative losses of the present calendar year and reduce that of the prior estimators of the expected ultimate cumulative losses.

More generally, the *Bornhuetter-Ferguson predictors* of order $m \in \mathbb{N}_0$ are defined by letting

$$\hat{S}_{i,k}^{(m)} := \begin{cases} S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\,\hat{\alpha}_i & \text{if } m = 0 \\ S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\,\hat{S}_{i,n}^{(m-1)} & \text{if } m \geq 1. \end{cases}$$

Then we have $\hat{S}_{i,k}^{(0)} = \hat{S}_{i,k}^{\text{BF}}$ and $\hat{S}_{i,k}^{(1)} = \hat{S}_{i,k}^{\text{BH}}$, and induction yields

$$\begin{aligned} \hat{S}_{i,k}^{(m)} &= \left(1-(1-\hat{\gamma}_{n-i})^m\right)\hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}} + (1-\hat{\gamma}_{n-i})^m\,\hat{S}_{i,k}^{\text{BF}} \\ &= \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}} + (1-\hat{\gamma}_{n-i})^m \left(\hat{S}_{i,k}^{\text{BF}} - \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}\right) \\ &= \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}} + (1-\hat{\gamma}_{n-i})^m (\hat{\gamma}_k - \hat{\gamma}_{n-i}) \left(\hat{\alpha}_i - \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}\right) \end{aligned}$$

for all $m \in \mathbb{N}_0$. In the particular case where $\hat{\alpha}_i = S_{i,n-i} / \hat{\gamma}_{n-i}$ or $\hat{\gamma}_{n-i} = 1$, the iteration is without interest since in that case the identity

$$\hat{S}_{i,k}^{(m)} = \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}$$

holds for all $m \in \mathbb{N}_0$. By contrast, the iteration is of considerable interest in the case where $0 < \hat{\gamma}_{n-i} < 1$ since in that case we obtain

$$\lim_{m \to \infty} \hat{S}_{i,k}^{(m)} = \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}$$

and convergence of the sequence of the iterated Bornhuetter-Ferguson predictors is monotone but may be increasing or decreasing.

Example B. The following table contains the prior estimators of the expected ultimate cumulative losses, the iterated Bornhuetter-Ferguson predictors

$$\hat{S}_{i,n}^{(m)} = \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}(1-\hat{\gamma}_{n-i})^{m+1}\left(\hat{\alpha}_i - \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}\right)$$

and their limits:

Accident	**Prior**	**Iterated Bornhuetter-Ferguson Predictors**									**Limit**
Year i	$\hat{\alpha}_i$	$\hat{S}_{i,5}^{(0)}$	$\hat{S}_{i,5}^{(1)}$	$\hat{S}_{i,5}^{(2)}$	$\hat{S}_{i,5}^{(3)}$	$\hat{S}_{i,5}^{(4)}$	$\hat{S}_{i,5}^{(5)}$	...	$\hat{S}_{i,5}^{(10)}$	...	
0	3517	*3483*	*3483*	*3483*	*3483*	*3483*	*3483*	...	*3483*	...	*3483*
1	3981	*4043*	*4046*	*4046*	*4046*	*4046*	*4046*	...	*4046*	...	*4046*
2	4598	*4621*	*4623*	*4624*	*4624*	*4624*	*4624*	...	*4624*	...	*4624*
3	5658	*5577*	*5553*	*5546*	*5544*	*5543*	*5543*	...	*5543*	...	*5543*
4	6214	*6306*	*6351*	*6373*	*6384*	*6389*	*6392*	...	*6394*	...	*6394*
5	6325	*6443*	*6528*	*6589*	*6633*	*6664*	*6687*	...	*6730*	...	*6746*

The iteration steps 0 and 1 correspond to the Bornhuetter-Ferguson method and to the Benktander-Hovinen method, respectively. The table illustrates that convergence is monotone but may be increasing or decreasing, and that convergence is usually fast for old accident years and slow for young accident years.

4.2 Loss-development Method

The loss-development method is based on the assumption that there exist parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$E[S_{i,k}] = \gamma_k E[S_{i,n}]$$

holds for all $i,k \in \{0,1,\ldots,n\}$. Then the parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ form a development pattern for cumulative quotas.

The loss-development method is also based on the additional assumption that *prior estimators*

$$\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_n$$

of the development pattern are given and that $\hat{\gamma}_n = 1$.

The *loss-development predictors* of the cumulative losses $S_{i,k}$ with $i+k \geq n$ are defined as

$$\hat{S}_{i,k}^{\mathrm{LD}} = \hat{\gamma}_k \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}}.$$

The definition of the loss-development predictors reminds of the identity

$$E[S_{i,k}] = \gamma_k \frac{E[S_{i,n}]}{\gamma_{n-i}}$$

which is a consequence of the model assumption.

When compared with the Bornhuetter-Ferguson predictors, the importance of the cumulative losses of the present calendar year and of the prior estimators of the development pattern is increased in the loss-development predictors since the latter do not involve any prior estimators of the expected ultimate cumulative losses.

Example C. We consider the following reduced run-off triangle for cumulative losses which contains the cumulative losses of the present calendar year and is complemented by the prior estimators of the development pattern:

Accident	Development Year k					
Year i	0	1	2	3	4	5
0						3483
1					3844	
2				3977		
3			3880			
4		3261				
5	1889					
$\hat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

Computing now the loss-development predictors, the run-off triangle is completed as follows:

Accident	Development Year k					
Year i	0	1	2	3	4	5
0						3483
1					3844	*4046*
2				3977	*4393*	*4624*
3			3880	*4767*	*5266*	*5543*
4		3261	*4476*	*5499*	*6074*	*6394*
5	1889	*3440*	*4722*	*5802*	*6409*	*6746*
$\hat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

The loss-development predictors can be written as

$$\hat{S}_{i,k}^{\mathrm{LD}} = S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\hat{S}_{i,n}^{\mathrm{LD}}.$$

This shows that the loss-development predictors are nothing else than the Bornhuetter-Ferguson predictors with respect to the prior estimators

$$\hat{\alpha}_i^{\mathrm{LD}} := \hat{S}_i^{\mathrm{LD}}$$

of the expected ultimate cumulative losses. In other words, the loss-development method is a particular case of the Bornhuetter-Ferguson method with prior estimators of the expected ultimate cumulative losses which are based on internal and external information.

Moreover, in the case where $0 < \hat{\gamma}_{n-i} < 1$, the loss-development predictors are precisely the limits of the sequences of the iterated Bornhuetter-Ferguson predictors with respect to arbitrary prior estimators of the expected ultimate cumulative losses, as has been shown in Section 4.1.

4.3 Chain-ladder Method

The chain-ladder method is based on the assumption that there exist parameters $\varphi_1, \ldots, \varphi_n$ such that the identity

$$E[S_{i,k}] = \varphi_k E[S_{i,k-1}]$$

holds for all $i \in \{0, 1, \ldots, n\}$ and $k \in \{0, 1, \ldots, n\}$. Then the parameters $\varphi_1, \ldots, \varphi_n$ form a development pattern for factors.

The *chain-ladder predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\hat{S}_{i,k}^{\mathrm{CL}} := S_{i,n-i} \prod_{l=n-i+1}^{k} \hat{\varphi}_l^{\mathrm{CL}}$$

where

$$\hat{\varphi}_k^{\mathrm{CL}} := \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

is the *chain-ladder* factor introduced in Section 3. The definition of the chain-ladder predictors reminds of the identity

$$E[S_{i,k}] = E[S_{i,n-i}] \prod_{l=n-i+1}^{k} \varphi_l$$

which is a consequence of the model assumption.

When compared with the loss-development predictors, it is remarkable that the chain-ladder predictors are not determined by the cumulative losses of the present calendar year but involve, via the chain-ladder factors, *all* cumulative losses of the run-off triangle.

Example D. We consider the following run-off triangle for cumulative losses:

Accident Year i	Development Year k: 0	1	2	3	4	5
0	1001	1855	2423	2988	3335	3483
1	1113	2103	2774	3422	3844	
2	1265	2433	3233	3977		
3	1490	2873	3880			
4	1725	3261				
5	1889					

Computing first the chain-ladder factors and then the chain-ladder predictors, the run-off triangle is completed as follows:

Accident Year i	Development Year k: 0	1	2	3	4	5
0	1001	1855	2423	2988	3335	3483
1	1113	2103	2774	3422	3844	*4013*
2	1265	2433	3233	3977	*4454*	*4650*
3	1490	2873	3880	*4780*	*5354*	*5590*
4	1725	3261	*4334*	*5339*	*5980*	*6243*
5	1889	*3587*	*4767*	*5873*	*6578*	*6867*
$\hat{\varphi}_k^{CL}$		*1.899*	*1.329*	*1.232*	*1.120*	*1.044*

It has been pointed out in Section 3 that the different development patterns and their estimators can be converted into each other. In particular, letting

$$\gamma_k := \prod_{l=k+1}^{n} \frac{1}{\varphi_l}$$

converts a development pattern for factors into a development pattern for cumulative quotas and letting

$$\hat{\gamma}_k := \prod_{l=k+1}^{n} \frac{1}{\hat{\varphi}_l}$$

converts the estimators of a development pattern for factors into estimators of a development pattern for cumulative quotas. Thus, letting

$$\hat{\gamma}_k^{\text{CL}} := \prod_{l=k+1}^{n} \frac{1}{\hat{\varphi}_l^{\text{CL}}}$$

the chain-ladder predictors can be written as

$$\hat{S}_{i,k}^{\text{CL}} = \hat{\gamma}_k^{\text{CL}} \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}^{\text{CL}}}.$$

This shows that the chain-ladder predictors are nothing else than the loss-development predictors with respect to the *chain-ladder cumulative quotas* $\hat{\gamma}_k^{\text{CL}}$ as prior estimators of the cumulative quotas. Furthermore, we have

$$\hat{S}_{i,k}^{\text{CL}} = S_{i,n-i} + (\hat{\gamma}_k^{\text{CL}} - \hat{\gamma}_{n-i}^{\text{CL}})\hat{S}_{i,n}^{\text{CL}}.$$

This shows that the chain-ladder predictors are precisely the Bornhuetter-Ferguson predictors with respect to the prior estimators $\hat{\gamma}_k^{\text{CL}}$ of the cumulative quotas and the prior estimators

$$\hat{\alpha}_i^{\text{CL}} := \hat{S}_{i,n}^{\text{CL}}$$

of the expected ultimate cumulative losses. In other words, the chain-ladder method is a particular case of the loss-development method and hence of the Bornhuetter-Ferguson method with prior estimators of the development pattern and the expected ultimate cumulative losses which are completely based on internal information.

The chain-ladder method can be modified by replacing the chain-ladder factors $\hat{\varphi}_k^{\text{CL}}$ by any other estimators of the form

$$\hat{\varphi}_k := \sum_{j=0}^{n-k} W_{j,k}\hat{\varphi}_{j,k}$$

with random variables (or constants) satisfying $\sum_{j=0}^{n-k} W_{j,k} = 1$.

4.4 Grossing-up Method

The grossing-up method is based on the assumption that there exist parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$E[S_{i,k}] = \gamma_k E[S_{i,n}]$$

holds for all $i,k \in \{0,1,\ldots,n\}$. Then the parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ form a development pattern for cumulative quotas.

The *grossing-up predictors* of the cumulative losses $S_{i,k}$ with $i+k \geq n$ are defined as

$$\hat{S}_{i,k}^{\text{GU}} := \hat{\gamma}_k^{\text{GU}} \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}^{\text{GU}}}$$

where

$$\hat{\gamma}_k^{\text{GU}} := \begin{cases} 1 & \text{if } k=n \\ \dfrac{\sum_{j=0}^{n-k-1} S_{j,k}}{\sum_{j=0}^{n-k-1} \hat{S}_{j,n}^{\text{GU}}} & \text{else} \end{cases}$$

is the *grossing-up cumulative quota* of development year k. The definition of the grossing-up predictors reminds of the identity

$$E[S_{i,k}] = \gamma_k \frac{E[S_{i,n-i}]}{\gamma_{n-i}}$$

which is a consequence of the model assumption.

The computation of the grossing-up cumulative quotas and of the grossing-up predictors for the ultimate cumulative losses proceeds by recursion along the accident years, which yields

$$\begin{array}{lll} \hat{\gamma}_n^{\text{GU}} = 1 & \text{and} & \hat{S}_{0,n}^{\text{GU}} = S_{0,n} \\ \hat{\gamma}_{n-1}^{\text{GU}} = \dfrac{S_{0,n-1}}{\hat{S}_{0,n}^{\text{GU}}} & \text{and} & \hat{S}_{1,n}^{\text{GU}} = \dfrac{S_{1,n-1}}{\hat{\gamma}_{n-1}^{\text{GU}}} \\ \hat{\gamma}_{n-2}^{\text{GU}} = \dfrac{S_{0,n-2}+S_{1,n-2}}{\hat{S}_{0,n}^{\text{GU}}+\hat{S}_{1,n}^{\text{GU}}} & \text{and} & \hat{S}_{2,n}^{\text{GU}} = \dfrac{S_{2,n-2}}{\hat{\gamma}_{n-2}^{\text{GU}}} \\ \vdots & & \vdots \end{array}$$

As can be seen from the definition, the grossing-up predictors are nothing else than the loss-development predictors with respect to the grossing-up cumulative quotas $\hat{\gamma}_k^{\text{GU}}$ as prior estimators of the cumulative quotas. Furthermore, we have

$$\hat{S}_{i,k}^{\text{GU}} = S_{i,n-i} + (\hat{\gamma}_k^{\text{GU}} - \hat{\gamma}_{n-i}^{\text{GU}})\hat{S}_{i,n}^{\text{GU}}$$

which shows that the grossing-up predictors are precisely the Bornhuetter-Ferguson predictors with respect to the prior estimators $\hat{\gamma}_k^{\text{GU}}$ of the cumulative quotas and the prior estimators

$$\hat{\alpha}_i^{\mathrm{GU}} := \hat{S}_{i,n}^{\mathrm{GU}}$$

of the expected ultimate cumulative losses. In other words, the grossing-up method is a particular case of the loss-development method and hence of the Bornhuetter-Ferguson method with prior estimators of the development pattern and the expected ultimate cumulative losses which are completely based on internal information.

Since the previous remark applies as well to the chain-ladder predictors, the question arises whether there is any difference between the grossing-up predictors and the chain-ladder predictors. The answer to this question is that there is no difference at all since it can be shown that the grossing-up cumulative quotas and the chain-ladder cumulative quotas are identical for all development years; see e. g. Lorenz and Schmidt [1999].

The grossing-up method thus provides a computational alternative to the chain-ladder method, but this alternative seems to be of little practical interest if any. The reformulation of the chain-ladder method provided by the grossing-up method is, however, of considerable interest with regard to the comparison of methods:

First, among all methods for cumulative losses considered here, the chain-ladder method appears to be somewhat singular since it uses estimators of a development pattern for factors instead of cumulative quotas, but its equivalence with the grossing-up method shows that this singularity is only due to the most intelligent formulation of an algorithm which avoids recursion and is hence more easily understood.

Second, the grossing-up method provides an substantial link between the chain-ladder method and the marginal-sum method; see Subsection 4.5.

4.5 Marginal-Sum Method

The marginal-sum method is based on the assumption that there exist parameters $\alpha_0, \alpha_1, \ldots, \alpha_n$ and $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ with $\sum_{l=0}^{n} \vartheta_l = 1$ such that the identity

$$E[Z_{i,k}] = \vartheta_k \alpha_i$$

holds for all $i, k \in \{0, 1, \ldots, n\}$. Summation yields

$$E[S_{i,n}] = \alpha_i$$

and hence

$$E[Z_{i,k}] = \vartheta_k E[S_{i,n}]$$

such that the parameters $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ form a development pattern for incremental quotas.

Observable random variables $\hat{\alpha}_0^{\text{MS}}, \hat{\alpha}_1^{\text{MS}}, \ldots, \hat{\alpha}_n^{\text{MS}}$ and $\hat{\vartheta}_0^{\text{MS}}, \hat{\vartheta}_1^{\text{MS}}, \ldots, \hat{\vartheta}_n^{\text{MS}}$ are said to be *marginal-sum estimators* if they are solutions to the *marginal-sum equations*

$$\sum_{l=0}^{n-i} \hat{\alpha}_i \hat{\vartheta}_l = \sum_{l=0}^{n-i} Z_{i,l}$$

for $i \in \{0, 1, \ldots, n\}$ and

$$\sum_{j=0}^{n-k} \hat{\alpha}_j \hat{\vartheta}_k = \sum_{j=0}^{n-k} Z_{j,k}$$

for $k \in \{0, 1, \ldots, n\}$ as well as

$$\sum_{l=0}^{n} \hat{\vartheta}_l = 1.$$

The marginal-sum equations remind of the identities

$$\sum_{l=0}^{n-i} \alpha_i \vartheta_l = \sum_{l=0}^{n-i} E[Z_{i,l}]$$

and

$$\sum_{j=0}^{n-k} \alpha_j \vartheta_k = \sum_{j=0}^{n-k} E[Z_{j,k}]$$

as well as

$$\sum_{k=0}^{n} \hat{\vartheta}_k = 1$$

which are immediate from the model assumptions.

The question arises whether marginal-sum estimators exist and are unique. The answer to this question is affirmative: Marginal-sum estimators exist and are unique, and they satisfy

$$\hat{\alpha}_i^{\text{MS}} = \hat{S}_{i,n}^{\text{GU}}$$

and

$$\hat{\vartheta}_k^{\text{MS}} = \begin{cases} \hat{\gamma}_0^{\text{GU}} & \text{if } k = 0 \\ \hat{\gamma}_k^{\text{GU}} - \hat{\gamma}_{k-1}^{\text{GU}} & \text{if } k \geq 1. \end{cases}$$

In view of the discussion of the grossing-up method, the previous identities imply that the marginal-sum estimators satisfy

$$\hat{\alpha}_i^{\text{MS}} = \hat{S}_{i,n}^{\text{CL}}$$

and

$$\hat{\vartheta}_k^{\text{MS}} = \begin{cases} \hat{\gamma}_0^{\text{CL}} & \text{if } k = 0 \\ \hat{\gamma}_k^{\text{CL}} - \hat{\gamma}_{k-1}^{\text{CL}} & \text{if } k \geq 1. \end{cases}$$

Thus, letting

$$\hat{\gamma}_k^{\text{MS}} := \sum_{l=0}^{k} \hat{\vartheta}_l^{\text{MS}}$$

we obtain

$$\hat{\gamma}_k^{\text{MS}} = \hat{\gamma}_k^{\text{CL}}$$

for all $k \in \{0, 1, \dots, n\}$.

The *marginal-sum predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\hat{S}_k^{\text{MS}} := \hat{\gamma}_k^{\text{MS}} \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}^{\text{MS}}}.$$

Then we have

$$\hat{S}_{i,k}^{\text{MS}} := \hat{S}_{i,k}^{\text{CL}}.$$

This shows that the marginal-sum method is equivalent to the chain-ladder method.

4.6 Cape-Cod Method

The Cape-Cod method is based on the assumption that there exist parameters $\gamma_0, \gamma_1, \dots, \gamma_n$ with $\gamma_n = 1$ such that the identity

$$E[S_{i,k}] = \gamma_k E[S_{i,n}]$$

holds for all $i, k \in \{0, 1, \dots, n\}$. Then the parameters $\gamma_0, \gamma_1, \dots, \gamma_n$ form a development pattern for cumulative quotas.

The Cape-Cod method is also based on the additional assumption that *premiums* or other *volume measures* $\pi_0, \pi_1, \dots, \pi_n \in (0, \infty)$ of the accident years are known, that the *expected* *ultimate cumulative loss ratios*

$$\kappa_i := E\left[\frac{S_{i,n}}{\pi_i}\right]$$

are identical for all accident years, and that *prior estimators* $\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_n$ of the development pattern are given and satisfy $\hat{\gamma}_n = 1$.

The *Cape-Cod predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\hat{S}_{i,k}^{\text{CC}} := S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\,\pi_i \hat{\kappa}^{\text{CC}}$$

where

$$\hat{\kappa}^{\text{CC}} := \frac{\sum_{j=0}^{n} S_{j,n-j}}{\sum_{j=0}^{n} \hat{\gamma}_{n-j} \pi_j}$$

is the *Cape-Cod loss ratio*, which is an estimator of the expected ultimate cumulative loss ratio (common to all accident years).

The Cape-Cod predictors are nothing else than the Bornhuetter-Ferguson predictors with respect to the prior estimators

$$\hat{\alpha}_i^{\text{CC}} := \pi_i \hat{\kappa}^{\text{CC}}$$

of the expected ultimate cumulative losses. In other words, the Cape-Cod method is a particular case of the Bornhuetter-Ferguson method with prior estimators of the expected ultimate cumulative losses which are based on both internal and external information.

Example E. We consider the following reduced run-off triangle for cumulative losses which contains the cumulative losses of the present calendar year and is complemented by the premiums and the prior estimators of the development pattern:

Accident		**Development Year** k					
Year i	π_i	0	1	2	3	4	5
0	4025						3483
1	4456					3844	
2	5315				3977		
3	5986			3880			
4	6939		4261				
5	8158	1889					
$\hat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

The previous triangle differs from those considered before since the value of $S_{4,1}$ is 4261 instead of 3261, which indicates that there might be an outlier in accident year 4. Using the table

i	$S_{i,5-i}$	$\hat{\gamma}_{5-i}$	π_i	$\hat{\gamma}_{5-i}\,\pi_i$
0	3483	1.000	4025	*4025*
1	3844	0.950	4456	*4233*
2	3977	0.860	5315	*4571*
3	3880	0.700	5986	*4190*
4	4261	0.510	6939	*3539*
5	1889	0.280	8158	*2284*
$\sum$	21334			*22842*

we obtain $\hat{\kappa}^{CC} = 0.934$. Computing now the prior estimators of the expected ultimate cumulative losses and the Cape-Cod predictors, the run-off triangle is completed as follows:

Accident		**Development Year k**					
Year i	$\hat{\alpha}_i$	0	1	2	3	4	5
0	*3758*						3483
1	*4162*					3844	*4052*
2	*4964*				3977	*4424*	*4672*
3	*5591*			3880	*4775*	*5278*	*5557*
4	*6481*		4261	*5492*	*6529*	*7113*	*7437*
5	*7619*	1889	*3641*	*5089*	*6308*	*6994*	*7375*
$\hat{\gamma}_k$		0.280	0.510	0.700	0.860	0.950	1.000

The previous table should be compared with the following one which is the same run-off triangle completed with the loss-development predictors:

Accident	**Development Year k**					
Year i	0	1	2	3	4	5
0						3483
1					3844	*4046*
2				3977	*4393*	*4624*
3			3880	*4767*	*5266*	*5543*
4		4261	*5848*	*7185*	*7937*	*8355*
5	1889	*3440*	*4722*	*5802*	*6409*	*6746*
$\hat{\gamma}_k$	0.280	0.510	0.700	0.860	0.950	1.000

The example indicates that the development of the Cape-Cod predictors over the accident years is much smoother than the development of the loss-development predictors which means that the Cape-Cod method reduces outlier effects. The smoothing effect is of course due to and depends on the premiums or other volume measures which are used instead.

The following considerations may help to understand the smoothing effect of the Cape-Cod method: Assume that, for every accident year i, the expected ultimate cumulative loss ratio is estimated by

$$\hat{\kappa}_i := \frac{\hat{S}_{i,n}^{\mathrm{LD}}}{\pi_i} = \frac{S_{i,n-i}}{\hat{\gamma}_{n-i}\,\pi_i}.$$

Then the Cape-Cod loss ratio can be written as a weighted mean

$$\hat{\kappa}^{\mathrm{CC}} = \frac{\sum_{j=0}^{n} S_{j,n-j}}{\sum_{j=0}^{n} \hat{\gamma}_{n-j}\,\pi_j} = \sum_{j=0}^{n} \frac{\hat{\gamma}_{n-j}\,\pi_j}{\sum_{h=0}^{n} \hat{\gamma}_{n-h}\,\pi_h}\,\hat{\kappa}_j$$

and the identity

$$S_{i,n-i} = \hat{\gamma}_{n-i}\,\pi_i\,\hat{\kappa}_i$$

suggests to decompose the cumulative loss $S_{i,n-i}$ of the present calendar year into its *regular part*

$$T_{i,n-i} := \hat{\gamma}_{n-i}\,\pi_i\,\hat{\kappa}^{\mathrm{CC}}$$

and its *outlier effect*

$$X_{i,n-i} := S_{i,n-i} - T_{i,n-i}$$

and then to apply the loss-development method to the regular part while keeping the outlier effect fixed over all subsequent development years. Since

$$\begin{aligned}
\hat{T}_{i,k}^{\mathrm{LD}} + X_{i,n-i} &= \hat{\gamma}_k \frac{T_{i,n-i}}{\hat{\gamma}_{n-i}} + (S_{i,n-i} - T_{i,n-i}) \\
&= S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\frac{T_{i,n-i}}{\hat{\gamma}_{n-i}} \\
&= S_{i,n-i} + (\hat{\gamma}_k - \hat{\gamma}_{n-i})\,\pi_i\,\hat{\kappa}^{\mathrm{CC}} \\
&= \hat{S}_{i,k}^{\mathrm{CC}}
\end{aligned}$$

we see that the resulting predictors are precisely the Cape-Cod predictors.

The Cape-Cod method can be modified by replacing the Cape-Cod loss ratio $\hat{\kappa}^{CC}$ by any other estimator of the form

$$\hat{\kappa} = \sum_{j=0}^{n} W_j \, \hat{\kappa}_j$$

with random variables (or constants) satisfying $\sum_{j=0}^{n} W_j = 1$.

4.7 Additive Method

The additive method is based on the assumption that there exist known parameters $\pi_0, \pi_1, \ldots, \pi_n \in (0,\infty)$ and unknown parameters $\zeta_0, \zeta_1, \ldots, \zeta_n$ such that the identity

$$E[Z_{i,k}] = \zeta_k \pi_i$$

holds for all $i,k \in \{0, 1, \ldots, n\}$.

If the parameters $\pi_0, \pi_1, \ldots, \pi_n$ are interpreted as *premiums* or other *volume* measures of the accident years, then the assumption means that, for every development year k, the *expected incremental loss ratios*

$$\zeta_{i,k} := E\left[\frac{Z_{i,k}}{\pi_i}\right]$$

are identical for all accident years. Letting

$$\alpha_i := \pi_i \sum_{k=0}^{n} \zeta_k$$

and

$$\gamma_k := \frac{\sum_{l=0}^{k} \zeta_l}{\sum_{l=0}^{n} \zeta_l}$$

we obtain

$$E[S_{i,k}] = \gamma_k \alpha_i$$

for all $i,k \in \{0, 1, \ldots, n\}$ such that $\alpha_i = E[S_{i,n}]$ and the parameters $\gamma_0, \gamma_1, \ldots, \gamma_n$ form a development pattern for cumulative quotas.

The *additive predictors* of the incremental losses $Z_{i,k}$ with $i + k \geq n$ are defined as

$$\hat{Z}_{i,k}^{AD} := \hat{\zeta}_k^{AD} \pi_i$$

and the *additive predictors* of the cumulative losses $S_{i,k}$ with $i + k \geq n$ are defined as

$$\hat{S}_{i,k}^{\text{AD}} := S_{i,n-i} + \sum_{l=n-i+1}^{k} \hat{Z}_{i,l}^{\text{AD}}$$

where

$$\hat{\zeta}_k^{\text{AD}} := \frac{\sum_{j=0}^{n-k} Z_{i,k}}{\sum_{j=0}^{n-k} \pi_j}$$

is the *additive incremental loss ratio* of development year k.

Example F. We consider the following run-off triangle for cumulative losses which is complemented by the premiums:

Accident		Development Year k					
Year i	π_i	0	1	2	3	4	5
0	4025	1001	1855	2423	2988	3335	3483
1	4456	1113	2103	2774	3422	3844	
2	5315	1265	2433	3233	3977		
3	5986	1490	2873	3880			
4	6939	1725	3261				
5	8158	1889					

We thus obtain the following run-off triangle for incremental losses which is complemented by the additive incremental loss ratios:

Accident		Development Year k					
Year i	π_i	0	1	2	3	4	5
0	4025	1001	854	568	565	347	148
1	4456	1113	990	671	648	422	
2	5315	1265	1168	800	744		
3	5986	1490	1383	1007			
4	6939	1725	1536				
5	8158	1889					
$\hat{\zeta}_k$		*0.243*	*0.222*	*0.154*	*0.142*	*0.091*	*0.037*

Computing now the additive predictors of the non-observable incremental losses, the run-off triangle of incremental losses is completed as follows:

Accident		Development Year k					
Year i	π_i	0	1	2	3	4	5
0	4025	1001	854	568	565	347	148
1	4456	1113	990	671	648	422	*165*
2	5315	1265	1168	800	744	*484*	*197*
3	5986	1490	1383	1007	*850*	*545*	*221*
4	6939	1725	1536	*1069*	*985*	*631*	*257*
5	8158	1889	*1811*	*1256*	*1158*	*742*	*302*
$\hat{\zeta}_k$		*0.243*	*0.222*	*0.154*	*0.142*	*0.091*	*0.037*

Accordingly, the run-off triangle of cumulative losses is completed as follows:

Accident		Development Year k					
Year i	π_i	0	1	2	3	4	5
0	4025	1001	1855	2423	2988	3335	3483
1	4456	1113	2103	2774	3422	3844	*4009*
2	5315	1265	2433	3233	3977	*4461*	*4658*
3	5986	1490	2873	3880	*4730*	*5275*	*5496*
4	6939	1725	3261	*4330*	*5315*	*5946*	*6203*
5	8158	1889	*3700*	*4956*	*6114*	*6856*	*7158*

Letting

$$\hat{\gamma}_k^{AD} := \frac{\sum_{l=0}^{k} \hat{\zeta}_l^{AD}}{\sum_{l=0}^{n} \hat{\zeta}_l^{AD}}$$

and

$$\hat{\alpha}_i^{AD} := \pi_i \sum_{l=0}^{n} \hat{\zeta}_l^{AD}$$

the additive predictors of the non-observable cumulative losses may be written as

$$\hat{S}_{i,k}^{AD} := S_{i,n-i} + (\hat{\gamma}_k^{AD} - \hat{\gamma}_{n-i}^{AD})\,\hat{\alpha}_i^{AD}.$$

This shows that the additive predictors of the cumulative losses are nothing else than the Bornhuetter-Ferguson predictors with respect to the *additive cumulative quotas* $\hat{\gamma}_k^{AD}$ and the prior estimators $\hat{\alpha}_i^{AD}$ of the expected ultimate cumulative losses. In other words, the additive method is a particular case of the Bornhuetter-Ferguson method with prior estimators of the cumulative quotas and of the expected ultimate cumulative losses which are based on both internal and external information.

The *expected cumulative loss ratios*

$$\kappa_i := E\left[\frac{S_{i,n}}{\pi_i}\right]$$

satisfy

$$\kappa_i = \sum_{l=0}^{n} \zeta_{i,l}.$$

Since the expected incremental loss ratios are identical for all accident years, it follows that also the expected cumulative loss ratios are identical for all accident years. Therefore, the *additive loss ratio*

$$\hat{\kappa}^{\mathrm{AD}} := \sum_{l=0}^{n} \hat{\zeta}_l^{\mathrm{AD}}$$

can be interpreted as an estimator of the expected ultimate cumulative loss ratio

$$\kappa = \sum_{l=0}^{n} \zeta_l$$

common to all accident years. Moreover, the prior estimators $\hat{\alpha}_i^{\mathrm{AD}}$ can be written as

$$\hat{\alpha}_i^{\mathrm{AD}} := \pi_i \, \hat{\kappa}^{\mathrm{AD}}$$

and it can be shown that

$$\hat{\kappa}^{\mathrm{AD}} = \frac{\sum_{j=0}^{n} S_{i,n-j}}{\sum_{j=0}^{n} \hat{\gamma}_{n-j}^{\mathrm{AD}} \pi_j}.$$

This shows that the additive predictors of the non-observable cumulative losses are nothing else than the Cape-Cod predictors with respect to the additive cumulative quotas $\hat{\gamma}_k^{\mathrm{AD}}$. In other words, the additive method is a particular case of the Cape-Cod method with prior estimators of the cumulative quotas which are based on both internal and external information.

The observation that the additive method is a special case of the Cape-Cod method is due to Zocher [2005].

4.8 Remarks

The following table compares the different methods of loss reserving considered in this section with regard to the choices of the prior estimators of the expected ultimate cumulative losses α_i and of the cumulative quotas γ_k:

Expected Ultimate	**Cumulative Quotas**		
Cumulative Losses	Arbitrary	$\hat{\gamma}_k^{CL}$	$\hat{\gamma}_k^{AD}$
Arbitrary	Bornhuetter-Ferguson Method		
$\hat{S}_{i,n}^{LD}$	Loss-Development Method	Chain-Ladder Method	
$\pi, \hat{\kappa}^{CC}$	Cape-Cod Method		Additive Method

Note that the prior estimators $\hat{S}_{i,n}^{LD}$ and $\pi_i \hat{\kappa}^{CC}$ depend on the choice of the prior estimators $\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_n$.

Of course, the four other combinations which apparently have not been given a name in the literature could be used as well, and even other choices of the prior estimators of the expected ultimate cumulative losses and of the cumulative quotas could be considered.

The discussion of the present section and, in particular, the above table shows that the Bornhuetter-Ferguson method provides a general principle under which several methods of loss reserving can be subsumed. The focus

- on prior estimators of the expected ultimate cumulative losses and
- on prior estimators of the cumulative quotas

provides a large variability of loss reserving methods. The above table contains important special cases but could certainly be enlarged. Moreover,

- any convex combination of prior estimators of the expected ultimate cumulative losses yields new prior estimators of the expected ultimate cumulative losses, and
- any convex combination of prior estimators of the development pattern for cumulative quotas yields new prior estimators of the development pattern.

This point is made precise in the following example:

Example G. Let $\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_n$ be prior estimators of $\alpha_0, \alpha_1, \ldots, \alpha_n$ and let

$\hat{\gamma}_0, \hat{\gamma}_1, \ldots, \hat{\gamma}_n$ be prior estimators of $\gamma_0, \gamma_1, \ldots, \gamma_n$ such that each of these prior estimators is completely based on external information. Then the prior estimators

$$\tilde{\alpha}_i := a_1\hat{\alpha}_i + a_2\hat{S}_{i,n}^{\mathrm{LD}} + a_3(\pi_i\hat{\kappa}^{\mathrm{CC}})$$

with $a_1 + a_2 + a_3 = 1$ and

$$\tilde{\gamma}_k := b_1\hat{\gamma}_k + b_2\hat{\gamma}_k^{\mathrm{CL}} + b_3\hat{\gamma}_k^{\mathrm{AD}}$$

with $b_1 + b_2 + b_3 = 1$ are prior estimators of $\alpha_0, \alpha_1, \ldots, \alpha_n$ and $\gamma_0, \gamma_1, \ldots, \gamma_n$, respectively, which through the weights a_1, a_2, a_3 and b_1, b_2, b_3 express the reliability attributed to the prior estimators $\hat{\alpha}_i$, $\hat{S}_{i,n}^{\mathrm{LD}}$, $\pi_i\hat{\kappa}^{\mathrm{CC}}$ and $\hat{\gamma}_k$, $\hat{\gamma}_k^{\mathrm{CL}}$, $\hat{\gamma}_k^{\mathrm{AD}}$, respectively.

5. LEAST-SQUARES PREDICTION

Least-squares prediction is one of the general principles of statistical inference. It is similar to least-squares estimation but differs from the latter since the target quantity is a non-observable random variable instead of a model parameter.

The main aspects of least-squares prediction are credibility prediction and Gauss-Markov prediction; in either case, the problem is to determine optimal predictors with respect to the expected squared prediction error.

An extension of Gauss-Markov prediction is conditional Gauss-Markov prediction in which unconditional first and second order moments are replaced by conditional moments.

5.1 Credibility Prediction

In the context of loss reserving, credibility prediction aims at predicting any linear combination T of (observable or non-observable) incremental losses by a predictor of the form

$$\hat{T} = a + \sum_{j=0}^{n}\sum_{l=0}^{n-j} a_{j,l}\, Z_{j,l}.$$

These predictors are said to be *admissible*. Note that

- the class of all admissible predictors does not depend on the sum to be predicted,
- the admissible predictors are not necessarily linear in the observable incremental losses since the coefficient a may be distinct from 0, and

– the admissible predictors are not assumed to be unbiased.

The general form of the prediction problem is reasonable since it includes, e. g., prediction of the ultimate cumulative losses $S_{i,n}$ which are sums of the observable incremental losses $Z_{i,0}, Z_{i,1}, \ldots, Z_{i,n-i}$ and the non-observable incremental losses $Z_{i,n-i+1}, \ldots, Z_{i,n}$.

For a sum T of incremental losses, an admissible predictor is said to be a *credibility predictor* of T if it minimizes the *expected squared prediction error*

$$E[(\hat{T} - T)^2]$$

over all admissible predictors $\hat{T}$.

The following results are well-known:

(1) For every sum T of incremental losses, there exists a credibility predictor $\hat{T}^{\mathrm{CR}}$ and the credibility predictor is unique.

(2) If T_1 and T_2 are sums of incremental losses and if c_1 and c_2 are real numbers, then the credibility predictor of

$$T := c_1 T_1 + c_2 T_2$$

satisfies

$$\hat{T}^{\mathrm{CR}} = c_1 \hat{T}_1^{\mathrm{CR}} + c_2 \hat{T}_2^{\mathrm{CR}}$$

which means that credibility prediction is linear.

(3) If T is a sum of incremental losses, then an admissible predictor $\hat{T}^{*}$ is the credibility predictor of T if and only if it satisfies the *normal equations*

$$E[\hat{T}^{*}] = E[T]$$

and

$$E[\hat{T}^{*} Z_{j,l}] = E[T\, Z_{j,l}]$$

for all $j, l \in \{0, 1, \ldots, n\}$ such that $j + l \leq n$.

(4) The credibility predictor of any sum of incremental losses is unbiased.

Because of (2) it is sufficient to determine the credibility predictors of the incremental losses $Z_{i,k}$. In the case where $i + k \leq n$, we have

$$\hat{Z}_{i,k}^{\mathrm{CR}} = Z_{i,k}.$$

In the case where $i+k \geq n+1$, we write

$$\hat{Z}_{i,k}^{\text{CR}} = a_{i,k} + \sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} Z_{h,m}$$

and determine the coefficients from the *normal equations*

$$E\left[a_{i,k} + \sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} Z_{h,m} \right] = E[Z_{i,k}]$$

and

$$E\left[\left(a_{i,k} + \sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} Z_{h,m} \right) Z_{j,l} \right] = E[Z_{i,k} Z_{j,l}]$$

which may equivalently be written as

$$a_{i,k} + \sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} E[Z_{h,m}] = E[Z_{i,k}]$$

and

$$\sum_{h=0}^{n} \sum_{m=0}^{n-h} a_{i,k,h,m} \operatorname{cov}[Z_{h,m}, Z_{j,l}] = \operatorname{cov}[Z_{i,k} Z_{j,l}]$$

for all $j,l \in \{0,1,\ldots,n\}$ such that $j+l \leq n$.

We thus see that the credibility predictor of a non-observable incremental loss is completely determined by the first and second order moments of the incremental losses. Solving the normal equations proceeds in two steps:

- The normal equations involving covariances form a system of linear equations for the coefficients $a_{i,k,h,m}$. The fact that a credibility predictor of $Z_{i,k}$ exists implies that this system of linear equations has at least one solution.

- Inserting any such solution into the normal equation involving expectations yields the coefficient $a_{i,k}$.

It should be noted that the system of linear equations may have several solutions (which is the case if and only if the covariance matrix of the observable cumulative losses is singular). This means that the credibility predictor of $Z_{i,k}$, which is known to be unique, can be *represented* in several ways.

In most credibility models for loss reserving which have been considered in the literature, it is assumed that any two incremental losses from different accident years are uncorrelated. In this case, the credibility predictor of a non-observable incremental loss $Z_{i,k}$ can be written as

$$\hat{Z}_{i,k}^{\text{CR}} = a_{i,k} + \sum_{m=0}^{n-i} a_{i,k,i,m} Z_{i,m}$$

and its coefficients can be determined from the reduced normal equations

$$a_{i,k} + \sum_{m=0}^{n-i} a_{i,k,b,m} E[Z_{i,m}] = E[Z_{i,k}]$$

and

$$\sum_{m=0}^{n-i} a_{i,k,i,m} \text{cov}[Z_{i,m}, Z_{j,l}] = \text{cov}[Z_{i,k} Z_{j,l}]$$

for all $l \in \{0, 1, \ldots, n-i\}$.

As an example, let us now consider credibility prediction in the credibility model of Witting, which is a model for claim counts:

Credibility Model of Witting:

(i) *Any two incremental losses of different accident years are uncorrelated.*

(ii) *There exist parameters* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n \in (0,1)$ *with* $\sum_{l=0}^{n} \vartheta_l = 1$ *such that, for every accident year* $i \in \{0, 1, \ldots, n\}$, *the conditional joint distribution of the family* $\{Z_{i,k}\}_{k \in \{0,1,\ldots,n\}}$ *with respect to the ultimate cumulative loss* $S_{i,n}$ *is the multinomial distribution with parameters* $S_{i,n}$ *and* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$.

For the remainder of this subsection we assume that the assumptions of the credibility model of Witting are fulfilled. Then we have

$$E(Z_{i,k} \mid S_{i,n}) = S_{i,n} \vartheta_k$$

$$\text{cov}(Z_{i,k}, Z_{i,l} \mid S_{i,n}) = \begin{cases} -S_{i,n} \vartheta_k^2 + S_{i,n} \vartheta_k & \text{if } k = l \\ -S_{i,n} \vartheta_k \vartheta_l & \text{else.} \end{cases}$$

Letting

$$\alpha_i := E[S_{i,n}]$$
$$\sigma_i := \text{var}[S_{i,n}]$$

we obtain

$$E[Z_{i,k}] = \alpha_i \vartheta_k$$
$$\operatorname{cov}[Z_{i,k}, Z_{i,l}] = \begin{cases} (\sigma_i - \alpha_i)\vartheta_k^2 + \alpha_i \vartheta_k & \text{if } k = l \\ (\sigma_i - \alpha_i)\vartheta_k \vartheta_l & \text{else.} \end{cases}$$

The first of the previous identities shows that the parameters $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ form a development pattern for incremental quotas. Inserting the previous identities into the normal equations, we obtain, for all $i, k \in \{0, 1, \ldots, n\}$ such that $i + k \geq n + 1$,

$$\hat{Z}_{i,k}^{\mathrm{CR}} = \vartheta_k \left(\frac{1}{1 + \gamma_{n-i}\tau_i} \alpha_i + \frac{\gamma_{n-i}\tau_i}{1 + \gamma_{n-i}\tau_i} \frac{S_{i,n-i}}{\gamma_{n-i}} \right)$$

and hence

$$\begin{aligned} \hat{S}_{i,k}^{\mathrm{CR}} &= S_{i,n-i} + \sum_{l=n-i+1}^{k} \hat{Z}_{i,k}^{\mathrm{CR}} \\ &= S_{i,n-i} + (\gamma_k - \gamma_{n-i}) \left(\frac{1}{1 + \gamma_{n-i}\tau_i} \alpha_i + \frac{\gamma_{n-i}\tau_i}{1 + \gamma_{n-i}\tau_i} \frac{S_{i,n-i}}{\gamma_{n-i}} \right) \end{aligned}$$

where $\gamma_k = \sum_{l=0}^{k} \vartheta_l$ and $\tau_i := (\sigma_i - \alpha_i) / \alpha_i$. This shows that the credibility predictor of the non-observable cumulative loss $S_{i,k}$ is the Bornhuetter-Ferguson predictor with respect to the prior estimators

$$\hat{\gamma}_k := \gamma_k$$

of the development pattern for cumulative quotas and the prior estimators

$$\hat{\alpha}_i^{\mathrm{CR}} := \frac{1}{1 + \gamma_{n-i}\tau_i} \alpha_i + \frac{\gamma_{n-i}\tau_i}{1 + \gamma_{n-i}\tau_i} \frac{S_{i,n-i}}{\gamma_{n-i}}$$

of the expected ultimate cumulative losses, which are weighted means of external information provided by the unknown parameter α_i and internal information provided by the loss-development predictor $\hat{S}_{i,n}^{\mathrm{LD}} = S_{i,n-i} / \gamma_{n-i}$.

Example H. If, in addition to the assumptions of the model of Witting, it is assumed that every ultimate cumulative loss $S_{i,n}$ has the Poisson distribution with expectation α_i, then we have $\tau_i = 0$ and the credibility predictors of every non-observable cumulative loss $S_{i,k}$ satisfy

$$\hat{S}_{i,k}^{\mathrm{CR}} = S_{i,n-i} + (\gamma_k - \gamma_{n-i}) \alpha_i$$

and are thus identical with the Bornhuetter-Ferguson estimators with respect to the prior estimators $\hat{\gamma}_k := \gamma_k$ and $\hat{\alpha}_i := \alpha_i$. In this case, the assumptions of the Poisson model are fulfilled and maximum-likelihood estimation could be used as an alternative to credibility prediction; see subsection 6.1 below.

Similar results obtain in the credibility model of Mack [1990] and in a special case of the credibility model of Hesselager and Witting [1998]; see Radtke and Schmidt [2004].

5.2 Gauss-Markov Prediction

A predictor $\hat{T}$ of a linear combination T of (observable or non-observable) incremental losses is said to be

- a *linear predictor* if there exists a family $\{a_{j,l}\}_{j,l\in\{0,1,\dots,n\},\, l+j\le n}$ of coefficients such that

$$\hat{T} = \sum_{j=0}^{n}\sum_{l=0}^{n-j} a_{j,l} Z_{j,l}$$

- an *unbiased predictor* of T if

$$E[\hat{T}] = E[T]$$

- a *Gauss-Markov predictor* of T if it is an unbiased linear predictor of T which minimizes the *expected squared prediction error*

$$E[(\hat{T}-T)^2]$$

  over all unbiased linear predictors $\hat{T}$ of T.

The existence of a Gauss-Markov predictor of T cannot be guaranteed in general. (For example, if $E[Z_{i,k}] = 0$ holds for every observable every incremental loss and if T is such that $E[T] \neq 0$, then there exists no unbiased linear estimator of T.) Therefore, we consider Gauss-Markov prediction only under the assumptions of the linear model.

Let $\boldsymbol{Z}_1$ denote a random vector consisting of the observable incremental losses and let $\boldsymbol{Z}_2$ denote a random vector consisting of the non-observable incremental losses (arranged in any order).

Linear Model:

(i) *There exist matrices* $\boldsymbol{A}_1$ *and* $\boldsymbol{A}_2$ *and a vector* $\boldsymbol{\beta}$ *such that*

$$E[\boldsymbol{Z}_1] = \boldsymbol{A}_1\boldsymbol{\beta}$$
$$E[\boldsymbol{Z}_2] = \boldsymbol{A}_2\boldsymbol{\beta}$$

(ii) *The matrix* A_1 *has full column rank.*

(iii) *The matrix*

$$\boldsymbol{\Sigma}_{11} := \text{var}[\boldsymbol{Z}_1]$$

is invertible.

For the remainder of this subsection, we assume that the assumptions of the linear model are fulfilled.

Under the assumptions of the linear model, the following results are well-known:

(1) For every sum T of incremental losses, there exists a Gauss-Markov predictor $\hat{T}^{\text{GM}}$ and the Gauss-Markov predictor is unique.

(2) If T_1 and T_2 are sums of incremental losses and if c_1 and c_2 are real numbers, then the Gauss-Markov predictor of

$$T := c_1 T_1 + c_2 T_2$$

satisfies

$$\hat{T}^{\text{GM}} = c_1 \hat{T}_1^{\text{GM}} + c_2 \hat{T}_2^{\text{GM}}$$

which means that Gauss-Markov prediction is linear.

Because of (2) it is sufficient to determine the Gauss-Markov predictors of the incremental losses $Z_{i,k}$. In the case where $i+k \leq n$, we have

$$\hat{Z}_{i,k}^{\text{GM}} = Z_{i,k}.$$

In the case where $i+k \geq n+1$, we obtain

$$\hat{Z}_{i,k}^{\text{GM}} = \boldsymbol{a}'_{i,k}\hat{\boldsymbol{\beta}}^{\text{GM}} + \text{cov}[Z_{i,k}, \boldsymbol{Z}_1]\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{Z}_1 - \boldsymbol{A}_1\hat{\boldsymbol{\beta}}^{\text{GM}})$$

where $\boldsymbol{a}'_{i,k}$ is the row vector of the matrix $\boldsymbol{A}_2$ satisfying $E[Z_{i,k}] = \boldsymbol{a}'_{i,k}\boldsymbol{\beta}$,

$$\hat{\boldsymbol{\beta}}^{\text{GM}} := (\boldsymbol{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{A}_1)^{-1}\boldsymbol{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{Z}_1$$

is the *Gauss-Markov estimator* of $\boldsymbol{\beta}$ (based on the observable incremental losses) and $\text{cov}[Z_{i,k}, \boldsymbol{Z}_1]$ is the row vector with entries $\text{cov}[Z_{i,k}, Z_{j,l}]$ with $j,l \in \{0,1,\ldots,n\}$ and $j+l \leq n$; see Goldberger [1962] and Rao and Toutenburg [1995] as well as Halliwell [1996, 1999], Hamer [1999] and Schmidt [1998, 1999a, 2004].

As an example, let us now consider Gauss-Markov prediction in the linear model of Mack:

Linear Model of Mack: *There exist parameters* $\pi_0, \pi_1, \ldots, \pi_n \in (0,\infty)$ *and* $\zeta_0, \zeta_1, \ldots, \zeta_n$ *as well as* $\sigma_0, \sigma_1, \ldots, \sigma_n \in (0,\infty)$ *such that*

$$E[Z_{i,k}] = \pi_i \zeta_k$$

and

$$\operatorname{cov}[Z_{i,k}, Z_{j,l}] = \begin{cases} \pi_i \sigma_k^2 & \textit{if } i = j \textit{ and } k = l \\ 0 & \textit{else} \end{cases}$$

holds for all $i, j, k, l \in \{0, 1, \ldots, n\}$.

For the remainder of this subsection we assume that the assumptions of the linear model of Mack are fulfilled. Define

$$\boldsymbol{\beta} := \begin{pmatrix} \zeta_0 \\ \zeta_1 \\ \vdots \\ \zeta_n \end{pmatrix}$$

and, for all $i, k \in \{0, 1, \ldots, n\}$,

$$a'_{i,k} := (0 \quad \ldots \quad 0 \quad \pi_i \quad 0 \quad \ldots \quad 0)$$

where π_i occurs in position $1 + k$. This shows that the linear model of Mack satisfies indeed the assumptions of the linear model. For the Gauss-Markov estimator of $\boldsymbol{\beta}$ we obtain

$$\hat{\boldsymbol{\beta}}^{\mathrm{GM}} = \begin{pmatrix} \dfrac{\sum_{j=0}^{n} Z_{j,0}}{\sum_{j=0}^{n} \pi_j} \\ \dfrac{\sum_{j=0}^{n} Z_{j,1}}{\sum_{j=0}^{n} \pi_j} \\ \vdots \\ \dfrac{Z_{0,n}}{\pi_0} \end{pmatrix}.$$

Since $\operatorname{cov}[Z_{i,k}, Z_{j,l}] = 0$ holds for all $i, j, k, l \in \{0, 1, \ldots, n\}$ such that $i + k \geq n + 1$ and $j + l \leq n$, it follows that the Gauss-Markov predictor of the non-observable incremental loss $Z_{i,k}$ satisfies

$$\hat{Z}_{i,k}^{\text{GM}} = \pi_i \frac{\sum_{j=0}^{n-k} Z_{j,k}}{\sum_{j=0}^{n-k} \pi_j}$$

and hence

$$\hat{Z}_{i,k}^{\text{GM}} = \hat{Z}_{i,k}^{\text{AD}}$$

and linearity of Gauss-Markov prediction yields

$$\hat{S}_{i,k}^{\text{GM}} = \hat{S}_{i,k}^{\text{AD}}.$$

This shows that the additive method is justified by Gauss-Markov prediction in the linear model of Mack.

5.3 Conditional Gauss-Markov Prediction

In the present subsection we consider a sequential model for the chain-ladder method. This model is a sequential model since it involves successive conditioning with respect to the σ-algebras $\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{n-1}$ where, for each $k \in \{1, \ldots, n\}$, the σ-algebra

$$\mathcal{G}_{k-1}$$

represents the information provided by the cumulative losses $S_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$, which is at the same time the information provided by the incremental losses $Z_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$.

Sequential Chain-Ladder Model: *For each $k \in \{1, \ldots, n\}$, there exists a random variable φ_k and a strictly positive random variable σ_k such that*

$$E^{\mathcal{G}_{k-1}}(S_{i,k}) = S_{i,k-1}\,\varphi_k$$

and

$$\text{cov}^{\mathcal{G}_{k-1}}(S_{i,k} S_{j,k}) = \begin{cases} S_{i,k-1}\sigma_k & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

holds for all $i, j \in \{0, 1, \ldots, n-k+1\}$.

In the case where the random variables $\varphi_1, \ldots, \varphi_n$ are all constant, integration yields $E[S_{i,k}] = \varphi_k\, E[S_{i,k-1}]$ such that the parameters $\varphi_1, \ldots, \varphi_n$ form a development pattern for factors. In the general case, the random parameters $\varphi_1, \ldots, \varphi_n$ may be interpreted as a *random development pattern* for factors.

The sequential chain-ladder model may be considered as a sequence of n conditional linear models corresponding to the development years $k \in \{1, \ldots, n\}$. Each of these conditional linear models consists of an observable part

$$\begin{pmatrix} E^{\mathcal{G}_{k-1}}(S_{0,k}) \\ E^{\mathcal{G}_{k-1}}(S_{1,k}) \\ \vdots \\ E^{\mathcal{G}_{k-1}}(S_{n-k,k}) \end{pmatrix} = \begin{pmatrix} S_{0,k-1} \\ S_{1,k-1} \\ \vdots \\ S_{n-k,k-1} \end{pmatrix} \varphi_k$$

and a non-observable part

$$E^{\mathcal{G}_{k-1}}(S_{n-k+1,k}) = S_{n-k+1,k-1}\,\varphi_k$$

Then $\mathcal{G}_{k-1}$-conditional Gauss-Markov estimator $\hat{\varphi}_k^{\text{GM}}$ of the random parameter φ_k satisfies

$$\hat{\varphi}_k^{\text{GM}} = \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

and hence coincides with the chain-ladder factor $\hat{\varphi}_k^{\text{CL}}$.

Furthermore, for every accident year $i \geq n-k+1$, the $\mathcal{G}_{k-1}$-conditional Gauss-Markov predictor $\hat{S}_{i,k}^{\text{GM}}$ of the non-observable cumulative loss $S_{i,k}$ satisfies

$$\begin{aligned} \hat{S}_{i,k}^{\text{GM}} &= S_{i,k-1}\,\hat{\varphi}_k^{\text{GM}} \\ &= S_{i,k-1}\,\hat{\varphi}_k^{\text{CL}}. \end{aligned}$$

The previous formula, however, is only useful when $S_{i,k-1}$ is observable, which is the case if and only if $i+k-1 \leq n$ and hence $i = n-k+1$.

Turning the point of view from development years to accident years, we see that the $\mathcal{G}_{(n-i)}$-conditional Gauss-Markov predictors of the first non-observable cumulative losses $S_{i,n-i+1}$ satisfy

$$\hat{S}_{i,n-i+1}^{\text{GM}} = S_{i,n-i}\,\hat{\varphi}_{n-i+1}^{\text{CL}}$$

and hence coincide with the chain-ladder predictors.

In the case $i+k = n+1$, the chain-ladder predictors are thus justified by conditional

Gauss-Markov estimation, but another justification is needed in the case $i+k \geq n+2$. This can be achieved by minimizing the $\mathcal{G}_{k-1}$-conditional expected prediction error

$$E^{\mathcal{G}_{k-1}}\left((\hat{S}_{i,k}-S_{i,k})^2\right)$$

over the collection of all predictors $\hat{S}_{i,k}$ of $S_{i,k}$ satisfying

$$\hat{S}_{i,k} = \hat{S}^{\mathrm{CL}}_{i,k-1}\hat{\varphi}_k$$

for some $\mathcal{G}_{k-1}$-conditionally unbiased linear estimator $\hat{\varphi}_k$ of φ_k and it turns out that the minimum over this restricted class of predictors is attained for the chain-ladder predictor $\hat{S}^{\mathrm{CL}}_{i,k}$. The sequential optimality criterion adopted here reflects very well the sequential character of the chain-ladder method and of the chain-ladder model. The criterion is also reasonable since prediction for the first non-observable calendar year is much more important than prediction for subsequent calendar years: Predictors for the first non-observable calendar year cannot be corrected later whereas predictors for subsequent calendar years will be corrected anyway since already one year later additional loss experience and hence a new run-off triangle will be available.

The sequential chain-ladder model is due to Schnaus and was proposed by Schmidt and Schnaus [1996] where it is studied in detail; see also Schmidt [1997, 1999b, 2006]. The sequential chain-ladder model is a slight but convenient extension of the chain-ladder model of Mack [1993]. A systematic comparison of several models for the chain-ladder method is given in Hess and Schmidt [2002].

5.4 Remarks

Although least-squares prediction is a central topic in econometrics, it appears that this method has been ignored in loss reserving until recently. It is the merit of Halliwell [1996] that least-squares prediction is by now considered as a most useful tool in loss reserving; see also Schmidt [1999a], Hamer [1999], Halliwell [1999], Radtke and Schmidt [2004], and Schmidt [2006].

6. MAXIMUM-LIKELIHOOD ESTIMATION

Another general principle of statistical inference is maximum-likelihood estimation. The maximum-likelihood principle is applicable only if the joint distribution of all observable random variables is known with the exception of certain parameters.

The models considered here are models for claim counts. The basic model is the Poisson model which is a special case of the general multinomial model.

6.1 Poisson Model

The Poisson model is a model for claim counts and consists of the following assumptions:

Poisson model:

(i) *The family* $\{Z_{i,k}\}_{i,k\in\{0,1,\ldots,n\}}$ *of all incremental losses is independent.*

(ii) *There exists parameters* $\alpha_0, \alpha_1, \ldots, \alpha_n \in (0,\infty)$ *and* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n \in (0,1)$ *with* $\sum_{l=0}^{n} \vartheta_l = 1$ *such that for all* $i, k \in \{0, 1, \ldots, n\}$ *the incremental loss* $Z_{i,k}$ *has the Poisson distribution with expectation* $\alpha_i \vartheta_k$.

We assume in this subsection that the assumptions of the Poisson model are fulfilled. Because of (ii) we have

$$E[Z_{i,k}] = \alpha_i \vartheta_k .$$

Summation yields

$$E[S_{i,n}] = \alpha_i$$

and hence

$$E[Z_{i,k}] = \vartheta_k E[S_{i,n}]$$

such that the parameters $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ form a development pattern for incremental quotas.

In the Poisson model the joint distribution of all incremental losses is known except for the parameters. In fact, we have

$$P\left[\bigcap_{i=0}^{n}\bigcap_{k=0}^{n}\{Z_{i,k} = z_{i,k}\}\right] = \prod_{i=0}^{n}\prod_{k=0}^{n}\left(e^{-\alpha_i\vartheta_k}\frac{(\alpha_i\vartheta_k)^{z_{i,k}}}{z_{i,k}!}\right).$$

To estimate the parameters we can thus use the maximum-likelihood method. The maximum-likelihood method is based in the joint distribution of *all observable incremental losses* which is given by

$$P\left[\bigcap_{i=0}^{n}\bigcap_{k=0}^{n-i}\{Z_{i,k}=z_{i,k}\}\right]=\prod_{i=0}^{n}\prod_{k=0}^{n-i}\left(e^{-\alpha_i\vartheta_k}\frac{(\alpha_i\vartheta_k)^{z_{i,k}}}{z_{i,k}!}\right).$$

It follows that the likelihood function L is given by

$$L(\alpha_0,\alpha_1,\ldots,\alpha_n,\ \vartheta_0,\vartheta_1,\ldots,\vartheta_n\mid Z):=\prod_{i=0}^{n}\prod_{k=0}^{n-i}\left(e^{-\alpha_i\vartheta_k}\frac{(\alpha_i\vartheta_k)^{Z_{i,k}}}{Z_{i,k}!}\right)$$

where $Z:=\{Z_{i,k}\}_{i,k\in\{0,1,\ldots,n\},\,i+k\le n}$. Interpreting the maximum-likelihood principle in a wide sense (which ignores the second order conditions for a maximum), observable random variables

$$\hat{\alpha}_0^{\text{ML}},\hat{\alpha}_1^{\text{ML}},\ldots,\hat{\alpha}_n^{\text{ML}}$$

and

$$\hat{\vartheta}_0^{\text{ML}},\hat{\vartheta}_1^{\text{ML}},\ldots,\hat{\vartheta}_n^{\text{ML}}$$

are said to be ***maximum-likelihood estimators*** if they annihilate all first order partial derivatives of the likelihood function (or, equivalently, of the log-likelihood function) and satisfy the constraint

$$\sum_{l=0}^{n}\hat{\vartheta}_l^{\text{ML}}=1.$$

Straightforward computation shows that the maximum-likelihood estimators satisfy the marginal sum equations

$$\sum_{l=0}^{n-i}\hat{\alpha}_i\hat{\vartheta}_l=\sum_{l=0}^{n-i}Z_{i,l}$$

with $i\in\{0,1,\ldots,n\}$ and

$$\sum_{l=0}^{n-k}\hat{\alpha}_i\hat{\vartheta}_k=\sum_{l=0}^{n-k}Z_{i,k}$$

with $k\in\{0,1,\ldots,n\}$ and, of course, the constraint

$$\sum_{l=0}^{n}\hat{\vartheta}_l=1.$$

Therefore, the maximum-likelihood estimators coincide with the marginal sum estimators. It now follows from the properties of the marginal sum estimators that in the Poisson model the maximum-likelihood estimators of the expected ultimate cumulative losses are identical with the chain-ladder predictors of the ultimate cumulative losses. This was first observed by Hachemeister and Stanard [1975].

However, if, in addition to the assumptions of the Poisson model, it is assumed that the expected ultimate cumulative losses are all identical such that

$$\alpha_i = \alpha$$

holds for all $i \in \{0, 1, \dots, n\}$, then maximum-likelihood estimation is still possible but the maximum-likelihood estimators turn out to satisfy

$$\hat{\alpha} = \sum_{l=0}^{n} \frac{1}{n-l+1} \sum_{j=0}^{n-k} Z_{j,l}$$

and

$$\hat{\vartheta}_k = \frac{\dfrac{1}{n-k+1} \sum_{j=0}^{n-k} Z_{j,k}}{\sum_{l=0}^{n} \dfrac{1}{n-l+1} \sum_{j=0}^{n-k} Z_{j,l}}.$$

In particular, the maximum-likelihood estimators of the expected ultimate cumulative losses are *not* identical with the chain-ladder estimators of the ultimate cumulative losses; see Schmidt and Zocher [2005].

6.2 Multinomial Model

The multinomial model is a model for claim counts and consists of the following assumptions:

Multinomial model:

(i) *The accident years are independent.*

(ii) *There exist parameters* $\vartheta_0, \vartheta_1, \dots, \vartheta_n \in (0,1)$ *with* $\sum_{l=0}^{n} \vartheta_l = 1$ *such that, for every accident year* $i \in \{0, 1, \dots, n\}$, *the conditional joint distribution of the family* $\{Z_{i,k}\}_{k \in \{0,1,\dots,n\}}$ *with respect to the ultimate cumulative loss* $S_{i,n}$ *is the multinomial distribution with parameters* $S_{i,n}$ *and* $\vartheta_0, \vartheta_1, \dots, \vartheta_n$.

We assume in this subsection that the assumptions of the multinomial model are fulfilled. Because of (ii) we have

$$E[Z_{i,k} \mid S_{i,n}] = \vartheta_k S_{i,n}$$

and hence

$$E[Z_{i,k}] = \vartheta_k E[S_{i,n}]$$

such that the parameters $\vartheta_0, \vartheta_1, \ldots, \vartheta_n$ form a development pattern for incremental quotas.

The multinomial model is appealing since it suggests that every claim of any accident year is reported or settled with probability ϑ_k in development year k. It thus reminds of the urn model in which $S_{i,n}$ balls are drawn with replacement from an urn consisting of balls with $1+n$ different colours corresponding to the development years.

Letting

$$\alpha_i := E[S_{i,n}]$$

it is easy to see that the multinomial model contains the Poisson model as the special case in which every ultimate cumulative loss $S_{i,n}$ has the Poisson distribution with expectation α_i. Moreover, under the assumptions of the multinomial model, it can be shown that the incremental losses of any accident year are independent if and only if the family of all incremental losses is independent and every incremental loss has the Poisson distribution with expectation $\alpha_i \vartheta_k$. Therefore, the main advantage of the multinomial model over the Poisson model is the fact that it allows for dependence between the incremental losses of a given accident year.

If, in addition to the assumptions of the multinomial model, the distributions of the ultimate cumulative losses are assumed to belong to a parametric family of distributions, then the joint distribution of all incremental losses is known except for the parameters and maximum-likelihood estimation can be used to estimate the expected ultimate cumulative losses.

In the case where each of the ultimate cumulative losses has a Poisson distribution, we are back to the Poisson model and the maximum-likelihood estimators of the expected ultimate cumulative losses are identical with the chain-ladder predictors of the ultimate cumulative losses.

The same result obtains in the case where each of the ultimate cumulative losses has a negativebinomial distribution; see Schmidt and Wünsche [1998]. Negativebinomial distributions are of interest since they are mixed Poisson distributions (with respect to a mixing gamma distribution), and mixed Poisson distributions in turn are of interest since their variances exceed their expectations, which is the case for most empirical claim count distributions.

In fact, a much more general result is true: If, in addition to the assumptions of the multinomial model, each of the ultimate cumulative losses has a Hofmann distribution, then the maximum-likelihood estimators of the expected ultimate cumulative losses are identical with the chain-ladder predictors of the ultimate cumulative losses; see Schmidt and Zocher [2005]. The definition and the discussion of Hofmann distributions are beyond the scope of this paper, but we remark that Hofmann distributions were introduced by Hofmann [1955] and that every Hofmann distribution is at the same time a mixed Poisson distribution and a compound Poisson distribution and can be computed by recursion; see e. g. Hess, Liewald and Schmidt [2002].

Since the class of all Hofmann distributions is a wide class of mixed Poisson distributions, the multinomial model with ultimate cumulative loss numbers having a Hofmann distribution is a very general model for claim counts in which the maximum-likelihood estimators of the expected ultimate cumulative losses are identical with the chain-ladder predictors of the ultimate cumulative losses.

6.3 Remarks

Alternatively, the Poisson model can be extended to a general stochastic model in which the family $\{Z_{i,k}\}_{i,k\in\{0,1,\dots,n\}}$ is independent and the distribution of every incremental loss belongs to an exponential family. In such models, the theory of generalized linear models can be applied.

7. CONCLUSIONS

The notion of a development pattern, which can be expressed in three different but equivalent ways, provides a powerful tool for the comparison of different methods and of different model of loss reserving.

The general Bornhuetter-Ferguson method provides a general framework into which several methods of loss reserving can be embedded via

- a particular choice of the prior estimators of the development pattern for cumulative quotas and/or
- a particular choice of the prior estimators of the expected ultimate cumulative losses.

Moreover, there are many stochastic models in which

- the credibility predictors or
- the Gauss Markov predictors or
- the maximum-likelihood estimators of the expected ultimate cumulative losses

can be interpreted as Bornhuetter-Ferguson predictors.

The choice of a stochastic model or of a method of prediction is a choice which has to be made by the actuary and which may have a considerable impact on the result. In the Poisson model, e. g., credibility prediction and maximum-likelihood estimation are possible but lead to different results; here the choice of the statistical method could be based on the judgement that either external information or internal information is more reliable. Still in the Poisson model, the form of the maximum-likelihood estimators of the expected ultimate cumulative losses depends on the assumption that the expected ultimate cumulative losses may be different or are identical.

We also remark that the chain-ladder method and the additive method can be extended to the multivariate case which corresponds to a portfolio consisting of several subportfolios representing dependent lines of business. Moreover, the multivariate chain-ladder method and the multivariate additive method can be justified by multivariate models extending the univariate models considered in the present paper. A detailed discussion of these multivariate methods and models may be found in Schmidt [2006].

8. REFERENCES

[1] Benktander, Gunnar [1976]: "An approach to credibility in calculating IBNR for casualty excess reinsurance", *The Actuarial Review*, April 7.

[2] Goldberger, Arthur S. [1962]: "Best linear unbiased prediction in the generalized linear regression model", *J. American Statistical Association, Vol. 57*, 369-375.

[3] Hachemeister, Charles A., and James N. Stanard [1975]: "IBNR claims count estimation with static lag functions", Unpublished.

[4] Halliwell, Leigh J. [1996]: "Loss prediction by generalized least-squares", *PCAS, Vol. LXXXIII*, 436-489.

[5] Halliwell, Leigh J. [1999]: "Loss prediction by generalized least-squares – Author's response", *PCAS, Vol. LXXXVI*, 764-769.

[6] Hamer, Michael D. [1999]: "Loss prediction by generalized least-squares – Discussion", *PCAS, Vol. LXXXVI*, 748-763.

[7] Hess, Klaus Th., Anett Liewald, and Klaus D. Schmidt [2002]: "An extension of Panjer's recursion", *ASTIN Bulletin, Vol. 32*, 283-297.

[8] Hess, Klaus Th., and Klaus D. Schmidt [2002]: "A comparison of models for the chain-ladder method", *Insurance Mathematics and Economics, Vol. 31*, 351-364.

[9] Hesselager, Ole, and Thomas Witting [1988]: "A credibility model with random fluctuations in delay

probabilities for the prediction of IBNR claims", *ASTIN Bulletin, Vol. 18*, 79-90.

[10] Hofmann, Martin [1955]: "Über zusammengesetzte Poisson-Prozesse und ihre Anwendungen in der Unfallversicherung", *Mitteilungen VSVM, Vol. 55*, 499-575.

[11] Hovinen, Esa [1981]: "Additive and continuous IBNR", Paper presented at the International ASTIN Colloquium in Loen.

[12] Lorenz, Holger, and Klaus D. Schmidt [1999]: "Grossing-up, chain-ladder and marginal-sum estimation", *Blätter DGVM, Vol. 24*, 195-200.

[13] Mack, Thomas [1990]: "Improved estimation of IBNR claims by credibility theory", *Insurance Mathematics and Economics, Vol. 9*, 51-57.

[14] Mack, Thomas [1993]: "Distribution-free calculation of the standard error of chain-ladder reserve estimates", *ASTIN Bulletin, Vol. 23*, 213-225.

[15] Mack, Thomas [2000]: "Credible claims reserves – The Benktander method", *ASTIN Bulletin, Vol. 30*, 333-347.

[16] Neuhaus, Walther [1992]: "Another pragmatic loss reserving method or Bornhuetter-Ferguson revisited", *Scandinavian Actuarial Journal*, 151-162.

[17] Radtke, Michael, and Klaus D. Schmidt (Eds.) [2004]: *Handbuch zur Schadenreservierung*, Karlsruhe, Verlag Versicherungswirtschaft.

[18] Rao, C. Radhakrishna, and Helge Toutenburg [1995]: *Linear Models – Least-squares and Alternatives*, Springer, Berlin - Heidelberg - New York.

[19] Schmidt, Klaus D. [1997]: "Non-optimal prediction by the chain-ladder method", *Insurance Mathematics and Economics, Vol. 21*, 17-24.

[20] Schmidt, Klaus D. [1998]: "Prediction in the linear model – A direct approach", *Metrika, Vol. 48*, 141-147.

[21] Schmidt, Klaus D. [1999a]: "Loss prediction by generalized least-squares – Discussion", *PCAS, Vol. LXXXVI*, 736-747.

[22] Schmidt, Klaus D. [1999b]: "Chain-ladder prediction and asset-liability management", *Blätter DGVM, Vol. 24*, 1-9.

[23] Schmidt, Klaus D. [2004]: "Prediction", *Encyclopedia of Actuarial Science, Vol. 3*, 1317-1321, Chichester, Wiley.

[24] Schmidt, Klaus D. [2006]: "Optimal and additive loss reserving for dependent lines of business", CAS 2006 Call Paper Program.

[25] Schmidt, Klaus D., and Anja Schnaus [1996]: "An extension of Mack's model for the chain-ladder method", *ASTIN Bulletin, Vol. 26*, 247-262.

[26] Schmidt, Klaus D., and Angela Wünsche [1998]: "Chain-ladder, marginal-sum and maximum-likelihood estimation", *Blätter DGVM, Vol. 23*, 289-307.

[27] Schmidt, Klaus D., and Mathias Zocher [2005]: "Loss reserving and Hofmann distributions", *Bulletin SAV*, 127-162.

[28] Taylor, Greg C. [2000]: *Loss Reserving – An Actuarial Perspective*, Kluwer, Boston/Dordrecht/London.

[29] Vaughn, Trent R., and Phoebe Tinney [2004]: "The modified Bornhuetter-Ferguson approach to IBNR allocation", *PCAS, Vol. XCI*, 1-13.

[30] Witting, Thomas [1987]: "Kredibilitätsschätzungen für die Anzahl IBNR-Schäden", *Blätter DGVM, Vol. 18*, 45-58.

[31] Zocher, Mathias [2005]: Personal communication.

Biography of the Author:

The author is professor for actuarial mathematics at Technische Universität Dresden. His main interests are mathematical models in loss reserving and reinsurance as well as linear and credibility models in non-life insurance. He has a degree in mathematics and economics from Universität Zürich and a PhD in mathematics from Universität Mannheim. He is an academic correspondent of the CAS and a member of the DAV (Deutsche Aktuar-Vereinigung) with particular engagement in working parties on education and loss reserving. Besides his publications on loss reserving, the author has compiled a *Bibliography on Loss Reserving* which is available on the Web under http://www.math.tu-dresden.de/sto/schmidt/dsvm/reserve.pdf.

CONTENTS

Optimal and Additive Loss Reserving for Dependent Lines of Business

Klaus D. Schmidt
Lehrstuhl für Versicherungsmathematik
Technische Universität Dresden

Abstract.
In the present paper we review and extend two stochastic models for loss reserving and study their impact on extensions of the additive method and of the chain-ladder method. The first of these models is a particular linear model while the second one is a sequential model which is composed of a finite number of conditional linear models. These models lead to multivariate extensions of the additive method and of the chain-ladder method, respectively, which turn out to resolve the problem of additivity.

Keywords. Loss reserving; dependent lines of business; additivity; multivariate additive method; multivariate chain-ladder method.

1. INTRODUCTION

For a portfolio consisting of several lines of business, it is well-known that the chain-ladder predictors for the aggregate portfolio usually differ from the sums of the chain-ladder predictors for the different lines of business; see Ajne [1994] and Klemmt [2004]. It is one of the purposes of the present paper to point out that the non-coincidence between a chain-ladder predictor for the aggregate portfolio and the sum of the corresponding chain-ladder predictors for the different lines of business has its origin in the univariate character of the chain-ladder method which neglects dependence between the different lines of business.

The problem of dependence between different lines of business has already been addressed by Holmberg [1994]. His paper is remarkable since it adopts a general point of view and considers

- correlation within accident years,
- correlation between accident years, and
- correlation between different lines of business.

Nevertheless, the major part of Holmberg's paper is devoted to correlation within and between accident years and the author expresses the opinion that, in practical applications, the great majority of the effects causing correlation between different lines of business are already captured in the correlation within and between accident years. It is another purpose of the present paper to show that correlation between different lines of business can be modelled and that the resulting models, combined with a general optimality criterion, lead to multivariate predictors which are superior to the univariate ones. Here and in the sequel, the term *univariate* refers to prediction for a single line of business and the term *multivariate* refers to simultaneous prediction for several lines of business or for different types of losses (like paid and incurred losses) of the same line of business.

The papers by Ajne [1994] and Holmberg [1994] were slightly preceded in time by a paper by Mack [1993] which, similar to the paper by Hachemeister and Stanard [1975], turned out to be path-breaking in the discussion of stochastic models for the chain-ladder method. In the model of Mack, dependence within accident years is expressed by conditioning, but it is also assumed that the accident years are independent. The assumption of independent accident years was subsequently relaxed in the model of Schnaus presented by Schmidt and Schnaus [1996]. Both of these models are univariate and hence do not reflect dependence between lines of business.

After the publication of the paper of Mack [1993], about a decade had to pass before the emergence of the first bivariate models related to the chain-ladder method. One of these models, due to Quarg and Mack [2004], expresses dependence between the paid and incurred losses of a single line of business (a topic which had already been studied before by Halliwell [1997] within the theory of linear models) and has been used as a foundation for the construction of certain bivariate predictors which are now known as Munich chain-ladder predictors. The other of these models, due to Braun [2004], expresses dependence between two lines of business and has been used to construct new estimators for the prediction errors of the univariate chain-ladder predictors, but it has not been used to construct bivariate predictors. Each of these models extends the model of Mack.

Quite recently, Pröhl and Schmidt [2005] as well as Hess, Schmidt and Zocher [2006] proposed multivariate models which reflect dependence between an arbitrary number of lines of business. The model of Pröhl and Schmidt extends the model of Braun in essentially the same way as the model of Schnaus extends the model of Mack, while the model of Hess, Schmidt and Zocher extends in a rather straightforward way the particular linear model which may be used to justify the additive method; see Radtke and Schmidt [2004]. These models, combined with a general optimality criterion, lead to multivariate versions of the chain-ladder method and of the additive method, respectively, which turn out to resolve the problem of additivity.

In the present paper we review these recent multivariate models and methods of loss reserving. In order to avoid the accumulation of technicalities, we start with a systematic review of the univariate case (Section 2) and of prediction in conditional linear models (Section 3). We then pass to the multivariate case (Section 4) and show that, the optimal multivariate predictors for the single lines of business sum up to the corresponding predictors for the aggregate portfolio (Section 5). We also show how the unbiased estimators of variances and covariances proposed by Braun [2004] can be adapted to the multivariate models considered here (Section 6). We conclude with some complementary remarks (Section 7) and a numerical example illustrating the multivariate chain-ladder method (Section 8).

Throughout this paper, let $(\Omega, \mathcal{F}, P)$ be a probability space on which all random variables, random vectors and random matrices are defined. We assume that all random variables are square integrable and that all random vectors and random matrices have square integrable coordinates. Moreover, all equalities and inequalities involving random variables are understood to hold almost surely with respect to the probability measure P.

2. UNIVARIATE LOSS PREDICTION

In the present section we review two univariate stochastic models which are closely related to two current methods of loss reserving.

We consider a single line of business which is described by a family $\{Z_{i,k}\}_{i,k\in\{0,1,\ldots,n\}}$ of random variables. We interpret $Z_{i,k}$ as the loss of *accident year* i which is reported or settled in *development year* k, and hence in *calendar year* $i+k$, and we refer to $Z_{i,k}$ as the *incremental loss* of accident year i and development year k.

We assume that the incremental losses $Z_{i,k}$ are *observable* for calendar years $i+k \le n$ and that they are *non-observable* for calendar years $i+k \ge n+1$. The observable incremental losses are represented by the following *run-off triangle*:

Accident Year	Development Year 0	1	…	k	…	$n-i$	…	$n-1$	n
0	$Z_{0,0}$	$Z_{0,1}$	…	$Z_{0,k}$	…	$Z_{0,n-i}$	…	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1,0}$	$Z_{1,1}$	…	$Z_{1,k}$	…	$Z_{1,n-i}$	…	$Z_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$Z_{i,0}$	$Z_{i,1}$	…	$Z_{i,k}$	…	$Z_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$Z_{n-k,0}$	$Z_{n-k,1}$	…	$Z_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$Z_{n-1,0}$	$Z_{n-1,1}$							
⋮	⋮								
n	$Z_{n,0}$								

Besides the incremental losses, we also consider the *cumulative losses* $S_{i,k}$ which are defined by

$$S_{i,k} := \sum_{l=0}^{k} Z_{i,l}.$$

Then the cumulative losses $S_{i,k}$ are observable for calendar years $i+k \leq n$ and they are non-observable for calendar years $i+k \geq n+1$. Just like the observable incremental losses, the observable cumulative losses are represented by a run-off triangle:

Accident Year	Development Year 0	1	…	k	…	$n-i$	…	$n-1$	n
0	$S_{0,0}$	$S_{0,1}$	…	$S_{0,k}$	…	$S_{0,n-i}$	…	$S_{0,n-1}$	$S_{0,n}$
1	$S_{1,0}$	$S_{1,1}$	…	$S_{1,k}$	…	$S_{1,n-i}$	…	$S_{1,n-1}$	
⋮	⋮	⋮		⋮		⋮			
i	$S_{i,0}$	$S_{i,1}$	…	$S_{i,k}$	…	$S_{i,n-i}$			
⋮	⋮	⋮		⋮					
$n-k$	$S_{n-k,0}$	$S_{n-k,1}$	…	$S_{n-k,k}$					
⋮	⋮	⋮							
$n-1$	$S_{n-1,0}$	$S_{n-1,1}$							
⋮	⋮								
n	$S_{n,0}$								

Of course, the incremental losses can be recovered from the cumulative losses.

2.1 Univariate Additive Model

Let us first consider the univariate additive model:

Univariate Additive Model: *There exist real numbers* $\vartheta_0, \vartheta_1, \ldots, \vartheta_n > 0$ *and*

$\sigma_0, \sigma_1, \ldots, \sigma_n > 0$ *as well as real parameters* $\zeta_0, \zeta_1, \ldots, \zeta_n$ *such that*

$$E[Z_{i,k}] = \nu_i \zeta_k$$

and

$$\mathrm{cov}[Z_{i,k}, Z_{j,l}] = \begin{cases} \nu_i \sigma_k & \text{if } i = j \text{ and } k = l \\ 0 & \text{else} \end{cases}$$

holds for all $i, j, k, l \in \{0, 1, \ldots, n\}$.

For $i, k \in \{0, 1, \ldots, n\}$ such that $i + k \geq n + 1$, the estimators and predictors

$$\hat{\zeta}_k^{\mathrm{AD}} := \frac{\sum_{j=0}^{n-k} Z_{j,k}}{\sum_{j=0}^{n-k} \nu_j}$$

$$\hat{Z}_{i,k}^{\mathrm{AD}} := \nu_i \, \hat{\zeta}_k^{\mathrm{AD}}$$

$$\hat{S}_{i,k}^{\mathrm{AD}} := S_{i,n-i} + \nu_i \sum_{l=n-i+1}^{k} \hat{\zeta}_l^{\mathrm{AD}}$$

are said to be the estimators and the predictors of the (*univariate*) *additive method*. Under the assumptions of the additive model, these estimators and predictors are indeed reasonable, as will be shown in Section 4 below.

2.2 Univariate Chain-Ladder Model

Let us now consider the univariate chain-ladder model due to Schnaus which was proposed by Schmidt and Schnaus [1996] and is a slight but convenient extension of the model of Mack [1993].

The chain-ladder model is a sequential model since it involves successive conditioning with respect to the σ-algebras $\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{n-1}$ where, for each $k \in \{0, 1, \ldots, n\}$, the σ-algebra

$$\mathcal{G}_{k-1}$$

represents the information provided by the cumulative losses $S_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$, which is at the same time the information provided by the incremental losses $Z_{i,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$.

We assume that $S_{i,k} > 0$ holds for all $i, k \in \{0, 1, \ldots, n\}$.

Univariate Chain-Ladder Model: *For each* $k \in \{1, \ldots, n\}$, *there exists a random variable* φ_k *and a strictly positive random variable* σ_k *such that*

$$E^{\mathcal{G}_{k-1}}[S_{i,k}] = S_{i,k-1}\,\varphi_k$$

and

$$\mathrm{cov}^{\mathcal{G}_{k-1}}(S_{i,k}, S_{j,k}) = \begin{cases} S_{i,k-1}\sigma_k & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

holds for all $i, j \in \{0, 1, \ldots, n-k+1\}$.

For $i, k \in \{0, 1, \ldots, n\}$ such that $i + k \geq n+1$, the estimators and predictors

$$\hat{\varphi}_k^{\mathrm{CL}} := \frac{\sum_{j=0}^{n-k} S_{j,k}}{\sum_{j=0}^{n-k} S_{j,k-1}}$$

$$\hat{S}_{i,k}^{\mathrm{CL}} := S_{i,n-i} \prod_{l=n-i+1}^{k} \hat{\varphi}_l^{\mathrm{CL}}$$

(such that $\hat{S}_{i,n-i}^{\mathrm{CL}} = S_{i,n-i}$) are said to be the estimators and the predictors of the (*univariate*) *chain-ladder method.* Under the assumptions of the chain-ladder model, these estimators and predictors are indeed reasonable, as will be shown in Section 4.

3. ESTIMATION AND PREDICTION IN THE CONDITIONAL LINEAR MODEL

In the present section we consider a random vector $\mathbf{X}$ and a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$. The σ-algebra $\mathcal{G}$ represents information which is provided by some other random quantities.

Conditional Linear Model: *There exists a* $\mathcal{G}$*-measurable random matrix* $\mathbf{A}$ *and a* $\mathcal{G}$*-measurable random vector* $\boldsymbol{\beta}$ *such that*

$$E^{\mathcal{G}}[\mathbf{X}] = \mathbf{A}\boldsymbol{\beta}.$$

The random matrix $\mathbf{A}$ is assumed to be observable and is said to be the *design matrix* and the random vector $\boldsymbol{\beta}$ is assumed to be non-observable and is said to be the *parameter vector* or the *parameter* for short.

In the subsequent discussion, we assume that the assumption of the conditional linear model is fulfilled.

We assume further that some of the coordinates of $\mathbf{X}$ are *observable* whereas some other coordinates are *non-observable.* Then the random vector $\mathbf{X}_1$ consisting of the observable coordinates of $\mathbf{X}$ and the random vector $\mathbf{X}_2$ consisting of the non-observable coordinates of $\mathbf{X}$ satisfy

$$E^{\mathcal{G}}[\mathbf{X}_1] = \mathbf{A}_1 \boldsymbol{\beta}$$
$$E^{\mathcal{G}}[\mathbf{X}_2] = \mathbf{A}_2 \boldsymbol{\beta}$$

for some submatrices $\mathbf{A}_1$ and $\mathbf{A}_2$ of $\mathbf{A}$.

We also assume that the matrix $\mathbf{A}_1$ has full column rank, that the random matrices

$$\boldsymbol{\Sigma}_{11} := \text{var}^{\mathcal{G}}[\mathbf{X}_1]$$
$$\boldsymbol{\Sigma}_{21} := \text{cov}^{\mathcal{G}}[\mathbf{X}_2, \mathbf{X}_1]$$

are known, and that $\boldsymbol{\Sigma}_{11}$ is (almost surely) invertible.

Since the random vector $\mathbf{X}_2$ is non-observable, only the random vector $\mathbf{X}_1$ can be used for the estimation of the parameter $\boldsymbol{\beta}$.

3.2 Gauss-Markov Estimation

Let us first consider the estimation problem for a random vector of the form $\mathbf{C}\boldsymbol{\beta}$, where $\mathbf{C}$ is a $\mathcal{G}$-measurable random matrix of suitable dimension.

A random variable $\hat{\mathbf{Y}}$ is said to be an *estimator* of $\mathbf{C}\boldsymbol{\beta}$ if it is a measurable transformation of the observable random vector $\mathbf{X}_1$. For an estimator $\hat{\mathbf{Y}}$ of $\mathbf{C}\boldsymbol{\beta}$, the random variable

$$E^{\mathcal{G}}\left[(\hat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta})'(\hat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta})\right]$$

is said to be the *$\mathcal{G}$-conditional expected squared estimation error* of $\hat{\mathbf{Y}}$. Since

$$E^{\mathcal{G}}\left[(\hat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta})'(\hat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta})\right] = \text{trace}\left(\text{var}^{\mathcal{G}}[\hat{\mathbf{Y}}]\right) + E^{\mathcal{G}}\left[\hat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta}\right]' E^{\mathcal{G}}\left[\hat{\mathbf{Y}} - \mathbf{C}\boldsymbol{\beta}\right]$$

the $\mathcal{G}$-conditional expected squared estimation error is determined by the $\mathcal{G}$-conditional variance of the estimator and the $\mathcal{G}$-conditional expectation of the estimation error. An observable random vector $\hat{\mathbf{Y}}$ is said to be

- a *linear estimator* of $\mathbf{C}\boldsymbol{\beta}$ if there exists a $\mathcal{G}$-measurable random matrix $\mathbf{Q}$ such that $\hat{\mathbf{Y}} = \mathbf{Q}\mathbf{X}_1$.
- a *$\mathcal{G}$-conditionally unbiased estimator* of $\mathbf{C}\boldsymbol{\beta}$ if $E^{\mathcal{G}}[\hat{\mathbf{Y}}] = E^{\mathcal{G}}[\mathbf{C}\boldsymbol{\beta}]$.

- a *Gauss-Markov predictor* of $\mathbf{C\beta}$ if it is a $\mathcal{G}$-conditionally unbiased linear estimator of $\mathbf{C\beta}$ and minimizes the $\mathcal{G}$-conditional expected squared estimation error over all $\mathcal{G}$-conditionally unbiased linear estimators of $\mathbf{C\beta}$.

We have the following result:

3.1 Proposition (Gauss-Markov Theorem for Estimators). *There exists a unique Gauss-Markov estimator* $\hat{\mathbf{Y}}^{GM}(\mathbf{C\beta})$ *of* $\mathbf{C\beta}$ *and it satisfies*

$$\hat{\mathbf{Y}}^{GM}(\mathbf{C\beta}) = \mathbf{C}(\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{X}_1.$$

In particular, $\hat{\mathbf{Y}}^{GM}(\mathbf{C\beta}) = \mathbf{C}\hat{\mathbf{Y}}^{GM}(\mathbf{\beta})$.

Proposition 3.1 implies that the coordinates of the Gauss-Markov estimator

$$\hat{\mathbf{\beta}}^{GM} := (\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{X}_1$$

of the parameter $\mathbf{\beta}$ coincide with the Gauss-Markov estimators of its coordinates.

3.2 Gauss-Markov Prediction

Let us now consider the prediction problem for a non-observable random vector of the form $\mathbf{DX}_2$, where $\mathbf{D}$ is a matrix of suitable dimension.

A random variable $\hat{\mathbf{Y}}$ is said to be a *predictor* of $\mathbf{DX}_2$ if it is a measurable transformation of the observable random vector $\mathbf{X}_1$. For a predictor $\hat{\mathbf{Y}}$ of $\mathbf{DX}_2$, the random variable

$$E^{\mathcal{G}}\left[(\hat{\mathbf{Y}} - \mathbf{DX}_2)'(\hat{\mathbf{Y}} - \mathbf{DX}_2)\right]$$

is said to be the $\mathcal{G}$*-conditional expected squared prediction error* of $\hat{\mathbf{Y}}$. Since

$$E^{\mathcal{G}}\left[(\hat{\mathbf{Y}} - \mathbf{DX}_2)'(\hat{\mathbf{Y}} - \mathbf{DX}_2)\right] = \text{trace}\left(\text{var}^{\mathcal{G}}[\hat{\mathbf{Y}} - \mathbf{DX}_2]\right) + E^{\mathcal{G}}\left[\hat{\mathbf{Y}} - \mathbf{DX}_2\right]' E^{\mathcal{G}}\left[\hat{\mathbf{Y}} - \mathbf{DX}_2\right]$$

the $\mathcal{G}$-conditional expected squared prediction error is determined by the $\mathcal{G}$-conditional variance and the $\mathcal{G}$-conditional expectation of the prediction error. An observable random vector $\hat{\mathbf{Y}}$ is said to be

- a *linear predictor* of $\mathbf{DX}_2$ if there exists a $\mathcal{G}$-measurable random matrix $\mathbf{Q}$ such that $\hat{\mathbf{Y}} = \mathbf{QX}_1$.
- $\mathcal{G}$*-conditionally unbiased predictor* of $\mathbf{DX}_2$ if $E^{\mathcal{G}}[\hat{\mathbf{Y}}] = E^{\mathcal{G}}[\mathbf{C\beta}]$.
- a *Gauss-Markov predictor* of $\mathbf{DX}_2$ if it is a $\mathcal{G}$-conditionally unbiased linear predictor of $\mathbf{DX}_2$ and minimizes the $\mathcal{G}$-conditional expected squared prediction error over all $\mathcal{G}$-

conditionally unbiased linear predictors of $\mathbf{DX}_2$.

We have the following result:

3.2 Proposition (Gauss-Markov Theorem for Predictors). There exists a unique Gauss-Markov predictor $\hat{\mathbf{Y}}^{\text{GM}}(\mathbf{DX}_2)$ of $\mathbf{DX}_2$ and it satisfies

$$\hat{\mathbf{Y}}^{\text{GM}}(\mathbf{DX}_2) = \mathbf{D}\left(\mathbf{A}_2\hat{\boldsymbol{\beta}}^{\text{GM}} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\left(\mathbf{X}_1 - \mathbf{A}_1\hat{\boldsymbol{\beta}}^{\text{GM}}\right)\right).$$

In particular, $\hat{\mathbf{Y}}^{\text{GM}}(\mathbf{DX}_2) = \mathbf{D}\hat{\mathbf{Y}}^{\text{GM}}(\mathbf{X}_2)$.

Proposition 3.2 shows that the Gauss-Markov predictor

$$\hat{\mathbf{X}}_2^{\text{GM}} := \mathbf{A}_2\hat{\boldsymbol{\beta}}^{\text{GM}} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\left(\mathbf{X}_1 - \mathbf{A}_1\hat{\boldsymbol{\beta}}^{\text{GM}}\right)$$

of the non-observable random vector $\mathbf{X}_2$ depends not only on the Gauss-Markov estimator $\hat{\boldsymbol{\beta}}^{\text{GM}}$ of the parameter $\boldsymbol{\beta}$ but also on the $\mathcal{G}$-conditional covariance $\boldsymbol{\Sigma}_{21}$ between the non-observable random vector $\mathbf{X}_2$ and the observable random vector $\mathbf{X}_1$. Moreover, the final assertion of Proposition 3.2 implies that the coordinates of the Gauss-Markov predictor of the non-observable random vector coincide with the Gauss-Markov predictors of its coordinates.

For a single non-observable random variable, the Gauss-Markov predictor has been determined by Goldberger [1962]; see also Rao and Toutenburg [1995]. We also refer to the paper by Halliwell [1996] and to the discussion of his paper by Schmidt [1999a] and Hamer [1999] and the author's response by Halliwell [1999]. Related results can also be found in Radtke and Schmidt [2004] and in Schmidt [1998, 2004].

The proof of Propositions 3.1 and 3.2 can be achieved in exactly the same way as in the unconditional case (which corresponds to the case $\mathcal{G} = \{\varnothing, \Omega\}$, where the $\mathcal{G}$-conditional expectations, variances and covariances are nothing else than the ordinary expectations, variances and covariances).

It is sometimes also of interest to predict a random vector of the form

$$\mathbf{DX} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{D}_2 \end{pmatrix}\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}.$$

An obvious candidate is the predictor

$$\hat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{DX}) := \begin{pmatrix} \mathbf{D}_1 & \mathbf{D}_2 \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 \\ \hat{\mathbf{X}}_2^{\mathrm{GM}} \end{pmatrix}.$$

Extending the definitions and repeating the discussion with $\mathbf{X}$ in the place of $\mathbf{X}_2$, it is easily seen that the predictor $\hat{\mathbf{Y}}^{\mathrm{GM}}(\mathbf{DX})$ is indeed the Gauss-Markov predictor of $\mathbf{DX}$; see also Hamer [1999] for the even more general case of Gauss-Markov estimation/prediction of the target quantity $\mathbf{D}_0\boldsymbol{\beta} + \mathbf{D}_1\mathbf{X}_1 + \mathbf{D}_2\mathbf{X}_2$.

4. MULTIVARIATE LOSS PREDICTION

We are now prepared to consider multivariate loss prediction.

We consider m lines of business all having the same number of development years. The m lines of business may be interpreted as subportfolios of an aggregate portfolio.

For the line of business $p \in \{1, \ldots, m\}$, we denote by

$$Z_{i,k}^{(p)}$$

and

$$S_{i,k}^{(p)}$$

the incremental loss and the cumulative loss, respectively, of accident year $i \in \{0, 1, \ldots, n\}$ and development year $k \in \{0, 1, \ldots, n\}$.

For $i, k \in \{0, 1, \ldots, n\}$, we thus obtain the m-dimensional random vectors

$$\mathbf{Z}_{i,k} := \left(Z_{i,k}^{(p)} \right)_{p \in \{1, \ldots, m\}}$$

and

$$\mathbf{S}_{i,k} := \left(S_{i,k}^{(p)} \right)_{p \in \{1, \ldots, m\}}$$

of incremental losses and cumulative losses of the combined subportfolios. The observable incremental losses and the observable cumulative losses are represented by the run-off triangles

Accident Year	0	1	...	k	...	$n-i$	...	$n-1$	n
0	$\mathbf{Z}_{0,0}$	$\mathbf{Z}_{0,1}$	...	$\mathbf{Z}_{0,k}$	...	$\mathbf{Z}_{0,n-i}$	...	$\mathbf{Z}_{0,n-1}$	$\mathbf{Z}_{0,n}$
1	$\mathbf{Z}_{1,0}$	$\mathbf{Z}_{1,1}$	...	$\mathbf{Z}_{1,k}$	...	$\mathbf{Z}_{1,n-i}$	...	$\mathbf{Z}_{1,n-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$			
i	$\mathbf{Z}_{i,0}$	$\mathbf{Z}_{i,1}$	...	$\mathbf{Z}_{i,k}$	...	$\mathbf{Z}_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$		$\vdots$					
$n-k$	$\mathbf{Z}_{n-k,0}$	$\mathbf{Z}_{n-k,1}$	...	$\mathbf{Z}_{n-k,k}$					
$\vdots$	$\vdots$	$\vdots$							
$n-1$	$\mathbf{Z}_{n-1,0}$	$\mathbf{Z}_{n-1,1}$							
$\vdots$	$\vdots$								
n	$\mathbf{Z}_{n,0}$								

and

Accident Year	0	1	...	k	...	$n-i$	...	$n-1$	n
0	$\mathbf{S}_{0,0}$	$\mathbf{S}_{0,1}$	...	$\mathbf{S}_{0,k}$	...	$\mathbf{S}_{0,n-i}$	...	$\mathbf{S}_{0,n-1}$	$\mathbf{S}_{0,n}$
1	$\mathbf{S}_{1,0}$	$\mathbf{S}_{1,1}$	...	$\mathbf{S}_{1,k}$	...	$\mathbf{S}_{1,n-i}$	...	$\mathbf{S}_{1,n-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$			
i	$\mathbf{S}_{i,0}$	$\mathbf{S}_{i,1}$	...	$\mathbf{S}_{i,k}$	...	$\mathbf{S}_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$		$\vdots$					
$n-k$	$\mathbf{S}_{n-k,0}$	$\mathbf{S}_{n-k,1}$	...	$\mathbf{S}_{n-k,k}$					
$\vdots$	$\vdots$	$\vdots$							
$n-1$	$\mathbf{S}_{n-1,0}$	$\mathbf{S}_{n-1,1}$							
$\vdots$	$\vdots$								
n	$\mathbf{S}_{n,0}$								

We can now present multivariate extensions of the models considered in Section 2:

4.1 Multivariate Additive Model

Let us first consider a multivariate extension of the additive model which applies to the combined subportfolios and was proposed by Hess, Schmidt and Zocher [2006].

Multivariate Additive Model: *There exist positive definite diagonal matrices* $\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_n$ *and positive definite symmetric matrices* $\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_n$ *as well as parameter vectors* $\boldsymbol{\zeta}_0, \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_n$ *such that*

$$E[\mathbf{Z}_{i,k}] = \mathbf{V}_i \boldsymbol{\zeta}_k$$

and

$$\text{cov}[\mathbf{Z}_{i,k}, \mathbf{Z}_{j,l}] = \begin{cases} \mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k \mathbf{V}_j^{1/2} & \textit{if } i = j \textit{ and } k = l \\ \mathbf{O} & \textit{else} \end{cases}$$

holds for all $i, j, k, l \in \{0, 1, \ldots, n\}$.

In the subsequent discussion, we assume that the assumption of the multivariate additive model is fulfilled and that the matrices $\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_n$ are known.

Because of the assumption on the expectations of the incremental losses, the multivariate additive model is a linear model. This can be seen as follows: Define

$$\boldsymbol{\beta} := \begin{pmatrix} \boldsymbol{\zeta}_0 \\ \boldsymbol{\zeta}_1 \\ \vdots \\ \boldsymbol{\zeta}_{k-1} \\ \boldsymbol{\zeta}_k \\ \boldsymbol{\zeta}_{k+1} \\ \vdots \\ \boldsymbol{\zeta}_n \end{pmatrix}$$

and, for all $i, k \in \{0, 1, \ldots, n\}$, define

$$\mathbf{A}_{i,k} := \begin{pmatrix} \mathbf{O} & \mathbf{O} & \ldots & \mathbf{O} & \mathbf{V}_i & \mathbf{O} & \ldots & \mathbf{O} \end{pmatrix}$$

where the matrix $\mathbf{V}_i$ occurs in position $1 + k$. Then we have

$$E[\mathbf{Z}_{i,k}] = \mathbf{A}_{i,k} \boldsymbol{\beta}$$

for all $i, k \in \{0, 1, \ldots, n\}$. Let $\mathbf{Z}_1$ and $\mathbf{A}_1$ denote a block vector and a block matrix consisting of the vectors $\mathbf{Z}_{i,k}$ and the matrices $\mathbf{A}_{i,k}$ with $i + k \leq n$ (arranged in the same order) and let $\mathbf{Z}_2$ and $\mathbf{A}_2$ denote a block vector and a block matrix consisting of the vectors $\mathbf{Z}_{i,k}$ and the matrices $\mathbf{A}_{i,k}$ with $i + k \geq n + 1$. Then we have

$$E[\mathbf{Z}_1] = \mathbf{A}_1 \boldsymbol{\beta}$$
$$E[\mathbf{Z}_2] = \mathbf{A}_2 \boldsymbol{\beta}.$$

Therefore, the multivariate additive model is indeed a linear model.

The following result provides formulas for the Gauss-Markov estimators of the parameters of the multivariate additive model:

4.1 Theorem. *For each* $k \in \{0, 1, \ldots, n\}$, *the Gauss-Markov estimator* $\hat{\boldsymbol{\zeta}}_k^{\text{GM}}$ *of* $\boldsymbol{\zeta}_k$ *satisfies*

$$\hat{\boldsymbol{\zeta}}_k^{\mathrm{GM}} = \left(\sum_{j=0}^{n-k} \mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k^{-1} \mathbf{V}_j^{1/2} \right)^{-1} \sum_{j=0}^{n-k} (\mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k^{-1} \mathbf{V}_j^{1/2}) \mathbf{V}_j^{-1} \mathbf{Z}_{j,k}.$$

Proof. Because of the diagonal block structure of $\boldsymbol{\Sigma}_{11} = \mathrm{var}[\mathbf{Z}_1]$ and the block structure of $\mathbf{A}_1$ we obtain

$$\mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_1 = \mathrm{diag}\left(\sum_{j=0}^{n-k} \mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k^{-1} \mathbf{V}_j^{1/2} \right)_{k \in \{0,\ldots,n\}}$$

and

$$\mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{Z}_1 = \left(\sum_{j=0}^{n-k} (\mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k^{-1} \mathbf{V}_j^{1/2}) \mathbf{V}_j^{-1} \mathbf{Z}_{j,k} \right)_{k \in \{0,\ldots,n\}}.$$

Now the Gauss-Markov Theorem for estimators yields

$$\hat{\boldsymbol{\beta}}^{\mathrm{GM}} = (\mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_1)^{-1} \mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{Z}_1 = \left(\left(\sum_{j=0}^{n-k} \mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k^{-1} \mathbf{V}_j^{1/2} \right)^{-1} \sum_{j=0}^{n-k} (\mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k^{-1} \mathbf{V}_j^{1/2}) \mathbf{V}_j^{-1} \mathbf{Z}_{j,k} \right)_{k \in \{0,\ldots,n\}}$$

and hence

$$\hat{\boldsymbol{\zeta}}_k^{\mathrm{GM}} = \left(\sum_{j=0}^{n-k} \mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k^{-1} \mathbf{V}_j^{1/2} \right)^{-1} \sum_{j=0}^{n-k} (\mathbf{V}_j^{1/2} \boldsymbol{\Sigma}_k^{-1} \mathbf{V}_j^{1/2}) \mathbf{V}_j^{-1} \mathbf{Z}_{j,k}$$

for all $k \in \{0, 1, \ldots, n\}$. □

The following result provides formulas for the Gauss-Markov predictors of the non-observable incremental losses and for the Gauss-Markov predictors of the non-observable cumulative losses:

4.2 Theorem. *For all* $i, k \in \{0, 1, \ldots, n\}$ *such that* $i + k \geq n + 1$, *the Gauss-Markov predictor* $\hat{\mathbf{Z}}_{i,k}^{\mathrm{GM}}$ *of* $\mathbf{Z}_{i,k}$ *satisfies*

$$\hat{\mathbf{Z}}_{i,k}^{\mathrm{GM}} = \mathbf{V}_i \, \hat{\boldsymbol{\zeta}}_k^{\mathrm{GM}}$$

and the Gauss-Markov predictor $\hat{\mathbf{S}}_{i,k}^{\mathrm{GM}}$ *of* $\mathbf{S}_{i,k}$ *satisfies*

$$\hat{\mathbf{S}}_{i,k}^{\mathrm{GM}} = \mathbf{S}_{i,n-i} + \mathbf{V}_i \sum_{l=n-i+1}^{k} \hat{\boldsymbol{\zeta}}_l^{\mathrm{GM}}.$$

Proof. Since $\boldsymbol{\Sigma}_{21} = \mathrm{cov}[\mathbf{Z}_1, \mathbf{Z}_2] = \mathbf{O}$, the first assertion is immediate from the Gauss-Markov Theorem for predictors and the second assertion follows from the final remark of

Section 3. □

The Gauss-Markov Theorem for predictors implies that

- the Gauss-Markov predictors of the sum of the non-observable incremental losses of a given accident year,
- the Gauss-Markov predictors of the sum of the non-observable incremental losses of a given calendar year, and
- the Gauss-Markov predictors of the sum of all non-observable incremental losses

are obtained by summation over the Gauss-Markov predictors of the corresponding single non-observable incremental losses.

For $i,k \in \{0,1,\ldots,n\}$ such that $i+k \geq n+1$, the estimators and predictors

$$\hat{\boldsymbol{\zeta}}_k^{\text{AD}} = \left(\sum_{j=0}^{n-k} \mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2}\right)^{-1} \sum_{j=0}^{n-k} (\mathbf{V}_j^{1/2}\boldsymbol{\Sigma}_k^{-1}\mathbf{V}_j^{1/2})\mathbf{V}_j^{-1}\mathbf{Z}_{j,k}$$

$$\hat{\mathbf{Z}}_{i,k}^{\text{AD}} = \mathbf{V}_i\hat{\boldsymbol{\zeta}}_k^{\text{AD}}$$

$$\hat{\mathbf{S}}_{i,k}^{\text{AD}} = \mathbf{S}_{i,n-i} + \mathbf{V}_i \sum_{l=n-i+1}^{k} \hat{\boldsymbol{\zeta}}_l^{\text{AD}}$$

are said to be the estimators and predictors of the *multivariate additive method.* Except for $m=1$ or $k=n$ they usually differ from the estimators and predictors

$$\hat{\boldsymbol{\zeta}}_k := \left(\sum_{j=0}^{n-k} \mathbf{V}_j\right)^{-1} \sum_{j=0}^{n-k} \mathbf{Z}_{j,k}$$

$$\tilde{\mathbf{Z}}_{i,k} := \mathbf{V}_i\hat{\boldsymbol{\zeta}}_k$$

$$\tilde{\mathbf{S}}_{i,k} := \mathbf{S}_{i,n-i} + \mathbf{V}_i \sum_{l=n-i+1}^{k} \tilde{\boldsymbol{\zeta}}_l$$

whose coordinates coincide with those of the univariate additive method.

4.2 Multivariate Chain-Ladder Model

Let us now consider a multivariate extension of the chain-ladder model which applies to the combined subportfolios and was proposed by Pröhl and Schmidt [2005]. This model is a slight but convenient extension of the model of Braun [2004]; see also Kremer [2005].

The multivariate chain-ladder model involves successive conditioning with respect to the σ-algebras $\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{n-1}$ where, for each $k \in \{0,1,\ldots,n\}$, the σ-algebra

$$\mathcal{G}_{k-1}$$

represents the information provided by the cumulative losses $\mathbf{S}_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$, which is at the same time the information provided by the incremental losses $\mathbf{Z}_{j,l}$ of accident years $j \in \{0, 1, \ldots, n-k+1\}$ and development years $l \in \{0, 1, \ldots, k-1\}$.

For all $i, k \in \{0, 1, \ldots, n\}$, we denote by

$$\mathbf{\Delta}_{i,k} := \text{diag}(\mathbf{S}_{i,k})$$

the diagonal random matrix whose diagonal elements are the coordinates of the random vector $\mathbf{S}_{i,k}$.

We assume that all coordinates of $\mathbf{S}_{i,k}$ are strictly positive. Then each $\mathbf{\Delta}_{i,k}$ is invertible and the identity

$$\mathbf{S}_{i,k} = \mathbf{\Delta}_{i,k-1}(\mathbf{\Delta}_{i,k-1}^{-1}\mathbf{S}_{i,k})$$

holds for all $i \in \{0, 1, \ldots, n\}$ and $k \in \{0, 1, \ldots, n\}$.

Multivariate Chain-Ladder Model: *For each* $k \in \{0, 1, \ldots, n\}$, *there exists a random parameter vector* $\mathbf{\Phi}_k$ *and a positive definite symmetric random matrix* $\mathbf{\Sigma}_k$ *such that*

$$E^{\mathcal{G}_{k-1}}[\mathbf{S}_{i,k}] = \mathbf{\Delta}_{i,k-1} - \mathbf{\Phi}_k$$

and

$$\text{cov}^{\mathcal{G}_{k-1}}[\mathbf{S}_{i,k}, \mathbf{S}_{j,k}] = \begin{cases} \mathbf{\Delta}_{i,k-1}^{1/2}\mathbf{\Sigma}_k\mathbf{\Delta}_{i,k-1}^{1/2} & \text{if } i = j \\ \mathbf{O} & \text{else} \end{cases}$$

holds for all $i, j \in \{0, 1, \ldots, n-k+1\}$.

In the subsequent discussion, we assume that the assumption of the multivariate chain-ladder model is fulfilled.

The multivariate chain-ladder model consists of n conditional linear models corresponding to the development years $k \in \{1, \ldots, n\}$. This can be seen as follows: Fix $k \in \{1, \ldots, n\}$, let $\mathbf{S}_1$ and $\mathbf{A}_1$ denote a block vector and a block matrix consisting of the random vectors $\mathbf{S}_{i,k}$ and the random matrices $\mathbf{\Delta}_{i,k}$ with $i \leq n-k$ (arranged in the same order) and let $\mathbf{S}_2 := \mathbf{S}_{n-k+1,k}$ and $\mathbf{A}_2 := \mathbf{\Delta}_{n-k+1,k}$. Then the random vectors $\mathbf{S}_1$ and $\mathbf{S}_2$ and the random matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ depend on k and we have

$$E^{\mathcal{G}_{k-1}}[\mathbf{S}_1] = \mathbf{A}_1\mathbf{\Phi}_k$$
$$E^{\mathcal{G}_{k-1}}[\mathbf{S}_2] = \mathbf{A}_2\mathbf{\Phi}_k.$$

Therefore, the multivariate chain-ladder model consists indeed of n conditional linear models.

The following result provides formulas for the Gauss-Markov estimators of the parameters in the multivariate chain-ladder model:

4.3 Theorem. *For each* $k \in \{1, \ldots, n\}$, *the Gauss-Markov estimator* $\hat{\mathbf{\Phi}}_k^{\text{GM}}$ *of* $\mathbf{\Phi}_k$ *satisfies*

$$\hat{\mathbf{\Phi}}_k^{\text{GM}} = \left(\sum_{j=0}^{n-k} \mathbf{\Delta}_{i,k-1}^{1/2}\mathbf{\Sigma}_k\mathbf{\Delta}_{i,k-1}^{1/2}\right)^{-1} \sum_{j=0}^{n-k} (\mathbf{\Delta}_{j,k-1}^{1/2}\mathbf{\Sigma}_k\mathbf{\Delta}_{j,k-1}^{1/2})\mathbf{\Delta}_{j,k-1}^{-1}\mathbf{S}_{j,k}.$$

Theorem 4.3 is immediate from the Gauss-Markov Theorem for estimators.

The following result provides formulas for the Gauss-Markov predictors of the cumulative losses of the first non-observable calendar year:

4.4 Theorem: *For each* $i \in \{1, \ldots, n\}$, *the Gauss-Markov predictor* $\hat{\mathbf{S}}_{i,n-i+1}^{\text{GM}}$ *of* $\mathbf{S}_{i,n-i+1}$ *satisfies*

$$\hat{\mathbf{S}}_{i,n-i+1}^{\text{GM}} = \mathbf{\Delta}_{i,n-i}\hat{\mathbf{\Phi}}_{i,n-i+1}^{\text{GM}}.$$

Theorem 4.4 is immediate from the Gauss-Markov Theorem for predictors.

For $i,k \in \{1, \ldots, n\}$ such that $i + k \geq n + 1$, the estimators and predictors

$$\hat{\mathbf{\Phi}}_k^{\text{CL}} := \left(\sum_{j=0}^{n-k} \mathbf{\Delta}_{j,k-1}^{1/2}\mathbf{\Sigma}_k^{-1}\mathbf{\Delta}_{i,k-1}^{1/2}\right)^{-1} \sum_{j=0}^{n-k} (\mathbf{\Delta}_{j,k-1}^{1/2}\mathbf{\Sigma}_k^{-1}\mathbf{\Delta}_{j,k-1}^{1/2})\mathbf{\Delta}_{j,k-1}^{-1}\mathbf{S}_{j,k}$$
$$\hat{\mathbf{S}}_{i,k}^{\text{CL}} := \hat{\mathbf{\Delta}}_{i,k-1}^{\text{CL}}\hat{\mathbf{\Phi}}_k^{\text{CL}}$$

with

$$\hat{\mathbf{\Delta}}_{i,k-1}^{\text{CL}} := \begin{cases} \text{diag}(\mathbf{S}_{i,n-i}) & \text{if } k = n - i + 1 \\ \text{diag}(\hat{\mathbf{S}}_{i,k-1}^{\text{CL}}) & \text{else} \end{cases}$$

are said to be the estimators and predictors of the *multivariate chain-ladder method.* Except for $m = 1$ or $k = n$ they usually differ from the estimators and predictors

$$\tilde{\mathbf{\Phi}}_k := \left(\sum_{j=0}^{n-k} \mathbf{\Delta}_{j,k-1}\right)^{-1} \sum_{j=0}^{n-k} \mathbf{S}_{j,k}$$
$$\tilde{\mathbf{S}}_{i,k} := \tilde{\mathbf{\Delta}}_{j,k}\tilde{\mathbf{\Phi}}_k$$

with

$$\hat{\mathbf{\Delta}}_{i,k-1} := \begin{cases} \text{diag}(\mathbf{S}_{i,n-i}) & \text{if } k = n-i+1 \\ \text{diag}(\tilde{\mathbf{S}}_{i,k-1}) & \text{else} \end{cases}$$

whose coordinates coincide with those of the *univariate chain-ladder method.*

In the case $i+k=n+1$, the multivariate chain-ladder predictors are justified by Theorem 4.4, but another justification is needed in the case $i+k \geq n+2$; this can be achieved by minimizing the $\mathcal{G}_{k-1}$-conditional expected prediction error over the collection of all predictors $\hat{\mathbf{S}}_{i,k}$ of $\mathbf{S}_{i,k}$ satisfying

$$\hat{\mathbf{S}}_{i,k} := \hat{\mathbf{\Delta}}_{i,k-1}^{\text{CL}} \hat{\mathbf{\Phi}}_k$$

for some $\mathcal{G}_{k-1}$-conditionally unbiased linear estimator $\hat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$; see Schmidt [1999b] for the univariate case. We have the following result:

4.5 Theorem. *For all $i,k \in \{1,\ldots,n\}$ such that $i+k \geq n+1$, the chain-ladder predictor $\hat{\mathbf{S}}_{i,k}^{\text{CL}}$ minimizes the $\mathcal{G}_{k-1}$-conditional expected prediction error over all predictors $\hat{\mathbf{S}}_{i,k}$ of $\mathbf{S}_{i,k}$ satisfying*

$$\hat{\mathbf{S}}_{i,k} := \hat{\mathbf{\Delta}}_{i,k-1}^{\text{CL}} \hat{\mathbf{\Phi}}_k$$

for some $\mathcal{G}_{k-1}$-conditionally unbiased linear estimator $\hat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$.

A proof of Theorem 4.5 will be given in the Appendix.

The optimality of the multivariate chain-ladder method guaranteed by Theorem 4.5 is sequential and one-step ahead. Of course, one would like to have a condition ensuring some kind of global optimality of the chain-ladder predictors; however, even in the univariate case, no such condition seems to be known.

To illustrate the situation without introducing additional notation, let us recall two results for the univariate case:

- The assumption of the univariate chain-ladder model is fulfilled in the model of Mack [1993] in which it is assumed that the accident years are independent and that the parameters φ_k and σ_k are non-random; see Schmidt and Schnaus [1996]. Under the assumptions of the model of Mack, it can be shown that all chain-ladder predictors are unbiased, but it can also be shown that many other predictors are unbiased as well. Therefore, unbiasedness does not distinguish the chain-ladder predictors among all other predictors.

– One might hope that the chain-ladder predictors minimize the $\mathcal{G}_{n-1}$-conditional expected squared predictor error over all predictors of the form

$$\hat{S}_{i,k} := S_{i,n-i} \prod_{l=n-i+1}^{k} \hat{\varphi}_l$$

where, for each $l \in \{n-i+1, \ldots, k\}$, $\hat{\varphi}_l$ is a $\mathcal{G}_l$-conditionally unbiased linear estimator of φ_l. Again, under the assumptions of the model of Mack, it has been shown in Schmidt [1997] that this kind of optimality may fail for the chain-ladder predictors.

Thus, even in the univariate case and under the stronger assumptions of the model of Mack, it remains an open question whether there exists a condition which is less restrictive than the sequential optimality criterion of Theorem 4.5 and still ensures some kind of global optimality of the chain-ladder predictors.

5. ADDITIVITY

Let $\mathbf{1}$ denote the m-dimensional vector with all coordinates being equal to 1. For $i,k \in \{0, 1, \ldots, n\}$ define

$$Z_{i,k} := \mathbf{1}'\mathbf{Z}_{i,k}$$
$$S_{i,k} := \mathbf{1}'\mathbf{S}_{i,k}.$$

We shall now study prediction of the non-observable incremental losses $Z_{i,k}$ and of the non-observable cumulative losses $S_{i,k}$ of the aggregate portfolio.

5.1 Multivariate Additive Model

In the multivariate additive model it is immediate from the Gauss-Markov Theorem for predictors that, for all $i,k \in \{0, 1, \ldots, n\}$ such that $i+k \geq n+1$, the Gauss-Markov predictor $\hat{Z}_{i,k}^{\text{GM}}$ of $Z_{i,k}$ and the Gauss-Markov predictor $\hat{S}_{i,k}^{\text{GM}}$ of $S_{i,k}$ satisfy

$$\hat{Z}_{i,k}^{\text{GM}} = \mathbf{1}'\hat{\mathbf{Z}}_{i,k}^{\text{AD}}$$
$$\hat{S}_{i,k}^{\text{GM}} = \mathbf{1}'\hat{\mathbf{S}}_{i,k}^{\text{AD}}.$$

This means that the Gauss-Markov predictors for the aggregate portfolio are obtained by summation over the Gauss-Markov predictors for the single lines of business. Therefore, the multivariate additive method is consistent in the sense that there is no problem of additivity.

Warning: One might believe that the Gauss-Markov predictors for the aggregate portfolio could also be obtained by applying the univariate additive method to the aggregate portfolio. This, however, is not the case since the multivariate additive model for the combined subportfolios does not lead to a univariate additive model for the aggregate portfolio.

5.2 Multivariate Chain-Ladder Model

In the multivariate chain-ladder model it is immediate from the Gauss-Markov Theorem for predictors that, for all $i \in \{1, \dots, n\}$, the Gauss-Markov predictor $\hat{S}^{GM}_{i,n-i+1}$ of $S_{i,n-i+1}$ satisfies

$$\hat{S}^{GM}_{i,n-i+1} = \mathbf{1}'\hat{\mathbf{S}}^{CL}_{i,n-i+1}.$$

This means that the Gauss-Markov predictors for the aggregate portfolio are obtained by summation over the multivariate Gauss-Markov predictors for the different lines of business. Moreover, it is easy to see that, for all $i,k \in \{0,1,\dots,n\}$ such that $i+k \geq n+2$, the predictor

$$\begin{aligned} S^{\bullet}_{i,k} &:= \mathbf{1}'\hat{\mathbf{S}}^{CL}_{i,k} \\ &= \mathbf{1}'\hat{\mathbf{\Delta}}^{CL}_{i,k-1}\hat{\mathbf{\Phi}}^{CL}_{k} \\ &= (\hat{\mathbf{S}}^{CL}_{i,k-1})'\hat{\mathbf{\Phi}}^{CL}_{k} \end{aligned}$$

minimizes the $\mathcal{G}_{k-1}$-conditional expected prediction error over all predictors $\hat{S}_{i,k}$ of $S_{i,k}$ satisfying

$$\begin{aligned} \hat{S}_{i,k} &:= \mathbf{1}'\hat{\mathbf{\Delta}}^{CL}_{i,k-1}\hat{\mathbf{\Phi}}_{k} \\ &= (\hat{\mathbf{S}}^{CL}_{i,k-1})'\hat{\mathbf{\Phi}}_{k} \end{aligned}$$

for some $\mathcal{G}_{k-1}$-conditionally unbiased linear predictor $\hat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$. Therefore, the multivariate chain-ladder method is consistent in the sense that there is no problem of additivity.

Warning: As in the case of the multivariate additive model, it would be a serious mistake to predict the non-observable cumulative losses of the aggregate portfolio on the basis of the observable cumulative losses of the aggregate portfolio since such an approach would ignore the correlation structure between the different lines of business; see Pröhl and Schmidt [2005].

6. ESTIMATION OF THE VARIANCE PARAMETERS

In the case $m = 1$, which is the univariate case, the variance parameters $\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_n$ drop out in the formulas for the Gauss-Markov predictors in the multivariate additive model and in the multivariate chain-ladder model.

In the case $m \geq 2$, only the variance parameter $\mathbf{\Sigma}_n$ drops out in the formulas for the Gauss-Markov predictors in the multivariate additive model and in the multivariate chain-ladder model; in this case, the variance parameters $\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_{n-1}$ must be estimated.

6.1 Multivariate Additive Model

Under the assumptions of the multivariate additive model and for $k \leq n-1$, the random matrix

$$\hat{\mathbf{\Sigma}}_k^{\mathrm{AD}} := \frac{1}{n-k} \sum_{j=0}^{n-k} \mathbf{V}_j^{-1/2} (\mathbf{Z}_{j,k} - \mathbf{V}_j \tilde{\boldsymbol{\zeta}}_k)(\mathbf{Z}_{j,k} - \mathbf{V}_j \tilde{\boldsymbol{\zeta}}_k)' \mathbf{V}_j^{-1/2}$$

is a positive semidefinite estimator of the positive definite matrix $\mathbf{\Sigma}_k$; moreover, its diagonal elements are unbiased estimators of the diagonal elements of $\mathbf{\Sigma}_k$ whereas its non-diagonal elements slightly underestimate the corresponding elements of $\mathbf{\Sigma}_k$.

Although unbiasedness of an estimator is usually considered to be desirable, this property would not be helpful in the present situation since any estimator of $\mathbf{\Sigma}_k$ has to be inverted and since the inverse of an unbiased estimator of $\mathbf{\Sigma}_k$ is very likely to be biased anyway. Moreover, the relative bias of the estimators proposed before can be shown to be very small.

By contrast, for any estimator of $\mathbf{\Sigma}_k$, the property of being positive semidefinite is a necessary, although not sufficient, condition for being positive definite and hence invertible. In fact, the estimator of $\mathbf{\Sigma}_k$ proposed before is always singular when $k \geq n - m + 2$ since in this case the dimension of the linear space generated by any realizations of the random vectors $\mathbf{V}_j^{1/2}(\mathbf{Z}_{i,k} - \mathbf{V}_j \tilde{\boldsymbol{\zeta}}_k)$ with $j \in \{0, 1, \ldots, n-k\}$ is at most $m - 1$ such that there exists at least one nonzero vector which is orthogonal to each of the realizations of these random vectors; moreover, the realizations of the random vectors $\mathbf{V}_j^{1/2}(\mathbf{Z}_{i,k} - \mathbf{V}_j \tilde{\boldsymbol{\zeta}}_k)$ may be linearly dependent also for some $k \leq n - m + 1$, which implies that the corresponding realization of the estimator of $\mathbf{\Sigma}_k$ proposed before may be singular also for some $k \leq n - m + 1$.

In practical applications, it is thus necessary to check whether the estimators proposed before are invertible or not, and to modify those estimators which are not invertible. Such modifications could be obtained by extrapolation or by the use of external information; see below.

6.2 Multivariate Chain-Ladder Model

Under the assumptions of the multivariate chain-ladder model and for $k \leq n-1$, the random matrix

$$\boldsymbol{\Sigma}_k^{\mathrm{CL}} := \frac{1}{n-k} \sum_{j=0}^{n-k} \boldsymbol{\Delta}_{j,k-1}^{1/2} (\mathbf{S}_{j,k} - \boldsymbol{\Delta}_{j,k-1} \hat{\boldsymbol{\Phi}}_k)(\mathbf{S}_{j,k} - \boldsymbol{\Delta}_{j,k-1} \hat{\boldsymbol{\Phi}}_k)' \boldsymbol{\Delta}_{i,k-1}^{1/2}$$

is a positive semidefinite estimator of the positive definite matrix $\boldsymbol{\Sigma}_k$; moreover, its diagonal elements are unbiased estimators of the diagonal elements of $\boldsymbol{\Sigma}_k$ whereas its non-diagonal elements slightly underestimate the corresponding elements of $\boldsymbol{\Sigma}_k$ and hence differ from the unbiased estimators proposed by Braun [2004].

The comments on the variance estimators proposed for the multivariate additive model apply as well to the variance estimators proposed for the multivariate chain-ladder model.

6.3 Extrapolation

In the case where the proposed estimators of the variances for late development years are singular or almost singular, it could be reasonable to replace these estimators with estimators obtained by extrapolation from the estimators for the first development years which are usually invertible.

6.4 Iteration

In both models, one may try to improve the estimators of the variances and hence the Gauss-Markov estimators of the parameters by iteration, as proposed by Kremer [2005]. However, the iterates of some of the estimators of the variances may again be singular, and it seems to be difficult to prove that the resulting empirical Gauss-Markov estimators of the parameters are indeed improved by iteration.

6.5 External Information

In both models, another possibility for the estimation of the variance parameters $\boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_{n-1}$ consists in the use of external information, which is not contained in the run-off triangle and could be obtained, e. g., from the run-off triangle of a similar portfolio or from market statistics.

7. REMARKS

Another bivariate model of loss reserving is the model of Quarg and Mack [2004]. Under the assumptions of their model, Quarg and Mack propose bivariate chain-ladder predictors for the paid and incurred cumulative losses of a single line of business with the aim of reducing the gap between the univariate chain-ladder predictors for the paid and incurred cumulative losses; see also Verdier and Klinger [2005] for a related model. None of these two models is contained in the multivariate models proposed in the present paper.

Since no conditions at all are imposed on the character of the different lines of business in the multivariate models presented here, the multivariate method and the multivariate chain-ladder method could, in principle, also be applied to the paid and incurred cumulative losses of a single line of business.

Let us finally note that the problem of additivity can also be solved in quite different models like credibility models; see Radtke and Schmidt [2004] and Schmidt [2004].

8. A NUMERICAL EXAMPLE

In this section we present a numerical example for the multivariate chain-ladder method in the case of $m = 2$ subportfolios and $n = 3$ development years.

8.1 The Data

The following run-off triangles contain the observable cumulative losses $S_{i,k}^{(1)}$, $S_{i,k}^{(2)}$, and $S_{i,k}$ of the two subportfolios and of the aggregate portfolio, respectively:

Subportfolio 1

AY	DY			
	0	1	2	3
0	2423	3123	3567	3812
1	2841	3422	3952	
2	3700	3977		
3	5231			

Subportfolio 2

AY	DY 0	1	2	3
0	3546	6578	7650	8123
1	4001	7566	8822	
2	4040	7813		
3	4300			

Aggregate Portfolio

AY	DY 0	1	2	3
0	5969	9701	11217	11935
1	6842	10988	12774	
2	7740	11790		
3	9531			

8.2 Univariate Chain-Ladder Method

Applying the univariate chain-ladder method to each of these run-off triangles yields the univariate chain-ladder factors (CLF) and the univariate chain-ladder predictors of the non-observable cumulative losses:

Subportfolio 1

AY	DY 0	1	2	3
0	2423	3123	3567	3812
1	2841	3422	3952	4223
2	3700	3977	4569	4883
3	5231	6140	7054	7538
CLF		1.1738	1.1488	1.0687

Subportfolio 2

AY	DY			
	0	1	2	3
0	3546	6578	7650	8123
1	4001	7566	8822	9367
2	4040	7813	9099	9662
3	4300	8148	9490	10076
CLF		1.8950	1.1646	1.0618

Aggregate Portfolio

AY	DY			
	0	1	2	3
0	5969	9701	11217	11935
1	6842	10988	12774	13592
2	7740	11790	13672	14547
3	9531	15063	17467	18585
CLF		1.8950	1.1646	1.0618

8.3 Multivariate Chain-Ladder Method

We now combine the run-off triangles of the two subportfolios into a single run-off triangle which contains the vectors $\mathbf{S}_{i,k}$ of cumulative losses:

Combined Subportfolios

AY	DY			
	0	1	2	3
0	$\begin{pmatrix}2423\\3546\end{pmatrix}$	$\begin{pmatrix}3123\\6578\end{pmatrix}$	$\begin{pmatrix}3567\\7650\end{pmatrix}$	$\begin{pmatrix}3812\\8123\end{pmatrix}$
1	$\begin{pmatrix}2841\\4001\end{pmatrix}$	$\begin{pmatrix}3422\\7566\end{pmatrix}$	$\begin{pmatrix}3952\\8822\end{pmatrix}$	
2	$\begin{pmatrix}3700\\4040\end{pmatrix}$	$\begin{pmatrix}3977\\7813\end{pmatrix}$		
3	$\begin{pmatrix}5231\\4300\end{pmatrix}$			

Transforming the vectors $\mathbf{S}_{i,k}$ of cumulative losses into diagonal matrices, we obtain the following run-off triangle for the matrices $\boldsymbol{\Delta}_{i,k} = \text{diag}(\mathbf{S}_{i,k})$ which is completed by the vectors $\tilde{\boldsymbol{\Phi}}_k$ of univariate chain-ladder factors:

Combined Subportfolios

AY	DY 0	DY 1	DY 2	DY 3
0	$\begin{pmatrix} 2423 & 0 \\ 0 & 3546 \end{pmatrix}$	$\begin{pmatrix} 3123 & 0 \\ 0 & 6578 \end{pmatrix}$	$\begin{pmatrix} 3567 & 0 \\ 0 & 7650 \end{pmatrix}$	$\begin{pmatrix} 3812 & 0 \\ 0 & 8123 \end{pmatrix}$
1	$\begin{pmatrix} 2841 & 0 \\ 0 & 4001 \end{pmatrix}$	$\begin{pmatrix} 3422 & 0 \\ 0 & 7566 \end{pmatrix}$	$\begin{pmatrix} 3952 & 0 \\ 0 & 8822 \end{pmatrix}$	
2	$\begin{pmatrix} 3700 & 0 \\ 0 & 4040 \end{pmatrix}$	$\begin{pmatrix} 3977 & 0 \\ 0 & 7813 \end{pmatrix}$		
3	$\begin{pmatrix} 5231 & 0 \\ 0 & 4300 \end{pmatrix}$			
$\tilde{\mathbf{\Phi}}_k$		$\begin{pmatrix} 1.1738 \\ 1.8950 \end{pmatrix}$	$\begin{pmatrix} 1.1488 \\ 1.1646 \end{pmatrix}$	$\begin{pmatrix} 1.0687 \\ 1.0618 \end{pmatrix}$

For the estimators of the variances we thus obtain

$$\hat{\mathbf{\Sigma}}_1^{\mathrm{CL}} = \begin{pmatrix} 35.4968 & -14.3861 \\ -14.3861 & 5.9200 \end{pmatrix}$$

$$\hat{\mathbf{\Sigma}}_2^{\mathrm{CL}} = \begin{pmatrix} 0.2637 & 0.0926 \\ 0.0926 & 0.0325 \end{pmatrix}$$

and hence

$$\left(\hat{\mathbf{\Sigma}}_1^{\mathrm{CL}}\right)^{-1} = \begin{pmatrix} 1.8616 & 4.5239 \\ 4.5239 & 11.1624 \end{pmatrix}$$

$$\left(\hat{\mathbf{\Sigma}}_2^{\mathrm{CL}}\right)^{-1} = \begin{pmatrix} 25876.4330 & -73727.6467 \\ -73727.6467 & 210097.0596 \end{pmatrix}$$

Note that estimators of the variances $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_3$ are not needed. Applying the multivariate chain-ladder method to the combined subportfolios yields the multivariate chain-ladder predictors of the non-observable cumulative losses:

Combined Subportfolios

AY	DY 0	1	2	3
0	$\begin{pmatrix}2423\\3546\end{pmatrix}$	$\begin{pmatrix}3123\\6578\end{pmatrix}$	$\begin{pmatrix}3567\\7650\end{pmatrix}$	$\begin{pmatrix}3812\\8123\end{pmatrix}$
1	$\begin{pmatrix}2841\\4001\end{pmatrix}$	$\begin{pmatrix}3422\\7566\end{pmatrix}$	$\begin{pmatrix}3952\\8822\end{pmatrix}$	$\begin{pmatrix}4223\\9367\end{pmatrix}$
2	$\begin{pmatrix}3700\\4040\end{pmatrix}$	$\begin{pmatrix}3977\\7813\end{pmatrix}$	$\begin{pmatrix}4569\\9099\end{pmatrix}$	$\begin{pmatrix}4883\\9661\end{pmatrix}$
3	$\begin{pmatrix}5231\\4300\end{pmatrix}$	$\begin{pmatrix}6105\\8167\end{pmatrix}$	$\begin{pmatrix}7013\\9512\end{pmatrix}$	$\begin{pmatrix}7495\\10100\end{pmatrix}$
$\hat{\boldsymbol{\Phi}}_k$		$\begin{pmatrix}1.1670\\1.8994\end{pmatrix}$	$\begin{pmatrix}1.1489\\1.1646\end{pmatrix}$	$\begin{pmatrix}1.0687\\1.0618\end{pmatrix}$

8.4 Comparison

Predictors for non-observable aggregate cumulative losses may be computed by the following three methods:

- **Method A:** Apply the univariate chain-ladder method to the aggregate portfolio.
- **Method B:** Apply the univariate chain-ladder method to each of the subportfolios and take sums of the univariate predictors.
- **Method C:** Apply the multivariate chain-ladder method to the combined subportfolios and take sums of the multivariate predictors.

For example, for the ultimate aggregate cumulative loss of accident year 3,

- Method A yields the value 18585.
- Method B yields the value $7538 + 10076 = 17614$.
- Method C yields the value $7495 + 10100 = 17595$.

The following table presents several reserves obtained by these three methods:

Reserve	Method A	Method B	Method C
Accident Year 1	818	817	817
Accident Year 2	2757	2754	2754
Accident Year 3	9054	8084	8064
Total	12628	11655	11635
Calendar Year 4	8231	7452	7436
Calendar Year 5	3279	3131	3129
Calendar Year 6	1118	1071	1070
Total	12628	11655	11635

Due to round-off errors, some of the total reserves differ slightly from the sums of the reserves over accident years of calendar years. In the present example, the results obtained by Methods B and C are quite similar, but they differ considerably from those obtained by Method A.

8.5 Preliminary Conclusions

Of course, one should not draw general conclusions from a single numerical example. Nevertheless, the present example and experience with other sets of data justify the following rule of thumb:

- Method C is optimal when the model assumptions and the optimality criteria for the multivariate chain-ladder method can be accepted.
- Method B may in many cases provide a reasonable approximation to Method C.
- Method A may be disastrous since it ignores correlation between the different lines of business.

Experience with other sets of data also indicates that the similarities and differences between the three methods may vary with

- the lines of business under consideration,
- the number of lines of business, and
- the number of development years.

It is therefore indispensable for the actuary to acquire practical experience for every combined portfolio of interest.

APPENDIX

Here we give a proof of Theorem 4.5.

Proof. Consider any $\mathcal{G}_{k-1}$-conditionally unbiased linear estimator $\hat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$. Then there exist $\mathcal{G}_{k-1}$-measurable matrices $\mathbf{Q}_{0,k-1}, \mathbf{Q}_{1,n-1}, \ldots, \mathbf{Q}_{n-k,k-1}$ satisfying

$$\hat{\mathbf{\Phi}}_k = \sum_{j=0}^{n-k} \mathbf{Q}_{j,k-1} \mathbf{S}_{j,k}$$

and $\sum_{j=0}^{n-k} \mathbf{Q}_{j,k-1} \mathbf{\Delta}_{j,k-1} = \mathbf{I}$. Also, letting

$$\mathbf{Q}_{j,k-1}^{\mathrm{CL}} := \left(\sum_{s=0}^{n-k} \mathbf{\Delta}_{s,k-1}^{1/2} \mathbf{\Sigma}_k^{-1} \mathbf{\Delta}_{s,k-1}^{1/2} \right)^{-1} (\mathbf{\Delta}_{j,k-1}^{1/2} \mathbf{\Sigma}_k^{-1} \mathbf{\Delta}_{j,k-1}^{1/2}) \mathbf{\Delta}_{j,k-1}^{-1}$$

we obtain

$$\hat{\mathbf{\Phi}}_k^{\mathrm{CL}} = \sum_{j=0}^{n-k} \mathbf{Q}_{j,k-1}^{\mathrm{CL}} \mathbf{S}_{j,k}$$

and $\sum_{j=0}^{n-k} \mathbf{Q}_{j,k-1}^{\mathrm{CL}} \mathbf{\Delta}_{j,k-1} = \mathbf{I}$. We thus obtain

$$\sum_{j=0}^{n-k} (\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}}) \mathbf{\Delta}_{j,k-1} = \mathbf{O}.$$

Since

$$\mathbf{Q}_{j,k-1}^{\mathrm{CL}} := \left(\sum_{s=0}^{n-k} \mathbf{\Delta}_{s,k-1}^{1/2} \mathbf{\Sigma}_k^{-1} \mathbf{\Delta}_{s,k-1}^{1/2} \right)^{-1} \mathbf{\Delta}_{j,k-1} \left(\mathrm{var}^{\mathcal{G}_{k-1}}[\mathbf{S}_{i,k}] \right)^{-1}$$

this yields

$$\begin{aligned}
\mathrm{cov}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{\Phi}}_k - \hat{\mathbf{\Phi}}_k^{\mathrm{CL}}, \hat{\mathbf{\Phi}}_k^{\mathrm{CL}} \right] &= \sum_{j=0}^{n-k} \sum_{l=0}^{n-k} \left(\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}} \right) \mathrm{cov}^{\mathcal{G}_{k-1}}\left[\mathbf{S}_{j,k}, \mathbf{S}_{l,k} \right] \left(\mathbf{Q}_{l,k-1}^{\mathrm{CL}} \right)' \\
&= \sum_{j=0}^{n-k} \left(\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}} \right) \mathrm{var}^{\mathcal{G}_{k-1}}\left[\mathbf{S}_{j,k} \right] \left(\mathbf{Q}_{j,k-1}^{\mathrm{CL}} \right)' \\
&= \sum_{j=0}^{n-k} \left(\mathbf{Q}_{j,k-1} - \mathbf{Q}_{j,k-1}^{\mathrm{CL}} \right) \mathbf{\Delta}_{j,k-1} \left(\sum_{s=0}^{n-k} \mathbf{\Delta}_{s,k-1}^{1/2} \mathbf{\Sigma}_k^{-1} \mathbf{\Delta}_{s,k-1}^{1/2} \right)^{-1} \\
&= \mathbf{O}.
\end{aligned}$$

Since $i + k \geq n + 1$, we also have $\mathrm{cov}^{\mathcal{G}_{k-1}}\left[\mathbf{S}_{j,k}, \mathbf{S}_{i,k} \right] = \mathbf{O}$ and thus

$$\begin{aligned}\mathrm{cov}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}-\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}},\mathbf{S}_{i,k}\right] &= \mathrm{cov}^{\mathcal{G}_{k-1}}\left[\hat{\boldsymbol{\Delta}}_{i,k-1}^{\mathrm{CL}}\hat{\boldsymbol{\Phi}}_k-\hat{\boldsymbol{\Delta}}_{i,k-1}^{\mathrm{CL}}\hat{\boldsymbol{\Phi}}_k^{\mathrm{CL}},\mathbf{S}_{i,k}\right]\\ &= \hat{\boldsymbol{\Delta}}_{i,k-1}^{\mathrm{CL}}\,\mathrm{cov}^{\mathcal{G}_{k-1}}\left[\hat{\boldsymbol{\Phi}}_k-\hat{\boldsymbol{\Phi}}_k^{\mathrm{CL}},\mathbf{S}_{i,k}\right]\\ &= \hat{\boldsymbol{\Delta}}_{i,k-1}^{\mathrm{CL}}\sum_{j=0}^{n-k}\left(\mathbf{Q}_{j,k-1}-\mathbf{Q}_{j,k-1}^{\mathrm{CL}}\right)\mathrm{cov}^{\mathcal{G}_{k-1}}\left[\mathbf{S}_{j,k},\mathbf{S}_{i,k}\right]\\ &= \mathbf{O}.\end{aligned}$$

Using the two identities established before, we thus obtain

$$\begin{aligned}\mathrm{cov}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}-\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}},\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}-\mathbf{S}_{i,k}\right] &= \mathrm{cov}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}-\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}},\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}\right]\\ &= \hat{\boldsymbol{\Delta}}_{i,k-1}^{\mathrm{CL}}\,\mathrm{cov}^{\mathcal{G}_{k-1}}\left[\hat{\boldsymbol{\Phi}}_k-\hat{\boldsymbol{\Phi}}_k^{\mathrm{CL}},\hat{\boldsymbol{\Phi}}_k^{\mathrm{CL}}\right]\hat{\boldsymbol{\Delta}}_{i,k-1}^{\mathrm{CL}}\\ &= \mathbf{O}\end{aligned}$$

and hence

$$\mathrm{var}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}-\mathbf{S}_{i,k}\right]=\mathrm{var}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}-\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}\right]+\mathrm{var}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}-\mathbf{S}_{i,k}\right].$$

We thus obtain

$$\begin{aligned}E^{\mathcal{G}_{k-1}}\left[\left(\hat{\mathbf{S}}_{i,k}-\mathbf{S}_{i,k}\right)'\left(\hat{\mathbf{S}}_{i,k}-\mathbf{S}_{i,k}\right)\right] &= \mathrm{trace}\left(\mathrm{var}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}-\mathbf{S}_{i,k}\right]\right)\\ &= \mathrm{trace}\left(\mathrm{var}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}-\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}\right]\right)+\mathrm{trace}\left(\mathrm{var}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}-\mathbf{S}_{i,k}\right]\right)\\ &\geq \mathrm{trace}\left(\mathrm{var}^{\mathcal{G}_{k-1}}\left[\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}-\mathbf{S}_{i,k}\right]\right)\\ &= E^{\mathcal{G}_{k-1}}\left[\left(\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}-\mathbf{S}_{i,k}\right)'\left(\hat{\mathbf{S}}_{i,k}^{\mathrm{CL}}-\mathbf{S}_{i,k}\right)\right]\end{aligned}$$

which proves the theorem. □

ACKNOWLEDGEMENT

I am grateful to Leigh Halliwell for his comments on earlier drafts of this paper and to Mathias Zocher who provided the numerical example.

REFERENCES

[1] Ajne, Björn [1994]: "Additivity of chain-ladder projections", *ASTIN Bulletin, Vol. 24*, 313-318.

[2] Braun, Christian [2004]: "The prediction error of the chain-ladder method applied to correlated run-off triangles", *ASTIN Bulletin, Vol. 34*, 399-423.

[3] Goldberger, Arthur S. [1962]: "Best linear unbiased prediction in the generalized linear regression

model", *J. American Statistical Association, Vol. 57*, 369-375.

[4] Hachemeister, Charles A., and James N. Stanard [1975]: "IBNR claims count estimation with static lag functions", Unpublished.

[5] Halliwell, Leigh J. [1996]: "Loss prediction by generalized least squares", *PCAS, Vol. LXXXIII*, 436-489.

[6] Halliwell, Leigh J. [1997]: "Conjoint prediction of paid and incurred losses", *CAS Forum Summer 1997, Vol. I*, 241-379.

[7] Halliwell, Leigh J. [1999]: "Loss prediction by generalized least squares – Author's response", *PCAS, Vol. LXXXVI*, 764-769.

[8] Hamer, Michael D. [1999]: "Loss prediction by generalized least squares – Discussion", *PCAS, Vol. LXXXVI*, 748-763.

[9] Hess, Klaus T., and Klaus D. Schmidt [2002]: "A comparison of models for the chain-ladder method", *Insurance Mathematics and Economics, Vol. 31*, 351-364.

[10] Hess, Klaus T., Klaus D. Schmidt and Mathias Zocher [2006]: "Multivariate loss reserving in the multivariate additive model", *Insurance Mathematics and Economics* (in press).

[11] Holmberg, Randall D. [1994]: "Correlation and the measurement of loss reserve variability", *CAS Forum Spring 1994*, 247-278.

[12] Klemmt, Heinz J. [2004]: "Trennung von Schadenarten und Additivität bei Chain-Ladder Prognosen", Paper presented at the 2004 Fall Meeting of the German ASTIN Group in München.

[13] Kremer, Erhard [2005]: "The correlated chain-ladder method for reserving in case of correlated claims developments", *Blätter DGVFM, Vol. 27*, 315-322.

[14] Mack, Thomas [1993]: "Distribution-free calculation of the standard error of chain-ladder reserve estimates", *ASTIN Bulletin, Vol. 23*, 213-225.

[15] Pröhl, Carsten, and Klaus D. Schmidt [2005]: "Multivariate chain-ladder", *Dresdner Schriften zur Versicherungsmathematik 3/2005*, Paper presented at the International ASTIN Colloquium 2005 in Zürich.

[16] Quarg, Gerhard, and Thomas Mack [2004]: "Munich chain-ladder – A reserving method that reduces the gap between IBNR projections based on paid losses and IBNR projections based on incurred losses", *Blätter DGVFM, Vol. 26*, 597-630.

[17] Radtke, Michael, and Klaus D. Schmidt (eds.) [2004]: *Handbuch zur Schadenreservierung*, Karlsruhe, Verlag Versicherungswirtschaft.

[18] Rao, C. Radharkrischna, and Helge Toutenburg [1995]: *Linear Models – Least Squares and Alternatives*, Berlin - Heidelberg - New York, Springer.

[19] Schmidt, Klaus D. [1997]: "Non-optimal prediction by the chain-ladder method", *Insurance, Mathematics and Economics, Vol. 21*, 17-24.

[20] Schmidt, Klaus D. [1998]: "Prediction in the linear model – A direct approach", *Metrika, Vol. 48*, 141-147.

[21] Schmidt, Klaus D. [1999a]: "Loss prediction by generalized least squares – Discussion", *PCAS, Vol. LXXXVI*, 736-747.

[22] Schmidt, Klaus D. [1999b]: "Chain ladder prediction and asset liability management", *Blätter DGVM, Vol. 24*, 1-9.

[23] Schmidt, Klaus D. [2004]: "Prediction", *Encyclopedia of Actuarial Science, Vol. 3*, 1317-1321, Chichester, Wiley.

[24] Schmidt, Klaus D., and Anja Schnaus [1996]: "An extension of Mack's model for the chain-ladder method", *ASTIN Bulletin, Vol. 26*, 247-262.

[25] Verdier, Bertrand, and Artur Klinger [2005]: "JAB chain – A model based calculation of paid and incurred loss development factors", Paper presented at the International ASTIN Colloquium 2005 in Zürich.

Biography of the Author:

The author is professor for actuarial mathematics at Technische Universität Dresden. His main interests are mathematical models in loss reserving and reinsurance as well as linear and credibility models in non-life insurance. He has a degree in mathematics and economics from Universität Zürich and a PhD in mathematics from Universität Mannheim. He is an academic correspondent of the CAS and a member of the DAV (Deutsche Aktuar-Vereinigung) with particular engagement in working parties on education and loss reserving. Besides his publications on loss reserving, the author has compiled a *Bibliography on Loss Reserving* which is available on the Web under http://www.math.tu-dresden.de/sto/schmidt/dsvm/reserve.pdf.

CONTENTS

A Nonlinear Regression Model of Incurred But Not Reported Losses

Scott Stelljes, ACAS, MAAA

Abstract

The process of loss development has been studied by casualty actuaries for many years. When an accident period is closed, the ultimate claim liabilities are unknown because many of the claims are still unreported and some that are reported remain unsettled. The difference between ultimate losses and reported losses is known as "Incurred But Not Reported" loss or IBNR. The reserve for IBNR losses is the largest liability on an insurer's balance sheet. Quantifying the uncertainty in estimates of IBNR is of great importance to the financial health of casualty insurance companies.

Most of the current methods for estimating ultimate losses focus on estimation of loss development factors which relate the emergence of losses to the amount of losses already reported. This paper presents a model for predicting incremental losses as a function of exposures, calendar period and development age.

A nonlinear regression model is used for estimating the 95% confidence interval of IBNR for an accident period. The model predicts the incremental pure premium for a development interval as a function of development age, calendar quarter and exposure. The estimated IBNR is the sum of forecasted incremental pure premiums. The regression model produces confidence interval estimates for the model parameters and for IBNR.

The regression model is applied to trended losses. We assume that the trend has been estimated by some reasonable time series method that produces confidence interval estimates of trend factors. Many good methods are available. We use the confidence interval estimate of the trend factors to adjust the IBNR estimates for uncertainty in loss trend.

The model presented here assumes normally distributed residuals. Although the underlying loss severities are probably not normal, the central limit theorem implies that this assumption would be appropriate if the number of claims is large. Thus, the model will most likely work well for high frequency lines of business such as personal auto.

We will present methods for estimating parameters, confidence intervals for the parameters, and the distribution of IBNR. These methods will be illustrated using simulated automobile bodily injury liability data. Model predictions will be compared to actual emerged losses.

Based on a comparison of predicted IBNR to the "actual" IBNR from the simulated data, the model appears to produce unbiased predictions and reasonable confidence interval estimates of IBNR. We conclude that the distribution of incremental pure premiums is close to normal and there is not a significant correlation between development age intervals. Thus, traditional regression methods can be used to estimate the distribution of forecasted incremental pure premiums and consequently, IBNR.

Keywords: Non-linear regression, IBNR, reserving.

1. INTRODUCTION

Many actuaries and their clients are unsatisfied with point estimates of IBNR reserves. Better decisions can be made if one has a range of possible outcomes and associated probabilities. Confidence interval estimates would satisfy this need. We will introduce a nonlinear regression model that produces confidence interval estimates of IBNR. The models are fitted to incremental pure premiums - the incremental change in case incurred (or paid) losses for an accident period during a development interval divided by the corresponding calendar period earned exposures. This approach was inspired by Buhlman's complementary loss ratio method as presented by Stanard [3].

1.1 Research Context

The context of this paper is reserving methods and reserving uncertainty and ranges.

1.2 Objective

The objective of this research is to produce a model of loss development that models losses as a function of exposures, can be applied to either paid or incurred losses, and produces a confidence interval estimate of IBNR.

The current literature includes some papers, e.g., Murphy [1] that present regression models to predict age-to-age loss development factors and measure the uncertainty in the predicted factors but there are very few that present models of loss dollars. Barnett and Zehnwirth [2] is an excellent example of a dollar based model, but it is applied to the logarithms of incremental losses and this becomes a problem when there is negative loss development. Recoveries lead to negative paid development and case reserve estimation errors can result in negative case development. In order to use a log link, it is necessary to discard information. Less information is discarded if the analysis is performed on paid losses but much of the data in the tail of a case incurred development triangle is negative. Many reserving actuaries believe that there is useful information in case incurred losses and they often compare estimates derived from paid and incurred data.

Furthermore, Narayan [4] remarks that dollar based regression models do not take into account changing levels of exposure. This is a serious flaw because the amount of loss in an accident period is highly correlated to the number of earned exposures.

Thus, there is a need for a dollar based regression model that can be applied without using a log link and that makes appropriate adjustments for changing levels of exposure.

In this paper, we present a nonlinear regression model that predicts incremental pure premiums as a function of development age. The model is applied to losses that have been adjusted for loss trend using a separate trend model. The trend model can be any time series model that produces confidence intervals for future trend factors. In the examples, we assume that future trend is represented by a geometric Brownian motion process but this is not necessarily the only model for future loss trend. Adjusting losses for trend is not necessary in a link ratio method because future development is predicted as a function of case or paid losses. The link ratios are multiplied by losses which are already stated at the appropriate cost level. The factors produced by our model are applied to exposures so it is necessary to adjust losses for trend.

The model presented in this paper does not require the use of any link function, so it can be applied either to paid or case incurred loss data. Furthermore, since we use pure premiums with exposure weights, the model relates losses to exposures.

1.3 Outline

The remainder of the paper proceeds as follows.

Section 2: Presentation of data. A simulated data set including a loss triangle and earned exposures is presented along with some observations. The nonlinear model is presented and the estimation of parameters is explained.

Section 3: The model is fitted to the simulated data and used to produce confidence interval estimates of ultimate incurred losses for each accident quarter. An analysis of residuals is presented.

Section 4: Conclusions.

Section 5: References.

2. BACKGROUND AND METHODS

A nonlinear regression model will be presented and used to analyze simulated loss development data. The model will be fitted to incremental pure premiums. The incremental pure premium for an accident quarter/development quarter is defined as the change in case incurred loss during the development quarter divided by the calendar quarter earned exposures.

In section 2.1, we will present the simulated loss development data. The data was simulated based on method 4 in Narayan [4] with some modifications. See Appendix B for a description of the method used to simulate the data. Narayan and other authors simulated thousands of sets of data for the purpose of comparing methods. We simulated a single triangle for the purpose of showing sample calculations. The simulation is not intended to validate the model. The simulated data is intended to resemble personal auto bodily injury data in accident quarter/development quarter format.

Section 2.2 is a presentation of the nonlinear regression model.

In section 2.3, we present the mathematics of estimating confidence intervals for the model parameters and IBNR.

2.1 Loss Development Data

Exhibit 2.1.1 shows a small portion of the simulated loss data in the traditional triangular array. The losses shown in Table 1 are cumulative case incurred losses. I.e., the amount shown for each development quarter is the sum of all paid losses from the beginning of the accident quarter through the end of the development quarter and the outstanding case reserves as of the end of the development quarter. The column to the left of the losses shows earned exposures. The second table shows the incremental pure premiums. These are the incremental losses divided by earned exposures. For example, the entry for accident quarter 1, development interval 1-2 is (1,713,179-1,244,722)/50,333.

EXHIBIT 2.1.1

Tabl 1. Cumulative Losses by Accident Quarter and Development Age

Accident Quarter	Earned Exposures	Development Age 1	2	3	4	5
1	50,333	1,244,722	1,713,179	1,996,372	2,065,006	2,166,446
2	50,801	1,417,101	2,004,222	2,341,886	2,437,727	
3	51,187	1,143,473	1,646,289	2,130,201		
4	51,146	1,055,290	2,268,788			
5	51,527	1,508,450				

Table 2. Incremental Pure Premiums

Accident Quarter	Earned Exposures	Development Interval 0-1	1-2	2-3	3-4	4-5
1	50,333	24.73	9.31	5.63	1.36	2.02
2	50,801	27.90	11.56	6.65	1.89	
3	51,187	22.34	9.82	9.45		
4	51,146	20.63	23.73			
5	51,527	29.27				

Exhibit 2.1.2 shows the averages and variances and Pearson correlations of incremental pure premiums by development age for some simulated data. The data exhibits a typical loss development pattern. We see that the average incremental pure premiums start high and decrease rapidly as the development age increases, converging to zero. There are some negative incremental losses resulting from recoveries, settling of claims for less than the case reserve, and reductions to case reserves. The table also shows that the variance decreases as development age increases. Thus, most of the uncertainty in loss development is in the early stages. The correlation matrix shows that the correlation of incremental pure premiums between different ages is usually insignificant.

EXHIBIT 2.1.2

Sample of Simulated Incremental Pure Premiums - Ages 1-7

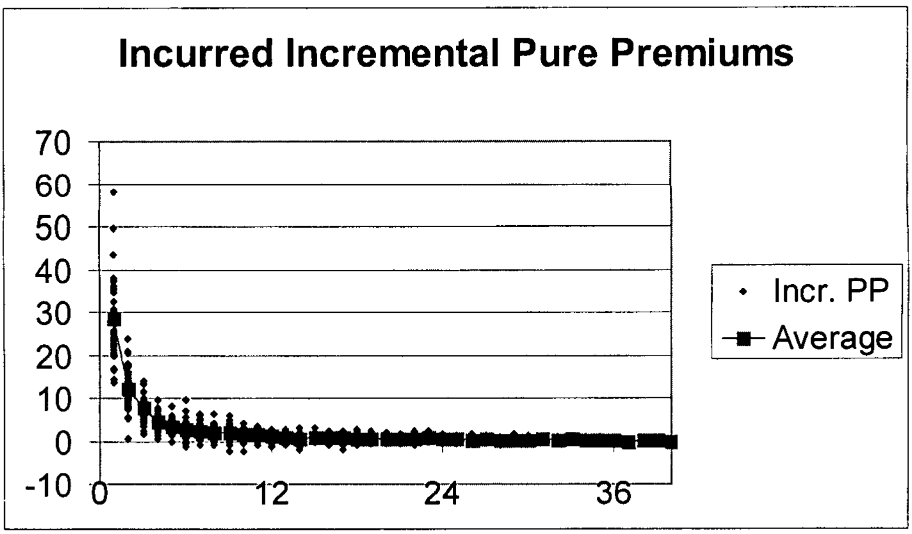
Age	0-1	1-2	2-3	3-4	4-5	5-6	6-7
Average	28.70	12.13	7.67	4.67	3.49	2.84	2.48
Variance	87.26	23.19	8.40	3.99	2.81	4.33	3.24

Pearson Correlations

	0-1	1-2	2-3	3-4	4-5	5-6	6-7
0-1	1.00	0.38	0.13	0.45	0.14	0.63	0.21
1-2		1.00	0.31	0.44	-0.01	0.29	0.25
2-3			1.00	0.15	0.11	-0.01	0.15
3-4				1.00	0.47	0.45	0.20
4-5					1.00	0.16	0.00
5-6						1.00	0.11
6-7							1.00

Exhibit 2.1.3 shows a scatter plot of the incremental pure premiums and the average incremental pure premiums by age.

EXHIBIT 2.1.3



In the scatter plot, the incremental pure premiums appear to be distributed around the average symmetrically. This and the fact that the correlations are not significant imply that the data fits the assumptions of regression models as stated in [5] reasonably well. The nonconstancy of the variances is a violation of the assumptions underlying ordinary regression but that problem can be solved by using a weighted regression model.

A weighted regression model is one in which a weight is assigned to each observation in the data. The more weight given to an observation, the more influence it has on the parameter estimates. We need to use a weight function that is inversely proportional to the variance of the data. It would also be advantageous to obtain exposure weighted parameter estimates. So, we will use weights that are a function of development age and exposures.

We will now define some of the variables that will be used in the analysis. First, the accident quarter will be represented by t which will take values of 1, 2, ..., 40. The development quarters will be represented by x which will be assigned the value of the development age (in quarters) at the end of the interval. For example, $x = 1$ will correspond to the 0-3 months development interval. Calendar quarters will be represented by u and will be calculated as $u = t + x - 1$. The incremental losses for accident quarter t and development interval x will be represented by $L_{t,x}$. Car months will be represented by c_t.

Appendix A shows the full set of simulated loss development data.

2.2 The Model

Our model of incremental pure premiums is a nonlinear regression model. Nonlinear regression models are statistical models of the form:

$$y = f\left(\vec{x}, \vec{\theta}\right) + \varepsilon \tag{2.2.1}$$

In (2.2.1), $\vec{x}$ is a vector of predictor variables, $\vec{\theta}$ is a vector of parameters, f is a nonlinear function, and ε is a normal random variable with mean 0. Usually, ε is assumed to have a constant variance σ^2. If the variance of the error term is not constant, a weight function that is inversely proportional to the variance may be specified.

The parameters of a nonlinear regression model are estimated by solving the normal equations. This usually requires using a numerical method such as the Gauss-Newton algorithm.

There are many commercial statistical software packages available that will perform the calculations and also provide approximate confidence intervals for the parameters and for predicted observations. The SAS system was used to perform the calculations to estimate confidence intervals for the model parameters and predicted IBNR.

We fit the following model to the incremental incurred pure premium data:

$$y = \left[\alpha \exp(\beta x) + \gamma \exp(\delta x)\right] + \frac{1}{w}\varepsilon \tag{2.2.2}$$

where $y = \frac{L_{t,x}}{c_t} \cdot \exp(ru)$ is the incremental pure premium for accident quarter t in development age interval x adjusted for loss trend. u represents the calendar quarter. r is the loss trend. α, β, γ, and δ are the model parameters. $w = x^{1.5} \cdot c_t$ is the weight function. This weight function was selected based on an analysis of the residuals from an unweighted regression model.

We assume here that the trend r has been estimated by some reasonable method and that we have confidence interval estimates for the trend factors that we will apply to the IBNR estimates. The confidence intervals for the IBNR estimates will be adjusted to reflect the uncertainty in the trend factors.

It is tempting to include loss trend as a fifth parameter in the model in order to obtain prediction intervals for trend-adjusted IBNR directly. The resulting model equation would be

$$y = \left[\alpha \exp(\beta x) + \gamma \exp(\delta x)\right] \cdot \exp(ru) + \frac{1}{w}\varepsilon$$

Unfortunately, there are two problems with this model. One is that the model sometimes produces unrealistic estimates of trend due to a lack of credibility. The other problem is that we would be extrapolating the model instead of interpolating it. Extrapolation can be misleading even in the case of linear models and it is strongly discouraged in the case of nonlinear models. Of course, we need to extrapolate the trend factors but there are mathematically sound time series models available for this purpose.

2.3 Estimation of Parameters

The SAS system used the Gauss-Newton method to estimate the least squares estimates of the model parameters. The following presentation of the mathematics of the Gauss-Newton method is based on Seber and Wild [5].

To estimate the least squares parameters, we need to minimize the sum of squared errors of the n observations:

$$S(\vec{\theta}) = \sum_{i=1}^{n} \left[y_i - f(x_i; \vec{\theta}) \right]^2 \tag{2.3.1}$$

In the case of our model, $\vec{\theta} = (\alpha, \beta, \gamma, \delta)$ and $f(x; \vec{\theta}) = \alpha \exp(\beta x) + \gamma \exp(\delta x)$. We find the minimum of $S(\vec{\theta})$ by setting all of its partial derivatives to 0.

Minimizing the sum of squared errors is a straightforward procedure for linear models but when f is nonlinear we must use numerical methods to estimate the parameters. One commonly used method is the Gauss-Newton method which works well in the case of normally distributed residuals.

We define the following matrices: $F\left(\bar{\theta}\right)=\left[\left(\dfrac{\partial f\left(x_i;\bar{\theta}\right)}{\partial\theta_j}\right)\right]$ and

$\mathbf{f}\left(\bar{\theta}\right)=\left(f\left(x_1;\bar{\theta}\right),f\left(x_2;\bar{\theta}\right),\ldots,f\left(x_n;\bar{\theta}\right)\right)'$.

F is an $n\times p$ matrix where n is the number of observations and p is the number of parameters. $\mathbf{f}\left(\bar{\theta}\right)$ has dimension $n\times 1$.

Suppose $\theta^{(a)}$ is an approximation to $\bar{\theta}$. We approximate $\mathbf{f}\left(\bar{\theta}\right)$ by the first order terms of its Taylor series in a small neighborhood near $\theta^{(a)}$:

$$\mathrm{f}\left(\bar{\theta}\right)\approx\mathrm{f}\left(\theta^{(a)}\right)+F\left(\bar{\theta}-\theta^{(a)}\right) \tag{2.3.2}$$

The residual vector is $r\left(\bar{\theta}\right)=y-\mathrm{f}\left(\theta^{(a)}\right)\approx r\left(\theta^{(a)}\right)-F\left(\bar{\theta}-\theta^{(a)}\right)$. Substituting $S\left(\bar{\theta}\right)=r'\left(\bar{\theta}\right)r\left(\bar{\theta}\right)$ leads to

$$S\left(\bar{\theta}\right)\approx r'\left(\theta^{(a)}\right)r\left(\theta^{(a)}\right)-2r'\left(\theta^{(a)}\right)F\left(\bar{\theta}-\theta^{(a)}\right)+\left(\bar{\theta}-\theta^{(a)}\right)'F'\left(\bar{\theta}-\theta^{(a)}\right)F\left(\bar{\theta}-\theta^{(a)}\right) \tag{2.3.3}$$

The right hand side of (2.3.3) is minimized with respect to $\bar{\theta}$ when

$$\bar{\theta}-\theta^{(a)}=F'\left(\bar{\theta}-\theta^{(a)}\right)F\left(\bar{\theta}-\theta^{(a)}\right)r\left(\theta^{(a)}\right)=\delta^{(a)}$$

This produces iterative approximations of $\theta^{(a)}$:

$$\theta^{(a+1)}=\theta^{(a)}+\delta^{(a)} \tag{2.3.4}$$

To use the Gauss-Newton method, one must provide $\theta^{(0)}$, the initial approximation to $\bar{\theta}$. The algorithm will converge provided the first approximation is sufficiently close to the fitted value, $\hat{\theta}$.

After fitting data to the model presented in Section 2.2, we estimated confidence intervals for the parameters and for the predicted observations. Seber and Wild [5] present formulas for approximate confidence intervals for the model parameters and for a predicted observation.

The 95% confidence interval for parameter θ_i is given by

$$\hat{\theta}_i \pm \left(s \cdot c_{ii}\right)^{1/2} \cdot t\left(N-P,.025\right) \tag{2.3.5}$$

where s^2 is the mean square error and c_{ii} is the i^{th} diagonal element of $\left(F'WF\right)^{-1}$.

The 95% confidence interval for a predicted observation corresponding to age x_i is given by

$$\hat{y}_i \pm s \cdot \left(\frac{1}{w_i} + f_i'\left(F'WF\right)^{-1} f_i\right)^{1/2} \cdot t\left(N-P,.025\right) \tag{2.3.6}$$

where f_i is the i^{th} row of F, i.e. the vector of estimated first derivatives evaluated at x_i and W is a $N \times N$ matrix with w_i as the i^{th} diagonal entry and all other entries equal to 0. $t\left(N-P,.025\right)$ is the value of Student's t distribution for $N-P$ degrees of freedom and probability .025.

The confidence intervals for predicted observations can be used to produce a confidence interval for IBNR. Based on the assumption that the incremental pure premiums for different development intervals are independent, the variance of IBNR pure premium is the sum of the variances of the incremental pure premiums for the remaining development intervals. From equation (2.3.6) we see that the variance of the incremental pure premium for one development interval is $s^2 \cdot \left(\frac{1}{w_i} + f_i'\left(F'WF\right)^{-1} f_i\right)$. The expected value of IBNR pure premium is the sum of the expected incremental pure premiums.

3. RESULTS AND DISCUSSION

The model presented in section 2 was fitted to the data presented in section 1. Only data for the latest 20 calendar quarters was used to estimate parameters. This is consistent with common actuarial practice of using recent calendar quarters rather than all of the available data so that predictions are responsive to recent changes in development patterns. We also used only data for age>1 since we do not need to estimate IBNR for that age interval. Thus, 590 observations were used to fit the model.

We used the estimated parameters to produce confidence interval estimates of IBNR for each accident quarter.

In section 3.1 we will show confidence intervals for the estimated parameters. The confidence intervals for predicted IBNR will be presented in section 3.2. In section 3.3 we present an analysis of the residuals.

3.1 Confidence Interval Estimates of Parameters

The estimated parameters and standard errors for our simulated data are:

$$\begin{aligned} \hat{\alpha} &= 3.1994, \ s(\hat{\alpha}) = 0.5807 \\ \hat{\beta} &= -0.0754, \ s(\hat{\beta}) = 0.0096 \\ \hat{\gamma} &= 29.4446, \ s(\hat{\gamma}) = 5.5549 \\ \hat{\delta} &= -0.5480, \ s(\hat{\delta}) = 0.0767 \end{aligned} \tag{3.1.1}$$

A 95% confidence interval for each parameter is of the form $\left(\hat{\theta} - s(\hat{\theta})t(.025, n-p), \hat{\theta} + s(\hat{\theta})t(.025, n-p)\right)$. There were 590 observations and we estimated 4 parameters. The resulting confidence intervals are:

$$\begin{aligned} \hat{\alpha}&: \ (2.0596, 4.3392) \\ \hat{\beta}&: \ (-0.0942, -0.0566) \\ \hat{\gamma}&: \ (18.5334, 40.3557) \\ \hat{\delta}&: \ (-0.6986, -0.3974) \end{aligned} \tag{3.1.2}$$

The Mean Square Error from the estimation is 2,987,236.

An advantage of having confidence interval estimates of the parameters is that when more data becomes available, we can test whether the current parameters should be rejected. We would reject the current estimates only if the new estimates lie outside the intervals in (3.1.2). This procedure will lead to more stable estimates of ultimate losses and IBNR.

3.2 Confidence Interval Estimates of IBNR

The estimation of IBNR was performed in two steps. First, we use equation (2.3.6) to calculate an expected value and standard error for the incremental pure premium for each development quarter until age 40 (for simplification, we assume that this is ultimate). This results in deflated IBNR estimates. The second step is to find a confidence interval for the inflation adjusted IBNR. This was done using a simulation.

<u>Step 1: Predicted incremental pure premiums</u>

The expected value of each predicted incremental pure premium is calculated by substituting the estimated parameters from (3.1.1) into the model equation,
$\hat{y} = \hat{\alpha} \cdot \exp(\hat{\beta}x) + \hat{\gamma} \cdot \exp(\hat{\delta}x)$ where x is the age of the development quarter.

In order to estimate the standard errors, we need the matrix defined in section 2.3:

$$\left(F'WF\right)^{-1}=\begin{bmatrix} 0.000000112875 & -0.000000001771 & 0.000000486556 & -0.000000010876 \\ -0.000000001771 & 0.000000000031 & -0.000000006728 & 0.000000000158 \\ 0.000000486556 & -0.000000006728 & 0.000010329411 & -0.000000127258 \\ -0.000000010876 & 0.000000000158 & -0.000000127258 & 0.000000001968 \end{bmatrix}$$

As an example, we will calculate the IBNR prediction interval for accident quarter 2. $x = 40$ for the remaining development quarter. The expected IBNR pure premium is

$$3.1994\times\exp\left(-0.0754\times 40\right)+29.446\times\exp\left(-0.5480\times 40\right)$$
$$=.15676.$$

We will need the above matrix and the derivatives of the model function evaluated at $x = 40$ to calculate the standard error of the predicted IBNR. The derivatives are:

$$\frac{\partial f}{\partial\alpha}=\exp\left(\beta x\right)$$
$$\frac{\partial f}{\partial\beta}=\alpha x\cdot\exp\left(\beta x\right)$$
$$\frac{\partial f}{\partial\gamma}=\exp\left(\delta x\right)$$
$$\frac{\partial f}{\partial\delta}=\gamma x\cdot\exp\left(\delta x\right)$$

Evaluating the derivatives at age 40 and the estimated parameters, we obtain:

$$\frac{\partial f}{\partial\alpha}\left(40\right)=0.0490$$
$$\frac{\partial f}{\partial\beta}\left(40\right)=6.2704$$
$$\frac{\partial f}{\partial\gamma}\left(40\right)=3.02\times 10^{-10}$$
$$\frac{\partial f}{\partial\delta}\left(40\right)=3.56\times 10^{-7}$$

Let the element in the j^{th} row and k^{th} column of $\left(F'WF\right)^{-1}$ be denoted m_{jk}. We calculate $f_i'\left(F'WF\right)^{-1}f_i$ from (2.3.6) as:

$$
\begin{aligned}
&f_i'\left(F'WF\right)^{-1} f_i \\
&= m_{11}\frac{\partial f}{\partial\alpha}\cdot\frac{\partial f}{\partial\alpha} + m_{12}\frac{\partial f}{\partial\alpha}\cdot\frac{\partial f}{\partial\beta} + m_{13}\frac{\partial f}{\partial\alpha}\cdot\frac{\partial f}{\partial\gamma} + m_{14}\frac{\partial f}{\partial\alpha}\cdot\frac{\partial f}{\partial\delta} \\
&+ m_{21}\frac{\partial f}{\partial\beta}\cdot\frac{\partial f}{\partial\alpha} + m_{22}\frac{\partial f}{\partial\beta}\cdot\frac{\partial f}{\partial\beta} + m_{23}\frac{\partial f}{\partial\beta}\cdot\frac{\partial f}{\partial\gamma} + m_{24}\frac{\partial f}{\partial\beta}\cdot\frac{\partial f}{\partial\delta} \\
&+ m_{31}\frac{\partial f}{\partial\gamma}\cdot\frac{\partial f}{\partial\alpha} + m_{32}\frac{\partial f}{\partial\gamma}\cdot\frac{\partial f}{\partial\beta} + m_{33}\frac{\partial f}{\partial\gamma}\cdot\frac{\partial f}{\partial\gamma} + m_{34}\frac{\partial f}{\partial\gamma}\cdot\frac{\partial f}{\partial\delta} \\
&+ m_{41}\frac{\partial f}{\partial\delta}\cdot\frac{\partial f}{\partial\alpha} + m_{42}\frac{\partial f}{\partial\delta}\cdot\frac{\partial f}{\partial\beta} + m_{43}\frac{\partial f}{\partial\delta}\cdot\frac{\partial f}{\partial\gamma} + m_{44}\frac{\partial f}{\partial\delta}\cdot\frac{\partial f}{\partial\delta} \\
&= 3.88134\times 10^{-10}
\end{aligned}
$$

The weight is $w_i = c_i x_i^{1.5} = 50{,}801\cdot 40^{1.5} = 12{,}851{,}749$. The mean square error is 2,987,236. $t\left(586,.05/2\right) = 1.96402$. Substituting this information into equation (2.3.6) we obtain 0.94925 as the width of the 95% confidence interval for the IBNR for accident quarter 2. Thus, the confidence interval for the IBNR pure premium is $\left(-0.79249, 1.10601\right)$. The confidence interval for the dollars of IBNR is $\left(-40259, 56186\right)$. For an accident quarter with more than one development quarter remaining, we would need to repeat these calculations for each remaining development quarter and sum the estimated expected values. Next, the estimated IBNR will be adjusted for loss trend.

Step 2: Including trend

Because we fitted the model to losses trended to the current calendar quarter, the dollars need to be adjusted to future cost levels. We also need to adjust the width of the confidence intervals for the uncertainty in the trend.

The trend was estimated from a time series method. The estimated trend had a mean of .005 per calendar quarter with a standard deviation of $.004\sqrt{t}$ where t is the number of quarters projected. We assume that the trend process is a Geometric Brownian Motion.

There are a number of ways to find the simultaneous confidence interval for loss development and trend. For example, we could use a Bonferronni confidence interval but this would result in an excessively wide confidence interval. Instead, we performed a simulation to estimate the variance of inflation adjusted IBNR.

We simulated incremental pure premiums before adjusting for inflation from a normal distribution with mean $\hat{\alpha}\cdot\exp\left(\hat{\beta}x\right)+\hat{\gamma}\cdot\exp\left(\hat{\delta}x\right)$ and variance given by equation (2.3.6). We simulated trend factors for each calendar quarter as a Geometric Brownian Motion with drift .005 and volatility .004. The inflation adjusted incremental pure premiums were calculated as the product of the simulated unadjusted pure premiums and the simulated trend factors. Next, the incremental pure premiums were summed over all remaining development quarters to obtain IBNR pure premium. The simulation was repeated 10000 times and the mean and standard deviation of the IBNR was calculated for each accident quarter. IBNR pure premium multiplied by exposures produces IBNR dollars.

Table 3.2.1 shows the results of the simulation. The actual IBNR is the difference between the age 40 evaluation (which we treat as ultimate here) and the evaluation at the end of the 40^{th} calendar quarter from the simulated loss development data. The expected total IBNR is 30105084. The standard deviation of the total IBNR is 1350093. The 95% confidence interval for total IBNR is (27458951 , 32751218). The actual total IBNR is 30120821.

Table 3.2.1

Accident Quarter	Exposures	Expected Value	Standard Deviation	95% Confidence Interval Lower	Upper	Actual IBNR
2	50,801	8,190	24,518	-39,864	56,244	-3,686
3	51,187	16,643	35,835	-53,593	86,879	20,450
4	51,146	26,310	44,192	-60,304	112,925	11,254
5	51,527	36,541	51,941	-65,262	138,344	73,738
6	52,348	49,099	58,839	-66,225	164,422	98,397
7	52,480	61,528	65,232	-66,325	189,381	37,099
8	53,148	75,340	71,800	-65,385	216,065	156,305
9	53,924	91,671	78,552	-62,287	245,629	237,876
10	54,403	109,065	85,433	-58,380	276,511	-95,408
11	54,557	124,874	91,436	-54,338	304,086	384,465
12	55,083	144,622	96,258	-44,040	333,284	260,118
13	55,292	168,450	103,341	-34,095	370,995	299,600
14	55,899	192,189	108,233	-19,944	404,322	175,632
15	56,067	215,948	115,108	-9,659	441,555	3,570
16	57,025	247,643	123,187	6,201	489,086	237,988
17	57,071	279,736	129,481	25,957	533,515	224,736
18	57,317	311,248	134,933	46,784	575,712	268,971
19	57,907	346,819	143,714	65,144	628,493	712,233
20	58,285	388,878	149,405	96,050	681,706	428,225
21	59,096	433,974	157,772	124,746	743,202	819,832
22	59,193	479,592	165,473	155,270	803,915	930,364
23	59,524	530,342	173,337	190,607	870,076	564,488
24	59,745	583,879	177,894	235,213	932,546	412,411
25	60,427	645,944	188,083	277,309	1,014,580	421,418
26	60,155	705,701	195,557	322,416	1,088,985	699,647
27	60,568	776,239	207,953	368,659	1,183,819	794,518
28	60,708	852,632	215,059	431,123	1,274,140	995,212
29	60,262	925,896	222,578	489,652	1,362,140	944,400
30	60,606	1,012,197	233,755	554,046	1,470,349	945,867
31	60,580	1,109,304	251,368	616,632	1,601,976	1,084,176
32	60,648	1,213,637	258,802	706,395	1,720,879	1,703,397
33	61,159	1,344,114	277,079	801,049	1,887,178	1,107,447
34	61,462	1,492,000	292,032	919,627	2,064,372	1,133,824
35	61,934	1,660,873	312,021	1,049,324	2,272,423	1,882,576
36	61,716	1,858,275	333,112	1,205,388	2,511,161	1,567,491
37	61,837	2,123,409	361,113	1,415,642	2,831,177	1,962,887
38	62,285	2,514,004	394,000	1,741,778	3,286,231	1,938,616
39	62,728	3,055,695	450,062	2,173,589	3,937,801	2,836,989
40	63,180	3,892,584	522,958	2,867,605	4,917,563	3,843,696

3.3 Analysis of Residuals

It is important to examine the residuals from a regression model to check the consistency of the data with the assumptions of the model. In this section we will look at plots of the residuals to look for patterns. We will also see the results of a Shapiro-Wilk test of normality, a histogram and a probability plot.

These tests are shown for demonstration purposes only. The data used to demonstrate the methodology in this paper is simulated and will pass the normality test. Real data might not pass tests of normality but if the deviation from normality is not too extreme, then the estimated confidence intervals are still reasonable.

The unmodified residuals, $r_i = y_i - \hat{y}_i$, do not have constant variance because the data do not have constant variance. The tests will be performed on studentized residuals, defined as $r_i / std(r_i)$. Seber and Wild [5] show the following formula for the standard errors of the residuals.

$$std(r_i) = s \cdot \left(\frac{1}{w_i} - f_i' \left(F'WF \right)^{-1} f_i \right) \tag{3.3.1}$$

Exhibits 3.3.1 through 3.3.3 show the scatter plots of the studentized residuals against predicted value, development age, and calendar quarter. The plots do not show any obvious patterns and the studentized residuals seem to have constant variance. Thus, the weight function appears to be appropriate and there does not appear to be any reason to modify the model.

Exhibit 3.3.4 is a histogram of the studentized residuals. Exhibit 3.3.5 is a normal probability plot (calculated using methodology from [6]). The shape of the histogram appears to be consistent with a normal distribution. The probability plot is nearly linear which supports the assumption that the residuals have a normal distribution. A Shapiro-Wilk test was performed on the residuals and produced a statistic of 0.9983 with a p-value of 0.8445. Thus, we cannot reject the hypothesis that the residuals have a normal distribution.

Exhibit 3.3.1

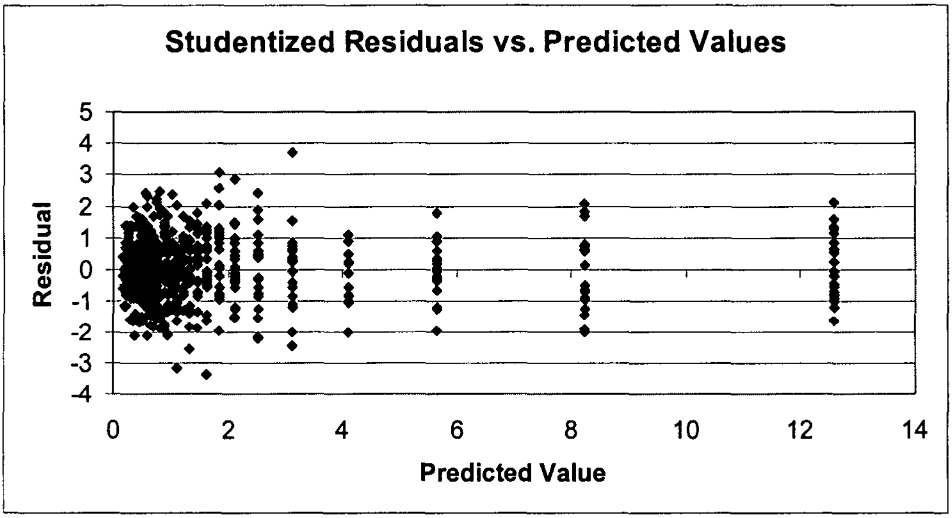


Exhibit 3.3.2

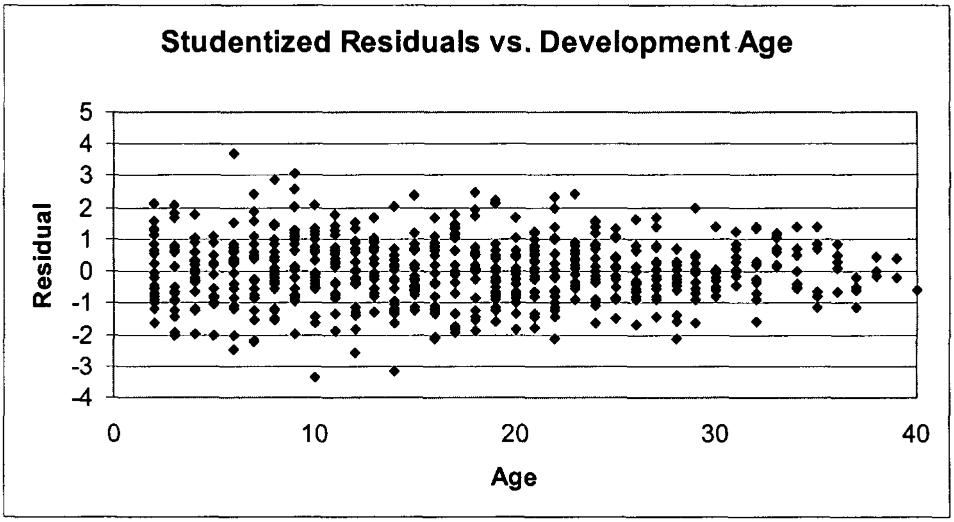


Exhibit 3.3.3

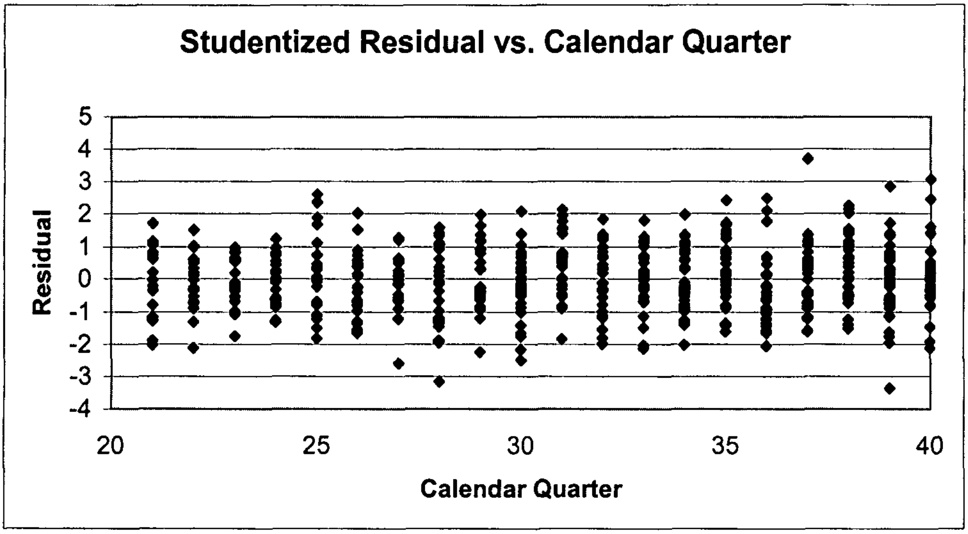


Exhibit 3.3.4

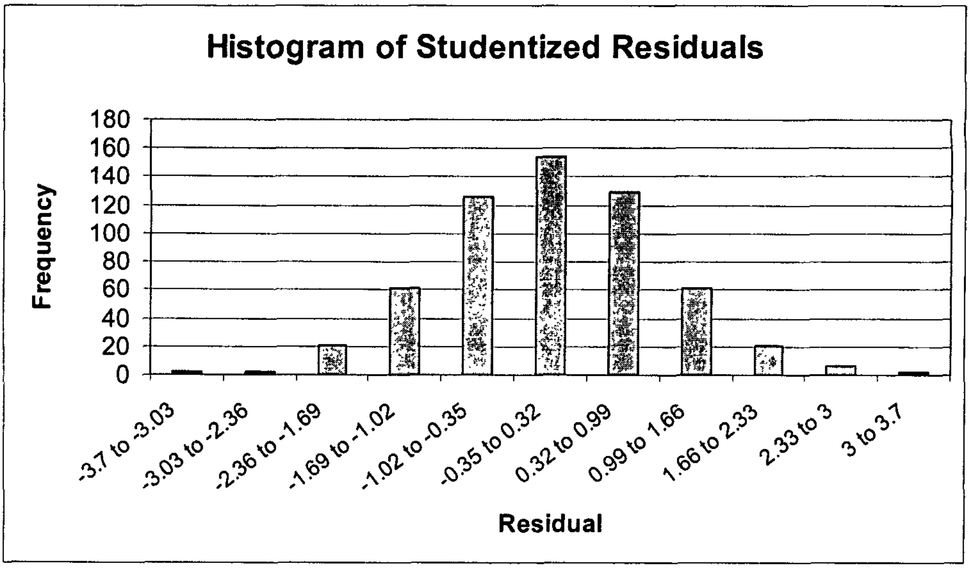
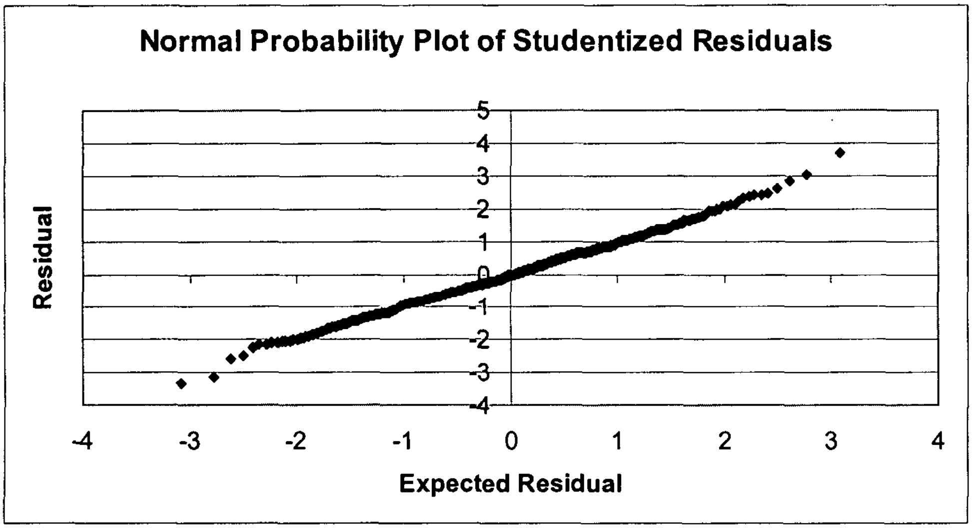


Exhibit 3.3.5



4. CONCLUSIONS

Our model has satisfied the objective stated in section 1.2. The model could be fitted either to paid or case incurred losses. Since the observations are incremental pure premiums and the weights are a function of exposures, the model makes appropriate adjustments for changing levels of exposures. By using nonlinear regression, we have avoided the need for a log link and we have been able to keep negative observations in the data. The model appears to produce unbiased estimates of IBNR and reasonable 95% confidence intervals.

The plots displayed in section 3.3 indicate that incremental pure premiums have an approximately normal distribution.

The assumptions we made work well with auto bodily injury data. We have assumed that the data satisfy the usual assumptions of nonlinear regression models including independent normal errors. We have also used a functional form that fits our data well but might not fit other lines. We would like to close with a few suggestions for fitting models to other lines.

The assumption of normal errors should be reasonable for high frequency lines of business. The assumption that the errors are uncorrelated should also be reasonable most of the time. If these assumptions are rejected, there are nonlinear models that may be used. Seber and Wild [5] discuss models with non-normal and autocorrelated errors.

Seber and Wild [5] has a chapter on growth models which lists many functional forms other than the form presented in this paper. Some of these models might fit the pure premiums of other lines of business. Some of the models could be applied to cumulative instead of incremental data.

Another class of models that will fit pure premium development is generalized linear models. In this type of model, the development age interval could be represented as a categorical variable. These models would allow the analyst to consider a great variety of error distributions and error correlation structures. One drawback to this approach is that there are more parameters to estimate which means that the confidence interval for IBNR will be wider. Dobson [7] is an excellent reference on generalized linear models.

Acknowledgment

The author acknowledges John Grogan for suggesting the function used to fit the data and Dale Porfilio for support and encouragement to submit this paper for publication.

5. REFERENCES

[1] D. Murphy, Unbiased Loss Development Factors, Casualty Actuarial Society Forum, 1994.
[2] G. Barnett and B. Zehnwirth, Best Estimates for Reserves, Casualty Actuarial Society Forum, 1998.
[3] J. Stanard, A Simulation Test of Prediction Errors of Loss Reserve Techniques, Proceedings of the Casualty Actuarial Society, 1985.
[4] P. Narayan, T. Warthen, A Comparative Study of the Performance of Loss Reserving Methods Through Simulation, Casualty Actuarial Society Forum, 1997.
[5] G. Seber, C. Wild, Nonlinear Regression, 1989 Wiley and Sons
[6] J. Neter, M. Kutner, C. Nachtsheim, W. Wasserman, Applied Linear Statistical Models, 1996, McGraw-Hill.
[7] A. Dobson, An Introduction to Generalized Linear Models Second Edition, 2001, Chapman and Hall/CRC

Appendix A – Simulated Loss Development Data – Earned Exposures and Incremental Case Incurred Losses

Accident Quarter	Exposures	Development Quarter 1	2	3	4	5	6	7	8	9	10
1	50,333	1,244,722	468,457	283,193	68,634	101,440	125,515	159,525	68,010	64,419	40,378
2	50,801	1,417,101	587,121	337,664	95,841	243,190	65,263	94,623	47,772	188,347	31,533
3	51,187	1,143,473	502,816	483,911	174,352	70,477	140,000	86,071	93,975	58,041	-13,132
4	51,146	1,055,290	1,213,499	496,421	393,358	173,547	116,010	101,531	115,414	31,102	64,744
5	51,527	1,508,450	834,730	302,691	409,622	423,103	143,476	67,506	146,727	146,017	38,057
6	52,348	1,192,515	1,074,133	478,010	172,336	69,719	-9,197	206,481	221,841	-2,220	170,525
7	52,480	1,067,318	438,952	368,962	214,718	274,665	86,113	39,208	121,683	62,838	8,398
8	53,148	758,275	513,455	255,088	159,199	60,798	102,129	97,269	107,077	134,067	86,292
9	53,924	1,664,156	31,358	491,390	196,506	202,088	27,387	48,404	-41,016	77,453	57,713
10	54,403	1,537,825	422,697	301,334	210,821	254,271	149,386	225,652	104,121	81,082	76,196
11	54,557	2,026,667	738,848	252,783	251,618	193,254	96,790	170,517	106,026	-121,070	152,682
12	55,083	1,296,855	495,927	439,098	125,722	96,000	111,872	213,356	188,106	123,309	67,820
13	55,292	1,995,401	821,508	552,503	226,522	145,695	389,146	333,936	118,922	136,678	104,544
14	55,899	2,078,843	786,272	529,533	329,974	179,953	135,260	180,141	106,218	165,608	63,277
15	56,067	1,952,667	859,868	632,198	264,522	231,370	216,887	22,205	117,571	81,594	149,627
16	57,025	1,258,033	650,068	387,309	183,986	152,797	244,063	68,088	103,095	56,465	80,728
17	57,071	1,627,621	320,911	303,800	327,057	236,332	161,152	205,081	147,898	288,069	127,213
18	57,317	1,681,507	446,643	359,407	248,809	270,162	229,530	58,483	-7,112	246,925	85,944
19	57,907	2,508,300	1,018,661	99,969	436,712	156,983	241,768	303,837	-9,729	194,850	181,037
20	58,285	1,238,641	812,792	542,250	329,100	246,551	61,085	173,928	17,813	183,213	64,235
21	59,096	1,793,043	482,793	546,164	313,044	353,857	327,614	90,275	235,255	32,150	-8,168
22	59,193	1,433,225	532,545	589,099	306,945	330,835	50,915	285,934	84,085	48,543	144,367
23	59,524	1,516,012	753,758	581,957	365,421	217,070	239,708	-63,906	191,485	107,079	181,666
24	59,745	1,803,164	519,924	295,180	225,283	222,089	122,650	-57,787	170,330	46,008	56,351
25	60,427	1,347,360	328,992	357,630	157,580	135,900	-80,274	117,487	208,666	121,095	179,925
26	60,155	810,643	604,364	214,555	155,748	114,459	93,877	397	95,471	40,225	82,597
27	60,568	1,850,892	980,308	613,338	369,051	298,488	272,379	196,454	107,684	175,612	251,126
28	60,708	3,006,298	1,044,056	843,024	581,194	261,303	209,512	255,504	255,482	95,327	-22,389
29	60,262	986,119	672,498	590,640	43,019	-8,877	-32,562	119,151	17,117	205,731	144,380
30	60,606	2,630,383	1,101,593	805,938	238,565	228,041	253,614	194,571	157,225	190,266	-153,391
31	60,580	1,515,313	601,512	511,685	304,718	142,590	143,733	206,446	79,617	-47,845	62,197
32	60,648	3,517,516	1,048,147	260,427	466,379	114,732	589,058	125,985	381,048	361,346	9,249
33	61,159	1,673,500	513,290	333,639	305,302	308,506	232,796	81,397	104,474	119,655	-44,314
34	61,462	1,207,813	739,162	524,302	392,092	363,797	230,703	398,177	149,413	19,749	166,569
35	61,934	2,202,629	528,671	378,846	162,033	150,420	225,454	174,285	201,873	193,199	7,944
36	61,716	1,051,422	470,986	415,687	433,785	152,892	263,161	151,337	-23,299	23,005	183,878
37	61,837	2,355,630	1,302,388	824,821	307,806	79,979	117,519	208,110	162,036	207,896	141,836
38	62,285	2,016,667	990,682	153,863	290,379	-10,658	120,430	7,588	208,320	71,027	123,875
39	62,728	1,468,675	925,175	157,502	266,297	252,103	380,364	226,500	76,920	43,621	26,222
40	63,180	1,952,713	712,475	446,253	551,239	361,511	276,575	355,898	-8,263	66,140	96,505

Accident		Development Quarter									
Quarter	Exposures	11	12	13	14	15	16	17	18	19	20
1	50,333	55,132	56,286	39,158	54,447	32,780	64,401	47,793	-30,285	60,984	-1,923
2	50,801	97,899	120,746	66,487	16,116	50,122	42,810	11,512	77,062	7,181	72,151
3	51,187	91,825	66,621	-46,239	10,535	-6,023	58,325	4,497	19,837	5,313	7,916
4	51,146	166,882	134,068	17,951	101,782	64,696	64,331	-98,606	-37,817	60,282	-3,207
5	51,527	119,089	-11,801	42,615	-73,816	72,867	-9,629	91,001	47,562	22,359	-16,392
6	52,348	16,662	112,356	-26,328	108,397	152,916	95,642	107,603	30,439	10,858	99,144
7	52,480	95,445	53,315	65,328	34,802	-4,403	-49,160	-32,400	40,083	-11,567	51,725
8	53,148	166,630	-14,316	19,746	-8,727	25,872	-1,430	53,704	73,004	-26,364	61,378
9	53,924	39,927	78,793	150,847	42,016	83,246	57,028	13,857	62,675	59,835	39,645
10	54,403	52,923	759	-8,435	45,026	37,059	126,682	-15,224	47,275	54,888	10,951
11	54,557	118,521	48,258	97,730	-6,982	168,831	30,198	100,201	-11,399	27,865	57,843
12	55,083	52,455	117,838	19,476	61,249	42,336	-6,884	-42,245	5,514	40,494	-37,779
13	55,292	108,005	101,991	50,016	-10,580	23,714	-14,118	101,221	66,648	131,158	33,186
14	55,899	31,057	-37,318	104,948	67,958	-7,386	95,217	-34,104	130,890	-6,796	28,246
15	56,067	94,894	-5,199	55,050	-107,620	33,005	35,708	113,029	-23,751	33,324	82,253
16	57,025	34,439	-82,833	-7,708	78,608	49,459	91,763	-36,547	48,994	3,417	39,090
17	57,071	119,713	127,702	120,055	98,655	33,349	36,053	79,890	72,189	80,971	2,954
18	57,317	-42,024	90,633	88,686	89,706	102,187	89,757	114,280	125,545	21,101	58,920
19	57,907	26,467	57,494	37,776	-1,643	120,996	-11,362	45,765	162,032	-7,833	9,218
20	58,285	54,460	168,172	60,942	33,469	43,582	95,786	136,815	7,129	146,101	8,598
21	59,096	108,307	118,119	130,671	12,719	66,407	-49,728	103,805	-23,377	13,446	24,913
22	59,193	170,747	121,252	122,821	-25,894	96,750	89,657	52,945	49,776	50,822	120,953
23	59,524	205,422	-10,765	87,080	7,915	20,942	82,590	86,042	-20,650	86,539	3,828
24	59,745	57,883	-9,148	57,563	76,990	72,755	53,851	37,035	75,261	-8,698	-11,311
25	60,427	98,081	67,455	30,353	184,721	26,902	16,334	62,868	80,394	2,462	19,806
26	60,155	-5,192	53,749	114,555	37,095	35,334	7,140	62,927	39,025	48,072	-549
27	60,568	166,693	82,674	70,339	70,978	79,356	-64,843	22,823	76,726	79,929	35,373
28	60,708	188,567	168,677	129,179	68,308	47,794	96,191	128,506	49,566	-28,134	58,887
29	60,262	104,592	80,640	95,315	34,245	48,974	81,604	39,399	32,106	55,537	-46,734
30	60,606	99,495	58,402	48,408	56,793	-4,451	19,091	-5,279	19,722	53,159	39,365
31	60,580	139,467	180,552	45,404	72,414	3,441	38,563	126,913	50,865	37,834	56,248
32	60,648	79,096	136,515	178,105	91,579	20,394	100,918	56,855	43,922	-7,463	34,194
33	61,159	54,143	-81,040	20,949	1,608	60,381	111,910	13,739	102,704	27,132	104,321
34	61,462	97,706	108,206	14,850	59,003	54,189	69,831	65,128	23,821	43,958	-11,047
35	61,934	59,384	69,087	148,809	136,211	71,394	4,055	126,075	52,993	84,082	56,630
36	61,716	139,013	56,147	92,609	125,058	7,067	90,151	101,031	27,566	-17,295	58,929
37	61,837	4,048	253,659	157,930	58,979	100,435	19,044	-20,740	51,891	112,978	-55,242
38	62,285	145,307	110,498	149,791	87,189	164,906	27,941	11,832	73,887	77,094	14,150
39	62,728	133,332	164,332	-10,897	106,455	136,006	141,784	83,994	79,801	71,479	-24,821
40	63,180	120,876	1,869	149,325	-46,560	52,798	85,751	68,371	100,236	-48,006	133,049

Accident Quarter	Exposures	Development Quarter 21	22	23	24	25	26	27	28	29	30
1	50,333	1,844	47,936	8,281	51,891	37,771	-4,045	27,774	14,035	871	56,156
2	50,801	21,390	63,117	23,327	-12,069	14,680	41,491	15,434	11,107	19,559	10,661
3	51,187	54,081	7,552	40,110	30,494	-5,797	730	29,749	25,467	30,419	7,913
4	51,146	76,821	112,076	30,277	20,160	57,926	74,676	-23,786	20,074	6,919	18,023
5	51,527	43,306	960	47,240	-7,478	-5,993	-31,465	45,344	28,740	26,218	10,956
6	52,348	8,403	24,816	14,994	66,326	7,418	22,099	3,600	-46,942	78,616	16,230
7	52,480	35,360	38,631	16,046	53,286	18,835	12,820	23,495	5,427	33,480	2,394
8	53,148	27,648	103,471	-2,524	-3,970	71,300	28,587	-3,460	9,452	-8,909	-6,737
9	53,924	54,178	71,192	59,018	52,434	-25,919	50,456	76,803	43,181	-2,099	17,733
10	54,403	33,351	24,839	42,521	26,870	17,470	10,409	-7,892	-29,828	2,882	200
11	54,557	65,973	-5,380	53,969	15,744	-3,427	4,913	8,390	-24,473	-32,538	62,557
12	55,083	63,685	-10,583	64,637	78,643	30,741	11,856	15,134	1,767	18,412	31,248
13	55,292	82,088	7,445	121,478	-32,097	41,168	47,156	49,125	9,276	44,273	23,737
14	55,899	50,701	-22,139	55,822	44,064	65,745	-5,697	71,653	59,301	-11,011	8,634
15	56,067	-18,731	-14,131	44,114	86,453	31,838	25,910	-15,473	17,799	-2,589	34,604
16	57,025	-22,239	89,127	13,948	20,393	-2,351	8,374	31,029	39,339	-8,451	1,222
17	57,071	57,279	36,946	39,534	90,362	14,387	-29,765	30,222	16,053	17,682	35,591
18	57,317	79,551	77,137	47,330	29,128	-40,416	46,964	9,795	23,656	43,627	-433
19	57,907	-36,342	-51,199	-629	38,859	-20,756	54,574	72,098	36,775	39,504	31,052
20	58,285	43,739	-8,134	54,269	25,913	-16,757	8,755	7,972	43,674	-3,448	64,314
21	59,096	75,294	-34,942	88,190	124,206	62,976	77,091	39,748	40,729	43,609	89,136
22	59,193	11,060	76,017	61,132	105,644	56,274	15,014	-2,897	80,213	53,917	118,331
23	59,524	29,866	45,800	38,868	68,925	7,687	-61,021	30,638	39,572	45,399	-11,739
24	59,745	46,041	33,062	16,682	40,849	-18,453	7,049	58,613	48,743	-17,040	8,158
25	60,427	25,489	-35,072	29,365	1,481	46,825	-43	39,986	80,497	51,650	-27,268
26	60,155	136,480	50,523	73,985	-15,999	21,991	43,033	32,821	8,902	29,994	41,090
27	60,568	-2,199	34,675	135,124	6,514	15,272	62,756	66,009	-10,230	-37,723	3,901
28	60,708	115,933	100,646	55,828	25,764	-3,515	9,366	-23,401	89,137	46,630	75,698
29	60,262	70,050	48,884	59,346	53,211	-3,141	6,048	29,235	13,746	38,350	43,614
30	60,606	100,823	-80,196	-23,695	19,793	20,686	-29,950	-5,204	99,580	36,328	56,872
31	60,580	42,546	19,448	19,949	-29,940	17,116	55,736	756	21,693	8,254	48,025
32	60,648	74,650	86,062	71,446	138,206	-8,941	75,564	27,495	84,913	-26,461	74,757
33	61,159	16,045	110,447	129,009	-45,715	68,666	7,394	20,046	33,159	7,386	18,884
34	61,462	4,309	-26,370	107,835	127,369	15,493	-50,769	-7,521	-25,623	-1,506	18,283
35	61,934	83,466	73,782	56,185	-32,328	-38,556	27,399	-11,618	54,166	26,555	-750
36	61,716	-27,140	93,574	66,551	13,086	30,072	-12,666	-11,496	-7,722	13,375	17,919
37	61,837	30,283	14,515	-30,671	-60,204	31,067	15,254	78,382	95,606	7,715	9,987
38	62,285	95,225	114,060	54,619	-67,884	7,563	-31,075	-36,590	9,379	78,245	14,113
39	62,728	4,900	-81	43,622	78,577	92,489	28,945	-16,724	67,108	17,473	-6,230
40	63,180	-21,730	80,710	55,218	15,476	39,584	3,858	18,112	22,462	13,209	33,635

Accident Quarter	Exposures	Development Quarter 31	32	33	34	35	36	37	38	39	40
1	50,333	34,276	3,835	43,711	-1,640	29,076	13,319	-2,630	8,678	3,188	-5,821
2	50,801	27,835	-5,203	29,996	-2,056	-5,783	18,323	-5,364	5,188	19,082	-3,686
3	51,187	17,345	51,332	18,327	1,471	34,161	24,299	-18,628	20,407	5,774	14,676
4	51,146	23,792	20,703	18,035	26,068	48,501	-5,982	5,236	30,317	-20,445	1,381
5	51,527	52,198	-30,522	43,073	32,051	-18,894	33,613	54,075	34,899	24,632	-39,868
6	52,348	35,800	-10,345	46,979	51,773	-9,409	-22,677	20,537	25,957	12,417	62,164
7	52,480	15,761	7,390	33,962	13,579	24,407	19,471	-3,988	-10,108	9,185	-1,869
8	53,148	3,334	24,196	20,303	24,861	40,319	12,113	-13,820	42,130	35,297	15,405
9	53,924	12,418	56,958	52,833	26,726	32,933	5,232	29,185	37,027	45,244	8,695
10	54,403	43,335	-31,285	-54,667	-16,646	15,567	-15,554	-9,313	-13,437	29,191	736
11	54,557	72,266	95,758	6,159	7,910	61,208	28,223	28,531	49,899	16,681	17,830
12	55,083	39,707	25,525	25,741	13,760	17,106	49,171	1,081	16,032	24,045	16,702
13	55,292	20,470	-17,379	26,954	5,100	45,000	38,726	-19,480	78,077	44,212	9,911
14	55,899	22,377	1,924	-9,766	47,354	18,223	454	6,891	20,907	2,586	7,758
15	56,067	-35,330	-18,262	-26,548	38,756	13,302	-9,275	-3,005	47,269	-34,312	-3,364
16	57,025	12,632	12,743	38,517	4,169	12,911	-1,785	-20,735	78,244	30,557	-779
17	57,071	38,585	45,591	11,315	-26,333	15,402	24,202	-25,715	11,080	9,528	36,913
18	57,317	20,486	28,330	4,385	-14,482	35,357	22,167	-4,089	2,423	1,184	60,888
19	57,907	68,328	82,874	121,115	-32,249	19,414	73,240	23,233	39,217	78,051	-12,467
20	58,285	42,726	28,751	58,837	7,617	7,172	72,846	1,544	-5,147	29,124	8,197
21	59,096	33,319	44,155	-39,904	29,104	50,092	10,920	13,243	23,352	35,124	14,389
22	59,193	-24,229	54,258	48,576	51,990	44,191	48,506	57,887	-6,941	-18,202	-21,329
23	59,524	48,463	18,459	7,095	13,631	6,314	16,901	46,450	-16,939	43,202	56,548
24	59,745	35,148	16,789	-7,315	-9,671	21,791	14,107	28,696	9,512	8,829	15,567
25	60,427	-16,598	26,696	11,564	14,065	-29,491	2,041	18,738	47,090	-8,041	-23,085
26	60,155	-8,915	52,310	915	813	56,718	-15,282	-26,165	20,384	20,458	18,977
27	60,568	22,553	42,332	9,009	-7,442	2,140	93,063	88,561	46,159	-5,060	-262
28	60,708	-14,237	65,453	25,751	12,368	-49,710	41,335	-49,919	-30,620	69,756	11,831
29	60,262	49,917	32,539	-6,986	48,452	15,625	28,630	15,743	23,349	-5,595	42,937
30	60,606	60,504	65,845	93,343	-27,623	-3,656	51,672	33,114	50,926	101,765	39,729
31	60,580	-47,046	46,028	24,302	56,096	-8,692	27,322	28,081	6,079	-23,284	20,008
32	60,648	63,634	122,152	-1,646	-37,185	-19,352	96,570	10,367	35,026	41,909	50,868
33	61,159	30,616	8,972	11,306	39,325	10,365	32,535	50,209	7,522	34,812	25,278
34	61,462	19,812	3,141	33,675	12,108	-21,363	18,639	44,897	46,331	-32,125	-14,164
35	61,934	-30,866	32,254	88,375	36,930	62,025	72,476	54,286	-50,512	12	-6,728
36	61,716	-35,620	-2,115	31,594	37,150	-2,481	26,166	14,732	19,708	19,340	5,106
37	61,837	35,752	60,881	-13,187	-21,121	39,280	-1,210	-23,822	11,761	42,508	39,751
38	62,285	-10,410	21,570	35,964	-13,033	-26,726	-20,093	-11,908	65,799	-11,499	-2,260
39	62,728	70,218	-25,147	46,379	5,606	39,349	-13,438	70,889	53,260	-18,834	-12,364
40	63,180	65,404	13,053	28,027	-40,448	-2,637	3,059	-7,238	41,295	-7,643	14,249

Appendix B – Simulation Model Used to Generate the Data

The simulation model used to generate the loss and exposure data is based on method 3 in Narayan [4] with some modifications. In this appendix, we will present an outline of the model and the SAS code used to produce the data. Note that the SAS program will not produce the same data every time it is run because the random number seeds were randomized.

Outline of the simulation methodology:

1. Initialize the values for exposures at 50,000 per quarter and the inflation index at unity.
2. For each of the 40 accident quarters:
   a. Generate a random number of exposures from a Brownian motion process.
   b. Generate a random frequency from a Normal distribution.
   c. Generate a random number of claims from a Poisson distribution with a parameter equal to the product of the exposures and the frequency.
   d. Generate an inflation index from a geometric Brownian motion.
   e. Initialize ultimate loss to zero. Then, for each claim
      i. Generate a random loss severity from a Lognormal distribution, multiply it by the inflation index and add it to the ultimate loss.
3. For each accident quarter,
   a. calculate 40 random increment factors from the formula: $incr = .33 \cdot age^{-1.25} + \left(.07 \cdot age^{-.7}\right) \cdot Normal\left(0,1\right)$. This is not guaranteed to add up to unity but the simulated values add up very close to unity. This procedure is similar to step (i) in Narayan's method 4 except that we are using a random decay pattern instead of a constant pattern.
   b. Multiply the ultimate loss by the increment factors to produce random incremental losses for 40 development quarters.

SAS code:

```
*random number seed;
%let seed=0;

*exposure parameters (Geometric Brownian Motion);
%let expostart = 50000;
%let grthmean = 0.005;
%let grthstdv = .005;

*frequency parameters (Normal);
%let frqmean = .01;
%let frqstdev = .001;

*untrended severity parameters (LogNormal);
%let mu = 8;
%let s = 1.4;

*inflation parameters (Geometric Brownian Motion);
%let cpi0 = 100;
%let cpimu = .006;
```

```
%let cpisig = .0035;

/* First data step - generate exposures and ultimate losses for 40
accident quarters.*/

data tri1;
      *initialize exposures and cpi;
      expos = &expostart;
      cpi = &cpi0;
      do aqtr=1 to 40;
            *generate exposures by brownian motion;
            expos = round(expos * (1 + &grthmean +
            &grthstdv*rannor(&seed)));

            *generate a normally distributed claim frequency;
            freq = &frqmean + &frqstdev*rannor(&seed);

            *generate a Poisson number of claims;
            clms = ranpoi(&seed,freq*expos);

            *generate an inflation index by geometric brownian motion;
            cpi = cpi*exp(&cpimu + &cpisig*rannor(&seed));

            *calculate aggregate loss (ultloss);
            ultloss = 0;
            do clmnum = 1 to clms;
            *calculate loss severity and add it to ultloss;
                        ultloss = ultloss +
            round(exp(&mu+&s*rannor(&seed))*cpi/&cpi0);
            end;
      output;
      end;


proc sort data=tri1; by aqtr;

/* Second data step - calculate incremental incurred losses for 40
development quarters for each accident quarter to produce a decumulated
loss development data set. */

data decumtri;
      set tri1;
      do age=1 to 40;
            decay = .33*age**(-1.25) + (.07*age**(-.7))*rannor(&seed);
            incr_inc = ultloss*decay;
            time = aqtr + age - 1;
            output;
      end;

run;
```

Multilevel Non-Linear Random Effects Claims Reserving Models And Data Variability Structures

Graciela Vera

Abstract

Characteristic of many reserving methods designed to analyse claims data aggregated by contract or sets of contracts, is the assumption that features typifying historical data are representative of the underwritten risk and of future losses likely to affect the contracts. Kremer (1982), Bornheutter and Ferguson (1972), de Alba (2002), and many others, consider models with development patterns common to all underwriting years and known mean-variance relationships. Data amenable to such assumptions are indeed rare. More usual are large variations in settlement speeds, exposure and claim volumes. Also typifying many published models are Incurred But Not Reported (*IBNR*) predictions limited to periods with known claims, frequently adjusted with "tail factors" generated from market statistics. Of concern could be analytical approach inconsistencies behind reserves for delay periods before and after the last known claims, under reserving and unfair reserve allocation at underwriting year, array or contract levels.

As applications of Markov Chain Monte Carlo (MCMC) methods, the models proposed in this paper depart from the neat assumptions of quasi-likelihood and extended quasi-likelihood, and introduce random effects models. The primary focus is the close dependency of the *IBNR* on data variability structures and variance models, built with reference to the generic model derived in Vera (2003). The models have been implemented in BUGS (http://www.mrc-bsu.cam.ac.uk/bugs)

Keywords: Markov Chain Monte Carlo, Non-linear Random Effects and GLM, Reserving.

1. INTRODUCTION

Insurance data reflect and react to financial uncertainty associated with external events, quantifiable time varying factors such as inflation and interest rate fluctuations, and non-quantifiable factors such as variations in litigation practices and underwriting policy terms. In an interesting historical account of legislative changes introduced in Israel to deal with inflation, Kahane (1987) illustrates how external events can be given functional interpretation in a reserving model. Further examples can be found in Taylor

(2000). Data distortions due to external events could undermine all stochastic assumptions. Concerned with the analysis of claims data, from the simplest aggregation levels, such as class of business, to multiple-nested groups, this paper deals with the construction of claims reserving models capable of capturing variability structures in a claims portfolio.

Hierarchical or multi-level claims reserving models are potential source of wide-ranging contribution to claims portfolio analysis beyond reserving. Identification of the causes of data variability with reference to hierarchical model structures could provide a statistical framework for parametric analyses of claims across a number of underwriting years. This would enhance our ability to construct more discriminating models, set initial parameter values, review and update our assumptions on risk premium calculations, related management strategies for commutations, portfolio composition, analysis, etc.

1.1 Research Context

As one of the simplest claims reserving methods, the chain ladder has motivated an extensive body of work intended to establish statistical basis for the problem of reserving. Models that fall within the category of generalized linear models (GLM) (McCullagh and Nelder (1989)), such as Renshaw (1989), Renshaw and Verrall (1998), Verrall (1991), Wright (1990), Mack (1991) and many others, have extended the research beyond assumptions of lognormality and explored applications from exponential family distributions. Carroll (2003) remarks "there are many instances where understanding the structure of variability is just as central as understanding the mean structure". The *IBNR* definition given in this paper is integral to the definition of the model itself, and its value is highly sensitive to model specification. Hence, the emphasis of this research is in the identification of suitable representations for the mean and data variability structures beyond assumptions of known and specific mean-variance relationships.

Reserving model structures depend on the intended use of the predicted reserves and on the sector of interest in the claims portfolio, such as insurance class, contract, specific loss, etc. The data assessment should determine the selection of the analytical approach.

For instance, an insurance contract provides cover against the hazards listed in the contract. Premium calculations reflect policy management expenses, expected returns and risk premiums for all the perils covered by the contract. Risk premium analyses, in general, are carried out by peril, ignoring the fact that a particular event could simultaneously hit more than one kind of cover. When reserve analysis of all perils with a single model is viable, it could deliver, for example, relative cost measures capable of generating more competitive commercial premiums, hence allowing cover assessment on statistical basis, identification of cross-subsidies and unexplored niches, etc.

Within the context of hierarchical models, claims data can be differently interpreted depending on their levels of aggregation. For instance:

- Each underwriting year data set could be described as a set or cohort of longitudinal data.
- A claims array could be considered single-level longitudinal data for more than one subject.
- A book of business segmented by class, type of loss and underwriting year, could be treated as multilevel longitudinal data or as multiple nested groups of single level longitudinal data.

Davinian and Giltinan (1993 and 1996) provide an introduction to the theory of non-linear random effects models and an overview of various techniques for the analysis of non-linear models with repeated observations. More recently, Pinheiro and Bates (2000) reviews the theory and applications of linear and non-linear mixed effect models to the analysis of grouped data.

In this paper it is shown that the generic model in Vera (2003), briefly outlined below, is key to the extension of random effect models to the analysis of reserves. If the claims process for underwriting year w is reported at times $t_1, t_2, \ldots, t_e$, such that $0 < t_1 < t_2 < \ldots < t_e$, and t_e is the final settlement period, the generic model is given in terms of a percentage cash flow and a ultimate claim amount functions, denoted respectively by P_{w,t_j} and C_w. $P_{w,t_j} = \int_0^{t_j} \pi(w,z)dz$, where $\pi(w,t)$ is a probability density function taking values from positive real numbers. $S_{w,t_j} = 1 - P_{w,t_j} = \int_{t_j}^{t_e} \pi(w,z)dz$, $P_{w,t_j} \le 1$ for $j < e$ and $P_{w,t_j} = 1$ otherwise. Finally, h_{w,t_j} and H_{w,t_j} are the instant and cumulative hazard rate functions, defined for underwriting year w and payment year τ (τ = w + delay time - 1) by

$$h_{w,\tau-w+1} = -\left(\frac{\partial\left(\ln\left(1-P_{w,z}\right)\right)}{\partial z}\right)_{z=\tau-w+1} = \left(\frac{C_w}{IBNR_{(w,\tau-w+1)}}\right)\left(\frac{\partial P_{w,z}}{\partial z}\right)_{z=\tau-w+1} \tag{1.1}$$

$$H_{w,\tau-w+1} = -\ln\left(1-P_{w,\tau-w+1}\right)$$

Hence, the following are alternative representations of the claims process for cumulative data $Y_{w,\tau-w+1}$:

$$Y_{w,\tau-w+1} = C_w P_{w,\tau-w+1} \tag{1.2}$$

$$Y_{w,\tau-w+1} = C_w\left(1-\exp\left(-H_{w,\tau-w+1}\right)\right) \tag{1.3}$$

$$Y_{w,\tau-w+1} = C_w\left(1-S_{w,\tau-w+1}\right) \tag{1.4}$$

Equivalently, for incremental data $y_{w,\tau-w+1}$

$$y_{w,\tau-w+1} = C_w * \left(P_{w,\tau-w+1} - P_{w,\tau-w+1-1}\right) \tag{1.5}$$

$$y_{w,\tau-w+1} = C_w\left(\exp\left(-H_{w,\tau-w}\right) - \exp\left(-H_{w,\tau-w+1}\right)\right) \tag{1.6}$$

$$y_{w,\tau-w+1} = C_w * \left(S_{w,\tau-w} - S_{w,\tau-w+1}\right) \tag{1.7}$$

The underwriting year and array *IBNR* and reported *IBNR* projections are respectively

$$\begin{aligned} IBNR_{(w,\tau-w+1)} &= C_w S_{w,\tau-w+1} \\ IBNR(\tau) &= \sum_{w=1}^{u} IBNR_{(w,\tau-w+1)} \end{aligned} \tag{1.8}$$

$$\begin{aligned} RIBNR_{(w,\tau-w+1)} &= IBNR_{(w,\tau-w+1)} + \left(C_w S_{w,\tau-w+1} - Y_{w,\tau-w+1}\right) \\ RIBNR(\tau) &= \sum_{w=1}^{u} RIBNR_{(w,\tau-w+1)} \end{aligned} \tag{1.9}$$

where u is the number of underwriting years in the array. *RIBNR* links the reserving analysis to the accounting processes, by adjusting the *IBNR* by the difference between the total claim amount incurred to date and its estimate. Due to the additional noise induced by the adjustment, (1.9) is only applied in the final stages of the reserving analysis. In contrast to many published reserving methods, an important aspect of the models is the unrestricted *IBNR* projection periods, since the period before the last claim is generally unknown. The above equations could make explicit, and potentially highlight, the sources of data variability. Settlement speeds differences between underwriting years should be captured by $P_{w,r-w+1}$, $H_{w,r-w+1}$ or $S_{w,r-w+1}$. Although exposure levels are largely determined by underwriting volumes and contract terms, neither necessarily random, to accelerate convergence and formulate the final model variance function, random effects are introduced in C_w. When more than one claims array are analyzed, the additional aggregation level and source of variability is *array*, indexed by subscript *r*.

1.2 Objective

The examples' aim is to show that more than one model could fit historical data, but not all may reliably predict the reserves. The reliability of the *IBNR* and ultimate claim amount predictions depends on the models' capacity to extract from the data claims volume and settlement speeds measures. This is possible when the variability of both can be represented parametrically and formulated into the variance model. The scope of the models is made evident by their formulation and by the data. As the variability in settlement speeds and claims volumes increase the underlying assumptions of GLM are no longer sustainable, and more complex variance models and random effect parameters for the mean response become essential. To illustrate the process of constructing variance models two data sets are selected. One is a claims array simulated from a mixed portfolio, and the second consists of three arrays simulated from a marine hull, marine cargo and aviation hull portfolios. The second, selected to exacerbate the variability encountered in the first, in addition to large claims volume differences between underwriting years, contains also 20 negative incremental claims entries.

Since the concepts of population models (Zeger, Liang and Albert (1988)) are intended to average random variability between subjects, they are implemented around the percentage cash flow function. They can be used to obtain average (or array) *IBNR* predictions for a given ultimate loss. Other array or average results are the weighted average array or portfolio hazard rates. They provide thresholds, useful to quantify the impact on the claims portfolio of excluding from it underwriting contracts associated with particular underwriting years or arrays.

1.3 Outline

The paper structure is as follows. Section 2 introduces random effect models for one array with a general formulation of non-linear random effects models, and translated into a Bayesian framework in section 2.1.1. Noted in section 2.2 are amendments necessary to formulate multi-array models.

The models selected to analyze the two data sets are presented in sections 3 and 4 respectively. Denoted 1.0 and 2.0, in section 3.1 two preliminary models for one array are given, followed by numerical examples in section 3.3. The examples identify 2.0 as the basis for further analysis to construct the final models. In section 3.4.5 the results from two validation and two final models are discussed. Also in two stages, in section 4 multi-array models are constructed for two mean response functions denoted respectively 7.0 and 8.0. The preliminary models, used to establish data variability structures, are introduced in section 4.1, followed by numerical examples in section 4.2. For mean response functions 7.0 and 8.0, results for precision parameters σ^2, σ_r^2 and σ_{rw}^2 are obtained, identifying the three model versions by (a), (b) and (c). The final models, defined in section 4.3, are analyzed in section 4.5. They emphasise the contribution the generic model makes to the analysis of reserves, and to random effects models and variance models in general.

Section 4.4 extends the claims array average percentage cash flow definition given in section 3.2 to introduce portfolio model average for the percentage cash flow. As immediate by-products of the reserving analysis, hazard rates are discussed in section 4.6. The claims' hazard rate profile, essential for further portfolio analyses, can be used also as a portfolio management template. Discussion on the contribution made by the models proposed is given in section 5.

For the models in section 3, the results are fully reported in appendix A. Given the size of the data used in section 4, the reported results in this section are restricted to *IBNR* and ultimate claim amount projections for the selected preliminary and final models.

2. GENERAL FORMULATION OF NON-LINEAR RANDOM EFFECTS MODELS

In non-linear hierarchical models, inter and intra-underwriting year variations are analysed as a *two-stage process*. In the first, the intra-underwriting year variation is defined by a non-linear regression model for the underwriting year covariance structure. In the second stage, the inter-underwriting year variation is represented by both, systematic and random variability. The models can be constructed within a Bayesian hierarchical structure by noting that the intra-underwriting variation is associated with the sampling distribution, while the prior distribution is relevant to the inter-underwriting variation. Because the models' notation will depend on the number of aggregation levels, in sections 2.1 and 2.2 the array and multi-array analytical frameworks are respectively given.

2.1 Analytical Framework For a Claims Array

For the purpose of defining the general model, ignoring whether claims are cumulative or incremental, the observation at development time *t* of response vector for underwriting year *w* is simply denoted by $y_{w,t}$, and the model is defined as follows:

$$y_{w,t} = \mu_{w,t}\left(\phi_{w}\right) + \varepsilon_{w,t} \tag{2.1}$$

where $\mu_{w,t}$ is a non-linear function common to the entire array, while parameter vector ϕ_{w} is specific to underwriting year *w*. $t = t_1, \ldots, t_{n_w}$; with t_{n_w} representing the last period with known claims to date, $w = 1, \ldots, u$, and *u* is the number of cohorts or underwriting years in the claims array. Hence

$$y_{w} = \left[y_{w,t_1}, \ldots, y_{w,t_{n_w}} \right]^T$$
$$\mu_{w} = \left[\mu_{w,t_1}, \ldots, \mu_{w,t_{n_w}} \right]^T$$
$$\varepsilon_{w} = \left[\varepsilon_{w,t_1}, \ldots, \varepsilon_{w,t_{n_w}} \right]^T$$

and

$$\operatorname{cov}\left(\varepsilon_{w}\right) = \sigma^2 R_{w} \tag{2.2}$$

R_w is the intra-underwriting year covariance matrix for underwriting year *w*.

Inter-underwriting year variation accounted by ϕ_w is assumed to be random and, rather than simply regarding $\phi_i \neq \phi_w$ for $i \neq w$, the model represents

$$\phi_w = A_w \beta + B_w b_w$$

where β is a p-dimensional fixed parameter effects vector, and b_w is a q-dimensional underwriting year or random effects vector. Parameters b_w are independent and identically distributed with zero mean and variance covariance matrix Σ. Finally, A_w and B_w are $(n_w \times p)$ and $(n_w \times q)$ design matrices for the fixed and random effects respectively. While missing data from the earliest payment years and irregular reporting time intervals are allowed by the model formulation, the code and model specification for data given at regular intervals are simpler. The length of the response vector for the array is $M = \sum_{w=1}^{u} n_w$ and

$$\begin{array}{lll} y = [y_1, \ldots, y_u]^T & \phi = [\phi_1, \ldots, \phi_u]^T & \Sigma = diag[\Sigma_1, \ldots, \Sigma_u] \\ \mu = [\mu_1, \ldots, \mu_u]^T & b = [b_1^T, \ldots, b_u^T]^T & R = diag[R_1, \ldots, R_u] \\ \varepsilon = [\varepsilon_1, \ldots, \varepsilon_u]^T & & B = diag[B_1, \ldots, B_u] \\ & & A = [A_1^T, \ldots, A_u^T]^T \end{array}$$

Hence, the overall model becomes

$$\begin{aligned} E(y) &= \mu(\phi) \\ \operatorname{var}(y) &= \sigma^2 R \\ \phi &= A\beta + Bb \\ b &\sim (0, \Sigma) \end{aligned} \tag{2.3}$$

Corresponding to the two stages in the hierarchical models are two possible types of inferences or derived results: array and underwriting year cohort. Parameters common to all underwriting years relate to the array inferences, while underwriting year parameters measure underwriting year deviations from the claims array mean. Array inferences are generic when they represent insurance classes, and can help reassess or draft underwriting contracts, for instance. Alternatively, underwriting year parametric structures can set foundations for more discriminating premium rates reflecting systematic trends evident in the losses experienced. The latter can be viewed as a continuous calibration process.

Unless a book of business is closed, the number of observations in the most recent underwriting years could restrict the choice of viable variance and covariance models, particularly with non-linear model structures. Inferences on parameters of non-linear mixed effects models implemented in S-Plus (Pinheiro, Bates and Lindstrom (1994)) are based on the linear mixed effect model approximation of the log-likelihood function. This relies on the restricted maximum likelihood estimates derived from asymptotic results and on the approximate distribution for the maximum likelihood estimates. Since the maximum likelihood estimates in the linear mixed effect models are assumed to be asymptotically normal (Pinheiro and Bates (2000), Lindstrom and Bates (1990) and others), implementations with NLME library have to be approached with care to meet the criteria of the generic reserving model. Alternative assumptions are also considered. In non-parametric models the distribution of the random effects is left unspecified, hence completely unrestricted. Escobar and West (1992) propose a non-parametric approach, where ϕ_w are taken from distribution classes provided by the Dirichlet processes. Wakefield and Walker (1994) consider a non-parametric approach when random effect parameters are suspected to be neither normal nor Student t distributed, and allow for multimodality and skewness. Beal and Sleiner (1992) use a mixture of normal distributions and Wakefield (1996) a multivariate t-distribution for the random effect parameters and lognormal distribution for the response. The heavier tails in the t-distribution accommodate outlying cohorts. To define the parameters it is necessary to establish the curve's behaviour with parameter value changes, categorising the conditions, if any, for convergence, divergence, discontinuities etc (Ratkowski (1990)). The models' capacity to predict reserves depends on the stability of the projected curves, which in turn depends on the variance model structure. The most complex are more easily implemented within a Bayesian framework, as outlined below.

2.1.1 Three-Stage Models With Heterogeneous Intra-Underwriting Year Variation: A Bayesian Approach

Gibbs sampler application to Bayesian hierarchical models removes obstacles associated with non-linear multi-parameter structures integration. First to consider the problem of fully Bayesian non-linear regression is Wakefield et al. (1994). Bayesian random effects models can be represented by the following three-stage structure:

First Stage: Intra-underwriting year variation:

It accounts for variability within underwriting years, through scale parameter σ^2 and, in some cases, through functions $V_w(\phi_w, \Im(t), \vartheta)$ or $\Gamma_w(\rho)$, or both, such that ϕ_w, ϑ and ρ are parameters, $\Im(t)$ is some function of t and $\Gamma_w(\rho)$ is a correlation matrix. Hence, given

$$y_w = \mu_w(\phi_w) + \varepsilon_w$$

for the most general case

$$R_w(\phi_w, \vartheta^T, \rho) = V_w^{1/2}(\phi_w, \Im(t), \vartheta)\Gamma_w(\rho)V_w^{1/2}(\phi_w, \Im(t), \vartheta) \quad (2.4)$$

So ε_w are independently and identically distributed with zero mean and

$$Cov(\varepsilon_w \mid \phi_w, \sigma, \vartheta^T, \rho) = \sigma^2 R_w(\phi_w, \vartheta^T, \rho) \quad (2.5)$$

The functional form of $R_w(\phi_w, \vartheta^T, \rho)$ and covariance parameters $\zeta = \left[\sigma, \vartheta^T, \rho\right]^T$ are the same for all underwriting years. Implicit in $V_w(\phi_w, \Im(t), \vartheta)$ are functions of $\mu_w(\phi_w)$ or t, and of some or all parameters in ϕ_w. If probability distribution function is denoted by f then

$$(y_w \mid \phi_w, \zeta) \sim f_{y_w \mid \phi_w, \zeta}(y_w \mid \phi_w, \zeta)$$

Second Stage: Inter- underwriting year variation:

The inter-underwriting year variation in the values of ϕ_w is represented by

$$\phi_w = A_w \beta + B_w b_w \quad (2.6)$$

The degree of complexity of design matrices A_w and B_w will depend on the data and the percentage cash flow function. Random effect parameters are assumed to be independent and identically distributed:

$$b_w \sim f_{b_w \mid \Sigma}(b_w \mid \Sigma) \quad (2.7)$$

Zero mean assumption for b_w is not essential and, with software packages such as BUGS, may not be attainable. Non parametric and semiparametric model specifications for $\phi_{w\cdot}$ can be considered.

Third Stage: Hyperprior distribution:

Definition of parameters β,ζ and Σ completes the model formulation.

$$(\beta,\zeta,\Sigma)\sim f_{\beta,\zeta,\Sigma}(\beta,\zeta,\Sigma) \tag{2.8}$$

The joint posterior distribution of all parameters upon which the Bayesian inferences are based is

$$f_{\beta,\zeta,b,\Sigma|y}(\beta,\zeta,b,\Sigma\,|\,y)=\frac{f_{y|\beta,\zeta,b}(y\,|\,\beta,\zeta,b)f_{b|\Sigma}(b\,|\,\Sigma)f_{\beta,\zeta,\Sigma}(\beta,\zeta,\Sigma)}{f_y(y)} \tag{2.9}$$

The marginal posterior distributions of interest are $f_{\beta|y}(\beta\,|\,y)$, $f_{b|y}(b\,|\,y)$ and $f_{\Sigma|y}(\Sigma\,|\,y)$. Implicit in the above are two simpler models:

- For uncorrelated intra-underwriting year observations $\Gamma_{w\cdot}(\rho)=I_{n_w\times n_w}$ and $\zeta=\left[\sigma,\vartheta^T\right]^T$.
- If the model is homoscedastic, then $\Gamma_{w\cdot}(\rho)=R_w(\phi_{w\cdot},\vartheta^T,\rho)=I_{n_w\times n_w}$ and $\zeta=\sigma$.

As a simple example, consider

$$b_w|\Sigma\sim N(0,\Sigma) \tag{2.10}$$

and

$$\begin{gathered}\beta\left|\beta^*,\Sigma_0\sim N\left(\beta^*,\Sigma_0\right)\right.\\ \Sigma^{-1}\left|\Sigma^*,\nu^*\sim Wi\left(\left(\nu^*\Sigma^*\right)^{-1},\nu^*\right)\right.\\ \frac{1}{\sigma^2}\left|\upsilon,\nu\sim Ga\left(\frac{\upsilon}{2},\frac{\upsilon\nu}{2}\right)\right.\end{gathered} \tag{2.11}$$

where parameters $\beta^*,\Sigma_0,\Sigma^*,\nu^*,\upsilon,\nu$ are known. When a linearization method is used Σ_0 could be replaced by $\sigma^2\left(\hat{X}^T\hat{X}\right)^{-1}$, such that $\hat{X}=\left.\frac{\partial\mu}{\partial\beta}\right|_{\hat{\beta},\hat{b}}$. If $\bar{\beta}=\frac{1}{u}\sum_{w=1}^{u}\beta_{w\cdot}$, the parameters' conditional distributions for correlated observations are:

$$(\beta \mid y,\sigma,\Sigma,\phi_w, w=1,\dots,u) \sim N\left(\begin{matrix}\left(u\Sigma^{-1}+\left(\Sigma_0^{-1}+\Sigma^{-1}\right)^{-1}\right)^{-1}\left(u\Sigma^{-1}\bar{\beta}+\left(\Sigma_0^{-1}+\Sigma^{-1}\right)^{-1}\beta^{*}\right),\\ \left(u\Sigma^{-1}+\left(\Sigma_0^{-1}+\Sigma^{-1}\right)^{-1}\right)^{-1}\end{matrix}\right)$$

$$\left(\Sigma^{-1} \mid y,\sigma,\beta,\phi_w, w=1,\dots,u\right) \sim Wi\left(\left[\sum_{w=1}^{u}(\phi_w-\beta)(\phi_w-\beta)^T+v^{*}\Sigma^{*}\right]^{-1}, u+v^{*}\right) \qquad (2.12)$$

$$\left(\sigma^{-2} \mid y,\beta,\Sigma,\phi_w, w=1,\dots,u\right) \sim$$
$$Ga\left(\left(\frac{\upsilon+M}{2}\right),\frac{1}{2}\left(\upsilon v+\sum_{w=1}^{u}\left(y_w-\mu_w(\phi_w)\right)^T R_w^{-1}(\phi_w,\vartheta^T,\rho)\left(y_w-\mu_w(\phi_w)\right)\right)\right)$$

Then, the conditional distributions of ϕ_w and ϑ are

$$f_{\phi_w \mid y,\beta,\Sigma,\sigma,\phi_i,w\neq i}\left(\phi_w \mid y,\beta,\Sigma,\sigma,\phi_i, w\neq i\right) \propto$$
$$\exp\left(-\frac{1}{2\sigma^2}\left(y_w-\mu_w(\phi_w)\right)^T R_w^{-1}(\phi_w,\vartheta^T,\rho)\left(y_w-\mu_w(\phi_w)\right)-\frac{1}{2}(\phi_w-\beta)\Sigma^{-1}(\phi_w-\beta)^T\right)\left|R_w(\phi_w,\vartheta^T,\rho)\right|^{-1/2}\sigma$$

$$f_{\vartheta \mid y,\beta,\Sigma,\sigma,\phi_w,w=1,\dots,u}\left(\vartheta \mid y,\beta,\Sigma,\sigma,\phi_w, w=1,\dots,u\right) \propto \prod_{w=1}^{u}\left(\frac{\exp\left(-0.5\sigma^{-2}\left(y_w-f_{wt}(\phi_w)\right)^T R_w^{-1}(\phi_w,\vartheta^T,\rho)\left(y_w-f_{wt}(\phi_w)\right)\right)}{\sigma^{-1}\left|R_w(\phi_w,\vartheta^T,\rho)\right|^{1/2}}\right)$$

Variations on the above general model, with $\Gamma_w(\rho)=I_{n_w \times n_w}$ can be found in Wakefield (1996). In relation to the purpose of this paper, in the first stage, where variability structures are established, the predicted values that contribute to the $IBNR_{(w,t_i)}$ (equation (1.8)) are simply defined as $C_w\int_{t_j}^{t_r}\pi(w,z)dz$. While for the final models, y_w^{*}, or predicted losses for underwriting year w, are sampled from the distribution $f_{y_w^{*} \mid y,\beta,\Sigma,\sigma,\phi_w}\left(y_w^{*} \mid y,\beta,\Sigma,\sigma,\phi_w\right)$ and applied to equation (1.9).

2.2 General Formulation Of Multi-Array Bayesian Models

Extending the general model for a single array, the three-stage multi-array hierarchical model requires the following notation. For underwriting year w in claims array r, where $r=1,\dots,r_1$ and $w=1,\dots,u_r$, let the response vector be

$$y_{r,w}=\left[y_{r,w,t_1},\dots,y_{r,w,t_{n_{rw}}}\right]^T$$

Hence, for the entire data set

$$y = \left[y_{1,1}, \ldots, y_{1,u_1}, \ldots, y_{r_1,1} \ldots, y_{r_1,u_{r_1}} \right]^T$$

$t = t_1, \ldots, t_{n_{rw}}$ are the reporting times, such that $t_{n_{rw}}$ denotes the last period with known claims for underwriting year w. $y_{r,w}$ will be replaced by $Y_{r,w}$ when the data analysed is cumulative. The length of the response vector is M, such that $n_r = \sum_{w=1}^{u_r} n_{rw}$ and $M = \sum_{r=1}^{r_1} n_r$. We write

$$y_{r,w} = \mu_{r,w}(\phi_{rw}) + \varepsilon_{r,w}$$

where

$$\phi_{rw} = A_{rw}\beta + B_{r,w}b_r + B_{rw}b_{rw} \tag{2.13}$$

β is a p-dimensional fixed effects parameter vector, b_r is a q_1-dimensional first level random effects vector and b_{rw} a q_2-dimensional second level random effects vector. b_r and b_{rw} could be defined to have zero mean and variance/covariance matrices Σ_1 and Σ_2 respectively. Through design matrices A_{rw}, $B_{r,w}$ and B_{rw} information specific to each underwriting year data set can be brought into the analysis. By replacing (2.6) by (2.13) the three-stage models accounts also for array variation.

The models in section 3, and those in section 4, show that more flexible covariance structures could provide insight into the data variability structures by exploring alternative definitions for $\zeta = \left[\sigma^2, \vartheta^T, \rho\right]^T$. However, to avoid degrading inferences on first moment components, the final model should assume common parameters ζ for all underwriting years and arrays. Hence, the problem consists of identifying any relationship evident between $\Im(t)$, ϕ_{rw}, $\mu_{r,w}$ or any other function of ϕ_{rw}, and the patterns of variability revealed by parameters $\zeta = \left[\sigma^2, \vartheta^T, \rho\right]^T$. Outliers could lead to incorrect inferences, possibly indicate that the claims distribution is in fact multimodal and the data should be segmented for analytical purposes. Although the models proposed do not include specific functions to capture payment year effects of the kind of systematic inflation, they can be easily amended to do so.

3. MODELS FOR ONE ARRAY

3.1 Examples Of Preliminary Models For One Array

Two preliminary models, denoted 1.0 and 2.0 respectively, are given below. Both have a power variance function. However, to assess variability assumptions and construct the final models, the power in model 2.0 is allowed to change with underwriting year. With the variance formulation of model 2.0 the standard variance parameter definition is disregarded, by using instead $\zeta_{w} = [\sigma, \vartheta_{w}^{T}, \rho]^{T}$, thereby weakening the inferential capability of the model. Hence, even if the *IBNR* and ultimate claim amount predictions for model 2.0 were satisfactory, model 2.0 should be treated as preliminary and used exclusively for exploratory purposes.

3.1.1 Model 1.0

The first heteroscedastic model is defined as follows:

$$Y_{w,t} = \mu_{w,t}(\phi_{w}) + \varepsilon_{w,t}$$

with

$$(Y_{w,t} | \phi_{w}, \zeta) \sim N\left(\mu_{w,t}(\phi_{w}), \sigma^{2} \mu_{w,t}(\phi_{w})^{\exp(\vartheta)}\right)$$

such that, $\zeta = [\sigma, \vartheta]^{T}$

$$\mu_{w,t}(\phi_{w}) = \frac{\exp(L + l_{w})}{\left\{1 + \exp\left(D + d_{w} - \exp(Kc + kc_{w})\ln(t^{*}) - \exp(Kd + kd_{w}) * t^{*}\right)\right\}} \tag{3.1}$$

and

$$\varepsilon_{w} | \phi_{w}, \zeta \sim N\left(0, \sigma^{2}\left(\mu_{w}(\phi_{w})^{\exp(\vartheta)}\right)^{T} I_{n_{w} \times n_{w}}\right)$$

where $t^{*} = \left(t + \frac{\exp(Ks_{1})}{t^{\exp(Ks_{2})}}\right)$ and

$$\phi_{w} = A_{w}\beta + B_{w}b_{w}$$

$$A_{w} = I_{7\times 7}$$

$$(B_{w})^{T} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$b_{w} = [l_{w}, d_{w}, kc_{w}, kd_{w}]^{T}, \qquad b_{w} | b_{w}^{*}, \Sigma_{w} \sim MVN(b_{w}^{*}, \Sigma_{w})$$

$$\beta = [Ks_{1}, Ks_{2}, \vartheta, L, D, Kc, Kd]^{T}, \quad \beta | \beta^{*}, \Sigma_{0} \sim MVN(\beta^{*}, \Sigma_{0})$$

The hyperprior distributions are $\frac{1}{\sigma^2} \sim Ga(0.001, 0.001)$, and

$$\beta^{\bullet} \big| \beta^{\bullet\bullet}, \Sigma_0^{\bullet\bullet} \sim MVN\left(\beta^{\bullet\bullet}, \Sigma_0^{\bullet\bullet}\right) \qquad \left(\Sigma_0\right)^{-1} \big| \Sigma_0^{\bullet} \sim Wi\left(\left(7\Sigma_0^{\bullet}\right)^{-1}, 7\right)$$

$$b_{w}^{\bullet} \big| b^{\bullet\bullet}, \Sigma_{w}^{\bullet\bullet} \sim MVN\left(b^{\bullet\bullet}, \Sigma_{w}^{\bullet\bullet}\right) \qquad \left(\Sigma_{w}\right)^{-1} \big| \Sigma_{w}^{\bullet} \sim Wi\left(\left(4\Sigma_{w}^{\bullet}\right)^{-1}, 4\right)$$

for given parameters $\beta^{\bullet\bullet}, b^{\bullet\bullet}, \Sigma_0^{\bullet}, \Sigma_0^{\bullet\bullet}, \Sigma_w^{\bullet}, \Sigma_w^{\bullet\bullet}$. Functions C_w, $P_{w,t}$ and $h_{w,t}$ describing the underlying claims process for model 2.0 are:

$$C_w = \exp\left(L + l_w\right)$$

$$P_{w,t} = \left\{1 + \exp\left(D + d_w - \exp\left(Kc + kc_w\right)\ln\left(t^{\bullet}\right) - \exp\left(Kd + kd_w\right) * t^{\bullet}\right)\right\}^{-1} \tag{3.2}$$

$$S_{w,t} = 1 - P_{w,t} \tag{3.3}$$

$$h_{w,t} = \left(\frac{\exp\left(Kc + kc_w\right)}{t^{\bullet}} + \exp\left(Kd + kd_w\right)\right)\left(1 - \frac{\exp\left(Ks_2 + Ks_1\right)}{t^{\exp(Ks_2)+1}}\right) P_{w,t} \tag{3.4}$$

3.1.2 Model 2.0

In model 2.0 $\mu_{w,t}\left(\phi_w\right)$ is given by (3.1), but

$$\left(Y_{w,t} \big| \phi_w, \zeta_w\right) \sim N\left(\mu_{w,t}\left(\phi_w\right), \sigma^2 \mu_{w,t}\left(\phi_w\right)^{\exp(\vartheta + \vartheta_w)}\right)$$

where $\zeta_w = \left[\sigma, \vartheta, \vartheta_w\right]^T$. Having included random effect parameters in the variance function, the following further amendments to model 1.0 are needed:

$$A_w = B_w = I_{5\times 5}$$

$$\varepsilon_w \mid \phi_w, \zeta_w \sim N\left(0, \sigma^2 \left(\mu_w\left(\phi_w\right)^{\exp(\vartheta + \vartheta_w)}\right)^T I_{n_w \times n_w}\right)$$

$$b_w = \left[\vartheta_w, l_w, d_w, kc_w, kd_w\right]^T, \quad b_w \big| b_w^{\bullet}, \Sigma_w \sim MVN\left(b_w^{\bullet}, \Sigma_w\right)$$

such that

$$b_{w}^{\bullet} \big| b^{\bullet\bullet}, \Sigma_{w}^{\bullet\bullet} \sim MVN\left(b^{\bullet\bullet}, \Sigma_{w}^{\bullet\bullet}\right) \qquad \left(\Sigma_{w}\right)^{-1} \big| \Sigma_{w}^{\bullet} \sim Wi\left(\left(5\Sigma_{w}^{\bullet}\right)^{-1}, 5\right)$$

The power parameters in models 1.0 and 2.0 are formulated as multivariate normal, together with the mean response parameters. The reason becomes evident in section 3.4, where the relationship between the parameters is analysed to construct the final models.

3.2 Claims Array Average Percentage Cash Flow Model

Non-linear mixed effect or population models (Zeger, Liang and Albert (1988)) are intended to deliver population parameter distributions to derive population inferences. The inter-subject variability allowed by the models assumes that the subject-specific parameters are identically and independently distributed. The generic claims reserving model describes the data as the product of functions for the percentage cash flow and the ultimate claim amount. In the best scenario the ultimate claim amount function would account for differences in claim and exposure volumes. Since both could be largely determined by underwriting contract terms, for array inferences to be representative of the type of peril the contract covers, they are better based on the percentage cash flow functions alone.

In the example that follows the general formulation of the random effects model, the random effects parameters are set to be $b_w|\Sigma \sim N(0,\Sigma)$, while observing that alternative definitions are feasible. In some applications or models it may not be possible to assume a zero mean for the random effects parameters, particularly when they are defined to belong to multivariate distributions. BUGS, for instance, cannot handle multivariate range restrictions, but can accommodate some simpler univariate centering forms.

Replacing design matrices A_w and B_w in models 1.0 and 2.0 by A and B respectively, the parameters for the claims array average percentage cash flow model have to be extracted from the parameter vector given by

$$\phi_A = A\beta + B\left(\frac{1}{u}\sum_{w=1}^{u} b_w\right)$$

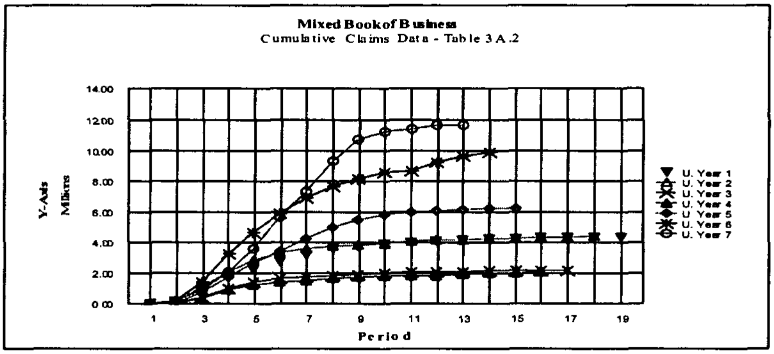
Hence, for model 2.0, the claims array average percentage cash flow curve is:

$$P_{A,t} = \left\{1+\exp\left(D+\left(\frac{1}{u}\sum_{w=1}^{u} d_w\right)-\exp\left(Kc+\left(\frac{1}{u}\sum_{w=1}^{u} kc_w\right)\right)\ln\left(t^*\right)-\exp\left(Kd+\left(\frac{1}{u}\sum_{w=1}^{u} kd_w\right)\right)*t^*\right)\right\}^{-1} \qquad (3.5)$$

with t^* defined as before. To ascertain if (3.5) is representative of the array, the curve is compared to the plots for the percentage cash flow for all underwriting years in the array.

3.3 Numerical Examples Of Preliminary Models 1.0 Aand 2.0

Extracted from a book of business containing more than one type of claim, the data selected for the examples display significant differences in the development patterns and exposure volumes across underwriting years, particularly evident in the last three underwriting years. (See graph 3.3.1 and tables A.1 and A.2). Another characteristic is the zero claims in the first reporting period. To ensure they are not interpreted as missing data, they have been set to one. This artifice is often necessary with non-linear models for the mean response or when the mean response is formulated into the variance.

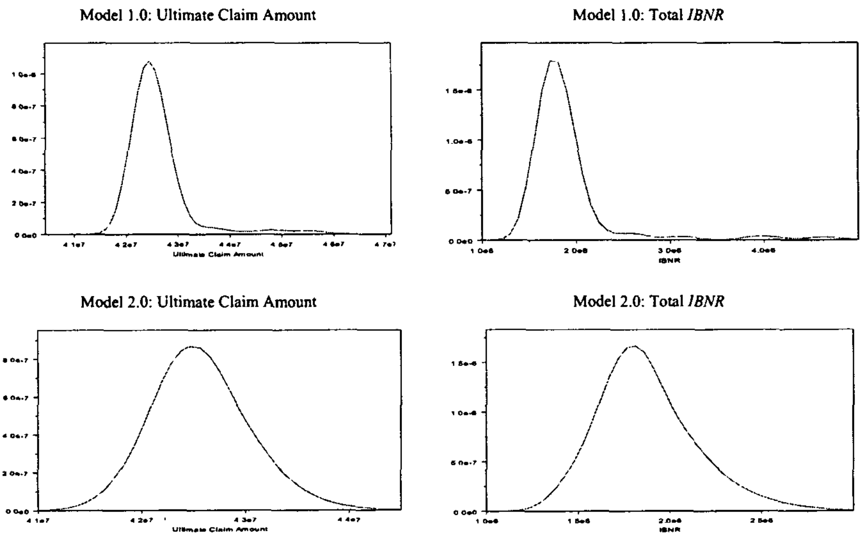


Graph 3.3.1 Cumulative paid claims data aggregated on annual basis.

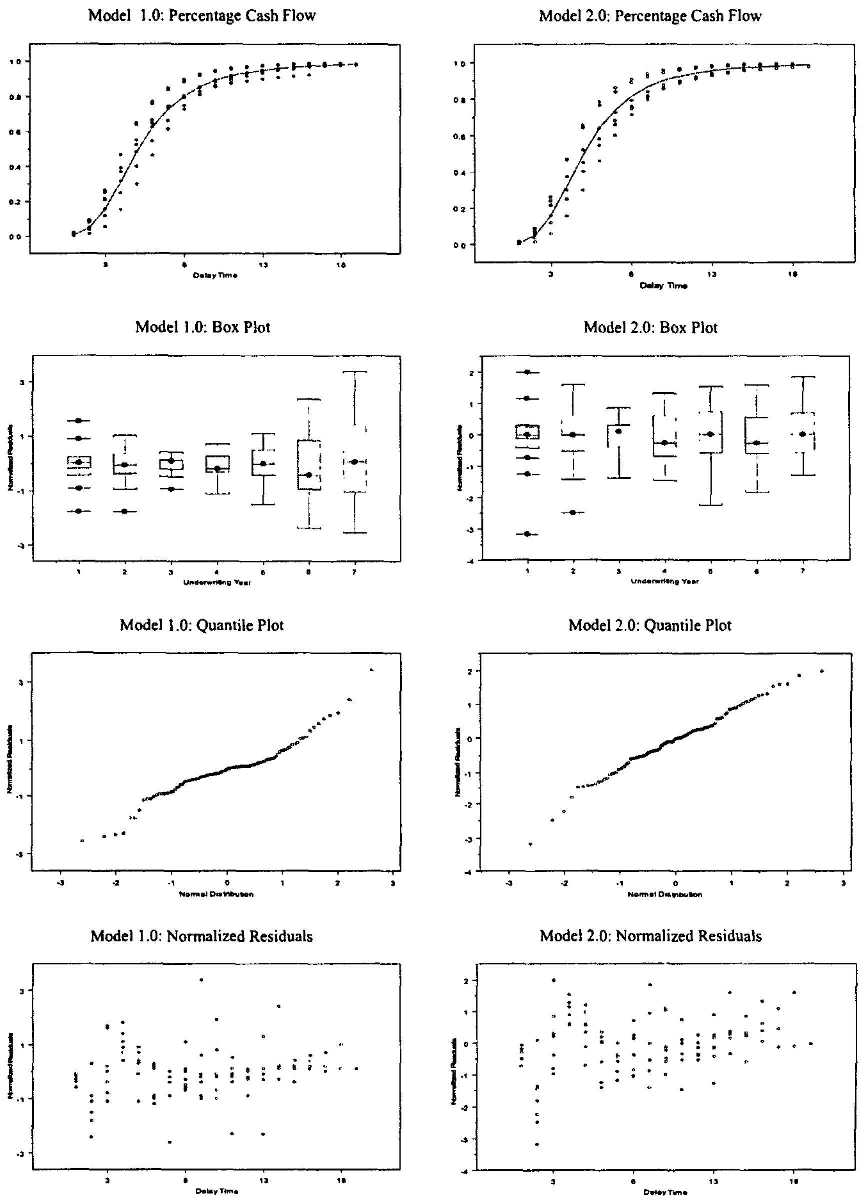
Of interest in the examples are the repercussions of hierarchical variance models. To facilitate the analysis of the preliminary models, the *IBNR* predictions do not include the accounting adjustment in (1.9) Graphs for observed claims and fitted values for the preliminary and final models would show that the fitted curves are almost indistinguishable and very close to the data. However, from table 3.3.1 and graph 3.3.3 observe that the *IBNR* predictions at underwriting year level for model 1.0 cannot be reliably used. The plot for the percentage cash flow for underwriting year 4 is unlikely to converge to 1. The model compensates by producing a higher *IBNR*. As graphical representations of spread, location and skewness for error distributions, the box plots show that, in contrast with model 1.0, with the introduction of parameter ϑ_w in the variance function, model 2.0 deals effectively with scale variability and with some of the outliers evident in the quantile plots.

		Ultimate Claim Amount				*IBNR* (1.8)			
		Mean	Standard Mean Sq. Pred. Error	Predictive Interval		Mean	Standard Mean Sq. Pred. Error	Predictive Interval	
				2.50%	97.50%			2.50%	97.50%
Model 1.0									
Und. Year	1	4,442,000	80,410	4,302,000	4,611,000	95,490	28,990	53,920	161,100
	2	4,342,000	73,480	4,219,000	4,487,000	46,260	18,980	19,770	87,900
	3	2,180,000	65,470	2,059,000	2,311,000	22,850	14,550	3,489	57,970
	4	2,179,000	539,200	1,889,000	4,383,000	173,100	467,000	4,487	2,116,000
	5	6,642,000	115,100	6,413,000	6,863,000	290,800	51,360	197,400	397,800
	6	10,170,000	153,100	9,892,000	10,500,000	607,200	85,100	457,700	791,100
	7	12,650,000	131,400	12,380,000	12,880,000	676,100	54,630	556,600	774,600
	Total	42,600,000	607,100	41,940,000	44,830,000	1,912,000	499,400	1,546,000	3,915,000
Model 2.0									
Und. Year	1	4,481,000	58,480	4,373,000	4,604,000	113,500	23,960	74,340	168,000
	2	4,327,000	48,600	4,233,000	4,424,000	40,920	11,260	21,990	66,410
	3	2,165,000	39,650	2,093,000	2,249,000	16,400	8,756	5,335	38,440
	4	2,007,000	54,090	1,909,000	2,128,000	31,570	17,560	8,824	79,130
	5	6,644,000	88,190	6,474,000	6,822,000	293,500	42,590	217,800	382,600
	6	10,160,000	257,300	9,668,000	10,700,000	606,400	144,000	348,000	933,100
	7	12,790,000	348,100	12,170,000	13,580,000	771,200	200,400	458,100	1,259,000
	Total	42,570,000	471,100	41,720,000	43,590,000	1,873,000	268,600	1,421,000	2,502,000

Table 3.3.1 Ultimate losses and *IBNR* predictive distributions for models 1.0 and 2.0



Graph 3.3.2 Kernel densities for ultimate losses and *IBNR* totals for preliminary models 1.0 and 2.0.



Graph 3.3.3 Percentage cash flow plots and normalized residuals for models 1.0 and 2.0.

Hence, model 2.0 in general, and ϑ_w in particular, should be analysed to formulate the final variance model. The Kernel densities for ultimate claim amount and *IBNR* projections in graph 3.3.2 suggest possible bi-modality, particularly for model 1.0.

Note that a variance derived directly from model 2.0 may not deal completely with the pattern evident in the plots for the normalized residuals (graph 3.3.3). Portfolio transfers or account consolidations often produce data sets where the settlement speeds of the new and old data differ significantly. The quantile and scatter plots point to the second observation in underwriting years 1, 2, 5 and 6 as possible outliers. These give an indication that the correction needed in the variance model may involve a function dependent on delay period *t*. In the next section the variance function for model 5.0 is derived form the output of model 2.0. With the variance function for model 6.0 it is aimed to deal with remaining outliers.

3.4. Final One-Array Models

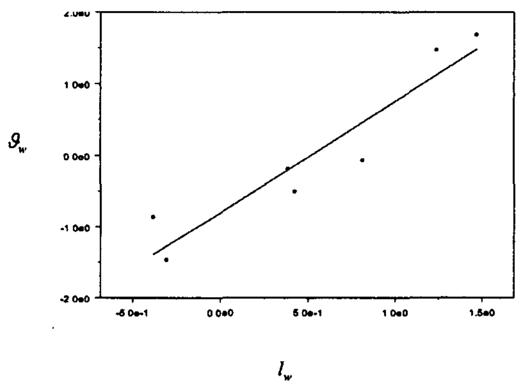
The generic model conveniently separates the percentage cash flow and the ultimate claim amount functions and, through the percentage cash flow function, can extract from the data settlement speed characteristics. Deviations induced by large differences in underwriting volumes between underwriting years may not be captured by random effect models, and the introduction of cluster structures may be necessary. The criteria needed to establish them remains to be determined.

	Model 1.0		Model 2.0	
	Fixed parameters		Fixed parameters	
	L	ϑ	L	ϑ
	15.7400	-6.2470	14.8900	-3.4710
Book Year	Random parameters		Random parameters	
	l_w	ϑ_w	l_w	ϑ_w
1	-0.4364		0.4214	-0.5061
2	-0.4592		0.3865	-0.1910
3	-1.1480		-0.3060	-1.4670
4	-1.1690		-0.3818	-0.8611
5	-0.0342		0.8153	-0.0693
6	0.3919		1.2400	1.4680
7	0.6101		1.4700	1.6780

Table 3.4.1 Parameter estimates for variance model and C_w function.

From tables 3.3.1 and 3.4.1 note the approximate correspondence between the order of magnitudes of c_w and ϑ_w for model 2.0. Hence consider the following regression line:

$$\vartheta_w = \delta_1 + \delta_2 l_w \tag{3.6}$$



Graph 3.4.1 Line $\vartheta_w = \delta_1 + \delta_2 l_w$ and scatter plot of l_w vs ϑ_w.

(3.6) gives $[\delta_1, \delta_2] = [-0.79914, 1.54866]$. From table A.6, there is no evident relationship, similar to (3.6), between ϑ_w and any of d_w, kc_w or kd_w, directly or through a suitable transformation. Graph 3.4.1 displays the scatter plot of ϑ_w versus l_w and regression line $\vartheta_w = \delta_1 + \delta_2 l_w$. $\exp(\vartheta + \vartheta_w)$ lies between 0.007 and 0.166, such that the minimum and maximum values correspond to $w = 3$ and $w = 7$ respectively. Had model 1.0 provided a better fit, the magnitudes of $\exp(\vartheta + \vartheta_w)$ could have influenced a decision to select a homoscedastic model. However, from table 3.3.1 note that $C_3 = 2.17$ million and $C_7 = 12.79$ million. Hence, it is justifiable to integrate $\exp(\vartheta + \vartheta_w) \cong \exp((\vartheta + \delta_1) + l_w * \delta_2)$ in the variance function definition as follows:

$$\text{var}(Y_{w,t}) = \sigma^2 \mu_{w,t} (\phi_w)^{\exp(\vartheta_1^* + l_w \vartheta_2^*)} \tag{3.7}$$

(3.7) satisfies covariance definition (2.5) because parameter vector $\zeta = \left[\sigma, \left[\vartheta_1^*, \vartheta_2^*\right]^T\right]^T$ is invariant with underwriting year. It is intuitively obvious that if the variance parameters for preliminary model 2.0 would have been defined as $\zeta_w = [\sigma_w, \vartheta^T]^T$ instead of $\zeta_w = [\sigma, \vartheta_w^T]^T$, a similar relationship to (3.6) would be evident between σ_w and l_w. This is further explored in section 4. Although the variance for final model 5.0 will be (3.7), to validate the model and explore alternative analytical approaches, the final model is preceded by other two. The first gives an appreciation of the *IBNR* reserve values that an analysis of the data segmented into *K* subsets would deliver, where subset membership criteria is determined by the values of c_w or ϑ_w. Hence, the variance function considered is

$$\text{var}(Y_{w,t}) = \sigma_k^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta)}$$

The second preliminary model assumes an autoregressive error structure. Variance function (3.7) may not successfully explain the variability evident in the normalized residual plot pattern of model 2.0 (graph 3.3.3) and function $\mu_{w,t}(\phi_w)^{\exp(\vartheta_1^* + l_w\vartheta_2^*)}$ may need to be adjusted. Hence, the function proposed for model 6.0 is

$$\text{var}(Y_{w,t}) = \sigma^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta_1^* + l_w\vartheta_2^*)} \exp\left(\vartheta_3^* (t^*)^{\vartheta_4^*}\right)$$

3.4.1 Model 3.0 – Validation Model

Model 3.0 is equal to model 1.0 in all respects, except that subset membership for each underwriting year is taken into account only at the point of calculating the variance, and for subset *k* σ_k^2 is estimated independently from the rest of the data. For underwriting year *w*, member of subset *k*

$$Y_{w,t} = \mu_{w,t}(\phi_w) + \varepsilon_{w,t}$$

with

$$\left(Y_{w,t} | \phi_w, \zeta_k\right) \sim N\left(\mu_{w,t}(\phi_w), \sigma_k^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta)}\right)$$

such that $\zeta_k = [\sigma_k, \vartheta]^T$, $\frac{1}{\sigma_k^2} \sim Ga(0.001, 0.001)$ and

$$\varepsilon_w | \phi_w, \zeta_k \sim N\left(0, \sigma_k^2 \left(\mu_w(\phi_w)^{\exp(\vartheta)}\right)^T I_{n_w \times n_w}\right)$$

3.4.2 Model 4.0 – Validation Model

In model 4.0 the option of using an autoregressive error structure is explored, to ascertain if this can effectively deal with scale variability between underwriting years:

$$Y_{w,t} = \mu_{w,t}(\phi_w) + W_{w,t}$$

such that

$$W_{w,t} = \rho W_{w,t-1} + \varepsilon_{w,t}$$

$$VAR(Y_{w,t}) = VAR(W_{w,t}) = \sigma^2$$

$$\varepsilon_w \mid \phi_w, \zeta_w \sim N\left(0, \sigma^2\left(1-\rho^2\right)\right)$$

and $\zeta = [\sigma, \rho]^T$.

3.4.3 Model 5.0 – Validation Model

Final model 5.0 integrates regression (3.6) into the variance model:

$$Y_{w,t} = \mu_{w,t}(\phi_w) + \varepsilon_{w,t}$$

with

$$\left(Y_{w,t} \middle| \phi_w, \zeta\right) \sim N\left(\mu_{w,t}(\phi_w), \sigma^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta_1^* + I_w \vartheta_2^*)}\right)$$

where $\vartheta^T = \left[\vartheta_1^*, \vartheta_2^*\right]$, $\varepsilon_w \mid \phi_w, \zeta \sim N\left(0, \sigma^2\left(\mu_w(\phi_w)^{\exp(\vartheta_1^* + I_w \vartheta_2^*)}\right)^T I_{n_w \times n_w}\right)$, $\frac{1}{\sigma^2} \sim Ga(0.001, 0.001)$ and $\zeta = \left[\sigma, \vartheta^T\right]^T$. Since $\mu_{w,t}$ is given by (3.1), in addition to the obvious changes in the design matrices, the other necessary amendments to model 2.0 are:

$$\beta = \left[Ks_1, Ks_2, \vartheta_1^*, \vartheta_2^*, L, D, Kc, Kd\right]^T, \quad \beta \middle| \beta^*, \Sigma_0 \sim MVN\left(\beta^*, \Sigma_0\right)$$

$$\beta^* \middle| \beta^{**}, \Sigma_0^{**} \sim MVN\left(\beta^{**}, \Sigma_0^{**}\right) \qquad \left(\Sigma_0\right)^{-1} \middle| \Sigma_0^* \sim Wi\left(\left(8\Sigma_0^*\right)^{-1}, 8\right)$$

3.4.4 Model 6.0 – Final Model

Final model 6.0 extends the variance model (3.7) as follows:

$$\left(Y_{w,t} \middle| \phi_w, \zeta\right) \sim N\left(\mu_{w,t}(\phi_w), \sigma^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta_1^* + I_w \vartheta_2^*)} \exp\left(\vartheta_3^*\left(t + \frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)^{\vartheta_4^*}\right)\right)$$

where $\zeta = [\sigma, \vartheta^T]^T$ and $\vartheta^T = [\vartheta_1^*, \vartheta_2^*, \vartheta_3^*, \vartheta_4^*]$. Hence the fixed effect parameter vector and related distributions are:

$$\beta = [Ks_1, Ks_2, \vartheta_1^*, \vartheta_2^*, \vartheta_3^*, \vartheta_4^*, L, D, Kc, Kd]^T, \quad \beta | \beta^*, \Sigma_0 \sim MVN(\beta^*, \Sigma_0)$$

$$\beta^* | \beta^{**}, \Sigma_0^{**} \sim MVN(\beta^{**}, \Sigma_0^{**}) \qquad (\Sigma_0)^{-1} | \Sigma_0^* \sim Wi\left((10\Sigma_0^*)^{-1}, 10\right)$$

3.4.5 Numerical Examples And Discussion For Validation Models 3.0 And 4.0 And Final Models 5.0 And 6.0

Data segmentation criteria described by the last column in table 3.4.5.1 and applied to model 3.0 is given by the values of l_w from model 2.0.

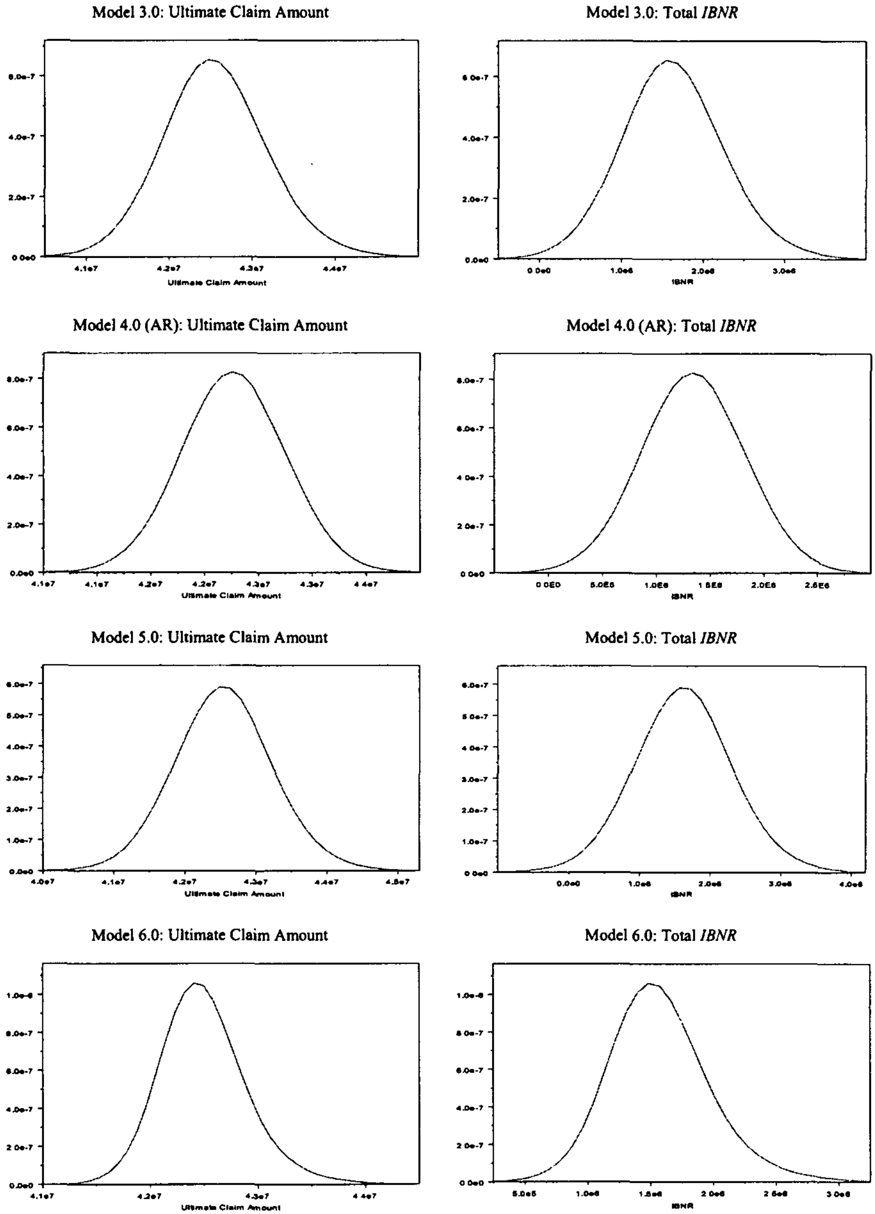
	Model 2.0		Model 3.0		Subset Membership For Model 3.0
	Fixed parameters		Fixed parameters		
	L	ϑ	L	ϑ	
	14.890	-3.471	14.290	-6.861	
Book Year	Random parameters		Random parameters		
	l_w	ϑ_w	l_w	ϑ_w	
1	0.4214	-0.5061	1.0300		1
2	0.3865	-0.1910	0.9940		1
3	-0.3060	-1.4670	0.3015		2
4	-0.3818	-0.8611	0.2259		2
5	0.8153	-0.0693	1.4240		3
6	1.2400	1.4680	1.8620		3
7	1.4700	1.6780	2.0660		3
Book Year	Combined Effect		Combined Effect		
	$L+l_w$	$\exp(\vartheta+\vartheta_w)$	$L+l_w$	$\exp(\vartheta+\vartheta_w)$	
1	15.3114	-3.9771	15.3200		1
2	15.2765	-3.6620	15.2840		1
3	14.5840	-4.9380	14.5915		2
4	14.5082	-4.3321	14.5159		2
5	15.7053	-3.5403	15.7140		3
6	16.1300	-2.0030	16.1520		3
7	16.3600	-1.7930	16.3560		3
σ^2	5.87E+09				
Subset 1 - σ_1^2			7.496E+09		
Subset 2 - σ_2^2			3.328E+09		
Subset 3 - σ_3^2			3.182E+10		
Deviance	2,936		2,937		

Table 3.4.5.1 Scale and deviance values and parameter estimates for models 2.0 and 3.0.

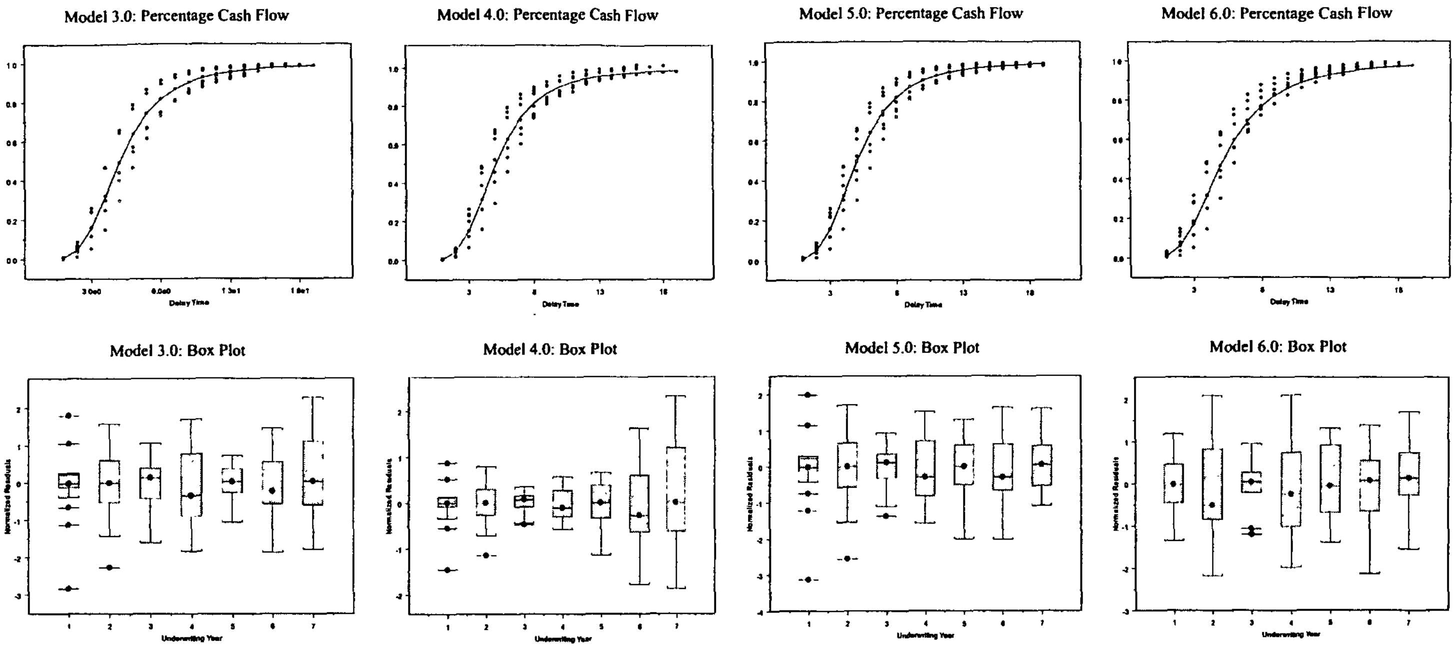
The table compares parameters L, l_w and $L+l_w$ for both models. Variance function power for model 3.0 is very small. For a model with variance σ_k^2, instead of $\sigma_k^2 \mu_{w,t}(\phi_w)^{\exp(\vartheta)}$, the values of $L+l_w$ are not significantly different. Although that model version is excluded from this paper, it is observed that its results indicate that in model 3.0 information on the data variability structures is mainly contained in σ_k^2, and that neither model successfully deals with claim volume differences between underwriting years.

		Ultimate Claim Amount				Reported *IBNR* (1.9)				Subset Membership For Model 3.0
		Mean	Standard Mean Sq. Pred. Error	Confidence Interval 2.50%	Confidence Interval 97.50%	Mean	Standard Mean Sq. Pred. Error	Confidence Interval 2.50%	Confidence Interval 97.50%	
Model 3.0		Validation Model								
Und. Year	1	4,480,000	122,100	4,239,000	4,721,000	115,200	122,100	-125,400	356,600	1
	2	4,324,000	115,800	4,097,000	4,556,000	-121,600	115,800	-347,900	110,200	1
	3	2,163,000	76,730	2,012,000	2,315,000	-21,340	76,730	-172,300	130,700	2
	4	2,004,000	80,470	1,847,000	2,164,000	-86,010	80,470	-243,900	73,750	2
	5	6,648,000	280,700	6,103,000	7,214,000	358,600	280,700	-186,000	924,700	3
	6	10,300,000	340,500	9,656,000	11,000,000	395,000	340,500	-251,900	1,088,000	3
	7	12,630,000	318,900	12,020,000	13,280,000	988,900	318,900	376,800	1,637,000	3
	Total	42,550,000	607,500	41,400,000	43,810,000	1,629,000	607,500	473,400	2,891,000	
Model 4.0		Validation Model - (AR) error structure								
Und. Year	1	4,451,000	164,000	4,131,000	4,776,000	86,840	164,000	-233,200	411,400	
	2	4,315,000	159,700	4,002,000	4,630,000	-130,300	159,700	-443,400	185,000	
	3	2,149,000	157,300	1,842,000	2,462,000	-34,560	157,300	-342,200	277,800	
	4	1,995,000	165,100	1,677,000	2,322,000	-94,990	165,100	-413,700	231,400	
	5	6,635,000	180,900	6,277,000	6,989,000	345,800	180,900	-11,950	699,800	
	6	10,130,000	211,700	9,732,000	10,570,000	226,500	211,700	-176,300	660,900	
	7	12,590,000	229,900	12,130,000	13,030,000	945,900	229,900	486,100	1,391,000	
	Total	42,270,000	469,500	41,340,000	43,190,000	1,345,000	469,500	420,800	2,264,000	
Model 5.0		Final Model								
Und. Year	1	4,477,000	114,300	4,254,000	4,703,000	112,700	114,300	-110,900	338,300	
	2	4,321,000	105,800	4,112,000	4,527,000	-124,800	105,800	-333,000	81,830	
	3	2,162,000	87,570	1,990,000	2,336,000	-22,300	87,570	-194,300	151,900	
	4	2,001,000	92,290	1,823,000	2,188,000	-89,490	92,290	-267,900	97,450	
	5	6,645,000	164,000	6,327,000	6,971,000	356,000	164,000	38,250	681,600	
	6	10,190,000	328,100	9,576,000	10,870,000	280,200	328,100	-332,600	965,000	
	7	12,750,000	513,700	11,770,000	13,800,000	1,110,000	513,700	127,000	2,162,000	
	Total	42,550,000	675,900	41,240,000	43,920,000	1,623,000	675,900	313,300	2,994,000	
Model 6.0		Final Model								
Und. Year	1	4,418,000	37,320	4,345,000	4,492,000	53,860	37,320	-19,210	127,700	
	2	4,426,000	42,000	4,338,000	4,504,000	-19,530	42,000	-107,100	58,360	
	3	2,201,000	26,940	2,146,000	2,253,000	16,980	26,940	-38,450	69,290	
	4	2,094,000	36,930	2,016,000	2,161,000	3,306	36,930	-74,000	70,160	
	5	6,525,000	70,370	6,395,000	6,673,000	235,700	70,370	106,000	384,200	
	6	10,470,000	231,800	10,040,000	10,960,000	560,000	231,800	136,600	1,054,000	
	7	12,370,000	283,300	11,880,000	13,010,000	724,400	283,300	235,500	1,365,000	
	Total	42,500,000	393,100	41,800,000	43,380,000	1,575,000	393,100	876,400	2,453,000	

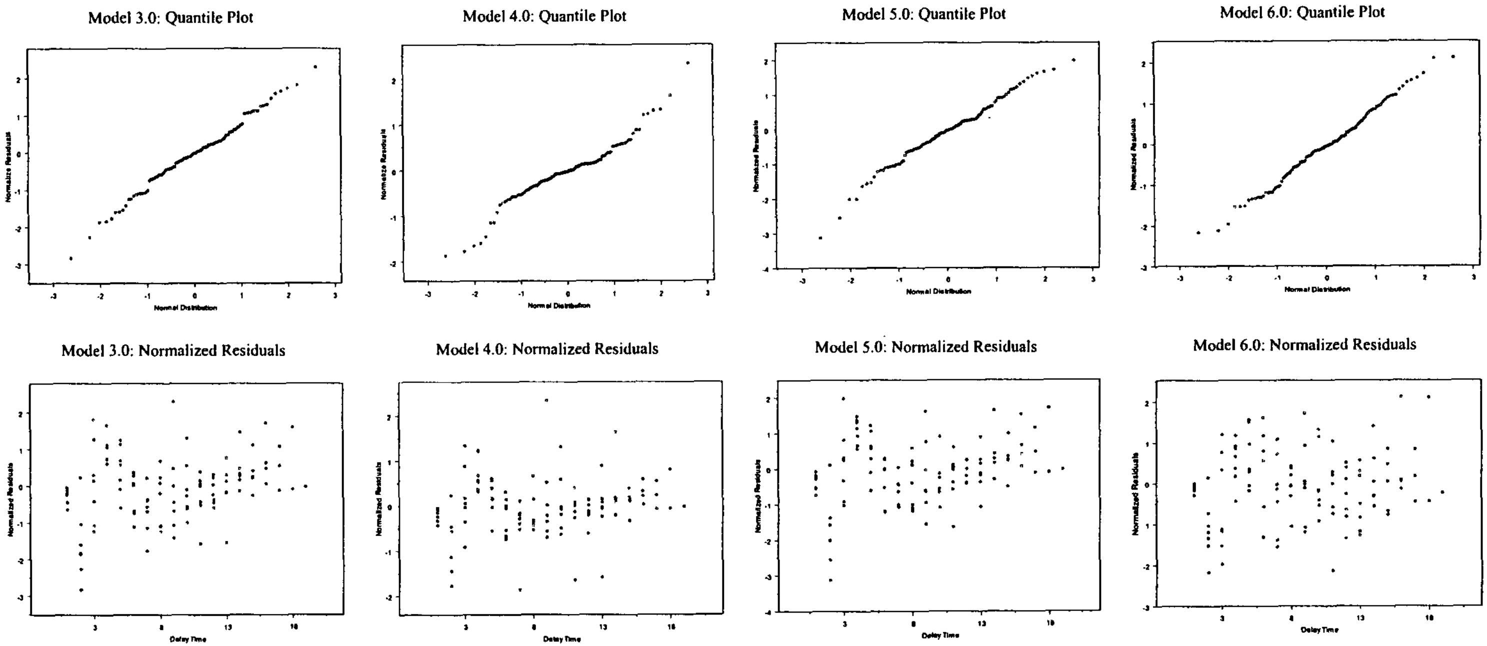
Table 3.4.5.2 Models 3.0 to 6.0: Ultimate losses and *IBNR* predictions and predictive distributions.



Graph 3.4.5.1 Models 3.0 to 6.0: Kernel densities and predictive distributions for ultimate losses and *IBNR* totals.



Graph 3.4.5.2 Models 3.0 to 6.0: Scatter plots and average array curve for percentage cash flow versus delay time and Box plots of normalized residuals.



Graph 3.4.5.3 Models 3.0 to 6.0: Quantile and scatter plots of normalized residuals.

Model	Preliminary models																	Devi.
	Predictive distributions								Log Likelihood			AIC			BIC			
	Ultimate Claim Amount				*IBNR* (1.8)													
	Mean	Standard Deviation	Predictive Interval		Mean	Standard Deviation	Predictive Interval		Mean	Confidence Interval		Mean	Confidence Interval		Mean	Confidence Interval		
			2.5%	97.5%			2.5%	97.5%		2.5%	97.5%		2.5%	97.5%		2.5%	97.5%	
1.0	42,600,000	607,100	41,940,000	44,830,000	1,912,000	499,400	1,546,000	3,915,000										2,980
2.0	42,570,000	471,100	41,720,000	43,590,000	1,873,000	268,600	1,421,000	2,502,000										2,930

Model	Validation models 3.0 and 4.0 and final models 5.0 and 6.0																	Devi.
	Predictive distributions								Log Likelihood			AIC			BIC			
	Ultimate Claim Amount				Reported *IBNR* (1.9)													
	Mean	Standard Deviation	Predictive Interval		Mean	Standard Deviation	Predictive Interval		Mean	Confidence Interval		Mean	Confidence Interval		Mean	Confidence Interval		
			2.5%	97.5%			2.5%	97.5%		2.5%	97.5%		2.5%	97.5%		2.5%	97.5%	
3.0	42,550,000	607,500	41,400,000	43,810,000	1,629,000	607,500	473,400	2,891,000	56.0	42.3	71.6	181.9	154.7	213.1	277.1	249.8	308.3	2,936
4.0	42,270,000	469,500	41,340,000	43,190,000	1,345,000	469,500	420,800	2,264,000	56.0	42.6	71.7	182.0	155.1	213.5	277.2	250.3	308.6	2,902
5.0	42,550,000	675,900	41,240,000	43,920,000	1,623,000	675,900	313,300	2,994,000	56.0	42.3	71.6	183.9	156.6	215.1	281.8	254.4	313.0	2,928
6.0	42,500,000	393,100	41,800,000	43,380,000	1,575,000	393,100	876,400	2,453,000	56.0	42.3	71.5	187.9	160.5	219.0	291.2	263.8	322.3	2,887

Table 3.4.5.3 Comparison of results for models 1.0 to 6.0.

The ultimate claim amount and *IBNR* predictions for final models 5.0 and 6.0 and preliminary models 3.0 and 4.0 are compared in table 3.4.5.2. The boxplots for model 5.0 are the most consistent with those of model 2.0. (see graphs 3.4.5.2 and 3.3.3), but the predictive intervals are slightly wider than for models 3.0 and 4.0. The autoregressive error structure in model 4.0 is insufficient to deal with scale variability. In contrast to model 6.0, models 2.0 to 5.0 do not resolve the downwards pattern in the quantile plots (see graph 3.4.5.3). From the percentage cash flow plots and the array average percentage cash flow curve it is evident that the curve is representative of the array. The additional variance parameters increase the AIC and BIC values with respect to model 3.0, but decrease the deviance (table 3.4.5.3). The slight skewness of the *IBNR* and the ultimate claim amount kernel densities for model 2.0 is no longer so evident in models 3.0 to 6.0 (see graphs 3.3.2 and 3.4.5.1).

4. MULTI-ARRAY MODELS

To explore data variability structures and illustrate the process of designing multiple-array models, two mean response functions are used. For the preliminary models the variance functions considered are σ^2, σ_r^2 and σ_{rw}^2, denoting the three model versions by a, b, and c respectively. In section 4.2, observations on the models and numerical examples highlight the motivation for their inclusion. In section 4.3 the values of σ_{rw}^2 are analysed and the final multi-array models are introduced. Numerical examples and assessment of the final models are given in sections 4.5 and 4.6.

4.1 Examples Of Preliminary Multi-Array Models

4.1.1 Models 7.0 (a), (b) and (c)

Model 7.0 is proposed as example of hierarchical reserving models with a limited number of parameters in the percentage cash flow function. It is followed by two amended versions selected to further explore variability patterns in the data.

Model 7.0(a)

For claims array r and underwriting year w, the first homoscedastic model at delay time t is defined as follows:

$$Y_{r,w,t} = \mu_{r,w,t}(\phi_{rw}) + \varepsilon_{r,w,t}$$

such that,

$$\mu_{r,w,t}(\phi_{rw}) = \exp\left(L + l^{*}_{rw}\right)\left\{1 - \frac{\left(1 + 2\left(\frac{\exp\left(D + d_r + d^{*}_{rw}\right)}{\exp\left(Kc + kc_r + kc^{*}_{rw}\right)}\right)\ln\left(t + \frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)}{\frac{\exp\left(\left(\exp\left(Kc + kc_r + kc^{*}_{rw}\right) + 1\right)\ln\left(t + \frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)}{\exp\left(-\exp\left(D + d_r + d^{*}_{rw}\right)\left(\ln\left(t + \frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)^2\right)}}\right\} \tag{4.1}$$

and

$$\begin{aligned} \phi_{rw} &= A_{rw}\beta + B_{r,w}b_r + B_{rw}b_{rw} \\ \varepsilon_{r,w}\,|\,\phi_{rw},\sigma &\sim N\left(0, \sigma^2 I_{n_{rw}\times n_{rw}}\right) \end{aligned} \tag{4.2}$$

where

$$\begin{aligned} b_{rw} &= \left[l^{*}_{rw}, d^{*}_{rw}, kc^{*}_{rw}\right]^T, & b_{rw}\,|\,b^{*}_{rw}, \Sigma^{o}_{r} &\sim MVN\left(b^{*}_{rw}, \Sigma^{o}_{r}\right) \\ b_r &= \left[d_r, kc_r\right]^T, & b_r\,|\,b^{*}_{r}, \Sigma_r &\sim MVN\left(b^{*}_{r}, \Sigma_r\right) \\ \beta &= \left[Ks_1, Ks_2 L, D, Kc\right]^T, & \beta\,|\,\beta^{*}, \Sigma_0 &\sim MVN\left(\beta^{*}, \Sigma_0\right) \end{aligned}$$

The configuration of the design matrices is determined by the order of the parameters in the fixed and random effect parameter vectors. For known parameters $\beta^{**}, b^{**}_{rw}, b^{**}_{r}, \Sigma^{*}_{0}, \Sigma^{*o}_{r}, \Sigma^{*}_{r}, \Sigma^{**}_{0}, \Sigma^{o**}_{r}, \Sigma^{**}_{r}$, the hyperprior distributions are:

$$\begin{aligned} \beta^{*}\,|\,\beta^{**}, \Sigma^{**}_{0} &\sim MVN\left(\beta^{**}, \Sigma^{**}_{0}\right) & \left(\Sigma_0\right)^{-1}\,|\,\Sigma^{*}_{0} &\sim Wi\left(\left(5\Sigma^{*}_{0}\right)^{-1}, 5\right) \\ b^{*}_{rw}\,|\,b^{**}_{rw}, \Sigma^{o**}_{r} &\sim MVN\left(b^{**}_{rw}, \Sigma^{o**}_{r}\right) & \left(\Sigma^{o}_{r}\right)^{-1}\,|\,\Sigma^{*o}_{r} &\sim Wi\left(\left(3\Sigma^{*o}_{r}\right)^{-1}, 3\right) \\ b^{*}_{r}\,|\,b^{**}_{r}, \Sigma^{**}_{r} &\sim MVN\left(b^{**}_{r}, \Sigma^{**}_{r}\right) & \left(\Sigma_r\right)^{-1}\,|\,\Sigma^{*}_{r} &\sim Wi\left(\left(2\Sigma^{*}_{r}\right)^{-1}, 2\right) \end{aligned}$$

and $\frac{1}{\sigma^2} \sim Ga(0.001, 0.001)$. The claims process functions $C_{r,w}$ and $P_{r,w,t}$ for model 4.1 and the related survival and hazard functions $S_{r,w,t}$ and $h_{r,w,t}$ are:

$$C_{r,w} = \exp\left(L + l^{*}_{rw}\right)$$

$$P_{r,w,t} = \left\{ 1 - \frac{\left(1 + 2\left(\dfrac{\exp\left(D + d_r + d^{*}_{rw}\right)}{\exp\left(Kc + kc_r + kc^{*}_{rw}\right)}\right)\ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)}{\dfrac{\exp\left(\left(\exp\left(Kc + kc_r + kc^{*}_{rw}\right) + 1\right)\ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)}{\exp\left(-\exp\left(D + d_r + d^{*}_{rw}\right)\left(\ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)^2\right)}} \right\}$$

$$S_{r,w,t} = 1 - P_{r,w,t}$$

$$h_{r,w,t} = \frac{-\left(1 - \dfrac{\left(2\exp\left(D + d_r + d^{*}_{rw}\right)\ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right) + \left(\exp\left(Kc + kc_r + kc^{*}_{rw}\right) + 1\right)\right)}{\left(\dfrac{\exp\left(Kc + kc_r + kc^{*}_{rw}\right)}{2\exp\left(D + d_r + d^{*}_{rw}\right)} + \ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)^{-1}}\right)}{\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\left(\dfrac{\exp\left(Kc + kc_r + kc^{*}_{rw}\right)}{2\exp\left(D + d_r + d^{*}_{rw}\right)} + \ln\left(t + \dfrac{\exp\left(Ks_1\right)}{t^{\exp(Ks_2)}}\right)\right)\left(1 - \dfrac{\exp\left(Ks_1 + Ks_1\right)}{t^{\exp(Ks_2)+1}}\right)^{-1}} \quad (4.3)$$

Amended Versions Of Model 7.0(a)

In the alternative versions of model 7.0(a), denoted by 7.0(b) and 7.0(c), $\varepsilon_{r,w} \mid \phi_{rw}, \sigma_r \sim N\left(0, \sigma_r^2 I_{n_{rw} \times n_{rw}}\right)$ and $\varepsilon_{r,w} \mid \phi_{rw}, \sigma_{rw} \sim N\left(0, \sigma_{rw}^2 I_{n_{rw} \times n_{rw}}\right)$ replace $\varepsilon_{r,w} \mid \phi_{rw}, \sigma \sim N\left(0, \sigma^2 I_{n_{rw} \times n_{rw}}\right)$.

4.1.2 Models 8.0 (a), (b) and (c)

In model 8.0 the percentage cash flow function has more parameters than model 7.0 to assess if a more flexible percentage cash flow function could produce more reliable *IBNR* predictions. As with model 7.0, three versions are considered.

Model 8.0(a)

For claims array r, underwriting year w and development time t the model is given by:

$$Y_{r,w,t} = \mu_{r,w,t}\left(\phi_{rw}\right) + \varepsilon_{r,w,t}$$

such that,

$$\mu_{r,w,t}(\phi_{rw}) = \exp\left(L + l^{*}_{rw}\right)\left(1 + \frac{\exp\left(D + d_r + d^{*}_{rw}\right)}{\exp\left(\frac{\ln t}{\exp\left(-Kc - kc_r - kc^{*}_{rw}\right)} + \frac{t}{\exp\left(-Kd - kd_r - kd^{*}_{rw}\right)}\right)}\right)^{-1} \tag{4.4}$$

and

$$\phi_{rw} = A_{rw}\beta + B_{r,w}b_r + B_{rw}b_{rw} = d\left(a_{rw}, \beta, b_r, b_{rw}\right)$$

$$\varepsilon_{r,w}\left|\phi_{rw}, \sigma \sim N\left(0, \sigma^2 I_{n_{rw} \times n_{rw}}\right)\right.$$

$$\begin{aligned} b_{rw} &= \left[l^{*}_{rw}, d^{*}_{rw}, kc^{*}_{rw}, kd^{*}_{rw}\right]^T, & b_{rw}\left|b^{*}_{rw}, \Sigma_{rw} \sim MVN\left(b^{*}_{rw}, \Sigma_{rw}\right)\right. \\ b_r &= \left[d_r, kc_r, kd_r\right]^T, & b_r\left|b^{*}_r, \Sigma_r \sim MVN\left(b^{*}_r, \Sigma_r\right)\right. \\ \beta &= \left[L, D, Kc, Kd\right]^T, & \beta\left|\beta^{*}, \Sigma_0 \sim MVN\left(\beta^{*}, \Sigma_0\right)\right. \end{aligned} \tag{4.5}$$

The hyperprior distributions are:

$$\begin{aligned} &\beta^{*}\left|\beta^{**}, \Sigma^{**}_0 \sim MVN\left(\beta^{**}, \Sigma^{**}_0\right)\right. && \left(\Sigma_0\right)^{-1}\left|\Sigma^{*}_0 \sim Wi\left(\left(4\Sigma^{*}_0\right)^{-1}, 4\right)\right. \\ &b^{*}_{rw}\left|b^{**}_{rw}, \Sigma^{**}_{rw} \sim MVN\left(b^{**}_{rw}, \Sigma^{**}_{rw}\right)\right. && \left(\Sigma_{rw}\right)^{-1}\left|\Sigma^{*}_{rw} \sim Wi\left(\left(4\Sigma^{*}_{rw}\right)^{-1}, 4\right)\right. \\ &b^{*}_r\left|b^{**}_r, \Sigma^{**}_r \sim MVN\left(b^{**}_r, \Sigma^{**}_r\right)\right. && \left(\Sigma_r\right)^{-1}\left|\Sigma^{*}_r \sim Wi\left(\left(3\Sigma^{*}_r\right)^{-1}, 3\right)\right. \end{aligned}$$

and $\frac{1}{\sigma^2} \sim Ga(0.001, 0.001)$, such that $\beta^{**}, b^{**}_{rw}, b^{**}_r, \Sigma^{*}_0, \Sigma^{*o}_r, \Sigma^{*}_r, \Sigma^{**}_0, \Sigma^{o**}_r, \Sigma^{**}_r$ are known. Functions $C_{r,w}$ and $P_{r,w,t}$ for model 8.0(a) and the related survival and hazard functions are:

$$C_{r,w} = \exp\left(L + l^{*}_{rw}\right)$$

$$P_{r,w,t} = \left\{1 + \frac{\exp\left(D + d_r + d^{*}_{rw}\right)}{\exp\left(\exp\left(Kc + kc_r + kc^{*}_{rw}\right)\ln t + \exp\left(Kd + kd_r + kd^{*}_{rw}\right)t\right)}\right\}^{-1}$$

$$S_{r,w,t} = 1 - \left\{1 + \frac{\exp\left(D + d_r + d^{*}_{rw}\right)}{\exp\left(\exp\left(Kc + kc_r + kc^{*}_{rw}\right)\ln t + \exp\left(Kd + kd_r + kd^{*}_{rw}\right)t\right)}\right\}^{-1}$$

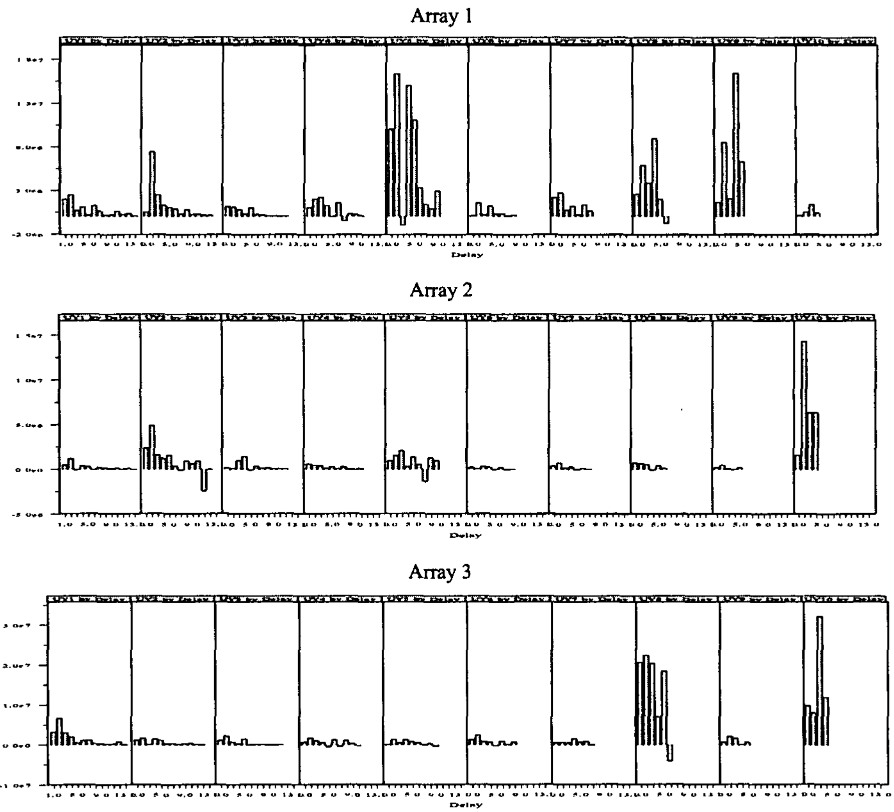
$$h_{r,w,t} = \left(\frac{\exp\left(Kc + kc_r + kc^{*}_{rw}\right)}{t} + \exp\left(Kd + kd_r + kd^{*}_{rw}\right)\right)P_{r,w,t} \tag{4.6}$$

Amended Versions Of Model 8.0(a)

Model versions 8.0(b) and 8.0(c) are derived from 8.0(a) as 7.0(b) and 7.0(c).

4.2 Numerical Examples And Discussion For Preliminary Models 7.0 And 8.0

The claims data selected to illustrate the models in section 4 are reported in tables B.1 to B.3. The data have been obtained through simulations based on a marine portfolio consisting of hull, cargo and aviation hull claims, labelled in graphs and tables as arrays 1, 2 and 3 respectively. Evident from graph 4.2.1 are the data variability and a large number of negative entries in the incremental claims data. Claims reserving models for multiple-array claims portfolios have to explain the variability emerging from the different array characteristics, settlement speeds and exposure levels. The broad range of the cumulative claim totals, from 1,013,800 to 85,287,218, suggest that such claim volume variability may not be effectively captured by the random effects parameters for the mean response model alone.

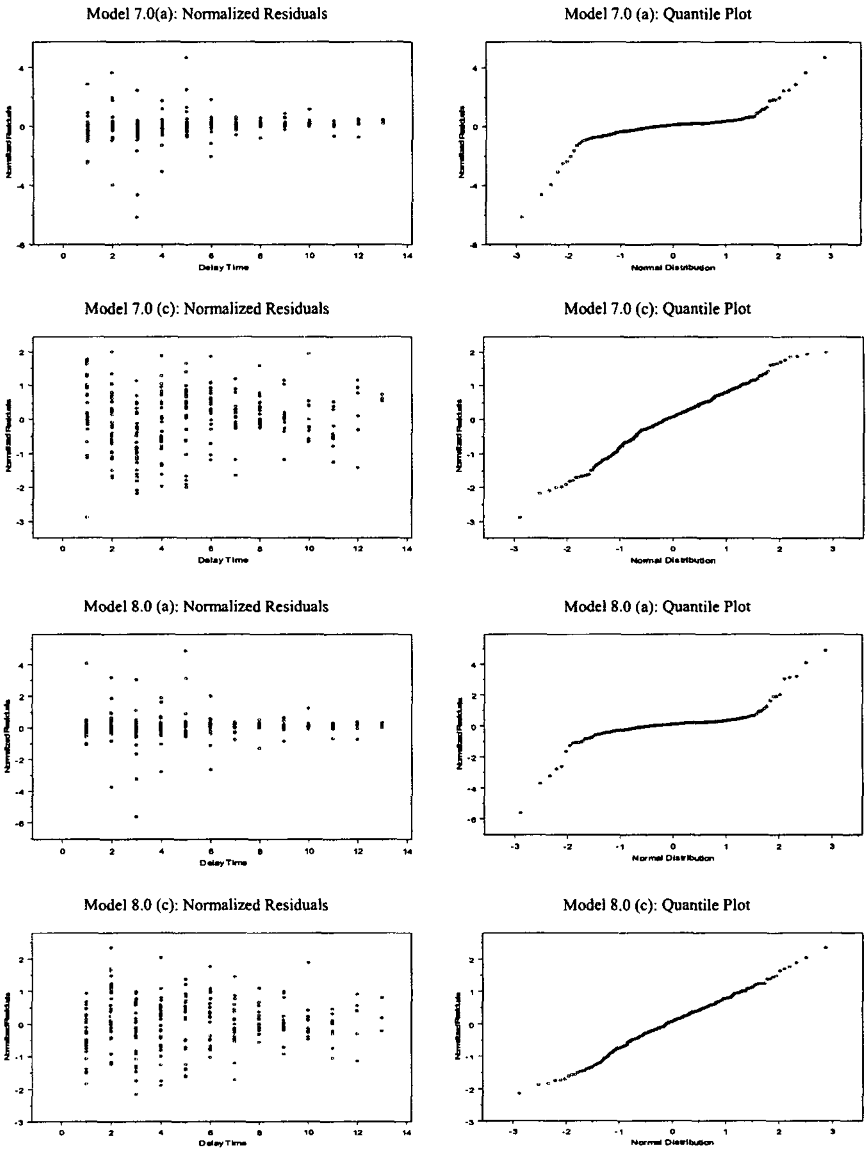


Graph 4.2.1 Incremental data bar plots by array and underwriting year for tables B.1, B.2 and B.3.

	Under. Year	Model 7.0 (c) Ultimate Claim Amount Mean	Standard Mean Sq. Predict Error	Predictive Interval 2.50%	Predictive Interval 97.50%	*IBNR* (1.8) Mean	Standard Mean Sq. Predict Error	Predictive Interval 2.50%	Predictive Interval 97.50%	σ^2_{rw}
Array 1	1	11,460,000	560,300	10,250,000	12,530,000	2,130,000	367,600	1,338,000	2,827,000	2.63E+11
	2	18,660,000	880,100	16,790,000	20,390,000	3,103,000	592,200	1,836,000	4,246,000	5.31E+11
	3	5,703,000	363,700	4,959,000	6,413,000	1,006,000	249,000	489,400	1,498,000	8.09E+10
	4	10,520,000	699,000	8,984,000	11,830,000	2,400,000	493,100	1,318,000	3,321,000	2.42E+11
	5	79,440,000	9,587,000	59,530,000	96,770,000	21,190,000	6,833,000	7,172,000	33,700,000	3.25E+13
	6	5,361,000	646,100	3,836,000	6,453,000	1,707,000	479,000	531,200	2,518,000	1.02E+11
	7	10,340,000	1,073,000	8,789,000	12,900,000	2,199,000	788,100	1,256,000	4,081,000	4.53E+11
	8	37,200,000	8,030,000	22,320,000	51,550,000	14,460,000	6,217,000	3,977,000	25,830,000	1.13E+13
	9	57,000,000	22,430,000	20,850,000	100,700,000	28,180,000	17,890,000	4,647,000	63,910,000	7.32E+13
	10	4,479,000	1,919,000	1,427,000	8,180,000	2,451,000	1,601,000	413,800	5,634,000	4.28E+11
Array 2	1	3,259,000	160,000	2,928,000	3,555,000	460,700	105,900	236,100	657,100	2.01E+10
	2	14,920,000	718,600	13,810,000	16,590,000	1,661,000	455,000	1,153,000	2,778,000	7.57E+11
	3	5,504,000	456,900	4,520,000	6,358,000	1,787,000	311,800	1,108,000	2,369,000	1.24E+11
	4	2,888,000	213,500	2,450,000	3,298,000	558,300	148,800	259,500	846,300	2.23E+10
	5	9,255,000	1,209,000	7,091,000	11,620,000	2,261,000	864,500	819,200	3,993,000	5.82E+11
	6	1,530,000	232,800	1,112,000	1,964,000	422,500	171,100	143,200	747,500	1.59E+10
	7	1,896,000	66,610	1,800,000	2,063,000	293,300	44,510	251,100	413,400	4.42E+09
	8	2,892,000	237,600	2,557,000	3,517,000	575,600	168,300	419,200	1,064,000	3.98E+10
	9	1,264,000	133,600	1,083,000	1,612,000	299,500	99,010	213,300	579,900	8.60E+09
	10	50,860,000	12,980,000	28,600,000	72,890,000	23,640,000	10,970,000	7,198,000	42,520,000	1.51E+13
Array 3	1	22,570,000	379,300	21,730,000	23,270,000	3,176,000	271,900	2,555,000	3,660,000	8.60E+10
	2	10,270,000	529,900	9,141,000	11,250,000	2,549,000	351,300	1,813,000	3,210,000	2.12E+11
	3	7,549,000	421,000	6,755,000	8,380,000	1,181,000	288,200	646,200	1,759,000	1.15E+11
	4	7,465,000	748,000	5,903,000	8,828,000	1,741,000	523,900	647,200	2,698,000	2.56E+11
	5	8,308,000	570,900	7,056,000	9,387,000	2,950,000	408,700	2,060,000	3,712,000	1.22E+11
	6	8,368,000	648,500	7,269,000	9,715,000	1,642,000	477,200	917,700	2,646,000	1.67E+11
	7	8,719,000	1,810,000	4,564,000	11,610,000	3,735,000	1,376,000	683,200	5,940,000	5.23E+11
	8	115,700,000	15,020,000	92,600,000	148,100,000	30,480,000	11,610,000	15,520,000	56,200,000	5.84E+13
	9	7,997,000	1,149,000	6,085,000	10,260,000	2,621,000	925,900	1,228,000	4,485,000	2.07E+11
	10	94,890,000	42,140,000	22,790,000	188,100,000	43,200,000	33,030,000	7,345,000	121,300,000	3.62E+14
By Array	Array 1	240,200,000	26,480,000	191,200,000	291,100,000	78,830,000	20,810,000	42,460,000	119,300,000	
	Array 2	94,270,000	13,240,000	71,450,000	117,200,000	31,960,000	11,160,000	14,750,000	51,400,000	
	Array 3	291,800,000	45,850,000	215,600,000	389,100,000	93,280,000	35,830,000	48,710,000	175,000,000	
Total		626,200,000	54,430,000	530,900,000	740,400,000	204,100,000	43,170,000	134,800,000	298,700,000	
Deviance		7,372								
Iterat.: Start +Sample		31,000								

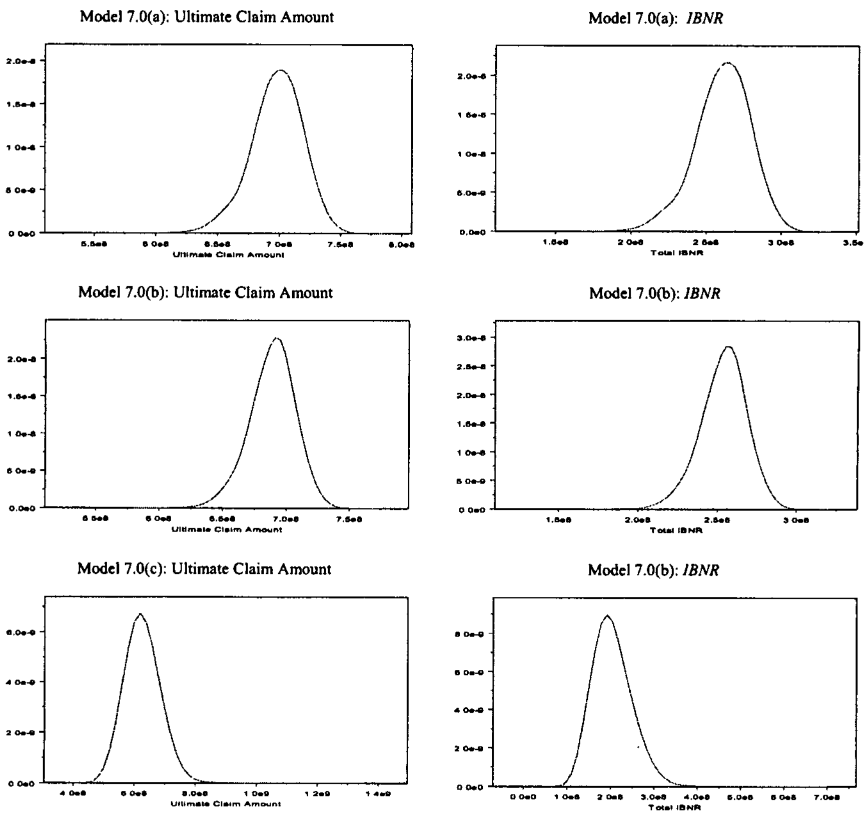
Table 4.2.1 Model statistics, ultimate losses and *IBNR* predictions, and respective predictive distributions for Model 7.0 (c).

Portfolios displaying large differences in exposure levels or claims magnitudes are not at all unusual, even in treaties where underwriting contracts remain unaltered. Cost limitations or timing restrictions may impede exploring methods, possibly able to deal with high variability in exposure volumes, such as analyses at transaction level. In the models proposed, a good fit to historical data as assessment criterion of the preliminary models, is as important as suitable variance models, as the latter determines the stability of *IBNR* and ultimate claim predictions. This is more likely to be achieved by models 7.0(c) and 8.0(c), as inspection of graph 4.2.2 and of actual and fitted claims confirm.

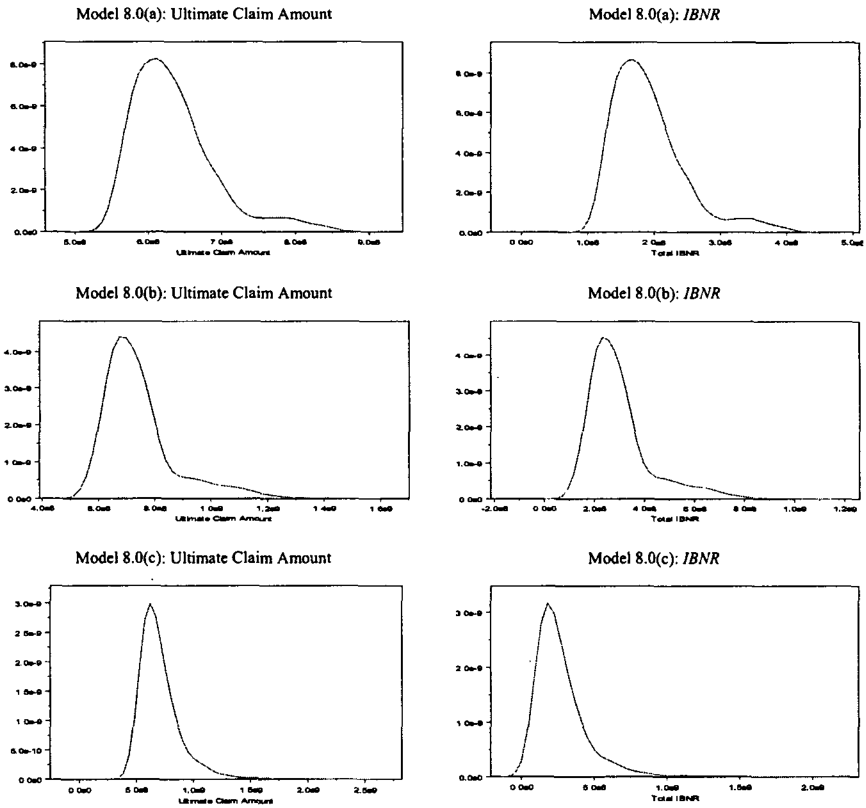


Graph 4.2.2 Normalized residuals and quantile plots for models 7.0 and 8.0 (a) and (c).

Plots for 7.0(b) and 8.0(b) were found to be uninformative, and for this reason were excluded from graph 4.2.2. While the rankings of σ_r^2 in models 7.0(b) and 8.0(b) are consistent, with $\sigma_2^2 < \sigma_3^2 < \sigma_1^2$, the claims volume variability within each array present similar problems to those encountered with model 1.0. According to the quantile plots only the residuals from models with variance function σ_{rw} may satisfy the Shapiro-Wilk test W for near-normality (Shapiro and Wilk (1965)). Model 7.0(c) gives narrower intervals for the mean *IBNR* at underwriting year, array levels and overall. (see table 4.5.3). The close equivalence of ranking orders for σ_{rw}^2 and $c_{r,w}$ (table 4.2.1) confirms the expectation that either $\zeta_{rw} = (\sigma_{rw}, \vartheta, \rho)$ or $\zeta_{rw} = (\sigma, \vartheta_{rw}, \rho)$ could reveal scale variability structures in the data. They do so more effectively than $\zeta_{rw} = (\sigma_{rw}, \vartheta_{rw}, \rho)$. In a variance model $\sigma_r^2 (\mu_{r,w,t}(\phi_{rw}))^{\exp(\vartheta + \vartheta_r)}$, parameters σ_r and ϑ_r are less informative.



Graph 4.2.3 Preliminary model 7.0: Kernel densities for ultimate losses and *IBNR* totals.



Graph 4.2.4 Preliminary model 8.0: Kernel densities for ultimate losses and *IBNR* totals.

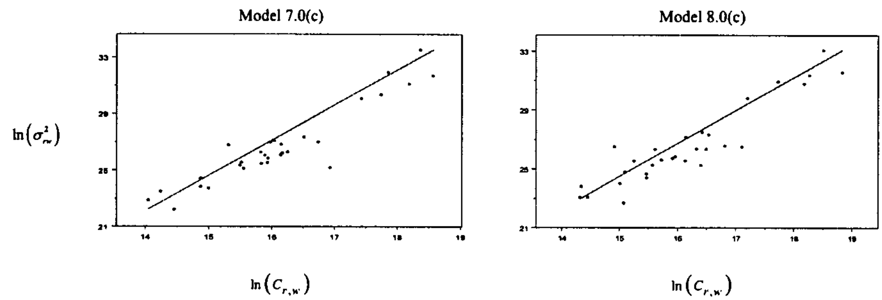
Graphs 4.2.3 and 4.2.4 and table 4.5.3 show that the kernels for mean *IBNR* and ultimate claim predictions are skewed. In the next section it is shown how σ^2_{rw} and $C_{r,w}$ can be used to construct the variance function for the final models.

4.3 Final Multi-Array Models

The preliminary models demonstrate that the data variability can be explored more freely when $\operatorname{var}(Y_{r,w}) = \sigma^2_{rw}$. The values of σ^2_{rw} and $C_{r,w}$ suggest a variability structure associated to scale differences between underwriting year data sets, around which a cluster structure could be constructed for analytical purposes. However, some management decisions, such as commutations, would require more precise *IBNR* and

$C_{r,w}$ predictions at underwriting year or contract levels. Reconciliation of reserves would be more difficult if the data of interest were not part of the same cluster. A better approach to deal with scale variability, and one that is totally coherent with the generic model, may involve formulating $C_{r,w}$ into the variance model. To assess this, the following regression is applied to the output of models 7.0(c) and 8.0(c):

$$\ln\left(\sigma_{r,w}^2\right) = \delta_1 + \delta_2 \ln\left(C_{r,w}\right) \tag{4.7}$$



Graph 4.3.1 Lines $\delta_1 + \delta_2 \ln\left(C_{r,w}\right)$ and scatter plots of $\ln\left(C_{r,w}\right)$ vs $\ln\left(\sigma_{rw}^2\right)$ on the y-axis.

(4.7) gives $[\delta_1, \delta_2] = [-8.7068, 2.1934]$ for model 7.0(c) and $[\delta_1, \delta_2] = [-6.355, 2.0347]$ for model 8.0(c). Graph 4.3.1 displays the regression lines and the scatter plots of $\ln\left(C_{r,w}\right)$ versus $\ln\left(\sigma_{rw}^2\right)$ for both models. Equation (4.7) suggests that the final models should be

$$\left(Y_{r,w,t} \middle| C_{r,w}, \zeta\right) \sim N\left(\mu_{r,w,t}\left(\phi_{rw}\right), \sigma^2 C_{r,w}^2\right) \tag{4.8}$$

such that $\zeta = \sigma^2$ and $\mu_{r,w,t}\left(\phi_{rw}\right)$ is given by equations (4.1) and (4.4) for models 7.0(d) and 8.0(d) respectively. From regression model (4.7) for model 7.0(c), $\exp(\delta_1) = 0.000165$ could set the initial value for σ^2. The outcome of the analysis is not unexpected. In fact, the inclusion of $C_{r,w}^2$ in the variance function has the effect of normalising the data, hence, reducing the reserving analysis with random effects models to a type of problem that is more consistent with the typical published examples, concerned with the analysis of repeated observations on subjects or trials that share some common characteristics.

See for instance Elashoff et al. (1982) and Aziz et al. (1978). For the final models the mean *IBNR* and ultimate claim amount predictions are replaced by estimates generated by their predictive distributions. The reported *IBNR* values are calculated along the lines of (1.9). Extending the definition of section 3.2, the portfolio average model for the percentage cash flow is given below.

4.4 Portfolio And Array Average Models For The Percentage Cash Flow

Section 3.2 identifies the percentage cash flow as the most suitable function in the reserving model where concepts on inferences on marginal distribution or population average models could be applied. Comparisons across a claims portfolio are more meaningful at percentage cash flow level. As observed in section 3, a model may be able to fit the data well even when the percentage cash flow function converges to a value different to 1. However, in such cases the ultimate claim amount and *IBNR* predictions would be incorrect.

To formulate the average models for the percentage cash flow the parameter vectors for the portfolio and array average models, ϕ_p and ϕ_A, are respectively defined:

$$\phi_p = A\beta + B_1\left(\frac{1}{r_1}\sum_{r=1}^{r_1} b_r\right) + B_2\left(\left(\sum_{r=1}^{r_1} u_r\right)^{-1}\sum_{r=1}^{r_1}\sum_{w=1}^{u_r} b_{rw}\right)$$

and

$$\phi_{A_r} = A\beta + B_1 b_r + B_2\left((u_r)^{-1}\sum_{w=1}^{u_r} b_{rw}\right)$$

such that, design matrices A_{rw}, $B_{r.w}$ and B_{rw} are replaced respectively by A, B_1 and B_2. Consider for example the mean response function for model 7.0(d). If

$$\begin{aligned} D_p &= D + \frac{1}{r_1}\sum_{r=1}^{r_1} d_r + \left(\sum_{r=1}^{r_1} u_r\right)^{-1}\sum_{r=1}^{r_1}\sum_{w=1}^{u_r} d^{*}_{rw} \\ Kc_p &= Kc + \frac{1}{r_1}\sum_{r=1}^{r_1} kc_r + \left(\sum_{r=1}^{r_1} u_r\right)^{-1}\sum_{r=1}^{r_1}\sum_{w=1}^{u_r} kc^{*}_{rw} \end{aligned} \tag{4.9}$$

such that $D, Kc, d_r, kc_r, d^{*}_{rw}$ and kc^{*}_{rw} are the percentage cash flow function parameters, then the portfolio average for the percentage cash flow function at time t is given by

$$P_{p,t} = 1 - \frac{\left(1 + 2\left(\frac{\exp(D_p)}{\exp(Kc_p)}\right)\ln\left(t + \frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)}{\exp\left(\left(\exp(Kc_p)+1\right)\ln\left(t + \frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)} \exp\left(-\exp(D_p)\left(\ln\left(t + \frac{\exp(Ks_1)}{t^{\exp(Ks_2)}}\right)\right)^2\right) \quad (4.10)$$

Additional insight may be gained by including in the plots a curve for the percentage cash flow average model for each array. Continuing with the example, for the array average model for array r in model 7.0(d)

$$D_{A_r} = D + d_r + \frac{1}{u_r}\sum_{w=1}^{u_r} d^{\bullet}_{rw}$$

$$Kc_{A_r} = Kc + kc_r + \frac{1}{u_r}\sum_{w=1}^{u_r} kc^{\bullet}_{rw}$$

should replace D_p Kc_p in equation (4.10).

4.5 Numerical Examples And Discussion For Models 7.0(d) And 8.0(d)

Models 7.0(d) and 8.0(d) provide close fit to the data. The portfolio reported *IBNR* and ultimate claim predictions for 7.0(d) and 8.0(d) are given on tables 4.5.1 and 4.5.2 and summarised on table 4.5.3. They show that the final models' predictive intervals are narrower than for their earlier model versions. At underwriting year level, the mean response function of model 7.0 is still the most useful of the two (see table 4.5.1).

Graph 4.5.2 compares scatter plots for the percentage cash flow values for both models and shows that model 8.0(d) is the least successful in separating the volume and development pattern elements in the data. Note that the graphs' scales are not the same and that the projection period for model 7.0(d) is longer than for model 8.0(d). The predictive interval for $\sum_{r=1}^{3}\sum_{w=1}^{10} C_{r,w}$ for model 8.0(d) is also wider. (See tables 4.5.2 and 4.5.3 and graph 4.5.1). Evident from graph 4.5.2 is the settlement speeds variability. A reduction in the reported *IBNR* predictive intervals is consistent with a reduction of the normalized residuals and the Bayesian Information Criterion. Particularly relevant to the claims process is the systematic correction of historical errors as claims evolve, since negative incremental entries frequently adjust earlier overstated claim entries. The box, scatter and quantile plots make apparent data anomalies generated by negative adjustments to paid claims and by large claim volume differences. As in section 3, neither can be addressed with autoregressive error structures. The negative incremental claim entries are responsible for the some of the outliers and, in particular, for the slight depression in the quantile plots, between −2 and −1 of the horizontal axis.

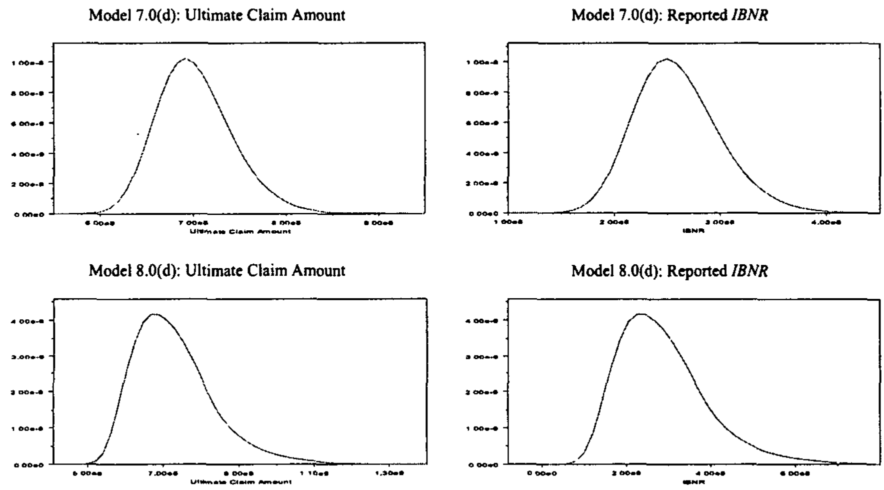
	Under. Year	Model 7.0(d)							
		Ultimate Claim Amount				Reported *IBNR* (1.9)			
		Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.50%	97.50%	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.50%	97.50%
Array 1	1	11,210,000	769,200	9,777,000	12,780,000	1,529,000	769,200	95,310	3,099,000
	2	17,710,000	1,249,000	15,420,000	20,320,000	2,353,000	1,249,000	56,370	4,962,000
	3	5,444,000	413,800	4,704,000	6,318,000	872,000	413,800	131,300	1,746,000
	4	10,080,000	837,600	8,526,000	11,800,000	2,233,000	837,600	681,600	3,954,000
	5	82,510,000	7,098,000	69,530,000	97,320,000	22,950,000	7,098,000	9,971,000	37,750,000
	6	5,419,000	523,300	4,480,000	6,556,000	1,814,000	523,300	874,500	2,951,000
	7	9,677,000	708,000	8,497,000	11,350,000	864,700	708,000	-315,500	2,534,000
	8	41,090,000	5,069,000	32,640,000	52,560,000	18,800,000	5,069,000	10,350,000	30,270,000
	9	77,960,000	12,940,000	57,800,000	108,600,000	43,250,000	12,940,000	23,090,000	73,890,000
	10	5,935,000	1,329,000	3,974,000	9,123,000	3,690,000	1,329,000	1,730,000	6,879,000
Array 2	1	2,954,000	159,900	2,655,000	3,284,000	82,640	159,900	-217,200	412,100
	2	14,730,000	867,600	13,140,000	16,560,000	2,603,000	867,600	1,004,000	4,431,000
	3	5,511,000	416,500	4,745,000	6,370,000	1,968,000	416,500	1,202,000	2,828,000
	4	2,795,000	234,300	2,377,000	3,289,000	388,300	234,300	-30,250	881,400
	5	9,708,000	836,200	8,198,000	11,490,000	1,928,000	836,200	418,600	3,708,000
	6	1,610,000	165,700	1,319,000	1,967,000	510,000	165,700	218,700	867,300
	7	1,903,000	110,600	1,695,000	2,130,000	315,000	110,600	106,900	541,700
	8	2,878,000	206,000	2,527,000	3,345,000	489,700	206,000	139,200	956,900
	9	1,256,000	130,000	1,089,000	1,527,000	242,700	130,000	74,990	513,500
	10	55,000,000	12,350,000	37,090,000	86,390,000	26,260,000	12,350,000	8,350,000	57,640,000
Array 3	1	21,060,000	1,339,000	18,570,000	23,870,000	1,494,000	1,339,000	-1,002,000	4,302,000
	2	9,372,000	709,500	8,070,000	10,840,000	1,601,000	709,500	298,900	3,064,000
	3	7,253,000	509,800	6,333,000	8,330,000	1,012,000	509,800	92,520	2,089,000
	4	7,430,000	611,500	6,300,000	8,685,000	1,447,000	611,500	316,900	2,702,000
	5	7,978,000	683,900	6,742,000	9,428,000	2,968,000	683,900	1,731,000	4,417,000
	6	7,794,000	585,200	6,785,000	9,090,000	509,400	585,200	-499,700	1,806,000
	7	9,066,000	992,900	7,317,000	11,150,000	3,852,000	992,900	2,103,000	5,936,000
	8	115,400,000	14,440,000	94,040,000	148,400,000	30,070,000	14,440,000	8,752,000	63,110,000
	9	7,423,000	1,083,000	6,060,000	10,240,000	1,912,000	1,083,000	549,400	4,731,000
	10	144,200,000	30,660,000	100,000,000	217,200,000	82,050,000	30,660,000	37,860,000	155,100,000
By Array	Array 1	267,000,000	15,930,000	239,600,000	302,600,000	98,350,000	15,930,000	70,950,000	133,900,000
	Array 2	98,350,000	12,500,000	80,030,000	130,000,000	34,790,000	12,500,000	16,460,000	66,480,000
	Array 3	336,900,000	34,040,000	284,200,000	416,100,000	126,900,000	34,040,000	74,190,000	206,100,000
Total		702,300,000	39,690,000	636,000,000	790,200,000	260,100,000	39,690,000	193,700,000	347,900,000
σ^2		0.002308							
Deviance		7,365							
Iterat.: Start +Sample		29,500							

Table 4.5.1 Model 7.0(d): statistics, ultimate loss and reported *IBNR* predictions, and predictive intervals.

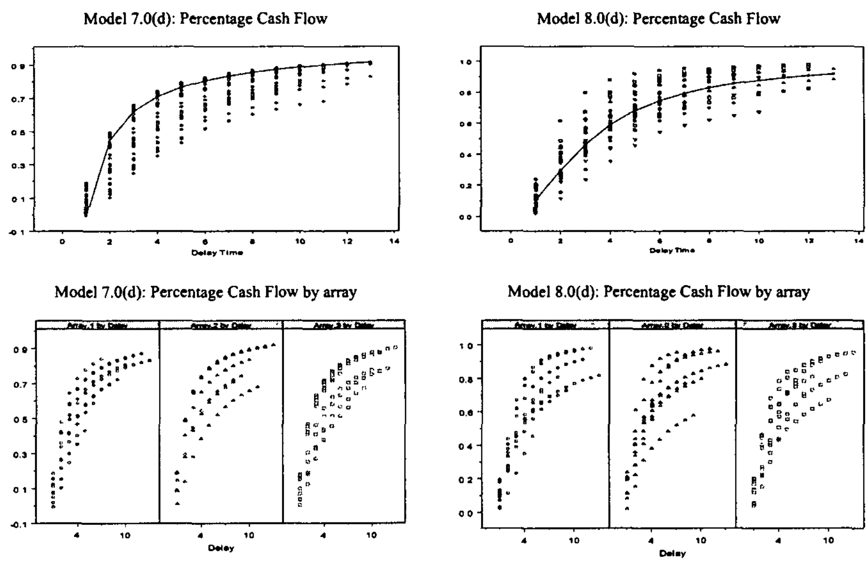
	Under. Year	Model 8.0(d)							
		Ultimate Claim Amount				Reported *IBNR* (1.9)			
		Mean	Standard Mean Sq. Predict. Error	Predictive Interval		Mean	Standard Mean Sq. Predict. Error	Predictive Interval	
				2.50%	97.50%			2.50%	97.50%
Array 1	1	11,650,000	1,272,000	9,664,000	14,610,000	1,964,000	1,272,000	-17,750	4,926,000
	2	15,330,000	925,100	13,650,000	17,290,000	-33,230	925,100	-1,713,000	1,931,000
	3	5,257,000	529,300	4,491,000	6,648,000	684,300	529,300	-81,330	2,075,000
	4	8,169,000	518,100	7,230,000	9,284,000	324,400	518,100	-614,400	1,439,000
	5	88,310,000	24,820,000	62,930,000	159,600,000	28,750,000	24,820,000	3,371,000	100,100,000
	6	5,174,000	2,195,000	3,529,000	11,800,000	1,569,000	2,195,000	-76,040	8,196,000
	7	12,640,000	4,024,000	8,691,000	24,190,000	3,827,000	4,024,000	-121,500	15,380,000
	8	43,080,000	20,980,000	24,940,000	94,960,000	20,790,000	20,980,000	2,646,000	72,670,000
	9	100,800,000	56,840,000	49,440,000	261,700,000	66,120,000	56,840,000	14,730,000	226,900,000
	10	3,732,000	2,999,000	2,169,000	11,610,000	1,488,000	2,999,000	-75,410	9,362,000
Array 2	1	3,378,000	642,300	2,746,000	5,221,000	506,300	642,300	-125,300	2,349,000
	2	13,820,000	1,035,000	12,190,000	16,230,000	1,690,000	1,035,000	58,680	4,094,000
	3	3,567,000	214,600	3,174,000	4,022,000	24,380	214,600	-368,000	479,200
	4	3,102,000	537,100	2,413,000	4,534,000	694,700	537,100	5,700	2,126,000
	5	13,150,000	5,018,000	7,527,000	26,000,000	5,372,000	5,018,000	-252,300	18,220,000
	6	1,679,000	615,700	1,125,000	3,443,000	579,000	615,700	24,490	2,343,000
	7	1,726,000	386,700	1,429,000	2,971,000	137,900	386,700	-159,100	1,383,000
	8	3,693,000	1,482,000	2,337,000	7,675,000	1,305,000	1,482,000	-51,290	5,287,000
	9	1,718,000	1,104,000	880,500	4,346,000	703,800	1,104,000	-133,300	3,332,000
	10	45,620,000	29,960,000	26,490,000	132,500,000	16,880,000	29,960,000	-2,253,000	103,800,000
Array 3	1	19,970,000	1,316,000	17,640,000	22,840,000	407,000	1,316,000	-1,924,000	3,271,000
	2	9,249,000	996,600	7,773,000	11,710,000	1,477,000	996,600	1,074	3,937,000
	3	7,050,000	1,353,000	5,927,000	11,690,000	809,000	1,353,000	-314,100	5,446,000
	4	9,812,000	2,500,000	6,652,000	16,390,000	3,829,000	2,500,000	668,500	10,410,000
	5	5,939,000	756,800	5,010,000	7,855,000	928,500	756,800	-607	2,844,000
	6	9,454,000	2,978,000	6,651,000	17,840,000	2,169,000	2,978,000	-633,000	10,560,000
	7	8,073,000	2,031,000	5,703,000	13,670,000	2,859,000	2,031,000	489,100	8,458,000
	8	125,800,000	37,370,000	89,280,000	229,700,000	40,550,000	37,370,000	3,992,000	144,400,000
	9	6,229,000	1,520,000	4,957,000	10,410,000	717,800	1,520,000	-554,400	4,897,000
	10	141,100,000	60,410,000	75,900,000	293,600,000	78,980,000	60,410,000	13,760,000	231,500,000
By Array	Array 1	294,200,000	66,190,000	217,800,000	475,200,000	125,500,000	66,190,000	49,160,000	306,500,000
	Array 2	91,450,000	30,130,000	68,590,000	177,000,000	27,890,000	30,130,000	5,030,000	113,400,000
	Array 3	342,700,000	72,200,000	252,700,000	525,000,000	132,700,000	72,200,000	42,750,000	315,000,000
Total		728,300,000	102,600,000	584,700,000	984,300,000	286,100,000	102,600,000	142,500,000	542,000,000
σ^2		0.002137							
Deviance		7,323							
Iterat.: Start +Sample		29,500							

Table 4.5.2 Model 8.0(d): statistics, ultimate loss and reported *IBNR* predictions, and predictive intervals.

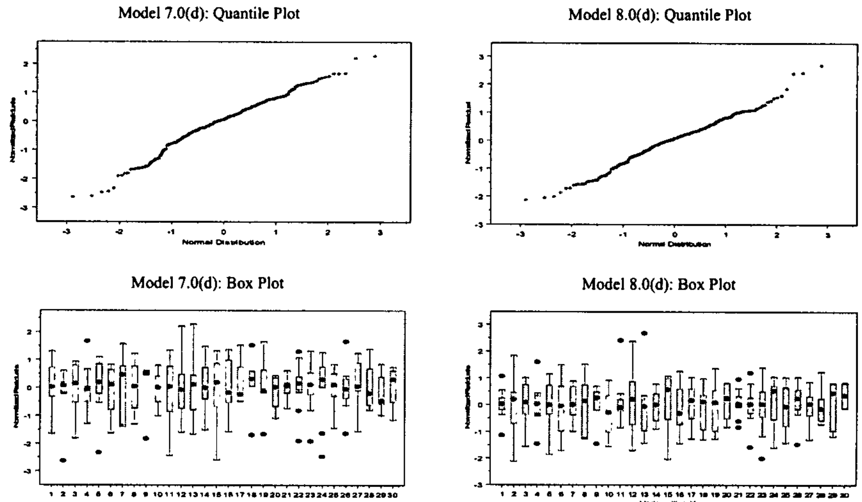
Historical claims add to 442,249,345. The difference between the ultimate claim amounts and the reported *IBNR* predictions for the final models are approximately 442 million. The order of accuracy in the WinBugs system prevents an exact reconciliation with the total claim amount to date. When model 7.0(d) is appraised for consistency with an analysis by array, the ultimate claim amount and reported *IBNR* predictions show respectively 1.1% and 3.1% overall difference from the predictions on table 4.5.1. In section 4.6 the hazard rate profile extracted from the model is discussed.



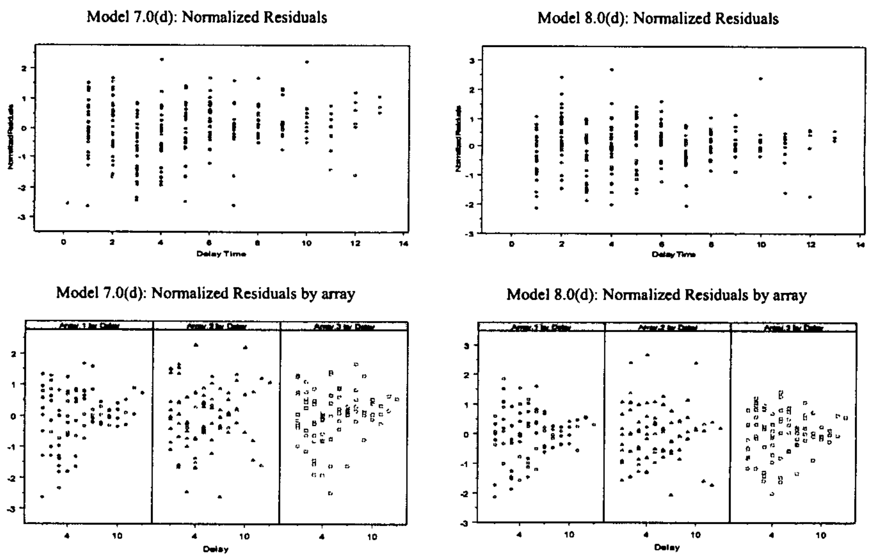
Graph 4.5.1 Kernel densities and predictive distributions for ultimate losses and reported *IBNR*.



Graph 4.5.2 Models 7.0(d) and 8.0(d): Scatter plots and average portfolio curve for percentage cash flow versus delay time.



Graph 4.5.3 Quantile plots and box plots by underwriting year. Underwriting years are labelled 1 to 30. The first 10 correspond to the marine hull, the next 10 to marine cargo and the last 10 to aviation cargo.



Graph 4.5.4 Models 7.0(d) and 8.0(d): Scatter plots versus delay time, overall and by array.

Model	Preliminary Models																	Devi.
	Distributions								Log Likelihood			AIC			BIC			
	Ultimate Claim Amount				*IBNR* (1.8)													
	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.5%	Predictive Interval 97.5%	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.5%	Predictive Interval 97.5%	Mean	Confidence Interval 2.5%	Confidence Interval 97.5%	Mean	Confidence Interval 2.5%	Confidence Interval 97.5%	Mean	Confidence Interval 2.5%	Confidence Interval 97.5%	
7.0 (a)	697,800,000	20,900,000	651,600,000	735,100,000	261,000,000	18,360,000	219,900,000	293,800,000										7942
7.0 (b)	688,900,000	17,810,000	650,100,000	721,400,000	253,800,000	14,350,000	222,600,000	280,000,000										7,866
7.0 (c)	626,200,000	54,430,000	530,900,000	740,400,000	204,100,000	43,170,000	134,800,000	298,700,000										7,372
8.0 (a)	635,000,000	56,760,000	558,600,000	793,200,000	192,800,000	55,590,000	120,200,000	350,300,000										7,912
8.0 (b)	742,600,000	126,600,000	587,000,000	1,107,000,000	298,100,000	125,500,000	144,500,000	661,200,000										7,816
8.0 (c)	711,200,000	186,900,000	505,300,000	1,190,000,000	274,700,000	182,000,000	88,270,000	746,500,000										7,335

Model	Final Models																	Devi.
	Predictive distributions								Log Likelihood			AIC			BIC			
	Ultimate Claim Amount				Reported *IBNR* (1.9)													
	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.5%	Predictive Interval 97.5%	Mean	Standard Mean Sq. Predict. Error	Predictive Interval 2.5%	Predictive Interval 97.5%	Mean	Confidence Interval 2.5%	Confidence Interval 97.5%	Mean	Confidence Interval 2.5%	Confidence Interval 97.5%	Mean	Confidence Interval 2.5%	Confidence Interval 97.5%	
7.0 (d)	702,300,000	39,690,000	636,000,000	790,200,000	260,100,000	39,690,000	193,700,000	347,900,000	127.0	106.0	150.0	452.1	409.5	498.8	802.7	760.1	849.4	7,365
8.0 (d)	728,300,000	102,600,000	584,700,000	984,300,000	286,100,000	102,600,000	142,500,000	542,000,000	127.0	106.0	150.0	520.1	477.6	565.7	991.1	948.6	1,037.0	7,323

Table 4.5.3 Comparison of results for models 7.0 and 8.0.

4.6 Average Hazard Rate For Model 4.1(d)

As a pure loss measure, hazard rate can help comparing underwriting year contracts, to formulate portfolio management strategies, determine future premiums, portfolio composition, commutation or closure policies, etc. Hazard rates by underwriting year, or weighted average hazard rates for each array or for the whole claims portfolio can be derived from a reserving analysis. For payment year τ these are respectively:

$$h_{r,w,\tau-w+1} = -\left(\frac{\partial\left(\ln\left(1-P_{r,w,z}\right)\right)}{\partial z}\right)_{z=\tau-w+1} = \frac{\left(\frac{\partial P_{r,w,z}}{\partial z}\right)_{z=\tau-w+1}}{1-P_{r,w,\tau-w+1}} = \frac{C_{r,w}}{IBNR_{(r,w,\tau-w+1)}}\left(\frac{\partial P_{r,w,z}}{\partial z}\right)_{z=\tau-w+1}$$

$$A\overline{h}_{r,\tau} = \sum_{w=1}^{u_r} h_{r,w,\tau-w+1}\left(\frac{IBNR_{(r,w,\tau-w+1)}}{\sum_{k=1}^{u_r} IBNR_{(r,k,\tau-k+1)}}\right) = \sum_{w=1}^{u_r} h_{r,w,\tau-w+1}\left(\frac{IBNR_{(r,w,\tau-w+1)}}{IBNR_r(\tau)}\right)$$

$$G\overline{h}_{\tau} = \sum_{r=1}^{\eta}\sum_{w=1}^{u_r} h_{r,w,\tau-w+1}\left(\frac{IBNR_{(r,w,\tau-w+1)}}{\sum_{r=1}^{\eta}\sum_{k=1}^{u_r} IBNR_{(r,k,\tau-k+1)}}\right) = \sum_{r=1}^{\eta}\sum_{w=1}^{u_r} h_{r,w,\tau-w+1}\left(\frac{IBNR_{(r,w,\tau-w+1)}}{IBNR(\tau)}\right)$$

Given in terms of the *IBNR*, the above equations make explicit the changing nature of the average hazard as claims evolve. Table 4.6.1 lists the hazard rate values for payment years 13, 15 and 17 for model 7.0(d). Since underwriting year losses are at different stages in their development, a similar table to 4.6.1 can be used to assess the impact on the claims portfolio of, for example, excluding from it underwriting year contracts related to underwriting year j_1 of array i_1 and underwriting year j_2 of array i_2. The average hazard rate for the reduced portfolio becomes:

$$g\overline{h}_{\tau} = G\overline{h}_{\tau} + \frac{\sum_{n=1}^{2}\left(G\overline{h}_{\tau} - h_{i_n,j_n,\tau-j_n+1}\right) IBNR_{(i_n,j_n,\tau-j_n+1)}}{IBNR(\tau) - \sum_{n=1}^{2} IBNR_{(i_n,j_n,\tau-j_n+1)}}$$

The exclusion of the contracts from the claims portfolio reduces the portfolio hazard rate only when

$$\sum_{n=1}^{2}\left(G\overline{h}_{\tau} - h_{i_n,j_n,\tau-j_n+1}\right) IBNR_{(i_n,j_n,\tau-j_n+1)} < 0$$

Table 4.6.1 shows that the exclusion of underwriting year 10 from any of the three arrays would reduce the portfolio average hazard rate. The removal of underwriting year data sets belonging to any of the first seven underwriting years would have the opposite effect. While not included in table 4.6.1, from the reserving analysis the full distribution for the hazard rates can be obtained.

Model 7.0(d)					
	Underwriting Year	Ultimate Loss	Hazard rates Payment Year 13	Payment Year 15	Payment Year 17
Array 1	1	11,230,000	0.0515	0.0456	0.0410
	2	17,660,000	0.0548	0.0482	0.0430
	3	5,458,000	0.0589	0.0513	0.0455
	4	9,996,000	0.0638	0.0550	0.0484
	5	82,800,000	0.0695	0.0593	0.0516
	6	5,340,000	0.0760	0.0640	0.0552
	7	9,662,000	0.0816	0.0681	0.0583
	8	40,970,000	0.0932	0.0761	0.0640
	9	77,950,000	0.1046	0.0842	0.0697
	10	5,835,000	0.1161	0.0933	0.0762
Array 2	1	2,963,000	0.0502	0.0446	0.0401
	2	14,650,000	0.0543	0.0478	0.0427
	3	5,535,000	0.0593	0.0517	0.0458
	4	2,787,000	0.0635	0.0548	0.0482
	5	9,692,000	0.0694	0.0592	0.0515
	6	1,623,000	0.0759	0.0639	0.0551
	7	1,900,000	0.0810	0.0677	0.0580
	8	2,866,000	0.0895	0.0739	0.0626
	9	1,238,000	0.0990	0.0812	0.0679
	10	55,250,000	0.1152	0.0929	0.0759
Array 3	1	21,130,000	0.0551	0.0489	0.0440
	2	9,844,000	0.0604	0.0532	0.0476
	3	7,251,000	0.0626	0.0548	0.0487
	4	7,515,000	0.0653	0.0564	0.0497
	5	7,943,000	0.0699	0.0597	0.0520
	6	7,859,000	0.0755	0.0641	0.0555
	7	9,352,000	0.0856	0.0711	0.0607
	8	115,200,000	0.0914	0.0756	0.0640
	9	7,536,000	0.1000	0.0818	0.0682
	10	148,000,000	0.1147	0.0928	0.0760
By Array	Array 1	266,900,000	0.0906	0.0738	0.0620
	Array 2	98,510,000	0.1034	0.0833	0.0685
	Array 3	341,600,000	0.1041	0.0846	0.0700
Overall		707,000,000	0.0992	0.0804	0.0668

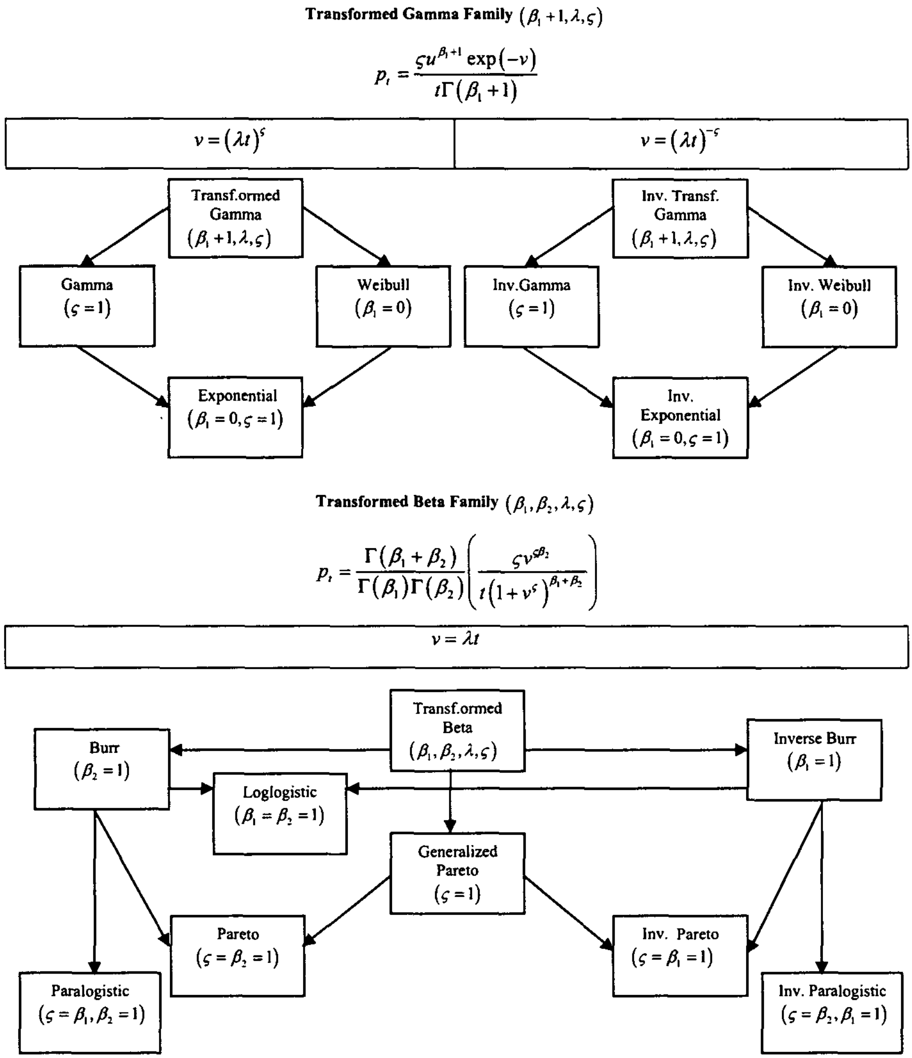
Table 4.6.1 Model 7.0(d): hazard rates for payment years 13,15 and 17.

5. Concluding Remarks

Reflective of the practical issues involved in the analysis of reserves, the related literature is extensive and explores a variety of theoretical frameworks. In general, having identified the salient data characteristics and gathered information on specific events that could have contributed to claims numbers and magnitudes, at the outset of every analysis a suitable analytical approach for the problem at hand has to be selected. Apart from any academic interest, it is likely that this search could have motivated some of the developments in reserving analysis, and will continue to do so. Hence, establishing the scope and limitations of each is important.

Through the generic model it is possible to give a functional interpretation to the claims data variability structure. As settlement speeds and scale variability increase, the assumptions and model structures encompassed by GLM models have to be replaced by more complex ones. The examples support remarks by Carroll (2003) with respect to the importance of the variance model. An inadequate variance model could lead to incorrect conclusions. The purpose of reserving analysis is not just to model historical claims data, but, more importantly, to predict *IBNR* and ultimate claim amounts. Both are strongly reliant on adequate variance definitions. Since claims records have to fulfil accounting requirements, corrections and adjustments to original entries are recorded as new transactions, and at unpredictable time lags. This could justify regarding measures of cumulative claims as repeated observations of an ongoing process. In this context, normal errors assumptions could be made tenable through suitable transformations or expectation functions, hence availing analytical approaches such as outlined in Lindstrom and Bates (1990). In the examples presented, and with the selected data, autoregressive error structures cannot be successfully used.

The generic model makes random effects models accessible to the problem of reserving. With the different variance model structures, it exponentially increases the analytical resources that can result from constructing families of reserving models around families of distributions. Graph 5.1 is an example of a template that can be used to identify the most suitable model structure for the data of interest and formulate the percentage cash flow function. With respect to the underlying assumptions for random effect parameters other alternatives are possible. Escobar and West (1992) propose a non-parametric approach, where the random parameter is taken from a rich class of distributions provided by the Dirichlet process. Lai and Shih (2003) leave the distribution of the random effects totally unspecified. The non-linear mixed effects models library (NLME) assumes that the random effects and the errors have Gaussian distributions. Using a matrix decomposition, Bates and Pinheiro (1998) shows that the random effects distribution expressed in terms of the relative precision factors can easily deliver the likelihood for the fixed and random effects. The flexibility of Gibbs sampling methods (Geman and Geman, 1984) has influenced the decision to implement the examples with BUGS (Spiegelhalter et al., 1995), as applications of Bayesian models and MCMC estimation methods. Nevertheless, other approaches in relation to analytical platforms, model structures and assumptions, beyond those explored, should be considered.



Graph 5.1 Examples of families of models for the percentage cash flow function.

ACKNOWLEGEMENTS

The author thanks Professor R. Verrall (Cass Business School) for his helpful remarks.

REFERENCES

Alba, de, E., (2002). Bayesian Estimation of Outstanding Claims Reserves. North American Actuarial Journal, Vol. 6, No. 4, 1-20.

Aziz, N. S., Gambertoglio, J. G., Lin, E. T., Grausz, H. and Benet, L. Z., (1978). Pharmacokinetics of Cefamandole using HPLC assay. Pharmacokinetics and Biopharmaceutics, 6, 153-164

Bates, D. M. and Pinheiro, J. C., (1998). Computational methods for multilevel models, Technical Memorandum BL0112140-980226-01TM, Bell Labs, Lucent Technologies, Murray Hill, NJ.

Beal, S. and Sheiner, L. B., (1992). NONMEM User's guides. NONMEM Project Group, University of California, San Francisco.

Bornheutter, R. L. and Ferguson, R. E., (1972). The actuary and IBNR. Proceedings of the Casualty Actuarial Society, LIX, 181-195.

Carroll, R. J., (2003). Variances are not always nuisance parameters. Biometrics, 59, 211-220.

Davinian M. and Giltinan, D. M., (1993). Some general estimation methods for non-linear mixed effects models. Journal of Biopharmaceutical Statistics, 3, 23-55.

Davinian M. and Giltinan, D. M., (1996). Nonlinear Models for Repeated Measurement Data. Chapman & Hall, London.

Elashoff, J. D., Reedy, T. J. and Meyer, J. H., (1982). Analysis of gastric emptying data. Gastroenterology 83, 1306-1312.

Escobar, M. and West, M., (1992). Computing Bayesian nonparametric hierarchical models. ISDS Discussion Paper 92-A20, Duke University.

Geman, S. and Geman, D., (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Tans. Pattn. Anal. Mach. Intell., 6(6), 721-741.

Kahane, Y., (1987). Ratemaking and reserving under decelerating inflation. (Unknown source).

Kremer, E., (1982). IBNR claims and the two way model of ANOVA. Scandinavian Actuarial Journal, 47-55.

Lai, T. L. and Shih, M., (2003). Nonparametric estimation in nonlinear mixed effects models. Biometrika, 90, 1, 1-13.

Lindstrom, M. J. and Bates, D. M., (1990). Nonlinear mixed effects models for repeated measures data. Biometrics 46, 673-687.

Mack, T., (1991). A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserves. ASTIN Bulletin 22, 93-109.

McCullagh, P. and Nelder, J., (1989). Generalised linear models, 2nd edition. Chapman and Hall.

Pinheiro, J. C., Bates, D.M., and Lindstrom, M. J., (1994). Model Building for Nonlinear Mixed Effects Models Technical Report 931, University of Wisconsin, Madison, Dept. of Statistics.

Pinheiro, J. C. and Bates, D. M., (2000). Mixed-Effects Models in S and S-Plus. Springer-Verlag New York, Inc.

Ratkowski, D. A., (1990). Handbook of Nonlinear Regression Models. Dekker, New York.

Renshaw, A. E., (1989). Chain-ladder and interactive modelling (claims reserving and GLIM). J.I.A. 116, 559-587.

Renshaw, A. E. and Verrall, R. J., (1998). A stochastic model underlying the chain-ladder technique. B.A.J. 4, 903-923.

Shapiro, S. S. and Wilk, M. B., (1965). An analysis of variance test for normality (complete samples), Biometrika, 52, 3 and 4, 591-611.

Spiegelhalter, D. J., Thomas, A., Best, N. G. and Gilks, W. R., (1995). BUGS 0.5: Bayesian inference using Gibbs sampling, MRC Biostatistics Unit, Cambridge, U.K.

Taylor, G. C., (2000). Loss reserving: an actuarial perspective. Kluwer Academic Publishers.

Verrall, R. J., (1991). On the Unbiased Estimation of Reserves from Loglinear Models. Insurance Mathematics and Economics, Vol. 10, No. 1, 75-80.

Vera, G. (2003). A generic claims reserving model. A fundamental risk analysis. Casualty Actuarial Society forum, 361-376.

Wakefield, J. C. and Walker, S., (1994). Population models with a non-parametric random coefficient distribution. Technical report, Department of Mathematics, Imperial College.

Wakefield, J. C., (1996). The Bayesian analysis of population pharmacokinetic models. Journal of the American Statistical Association, 91, 62-75.

Wright, T. S., (1990). A stochastic method for claims reserving in general insurance. J.I.A. 117, 677-731.

Zeger, S., L., Liang, K., Y. and Albert P., S., (1988). Models for Longitudinal Data: A generalized estimating equation approach, Biometrics, 44, 1049-1060.

APPENDIX A

A.1 INCREMENTAL PAID CLAIMS DATA

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	0	94,984	1,049,297	625,878	541,108	427,352	476,477	354,258	188,400	144,987
2	0	147,751	999,224	937,426	811,294	436,866	264,148	143,616	102,416	132,920
3	0	45,751	442,168	588,627	390,301	231,257	119,690	64,365	73,641	93,371
4	0	20,252	340,320	596,633	336,142	183,473	90,574	114,241	99,467	51,950
5	0	21,655	787,440	992,505	893,315	772,514	795,088	718,526	504,213	321,630
6	0	221,177	1,212,010	1,867,718	1,372,904	1,254,084	1,003,612	696,973	534,547	409,845
7	0	192,144	749,425	1,174,401	1,500,585	2,079,434	1,675,154	1,972,712	1,372,848	491,984

(cont.)

Und. Year	Development Period								
	11	12	13	14	15	16	17	18	19
1	124,614	111,642	56,210	64,259	33,893	15,440	8,255	22,300	25,173
2	109,996	58,163	53,679	54,255	25,631	51,443	56,702	59,857	
3	53,678	29,044	12,259	10,267	11,264	9,515	8,859		
4	45,692	21,824	36,117	54,185	52,194	47,355			
5	183,470	85,610	73,300	97,350	42,620				
6	111,090	529,552	403,242	291,414					
7	212,273	191,729	28,340						

Table A.1 Simulated data based on the claims experience of a mixed portfolio.

A.2 CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	0	94,984	1,144,281	1,770,159	2,311,267	2,738,619	3,215,096	3,569,354	3,757,754	3,902,741
2	0	147,751	1,146,975	2,084,401	2,895,695	3,332,561	3,596,709	3,740,325	3,842,741	3,975,661
3	0	45,751	487,919	1,076,546	1,466,847	1,698,104	1,817,794	1,882,159	1,955,800	2,049,171
4	0	20,252	360,572	957,205	1,293,347	1,476,820	1,567,394	1,681,635	1,781,102	1,833,052
5	0	21,655	809,095	1,801,600	2,694,915	3,467,429	4,262,517	4,981,043	5,485,256	5,806,886
6	0	221,177	1,433,187	3,300,905	4,673,809	5,927,893	6,931,505	7,628,478	8,163,025	8,572,870
7	0	192,144	941,569	2,115,970	3,616,555	5,695,989	7,371,143	9,343,855	10,716,703	11,208,687

(cont.)

Und. Year	Development Period								
	11	12	13	14	15	16	17	18	19
1	4,027,355	4,138,997	4,195,207	4,259,466	4,293,359	4,308,799	4,317,054	4,339,354	4,364,527
2	4,085,657	4,143,820	4,197,499	4,251,754	4,277,385	4,328,828	4,385,530	4,445,387	
3	2,102,849	2,131,893	2,144,152	2,154,419	2,165,683	2,175,198	2,184,057		
4	1,878,744	1,900,568	1,936,685	1,990,870	2,043,064	2,090,419			
5	5,990,356	6,075,966	6,149,266	6,246,616	6,289,236				
6	8,683,960	9,213,512	9,616,754	9,908,168					
7	11,420,960	11,612,689	11,641,029						

Table A.2 Cumulative data based on table A.1.

A.3 PRELIMINARY MODEL 1.0

A.3.1 MODEL 1.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	55,360	348,100	920,800	1,640,000	2,322,000	2,871,000	3,278,000	3,570,000	3,778,000	3,928,000
2	52,840	402,300	1,143,000	2,030,000	2,772,000	3,289,000	3,624,000	3,840,000	3,980,000	4,075,000
3	21,410	178,700	547,000	1,017,000	1,413,000	1,681,000	1,848,000	1,953,000	2,019,000	2,063,000
4	41,190	183,000	474,200	854,300	1,198,000	1,448,000	1,617,000	1,730,000	1,808,000	1,863,000
5	27,850	236,600	774,400	1,641,000	2,657,000	3,610,000	4,386,000	4,970,000	5,393,000	5,698,000
6	82,280	563,200	1,597,000	3,037,000	4,550,000	5,878,000	6,928,000	7,713,000	8,290,000	8,711,000
7	10,640	152,100	695,400	1,911,000	3,769,000	5,864,000	7,742,000	9,191,000	10,220,000	10,930,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,037,000	4,118,000	4,179,000	4,226,000	4,262,000	4,290,000	4,313,000	4,332,000	4,347,000
2	4,140,000	4,186,000	4,219,000	4,243,000	4,262,000	4,276,000	4,287,000	4,296,000	
3	2,093,000	2,113,000	2,128,000	2,139,000	2,147,000	2,153,000	2,157,000		
4	1,903,000	1,934,000	1,958,000	1,977,000	1,993,000	2,006,000			
5	5,917,000	6,077,000	6,195,000	6,284,000	6,351,000				
6	9,022,000	9,254,000	9,428,000	9,562,000					
7	11,410,000	11,740,000	11,970,000						

Table A.3 Fitted claims computed by Monte Carlo simulations estimated over 5000 independent samples.

A.3.2 MODEL 1.0 FIXED EFFECTS PARAMETER ESTIMATES, VARIANCE AND DEVIANCE

Und. Year	Fixed effect parameters						
	L	D	Kc	Kd	ϑ	K_{ε_1}	K_{ε_2}
	15.7400	4.8810	1.5470	-15.3600	-6.2470	-5.0680	-2.8080
	Underwriting year random effect parameters						
	l_w	d_w	kc_w	kd_w			
1	-0.4364	-0.4050	-0.5091	-0.1369			
2	-0.4592	-0.3512	-0.3995	-0.0313			
3	-1.1480	-0.0200	-0.3330	0.0317			
4	-1.1690	-0.1730	-0.4673	-0.4279			
5	-0.0342	0.6846	-0.3858	-0.0858			
6	0.3919	0.0078	-0.4838	-0.0484			
7	0.6101	2.3000	-0.1819	0.1548			
σ^2	1.85E+10						
Deviance	2,980						

Table A.4 Model 1.0 parameters and diagnostics.

A.4 PRELIMINARY MODEL 2.0

A.4.1 MODEL 2.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	65,520	381,400	962,800	1,665,000	2,323,000	2,855,000	3,255,000	3,547,000	3,759,000	3,914,000
2	47,310	385,000	1,127,000	2,026,000	2,778,000	3,299,000	3,634,000	3,847,000	3,985,000	4,077,000
3	15,680	155,300	519,400	1,006,000	1,420,000	1,696,000	1,864,000	1,965,000	2,028,000	2,068,000
4	17,190	143,400	444,600	847,300	1,211,000	1,473,000	1,645,000	1,756,000	1,828,000	1,876,000
5	28,240	238,800	780,300	1,649,000	2,662,000	3,612,000	4,385,000	4,967,000	5,390,000	5,695,000
6	81,750	558,800	1,592,000	3,036,000	4,555,000	5,885,000	6,932,000	7,715,000	8,288,000	8,706,000
7	14,370	176,900	753,200	1,981,000	3,810,000	5,856,000	7,699,000	9,140,000	10,180,000	10,910,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,029,000	4,115,000	4,181,000	4,232,000	4,272,000	4,303,000	4,329,000	4,350,000	4,367,000
2	4,139,000	4,183,000	4,214,000	4,237,000	4,255,000	4,268,000	4,278,000	4,286,000	
3	2,094,000	2,112,000	2,125,000	2,134,000	2,140,000	2,145,000	2,149,000		
4	1,908,000	1,931,000	1,948,000	1,960,000	1,969,000	1,976,000			
5	5,915,000	6,075,000	6,194,000	6,282,000	6,350,000				
6	9,014,000	9,244,000	9,417,000	9,550,000					
7	11,420,000	11,770,000	12,020,000						

Table A.5 Fitted claims computed by Monte Carlo simulations estimated over 7000 independent samples.

A.4.2 MODEL 2.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters						
	L	D	Kc	Kd	ϑ	K_{s_1}	K_{s_2}
	14.8900	5.0140	0.2660	-13.4400	-3.4710	-5.5230	-2.5030
	Underwriting year random effect parameters						
	l_w	d_w	kc_w	kd_w	ϑ_w		
1	0.4214	-0.7378	0.7250	-0.0208	-0.5061		
2	0.3865	-0.4262	0.8981	0.0177	-0.1910		
3	-0.3060	0.0697	0.9973	0.0297	-1.4670		
4	-0.3818	-0.0873	0.9274	-0.0648	-0.8611		
5	0.8153	0.5157	0.8900	-0.0115	-0.0693		
6	1.2400	-0.1151	0.8002	-0.1619	1.4680		
7	1.4700	1.9440	1.0560	0.3943	1.6780		
σ^2	5.87E+09						
Deviance	2,930						

Table A.6 Model 2.0 parameters and diagnostics.

A.5 VALIDATION MODEL 3.0

A.5.1 MODEL 3.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	65,890	381,500	961,700	1,663,000	2,321,000	2,854,000	3,255,000	3,547,000	3,760,000	3,915,000
2	45,370	377,000	1,117,000	2,023,000	2,782,000	3,305,000	3,640,000	3,852,000	3,989,000	4,079,000
3	14,930	152,700	516,400	1,005,000	1,421,000	1,698,000	1,866,000	1,967,000	2,029,000	2,068,000
4	16,700	142,400	443,800	846,500	1,210,000	1,473,000	1,646,000	1,757,000	1,828,000	1,876,000
5	29,470	241,700	782,300	1,649,000	2,664,000	3,614,000	4,386,000	4,967,000	5,389,000	5,692,000
6	97,230	611,800	1,662,000	3,081,000	4,555,000	5,852,000	6,887,000	7,673,000	8,259,000	8,695,000
7	10,310	147,300	679,700	1,884,000	3,745,000	5,856,000	7,748,000	9,202,000	10,230,000	10,940,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,030,000	4,116,000	4,182,000	4,233,000	4,273,000	4,304,000	4,330,000	4,351,000	4,368,000
2	4,140,000	4,183,000	4,214,000	4,236,000	4,253,000	4,266,000	4,276,000	4,283,000	
3	2,094,000	2,112,000	2,124,000	2,132,000	2,139,000	2,144,000	2,147,000		
4	1,909,000	1,931,000	1,948,000	1,960,000	1,968,000	1,975,000			
5	5,911,000	6,071,000	6,190,000	6,279,000	6,347,000				
6	9,020,000	9,266,000	9,454,000	9,600,000					
7	11,410,000	11,740,000	11,970,000						

Table A.7 Fitted claims computed by Monte Carlo simulations estimated over 25,500 independent samples.

A.5.2 MODEL 3.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters						
	L	D	Kc	Kd	ϑ	K_{s_1}	K_{s_2}
	14.2900	5.6670	0.3595	-13.6300	-6.8610	-5.3350	-2.4080
	Underwriting year random effect parameters						
	l_w	d_w	kc_w	kd_w			
1	1.0300	-1.3860	0.6305	0.1420			
2	0.9940	-1.0300	0.8152	0.0310			
3	0.3015	-0.5548	0.9109	0.0018			
4	0.2259	-0.7370	0.8359	-0.0281			
5	1.4240	-0.1191	0.7978	-0.0700			
6	1.8620	-0.9282	0.6661	-0.0521			
7	2.0660	1.5820	1.0150	0.2053			
Subset 1 - σ_1^2	9.53E+09						
Subset 2 - σ_2^2	4.21E+09						
Subset 3 - σ_3^2	4.05E+10						
Deviance	2,936						

Table A.8 Model 3.0 parameters and diagnostics.

A.6 VALIDATION MODEL 4.0

A.6.1 MODEL 4.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	68,260	386,900	967,500	1,667,000	2,323,000	2,854,000	3,253,000	3,545,000	3,757,000	3,912,000
2	47,650	385,200	1,126,000	2,026,000	2,779,000	3,300,000	3,635,000	3,848,000	3,986,000	4,077,000
3	14,160	149,000	512,300	1,005,000	1,424,000	1,702,000	1,869,000	1,969,000	2,030,000	2,069,000
4	15,730	138,100	438,100	844,200	1,212,000	1,477,000	1,649,000	1,759,000	1,830,000	1,877,000
5	28,640	236,300	771,100	1,637,000	2,656,000	3,613,000	4,389,000	4,971,000	5,392,000	5,695,000
6	80,580	552,900	1,581,000	3,023,000	4,545,000	5,881,000	6,933,000	7,719,000	8,293,000	8,713,000
7	10,650	149,800	685,900	1,892,000	3,749,000	5,853,000	7,740,000	9,195,000	10,230,000	10,930,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,028,000	4,114,000	4,181,000	4,232,000	4,272,000	4,305,000	4,331,000	4,352,000	4,369,000
2	4,139,000	4,183,000	4,214,000	4,237,000	4,254,000	4,268,000	4,278,000	4,286,000	
3	2,094,000	2,111,000	2,122,000	2,131,000	2,137,000	2,141,000	2,145,000		
4	1,909,000	1,931,000	1,946,000	1,958,000	1,966,000	1,973,000			
5	5,912,000	6,071,000	6,188,000	6,276,000	6,343,000				
6	9,021,000	9,251,000	9,424,000	9,556,000					
7	11,420,000	11,750,000	11,980,000						

Table A.9 Fitted claims computed by Monte Carlo simulations estimated over 25,500 independent samples.

A.6.2 MODEL 4.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters					
	L	D	Kc	Kd	K_{s_1}	K_{s_2}
	15.2900	5.6850	0.6367	-14.4900	-4.4210	-2.4880
	Underwriting year random effect parameters					
	l_w	d_w	kc_w	kd_w		
1	0.0272	-1.3800	0.3547	0.0661		
2	-0.0102	-0.9854	0.5449	-0.0122		
3	-0.7086	-0.2573	0.6872	-0.0329		
4	-0.7808	-0.5506	0.5905	0.0138		
5	0.4234	-0.1890	0.5087	0.0121		
6	0.8493	-0.8225	0.4171	0.0846		
7	1.0600	1.6470	0.7479	0.0277		
σ^2	1.06E+10					
ρ	0.0025					
Deviance	2,902					

Table A.10 Model 4.0 parameters and diagnostics.

A.7 FINAL MODEL 5.0

A.7.1 MODEL 5.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	65,470	380,000	959,800	1,662,000	2,322,000	2,855,000	3,256,000	3,549,000	3,761,000	3,915,000
2	45,730	377,900	1,118,000	2,023,000	2,782,000	3,305,000	3,640,000	3,853,000	3,989,000	4,079,000
3	14,630	150,600	513,300	1,004,000	1,423,000	1,700,000	1,868,000	1,969,000	2,030,000	2,069,000
4	16,030	138,600	438,300	844,500	1,212,000	1,477,000	1,649,000	1,759,000	1,830,000	1,877,000
5	28,820	240,200	780,900	1,648,000	2,661,000	3,611,000	4,385,000	4,968,000	5,391,000	5,696,000
6	84,690	567,600	1,601,000	3,038,000	4,549,000	5,875,000	6,924,000	7,710,000	8,288,000	8,711,000
7	13,570	171,000	738,800	1,964,000	3,803,000	5,864,000	7,714,000	9,153,000	10,190,000	10,910,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,030,000	4,115,000	4,181,000	4,231,000	4,271,000	4,303,000	4,328,000	4,349,000	4,365,000
2	4,141,000	4,183,000	4,214,000	4,236,000	4,253,000	4,266,000	4,276,000	4,283,000	
3	2,094,000	2,112,000	2,124,000	2,132,000	2,138,000	2,143,000	2,146,000		
4	1,908,000	1,930,000	1,946,000	1,958,000	1,966,000	1,973,000			
5	5,916,000	6,076,000	6,195,000	6,284,000	6,351,000				
6	9,023,000	9,256,000	9,433,000	9,568,000					
7	11,410,000	11,750,000	12,000,000						

Table A.11 Fitted claims computed by Monte Carlo simulations estimated over 25,500 independent samples.

A.7.2 MODEL 5.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters							
	L	D	Kc	Kd	ϑ_1^*	ϑ_2^*	K_{ϑ_1}	K_{ϑ_2}
	14.8300	5.2000	1.8770	-13.5100	-4.1220	1.6200	-4.4750	-2.2010
	Underwriting year random effect parameters							
	l_w	d_w	kc_w	kd_w				
1	0.4881	-0.8916	-0.8796	0.0577				
2	0.4528	-0.5451	-0.6994	-0.0371				
3	-0.2402	-0.0258	-0.5958	0.0205				
4	-0.3172	-0.1816	-0.6650	0.0317				
5	0.8829	0.3530	-0.7180	0.0166				
6	1.3100	-0.3071	-0.8147	-0.0499				
7	1.5340	1.8580	-0.5337	0.1329				
σ^2	5.05E+09							
Deviance	2,928							

Table A.12 Model 5.0 parameters and diagnostics.

A.8 FINAL MODEL 6.0

A.8.1 MODEL 6.0 FITTED VALUES OF CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Periods									
	1	2	3	4	5	6	7	8	9	10
1	66,200	383,900	964,400	1,663,000	2,317,000	2,848,000	3,249,000	3,542,000	3,755,000	3,912,000
2	135,300	641,200	1,389,000	2,138,000	2,748,000	3,197,000	3,517,000	3,745,000	3,908,000	4,027,000
3	35,860	237,300	618,100	1,046,000	1,401,000	1,652,000	1,820,000	1,931,000	2,006,000	2,057,000
4	59,230	265,600	575,400	903,400	1,188,000	1,410,000	1,575,000	1,697,000	1,787,000	1,854,000
5	24,320	220,100	747,500	1,618,000	2,651,000	3,620,000	4,404,000	4,988,000	5,407,000	5,704,000
6	165,500	810,000	1,909,000	3,252,000	4,598,000	5,790,000	6,773,000	7,554,000	8,166,000	8,642,000
7	8,110	125,600	616,800	1,796,000	3,698,000	5,889,000	7,832,000	9,285,000	10,280,000	10,940,000

(cont.)

Und. Year	Development Periods								
	11	12	13	14	15	16	17	18	19
1	4,028,000	4,116,000	4,182,000	4,234,000	4,275,000	4,307,000	4,334,000	4,355,000	4,372,000
2	4,116,000	4,183,000	4,235,000	4,275,000	4,308,000	4,334,000	4,355,000	4,372,000	
3	2,093,000	2,119,000	2,139,000	2,153,000	2,164,000	2,173,000	2,180,000		
4	1,905,000	1,945,000	1,975,000	2,000,000	2,019,000	2,035,000			
5	5,917,000	6,070,000	6,182,000	6,266,000	6,329,000				
6	9,015,000	9,309,000	9,543,000	9,731,000					
7	11,380,000	11,670,000	11,870,000						

Table A.13 Fitted claims computed by Monte Carlo simulations estimated over 25,500 independent samples.

A.8.2 MODEL 6.0 DIAGNOSTICS AND PARAMETER ESTIMATES

Und. Year	Fixed effect parameters									
	L	D	Kc	Kd	ϑ_1^*	ϑ_2^*	ϑ_3^*	ϑ_4^*	K_{ε_1}	K_{ε_2}
	15.8000	4.5610	1.2320	-12.9600	-1.9120	1.1780	15.6300	-0.1581	-6.3830	-1.3110
	Underwriting year random effect parameters									
	l_w	d_w	kc_w	kd_w						
1	-0.4813	-0.3208	-0.2500	-0.0844						
2	-0.4805	-1.0340	-0.3300	0.0070						
3	-1.1870	-0.3798	-0.1589	0.0487						
4	-1.2220	-0.9034	-0.3593	-0.0150						
5	-0.0970	1.0910	-0.0505	-0.0292						
6	0.3991	-0.3295	-0.3438	0.0398						
7	0.5364	2.9920	0.1889	0.0295						
σ^2	109.600									
Deviance	2.887									

Table A.14 Model 6.0 parameters and diagnostics.

B.1 ARRAYS 1 TO 3: CUMULATIVE PAID CLAIMS DATA

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	1,965,120	4,455,720	5,125,260	6,208,080	6,365,400	7,566,780	8,134,380	8,300,640	8,491,200	9,072,840
2	508,829	7,957,659	10,395,008	11,627,118	12,659,049	13,512,509	13,813,936	14,609,422	14,836,855	15,095,843
3	1,070,272	2,117,478	2,876,979	3,141,005	4,127,612	4,337,374	4,503,876	4,522,524	4,543,644	4,560,720
4	983,295	2,957,869	5,140,518	6,369,315	6,326,691	7,867,792	7,356,575	7,656,758	7,817,554	7,844,431
5	9,979,594	26,286,414	25,263,483	40,239,973	51,246,513	54,472,139	55,800,837	56,658,302	59,561,780	
6	55,668	1,586,037	1,764,809	2,888,328	3,158,562	3,445,626	3,459,794	3,604,995		
7	2,128,880	4,827,030	5,552,365	6,725,420	6,895,850	8,197,345	8,812,245			
8	2,528,789	8,400,695	12,219,988	21,139,396	23,109,446	22,292,555				
9	1,613,864	10,075,000	12,091,140	28,449,598	34,707,350					
10	110,580	576,687	1,887,649	2,244,074						

(cont.)

Und. Year	Development Period								
	11	12	13	14	15	16	17	18	19
1	9,298,740	9,640,380	9,681,600						
2	15,216,319	15,361,081							
3	4,572,331								
4									
5									
6									
7									
8									
9									
10									

Table B.1 Array 1: Simulated data based on the claims experience of a marine hull portfolio.

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	445,841	1,654,609	1,605,300	2,004,723	2,299,800	2,275,241	2,470,159	2,579,168	2,641,868	2,744,127
2	2,426,373	7,352,166	8,950,532	10,167,845	11,756,358	12,137,425	12,030,761	12,970,026	13,607,600	14,532,427
3	184,480	318,830	1,296,062	2,733,415	2,650,811	3,010,199	3,168,834	3,349,023	3,431,900	3,493,316
4	601,693	1,084,468	1,510,596	1,606,829	1,910,257	1,973,043	2,274,886	2,320,886	2,304,771	2,407,211
5	968,366	2,530,871	4,608,428	4,912,525	6,271,612	6,799,211	5,466,992	6,770,634	7,779,669	
6	239,105	326,614	670,735	905,967	913,090	1,131,129	1,090,519	1,100,114		
7	361,246	1,078,435	1,205,746	1,478,083	1,461,000	1,547,928	1,587,804			
8	742,335	1,392,375	1,937,655	1,865,055	2,282,175	2,388,270				
9	228,341	704,498	798,463	815,165	1,013,800					
10	1,589,527	15,936,500	22,331,091	28,741,729						

(cont.)

Und. Year	Development Period								
	11	12	13	14	15	16	17	18	19
1	2,786,141	2,870,495	2,871,777						
2	12,130,088	12,131,811							
3	3,542,361								
4									
5									
6									
7									
8									
9									
10									

Table B.2 Array 2: Simulated data based on the claims experience of a marine cargo portfolio.

Und. Year	Development Period									
	1	2	3	4	5	6	7	8	9	10
1	3,232,205	9,881,808	12,905,347	14,832,451	15,314,642	16,405,052	17,591,239	17,993,791	18,352,169	18,500,158
2	1,262,978	2,979,101	3,119,301	4,617,621	5,884,276	6,241,315	6,659,795	6,818,432	7,164,468	7,544,540
3	1,099,101	3,367,582	4,078,680	4,335,973	5,855,806	5,875,321	5,977,392	6,129,768	6,185,657	6,205,646
4	731,599	2,554,623	3,586,046	4,168,936	3,762,376	5,081,203	4,686,750	5,777,092	6,108,501	5,983,210
5	175,251	1,581,689	2,116,488	3,455,030	4,284,402	4,794,982	4,848,263	5,275,530	5,010,701	
6	1,339,210	3,824,400	4,704,740	5,565,181	5,412,190	6,389,658	6,517,524	7,284,369		
7	590,921	1,263,060	1,812,106	3,441,064	4,204,651	5,155,490	5,213,434			
8	20,698,911	43,203,529	63,631,429	70,713,997	89,249,817	85,287,218				
9	790,164	2,944,098	4,609,088	4,747,316	5,511,068					
10	9,900,060	17,916,636	50,167,809	62,132,660						

(cont.)

Und. Year	Development Period									
	11	12	13	14	15	16	17	18	19	20
1	18,733,437	19,496,439	19,567,234							
2	7,689,022	7,771,566								
3	6,240,897									
4										
5										
6										
7										
8										
9										
10										

Table B.3 Array 3: Simulated data based on the claims experience of an aviation hull portfolio.

A Least Squares Method of Producing Bornhuetter-Ferguson Initial Loss Ratios

Paul Brehm, FCAS, MAAA

Actuaries have relied on the Bornhuetter-Ferguson methodology in loss reserving since the "The Actuary and IBNR" [2] was published in 1972. The methodology is an intuitively appealing, credibility-weighted compromise between link ratio and expected loss ratio methods, where 'credibility' is inversely proportional to the remainder of the loss development tail. However, for almost as long as this method has been in existence, practitioners have been asking, "What do I use for my expected loss ratios?" Answers to this question (often unsatisfying ones) include industry data, company data for comparable classes of business, loss ratio pricing targets, planned loss ratios, and more.

This paper addresses the above question by offering a methodology for producing underlying loss ratios for use in the Bornhuetter-Ferguson method that are derived from the data itself. In particular, this paper addresses how to determine the underlying loss ratio for the initial time period in the analysis using a least squares methodology. The initial loss ratio is then used as the seed value for all subsequent loss ratios.

1. Derivation

Let:

L_{ij} = loss ratio for accident period i (i=1,...,n) evaluated cumulatively at j (j=1,...,m)

F_j = development factor from age j to ultimate

U_i^{BF} = Bornhuetter-Ferguson estimate of ultimate for accident period i

U_i^* = underlying ultimate loss ratio for accident period i (used in the Bornhuetter-Ferguson formula)

T_i = trend from time (i-1) to i – accident year dimension

P_i = earned effect of pricing from time (i-1) to i

Then the Bornhuetter-Ferguson estimate of the ultimate loss ratio for accident period i, with cumulative losses evaluated at j is:

$$U_i^{BF} = L_{ij} + U_i^*\left(1 - \frac{1}{F_j}\right) \tag{1.1}$$

An alternative estimate for accident period i can be derived by using losses evaluated one period earlier with the appropriate development factor, F_{j-1}:

$$U_i^{BF} = L_{i,j-1} + U_i^*\left(1 - \frac{1}{F_{j-1}}\right) \tag{1.2}$$

If (1.2) is subtracted from (1.1), the difference in estimated ultimates should be zero, but for some estimation error:

$$U_i^{BF} - U_i^{BF} = L_{ij} + U_i^*\left(1 - \frac{1}{F_j}\right) - L_{i,j-1} - U_i^*\left(1 - \frac{1}{F_{j-1}}\right) = 0 + \varepsilon_{ij}$$

Alternatively, after some manipulation:

$$\left(L_{ij} - L_{i,j-1}\right) = U_i^*\left[\left(1 - \frac{1}{F_{j-1}}\right) - \left(1 - \frac{1}{F_j}\right)\right] + \varepsilon_{ij} \tag{1.3}$$

Note that the term on the left of equation (1.3) is simply the incremental loss ratio. The term on the right is its expectation, conditioned on the underlying or expected loss ratio and the selected development pattern. If j = 1, that is, the first evaluation, then the term $\frac{1}{F_{j-1}}$ is undefined. For this initial condition, let $\frac{1}{F_{j-1}} = 0$, and the bracketed term on the right becomes $\left[\frac{1}{F_j}\right]$.

Now assume that the accident period loss ratios can be linked together over time by periodic trend (T_i) and pricing (P_i) factors according to:

$$U_i^* = U_{i-1}^* \frac{(1 + T_i)}{(1 + P_i)} \tag{1.4}$$

By successive substitutions, all underlying ultimate loss ratios can be linked back to the initial underlying loss ratio:

$$U_i^* = U_1^* \prod_{k=2}^{i} \frac{(1+T_k)}{(1+P_k)} \tag{1.5}$$

Substituting (1.5) into (1.3) for U_i^* yields the general formula

$$(L_{ij} - L_{i,j-1}) = U_1^* \prod_{k=2}^{i} \frac{(1+T_k)}{(1+P_k)} \left[\left(1 - \frac{1}{F_{j-1}}\right) - \left(1 - \frac{1}{F_j}\right) \right] + \varepsilon_{ij} \tag{1.6}$$

Note that (1.6) is of the form $Y_{ij}=\beta X_{ij}$, where the Y_{ij} are incremental loss ratios, the X_{ij} are the 'independent variables,' and β is the initial underlying loss ratio seed for the Bornhuetter-Ferguson model, U^*_1. The independent variables are simply derived values constructed from trend and pricing factors in the accident period dimension and loss development factors in the development dimension. Formula (1.6), then, can be estimated as a simple linear regression through the origin.

The parameter estimate, $\hat{\beta}$, is the initial loss ratio we are solving for. Think of the result as the initial loss ratio that is the least squares best estimate based on the data and conditioned on all the assumptions concerning pricing, trend, and loss development.

If it is assumed that trend is constant over the experience period, i.e., $T_i = T$ for all i, the formula (1.6) simplifies to:

$$(L_{ij} - L_{i,j-1}) = U_1^* (1+T)^{i-1} \left[\prod_{k=2}^{i} \frac{1}{(1+P_k)} \right] \left[\left(\frac{1}{1-F_{j-1}}\right) - \left(\frac{1}{1-F_j}\right) \right] + \varepsilon_{ij} \tag{1.7}$$

The functional form of (1.7) is particularly useful. Given a set of loss development factors and an earned price index, the above regression can be iterated over a range of annual trend assumptions. The final model can be chosen based on the underlying trend that maximizes R^2. (I know, it's data mining.)

Once U_1^* has been estimated, subsequent underlying loss ratios can be estimated as

$$U_i^* = U_1^* \prod_{k=2}^{i} \frac{(1+T_k)}{(1+P_k)} \tag{1.8}$$

or

$$U_i^* = U_1^*(1+T)^{i-1}\prod_{k=2}^{i}\frac{1}{(1+P_k)} \qquad (1.9)$$

for the constant trend case.

Since this is a regression model, the estimate $\hat{\beta}$ of U_1^* has an associated standard error, and a confidence interval can be established. The variance of $\hat{\beta}$ is

$$\sigma_{\hat{\beta}}^2 = \frac{\sigma^2}{\sum x_{ij}^2} \qquad (1.10)$$

where x_{ij} are the independent variables.

An unbiased estimate of σ^2 is

$$S^2 = \frac{\sum \varepsilon_{ij}^2}{(n-1)} \qquad (1.11)$$

where the ε_{ij} are the residuals from the regression, and n is the number of terms in the regression. There are n-1 degrees of freedom, as we are only estimating one parameter. The standard error of the coefficient -- the square root of the variance -- can be calculated as

$$S_{\hat{\beta}} = \left[\frac{\sum \varepsilon_{ij}^2}{(n-1)\sum x_{ij}^2}\right]^{1/2} \qquad (1.12)$$

The 100-α% confidence interval around $\hat{\beta}$ is

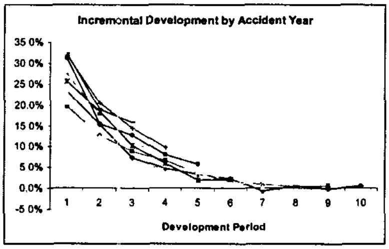
$$\hat{\beta} \pm t_{\alpha/2} S_{\hat{\beta}} \qquad (1.13)$$

The confidence interval (1.13) can be used to establish a range of estimates and thereby gauge the sensitivity of the reserve indication. For example, in the case where a trend factor is also estimated by ordinary least squares, a confidence interval can be estimated for the trend factor, as well. If a low estimate of trend is paired with the lower bound of β in formula (1.9) and a high estimate of trend is likewise paired with the upper bound of the confidence interval for β, both a low and a high loss ratio pattern can be traced over accident periods and used in the Bornhuetter-Ferguson estimation to derive low and high reserve estimates.

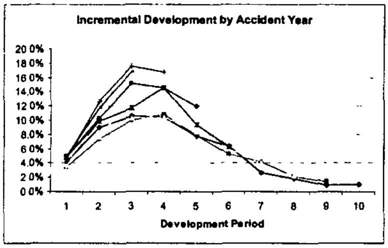
2. Example

Following is an example based on general liability data. The graphs below show the incremental loss ratios by accident period over time (development period) – case incurred on the left, paid data on the right.

Graph 2.1
Incurred Data
Accident Years over (Development) time

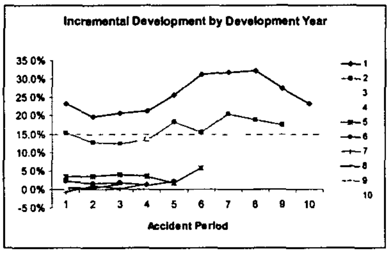


Graph 2.2
Paid Data
Accident Years over (Development) time

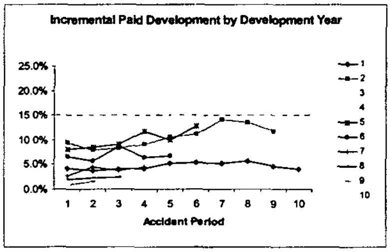


When viewed by development period over accident period below, the incremental loss ratios by evaluation would ideally behave like random pattern of points about a smooth trend line, if a constant trend and on-level factors truly picked up all the sources of systematic change over time. However, the data shows a departure in the pattern over accident periods starting in accident period 5 (see Graphs 2.3 and 2.4, below). This suggests a non-constant trend parameter or, alternatively, something affecting the loss ratios other than trend, e.g., underwriting or mix changes. In reality the departure associated with accident period 5 may well be better characterized as a calendar period distortion. Barnett and Zehnwirth's model [1] may be a good alternative in this case.

Graph 2.3
Incurred Data
Development Periods over (Accident) time

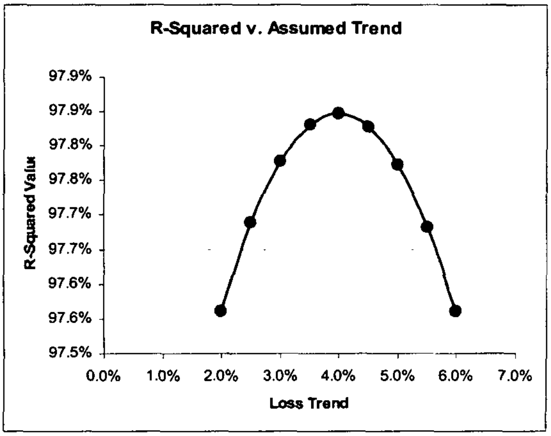


Graph 2.4
Paid Data
Development Periods over (Accident) time



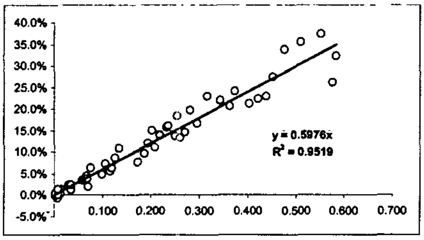
In this example, the constant trend case was modeled first for illustration purposes. In the example data, the least squares trend estimate using an exponential trend fit to pure premium was 3.5% (with an associated standard error of 0.016). However, R^2 was maximized using a trend of 4%:

Graph 2.5

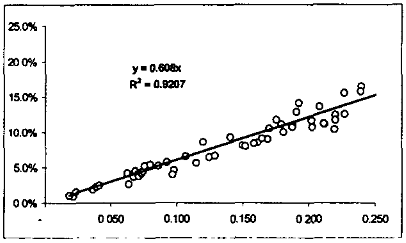


The resulting regressions can be seen below.

Graph 2.6
Incurred Data Regression



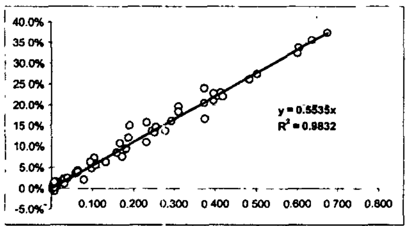
Graph 2.7
Paid Data Regression



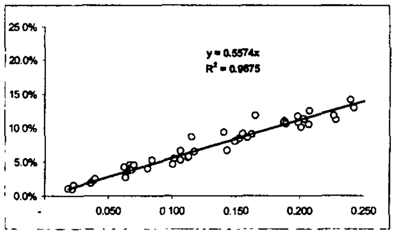
It can be seen from the above regressions that the paid and incurred data yield consistent results (initial loss ratios of 59.8% and 60.8%) from models with a strong goodness of fit (R^2 values of 96% and 92%, respectively). It would be appropriate, and more thorough, at this point to examine the residuals for serial correlation and non-constant error variance (heteroscedasticity). If either was a problem, the regressions could be adjusted accordingly.

To continue this example, trend was next assumed to vary over time. Underlying annual trend was set to 2% (rather than 4.0% overall), with additional period-on-period changes added to accident periods 5 through 9 to account for the calendar period distortion or "surprise" trend[1] (much like the industry observed in liability coverages in the late 90's). The resulting regressions are shown below.

Graph 2.8
Incurred Data Regression 2



Graph 2.9
Paid Data Regression 2



[1] I've never tried it, but it occurs to me that the regression could simply be augmented with 'distortion dummies' to automatically estimate the degree of departure from an underlying trend. This will be a subject of future research.

In the revised regressions, the case incurred estimate of the initial loss ratio is 55.3% (R^2 of 98.3%) and the paid loss estimate of the initial loss ratio is 55.7% (R^2 of 96.8%). $S_{\hat{\beta}}$ was 0.7% for both the paid and case incurred data, yielding a 95% confidence interval of roughly +/-1.5 loss ratio points at time 1.

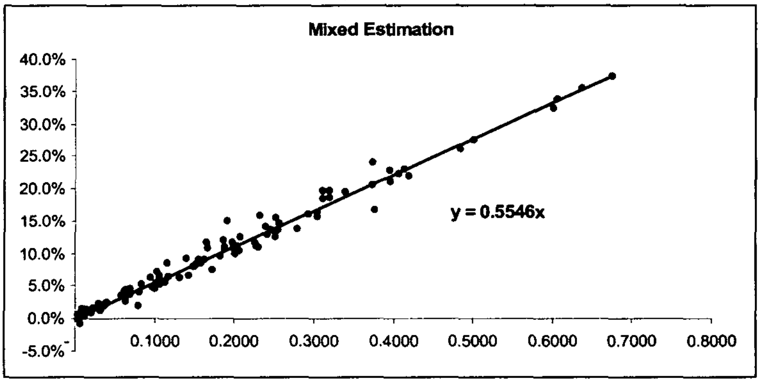
Given two estimates of the seed loss ratio, along with their respective error variances, we can credibility weight the two together to get one estimate. The formula for the paid data credibility parameter is:

$$Z_{Paid} = \frac{\left(\frac{1}{S^2_{Paid}}\right)}{\left(\left(\frac{1}{S^2_{Paid}}\right)+\left(\frac{1}{S^2_{Incurred}}\right)\right)} \qquad (2.1)$$

where S^2 is shown above in (1.11).

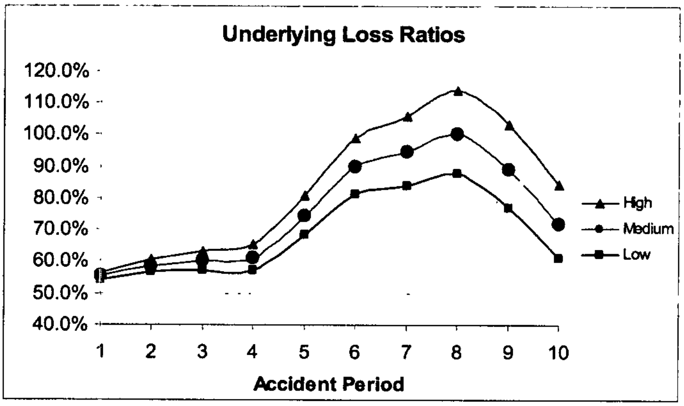
In practice, the credibility weighted solution can be derived directly by combining the paid and incurred regression matrices and doing a single, mixed regression [3]. The mixed estimate for this example is shown graphically in Graph 2.10. The mixed estimate of the initial loss ratio is 55.5% with an R^2 of 99.2%. $S_{\hat{\beta}}$ is 0.46% for both the paid and case incurred data, yielding a 95% confidence interval of roughly +/-1.1 loss ratio points at time 1.

Graph 2.10



The mixed estimate initial loss ratio, U^*_1, and the trend assumptions applied in the regression model substituted into formula (1.8) yields a pattern of underlying loss ratios as shown below. For the sake of this graph, the low and high loss ratios were calculated according to formula (1.13) for accident period 1 using a 95% confidence level. Subsequent accident period ultimate loss ratios were calculated with the selected trend plus and minus 1.6% respectively – one standard deviation around the least squares annual loss cost trend.

Graph 2.11



3. Conclusion

The above method has a strong appeal. Its strengths include utilizing all readily available data (dollars, counts, trends, premiums, exposures, pricing) and utilizing paid and incurred losses simultaneously to produce a 'best' (least squares) answer, in a computationally tractable manner, while still allowing the flexibility for ample actuarial judgment.

This method has always served me well, even with 'misbehaved' or sparse data. I hope it fills a need in your actuarial tool box.

References

[1] Barnett, G., and Zehnwirth, B. "Best Estimates for Reserves," 2000, Proceedings of the Casualty Actuarial Society, Vol. LXXXVII, p. 245

[2] Bornhuetter, R., and Ferguson, R. "The Actuary and IBNR," PCAS LIX, 1972, p. 181

[3] Brehm, P., and Guenthner D., "The Econometric Method of Mixed Estimation, An Application to the Credibility of Trend," CAS Discussion Paper Program, 1990.

Trending Entry Ratio Tables

Dan Corro*, NCCI

September, 2005

1 Background

Entry ratio tables are often a convenient mechanism for capturing information that is subject only to scale transforms. For example, the National Council on Compensation Insurance, Inc. (NCCI) stores excess loss factors (ELFs) in entry ratio tables. To determine an ELF at an attachment point, you simply divide the attachment point by the mean loss, and use that "entry ratio" value to look up the ELF in the table. A key assumption is that the underlying size of loss distribution changes only by a uniform scale transform over time (or by a transform that is close enough to a scale transform; c.f. Venter [3] for a discussion of scale adjustments and excess losses).

In fact, there can be forces at work that change the shape of size of loss distributions in ways that are not captured by scale transforms. For example, large claims might have greater trend factors than small claims (differential severity trend). Also, the frequency of small claims might decrease more than the frequency of large claims over some period of time (differential frequency trend). Not surprisingly, both of these possible effects act to "stiffen" the size of loss distribution, that is, increase the probability that a claim is "large," given that a claim occurs. A surprising result of our analysis is that the adjustments to entry ratio tables to take these phenomena into account, when they occur, often work in opposite directions. When large claims have

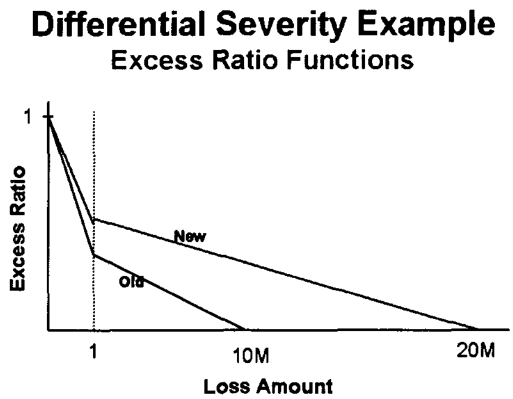
*Much thanks goes to Greg Engl and John Robertson, also of NCCI. Greg reviewed numerous drafts and his input improved the work throughout. John was key in promoting the topic within NCCI's actuarial research agenda. Both made direct and significant contributions to the paper.

greater trend factors than small claims, it might be necessary to increase the entry ratio table ELFs for large entry ratios. But when small claim frequency declines more rapidly than large claim frequency over a period of time, it might be necessary to *reduce* the tabular ELFs for large entry ratios.

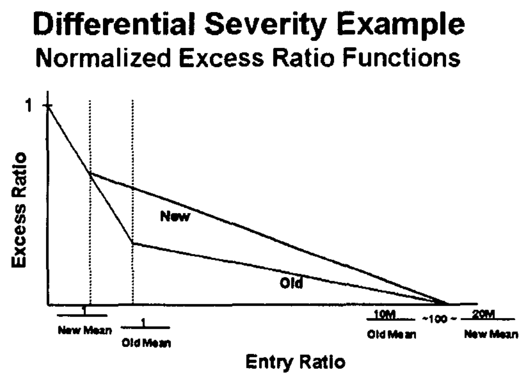
In this note we specify a generic, spreadsheet-friendly, format for an entry ratio table and consider the effects of differential trend and differential frequency changes. Each is illustrated by a real world Workers Compensation (WC) case study. We then describe general techniques for modifying an entry ratio table to account for not only a change in scale but also a change in the relativity between the mean and the median loss (or any fixed percentile loss) or a proportional shift in the hazard rate function of the loss distribution. The findings suggest that entry ratio tables work surprisingly well even for non-uniform trend and that in some important instances just a small adjustment can extend the shelf life of an entry ratio table.

2 Background

Before we get into the details of the paper, we present a thought experiment to illustrate some of the issues. Suppose we have 100 claims, 99 of which are for $1 and the other is a $10M claim. Consider what happens if over the next year inflation is expected to double the cost of the $10M claim, but leave the other 99 unchanged. Observe that the mean cost per claim is expected to roughly double, going from about $100K to about $200K. Recall that the **excess ratio** is simply the ratio of the sum of losses in excess of a per claim loss limitation to the total of all first dollar and up losses. The following is a sketch of the graph of the old and new excess ratios, expressed as functions of the loss limitation amount:

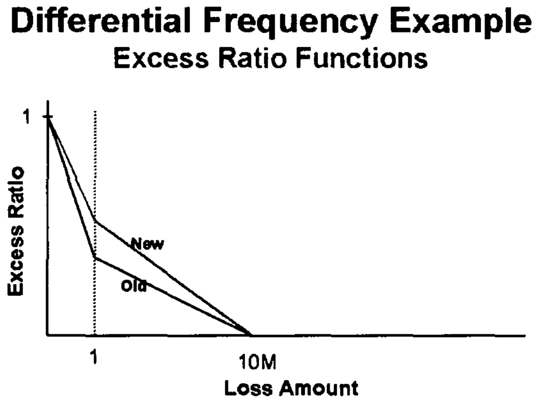


In practice, excess ratios are often captured in "entry ratio" tables, i.e. tables in which losses have been normalized to a mean value of 1. In this case, when we normailize the old and new losses by dividing by their respective means, the graph of the tabular excess ratios looks something like:

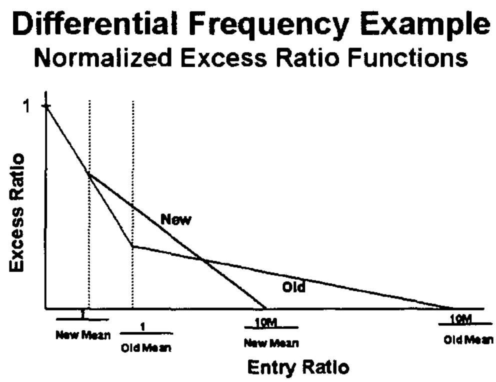


Observe that the new tabular values all lie at or above the old, which makes intuitive sense. Indeed, the inflation targeted the big claim, thereby "thickening" the tail of the loss distribution and necessitating the use of higher excess ratios next year. Because inflation changed the cost of claims selectively by size, this is a case of what the paper calls "Differential Severity".

Now suppose we begin with those same old 100 claims, but this time we consider what happens when, due to a safety initiative, half the \$1 claims do not emerge the next year. Because the change impacts claim frequency selectively by size, this is a case of what the paper calls "Differential Frequency". Notice that this experience change again roughly doubles the mean cost per case. Here the chart of the old and new excess ratio as a function of the loss limitation amount looks like:



and when normalized to entry ratio tabular values becomes:



Because the safety initiative is expected to be successful only for small claims, intuition again suggests a thickening of the tail. Observe, however, that the new tabular excess ratio values start out equal, then lie above, and eventually fall below the old. This suggests that, despite the similar impact on the mean cost per claim, something genuinely different is happening in the two scenarios. Actuaries should take heed that intuition can be a fallible guide to updating entry ratio tables.

3 Notation and Terminology

We start with a definition and, to keep the discussion self-contained, we derive some straightforward and familiar formulas:

Definition 1 *A random variable X is a* ***loss variable*** *if it has finite mean $\mu = E[X] > 0$ and has a density [PDF] f that is continuous when restricted to $[0, +\infty)$ and whose support is contained in $[0, +\infty)$. We denote the distribution function of X by $F(x) = \int_0^x f(y)dy$, whence $\frac{dF}{dx} = f(x)$ on $[0, +\infty)$. The survival function of X is $S = 1 - F$. The excess ratio function of X is given by $R(x) = \frac{E[Max(X-x,0)]}{\mu} = \frac{\int_x^\infty (y-x)f(y)dy}{\mu}$ for $x \geq 0$. We denote by $\widehat{F}$ the function given by $\widehat{F}(x) = \frac{\int_0^x yf(y)dy}{\mu}$ for $x \geq 0$. We use subscripts on μ_X, f_X, F_X, S_X, R_X, and $\widehat{F}_X$ when necessary to indicate dependence on X.*

The following proposition expresses the excess ratio function in terms of F and $\widehat{F}$.

Proposition 2 $R(x) = 1 - \widehat{F}(x) - \frac{x}{\mu}[1 - F(x)]$, *for all* $x \geq 0$.

Proof. From the definition of $R(x)$ we have

$$\begin{aligned} R(x) &= \frac{1}{\mu}\int_x^\infty (y-x)f(y)dy \\ &= \frac{1}{\mu}\left[\int_x^\infty yf(y)dy - x\int_x^\infty f(y)dy\right] \\ &= \frac{1}{\mu}\left[\mu - \int_0^x yf(y)dy - xS(x)\right] \\ &= 1 - \frac{1}{\mu}\int_0^x yf(y)dy - \frac{x}{\mu}S(x) \\ &= 1 - \widehat{F}(x) - \frac{x}{\mu}[1 - F(x)]. \end{aligned}$$

as required. This completes the proof. ∎

It is well known that the mean of a nonnegative random variable, X, can be expressed in terms of its survival function as $E[X] = \int_0^\infty S(x)dx$. It is easy to see that a similar result also holds for excess ratio functions.

Proposition 3 *Let X be a loss variable with survival function S and excess ratio function R, then*

$$R(x) = \frac{\int_x^\infty S(y)dy}{\int_0^\infty S(y)dy}, \text{ for all } x \geq 0.$$

Proof. Let X have density f, then noting that $\frac{dS}{dy} = -f(y)$ and using integration by parts, we have

$$\begin{aligned} \int_x^\infty S(y)dy &= yS(y)]_x^\infty + \int_x^\infty yf(y)dy \\ &= -xS(x) + \int_x^\infty yf(y)dy \\ &= -x\int_x^\infty f(y)dy + \int_x^\infty yf(y)dy \\ &= \int_x^\infty (y-x)f(y)dy, \end{aligned}$$

where the second equality follows from:

$$\begin{aligned} \mu &= E[X] < \infty \Rightarrow \text{[read "implies"]} \\ xS(x) &= x\int_x^\infty f(y)dy \leq \int_x^\infty yf(y)dy \to 0 \text{ as } x \to \infty. \end{aligned}$$

Thus

$$R(x) = \frac{\int_x^\infty (y-x)f(y)dy}{E[X]} = \frac{\int_x^\infty S(y)dy}{\int_0^\infty S(y)dy}.$$

as required. ∎

Corollary 4 $\frac{dR}{dx}(x) = \frac{-S(x)}{\mu}$, *for all $x \geq 0$.*

Proof. By the Fundamental Theorem of Calculus:

$$\frac{dR}{dx}(x) = \frac{d}{dx}\left(\frac{\int_x^\infty S(y)dy}{\mu}\right) = \frac{-S(x)}{\mu}.$$

as required. ∎

Proposition 5 *Let X be a loss variable with density f_X and distribution function F_X and let $\alpha, \beta > 0$ be any two positive constants. Set:*

$$Y = \alpha X^{\beta}$$

then for every $x, y > 0$:

1. $f_Y(y) = \frac{1}{\alpha^{\frac{1}{\beta}}\beta} y^{\frac{1-\beta}{\beta}} f_X\left(\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}\right)$

2. $F_Y(y) = F_X\left(\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}\right)$

3. $\widehat{F}_Y(y) = \frac{\int_0^{\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}} w^{\beta} f_X(w)dw}{\mu_{X^{\beta}}}$

4. $R_X(x) = R_{Y^{\frac{1}{\beta}}}(\alpha^{\frac{1}{\beta}} x)$

Proof. We note that

$$F_Y(y) = \Pr(Y \leq y) = \Pr(\alpha X^{\beta} \leq y) = \Pr(X \leq \left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}) = F_X\left(\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}\right)$$

proving 2.

For 1, just differentiate 2, using the change of variable $z = \left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}} \Rightarrow \frac{dz}{dy} = \frac{1}{\beta}\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}-1}\frac{1}{\alpha} = \frac{1}{\alpha^{\frac{1}{\beta}}\beta} y^{\frac{1-\beta}{\beta}}$:

$$f_Y(y) = \frac{dF_Y}{dy} = \frac{dF_X\left(\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}\right)}{dy} = \frac{dF_X(z)}{dz}\frac{dz}{dy} = f_X\left(\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}\right)\frac{1}{\alpha^{\frac{1}{\beta}}\beta} y^{\frac{1-\beta}{\beta}}$$

And for 3 just integrate using the change of variable

$$w = \left(\frac{z}{\alpha}\right)^{\frac{1}{\beta}} \Leftrightarrow \alpha w^{\beta} = z \Rightarrow \frac{dw}{dz} = \frac{1}{\alpha^{\frac{1}{\beta}}\beta} z^{\frac{1-\beta}{\beta}}$$

we have:

$$\begin{aligned}
\widehat{F}_Y(y) &= \frac{\int\limits_0^y z f_Y(z)dz}{\mu_Y} \\
&= \frac{\int\limits_0^{z=y} \alpha w^\beta \frac{1}{\alpha^{\frac{1}{\beta}}\beta} z^{\frac{1-\beta}{\beta}} f_X\left(\left(\frac{z}{\alpha}\right)^{\frac{1}{\beta}}\right) dz}{\alpha \mu_{X^\beta}} \\
&= \frac{\int\limits_0^{z=y} w^\beta f_X(w) \frac{dw}{dz} dz}{\mu_{X^\beta}} \\
&= \frac{\int\limits_0^{\left(\frac{y}{\alpha}\right)^{\frac{1}{\beta}}} w^\beta f_X(w) dw}{\mu_{X^\beta}}
\end{aligned}$$

Finally:

$$\begin{aligned}
Y &= \alpha X^\beta \Rightarrow Y^{\frac{1}{\beta}} = \alpha^{\frac{1}{\beta}} X \Rightarrow \mu_{Y^{\frac{1}{\beta}}} = \alpha^{\frac{1}{\beta}} \mu_X \\
\alpha^{\frac{1}{\beta}} \mu_X R_X(x) &= \alpha^{\frac{1}{\beta}} E[Max(X - x, 0)] \\
&= \alpha^{\frac{1}{\beta}} E[Max(\left(\frac{Y}{\alpha}\right)^{\frac{1}{\beta}} - x, 0)] \\
&= E[Max(Y^{\frac{1}{\beta}} - \alpha^{\frac{1}{\beta}} x, 0)] \\
&= \mu_{Y^{\frac{1}{\beta}}} R_{Y^{\frac{1}{\beta}}}(\alpha^{\frac{1}{\beta}} x) \\
&= \alpha^{\frac{1}{\beta}} \mu_X R_{Y^{\frac{1}{\beta}}}(\alpha^{\frac{1}{\beta}} x) \\
&\Rightarrow R_X(x) = R_{Y^{\frac{1}{\beta}}}(\alpha^{\frac{1}{\beta}} x)
\end{aligned}$$

completing the proof. ■

The special case $\beta = 1$ applies when normalizing losses, in particular when dividing by the mean loss to get entry ratios:

Corollary 6 *Let X be a loss variable with density f_X and distribution function F_X, and let $\alpha > 0$, then*

1. $f_{\alpha X}(y) = \frac{f_X(\frac{y}{\alpha})}{\alpha}$

2. $F_{\alpha X}(y) = F_X(\frac{y}{\alpha})$
3. $\widehat{F}_{\alpha X}(y) = \widehat{F}_X(\frac{y}{\alpha})$
4. $R_{\alpha X}(y) = R_X(\frac{y}{\alpha})$

Proof. All but number 3 are clear from Proposition 5, and 3 is very nearly so:

$$\widehat{F}_{\alpha X}(y) = \frac{\int_0^{\frac{y}{\alpha}} w f_X(w)dw}{\mu_X} = \widehat{F}_X(\frac{y}{\alpha})$$

as required. ∎

We associate to a loss variable X with (finite) mean $\mu = \mu_X = E[X]$ an entry ratio table, which we term the $rAB = rAB_X$ table. The table consists of the two functions:

$$\begin{aligned} A_X(r) &= F_{X/\mu}(r) = \mu \int_0^r f(\mu x)dx = F_X(\mu r) \\ B_X(r) &= \widehat{F}_{X/\mu}(r) = \mu \int_0^r x f(\mu x)dx = \widehat{F}_X(\mu r) \end{aligned}$$

Clearly, for any positive scalar $\alpha > 0$ if $Y = \alpha X$, then

$$\frac{Y}{\mu_Y} = \frac{\alpha X}{\alpha \mu_X} = \frac{X}{\mu} \Rightarrow A_X = A_Y \text{ and } B_X = B_Y \Rightarrow rAB_Y = rAB_X$$

and indeed the entry ratio table is invariant under such a transformation of scale.

The dependent variable r is termed an "entry ratio" and corresponds to losses (but has applications to any positive real valued distribution, e.g. a wage distribution) normalized to a mean of 1. We often speak of these two functions as determining the A and B "columns" of the entry ratio table. Note that:

$$\begin{aligned} A_X(\infty) &= \lim_{r\to\infty} A_X(r) = 1 \\ B_X(\infty) &= \lim_{r\to\infty} B_X(r) = 1 \end{aligned}$$

Column A is sometimes described as the percent of claims at or below the corresponding entry ratio (r), while column B is described as the percent of losses corresponding to the claims in column A. This rAB setup is employed in WC benefit on-level calculations, and is especially practical for spreadsheets that deal with calculations that involve normalized loss variables.

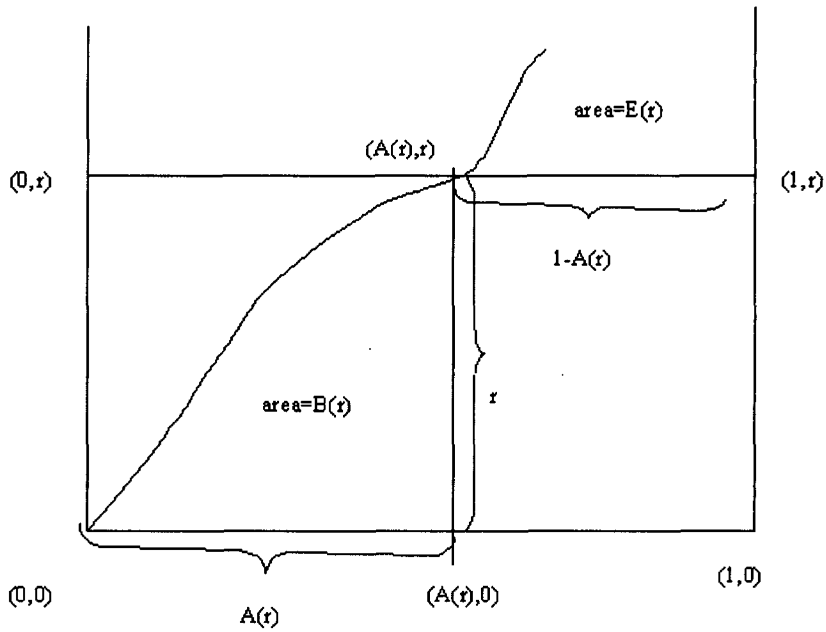
We are particularly interested in determining how $E_X(r)$, which we also refer to as the normalized excess ratio, behaves subject to a non-scale "trend" adjustment. For convenience we often expand the entry ratio table to include a third column E, readily derived from the others by applying Proposition 2 and Corollaries 4 and 6 to X/μ:

$$\begin{aligned} E_X(r) &= R_{X/\mu}(r) = 1 - B_X(r) - r(1 - A_X(r)) \\ \frac{dE_X}{dr}(r) &= \frac{dR_{X/\mu}}{dr}(r) = \frac{-S_{X/\mu}(r)}{1} = -S_X(\mu r) = F_X(\mu r) - 1 \end{aligned}$$

The following picture, reminicent of the area interpretation of integration by parts (c.f. Lee [2]), illustrates the usual way of visualing the rAB table and illustrates the formula for the normalized excess ratio:

$$E(r) = 1 - B(r) - r(1 - A(r))$$

in terms of r, A, and B:



We will let Y denote a loss variable that captures the effect of applying "trend" to X. We also set:

$$\begin{aligned} G &= F_Y \\ g &= f_Y \\ \nu &= E[Y]. \end{aligned}$$

Our goal is to determine rAB_Y from rAB_X. We are particularly interested in the absolute and relative impacts on the normalized excess ratio:

$$\begin{aligned} \delta(r) &= \delta_{XY}(r) = E_Y(r) - E_X(r) \\ \rho(r) &= \frac{\delta(r)}{E_X(r)}. \end{aligned}$$

We clearly have:

$$\begin{aligned} \delta(0) &= \rho(0) = 0 \\ \lim_{r \to \infty} \delta(r) &= 0 \end{aligned}$$

Taking derivatives and applying L'Hôpital (twice), we have:

$$
\begin{aligned}
\frac{d\delta}{dr} &= G(\nu r) - F(\mu r) \\
\frac{d^2\delta}{dr^2} &= \nu g(\nu r) - \mu f(\mu r) \\
\frac{d\rho}{dr} &= \frac{E_X(r)G(\nu r) - E_Y(r)F(\mu r) + \delta(r)}{E_X(r)^2} \\
1 + \lim_{r\to\infty} \rho(r) &= 1 + \lim_{r\to\infty} \frac{\delta(r)}{E_X(r)} = 1 + \lim_{r\to\infty} \frac{G(\nu r) - F(\mu r)}{F(\mu r) - 1} \\
&= 1 + \lim_{r\to\infty} \frac{G(\nu r) - 1 - (F(\mu r) - 1)}{F(\mu r) - 1} \\
&= 1 + \lim_{r\to\infty} \left(\frac{G(\nu r) - 1}{F(\mu r) - 1} - 1 \right) \\
&= \lim_{r\to\infty} \frac{G(\nu r) - 1}{F(\mu r) - 1} = \lim_{r\to\infty} \frac{\nu g(\nu r)}{\mu f(\mu r)} \\
&= \frac{\nu}{\mu} \lim_{s\to\infty} \frac{g(\frac{\nu}{\mu} s)}{f(s)} \text{ (since } r \to \infty \Leftrightarrow s = \mu r \to \infty \text{).}
\end{aligned}
$$

For large entry ratios, the impact of trend on the normalized excess ratio column, $E_X(r)$ vs. $E_Y(r)$, is dictated by the impact of trend on the mean and on the largest losses. For any loss variable X let M_X denote the maximum loss (in the case of no finite maximum loss amount, we set $M_X = \infty$).

Proposition 7 *Suppose X and Y are two loss variables with $M_X, M_Y < \infty$ and $\frac{M_X}{\mu_X} > \frac{M_Y}{\mu_Y}$, then there exists $b > 0$ such that $E_Y(b) < E_X(b)$ and $0 = E_Y(r) \leq E_X(r)$ for $r \geq b$.*

Proof. Setting $b = \frac{M_Y}{\mu_Y} < \frac{M_X}{\mu_X}$ we have

$$
\begin{aligned}
b\mu_X &< M_X \Rightarrow \\
E_Y(b) &= R_{Y/\mu_Y}(b) = R_Y(\mu_Y b) \\
&= R_Y(M_Y) = 0 < R_X(b\mu_X) = R_{X/\mu_X}(b) = E_X(b) \\
\text{and } r &\geq b \Rightarrow \\
r\mu_Y &\geq b\mu_Y = M_Y \Rightarrow E_Y(r) = R_{Y/\mu_Y}(r) \\
&= R_Y(\mu_Y r) = 0 \leq E_X(r)
\end{aligned}
$$

as required. ∎

We will find a use for the following later in Section 4:

Proposition 8 *Suppose X and Y are two loss variables with the same maximum loss $M_X = M_Y < \infty$ and with $\mu_Y > \mu_X$, then there exists $a > 0$ such that $R_Y(r) > R_X(r)$ for $0 < r \leq a$ and there exists $b > 0$ such that $E_Y(b) < E_X(b)$ and $0 = E_Y(r) \leq E_X(r)$ for $r \geq b$.*

Proof. Since $\mu_Y > \mu_X$, the existence of b follows from Proposition 7. For the existence of a, we have from Corollary 4:

$$\frac{dR_X}{dx}(0) = \frac{-1}{\mu_X} < \frac{-1}{\mu_Y} = \frac{dR_Y}{dy}(0)$$

Now clearly $R_Y(0) = R_X(0) = 1$ and since R_Y and R_X are continuously differentiable there exists $a > 0$ with

$$\begin{aligned}\frac{R_X(x) - 1}{x} &= \frac{R_X(x) - R_X(0)}{x - 0} \\ &< \frac{R_Y(y) - R_Y(0)}{y - 0} = \frac{R_Y(y) - 1}{y} \text{ for every } x, y \in (0, a].\end{aligned}$$

In particular:

$$\begin{aligned}0 &< r \leq a \Rightarrow \frac{R_X(r) - 1}{r} < \frac{R_Y(r) - 1}{r} \\ &\Rightarrow R_X(r) - 1 < R_Y(r) - 1 \Rightarrow R_X(r) < R_Y(r).\end{aligned}$$

This completes the proof. ∎

4 Differential Severity Trend

Let the function $h(x)$ defined on $[0, \infty)$ be such that $h(x) \geq 1$ and $\frac{dh}{dx} > 0$ on $[0, \infty)$. In this section we assume $f(x) > 0$ for $x > 0$. Think of $h(x)$ as a severity trend factor that increases with the size of loss x. The random variable of the trended loss is $Y = \psi(X)$, where the transformation $\psi(x) = h(x)x$ has $\frac{d\psi}{dx} = h(x) + x\frac{dh}{dx} > 1$ for $x > 0$ and is order preserving and invertible (and expands distances). Thus:

$$G(\psi(x)) = \Pr(Y \leq \psi(x)) = \Pr(\psi(X) \leq \psi(x)) = \Pr(X \leq x) = F(x).$$

We clearly have $\psi(x) \geq x \Rightarrow \nu = E[Y] = E[\psi(X)] \geq E[X] = \mu$ and $F(x) = G(\psi(x) \geq G(x)$ Observe that:

$$\begin{aligned}
x &\geq a \Leftrightarrow \psi(x) \geq \psi(a) \\
x &\geq a \Rightarrow \psi(x) - \psi(a) = h(x)x - h(a)a \geq h(a)x - h(a)a \\
&= h(a)\,(x-a) \geq x - a \\
&\Rightarrow \nu R_Y(\psi(x)) = E[Max(Y - \psi(x), 0)] \\
&= E[Max(\psi(X) - \psi(x), 0)] \geq E[Max(X - x, 0)] = \mu R_X(x) \\
&\Rightarrow R_Y(\psi(x)) \geq \left(\frac{\mu}{\nu}\right) R_X(x)
\end{aligned}$$

But the relationship between the normalized excess ratios $E_Y(r)$ and $E_X(r)$ is more subtle.

Let $h_M = \lim_{x\to\infty} h(x)$ and $h_m = h(0)$, then $1 \leq h_m < h_M \leq \infty$ and we have:

$$\begin{aligned}
h_m\mu &= h_m E[X] < E[h(X)X] = E[\psi(X)] = E[Y] = \nu < h_M E[X] = h_M\mu \\
&\Rightarrow h_m < \frac{\nu}{\mu} < h_M \\
&\Rightarrow \text{there exists exactly one } a > 0 \text{ such that } h(a) = \frac{\nu}{\mu}.
\end{aligned}$$

However, we see that since F and ψ are both monotonic increasing, whence invertible, and so too is $G = F \circ \psi^{-1}$. Whence for $r > 0$ we have the equivalence:

$$\begin{aligned}
0 &= \frac{d\delta}{dr} = G(\nu r) - F(\mu r) \Leftrightarrow G(\nu r) = F(\mu r) \Leftrightarrow \nu r = \psi(\mu r) \\
&\Leftrightarrow h(\mu r)\mu r = \nu r \\
&\Leftrightarrow h(\mu r)\mu = \nu \Leftrightarrow h(\mu r) = \frac{\nu}{\mu} = h(a) \Leftrightarrow a = \mu r \Leftrightarrow r = \frac{a}{\mu}.
\end{aligned}$$

Now $0 = \delta(0) = \lim_{r\to\infty} \delta(r)$ and so it follows that, unless $\delta(r) = 0$ for every $r \geq 0$, the function $\delta(r)$ has either a unique minimum or a unique maximum on $(0, \infty)$, and consequently $\delta(r)$ is either always ≥ 0 or always ≤ 0, for all $r \geq 0$. We claim that $\delta(r) \geq 0$ for all $r \geq 0$. To verify this, select β such that

$h_m < \beta < \frac{\nu}{\mu}$ and let $b = \mu s = h^{-1}(\beta) > 0$; then:

$$\begin{aligned}
a &= r\mu, b = s\mu, 1 < \beta = h(b) < \frac{\nu}{\mu} = h(a) \Rightarrow b < a \\
\psi(s\mu) &= \psi(b) = h(b)b = \beta b < \frac{\nu}{\mu}b = \frac{\nu}{\mu}s\mu = \nu s \\
&\Rightarrow F(s\mu) = G(\psi(s\mu)) < G(\nu s) \\
&\Rightarrow \frac{d\delta}{dr}(s) = G(\nu s) - F(s\mu) > 0
\end{aligned}$$

It follows that $\delta(r)$ is increasing at $s = \frac{b}{\mu}$ and therefore on the entire interval $(0, \frac{a}{\mu})$. Since $\delta(0) = 0$, this clearly forces $\delta(\frac{a}{\mu}) > 0$ and consequently $\delta(r) \geq 0$ for all $r \geq 0$, as claimed.

We see that the graph of $\delta(r)$ is $\cap$-shaped, i.e. is concave with $0 = \delta(0) = \lim_{r\to\infty} \delta(r)$, with a unique maximum at $r = \frac{a}{\mu}$. We have established:

Proposition 9 *In the case of the differential severity trend model $G(\psi(x)) = F(x)$ and $f(x) > 0$ for $x > 0$, as defined above, $E_Y(r) - E_X(r) > 0$ for all $r > 0$.*

Let $r_0 = 0 < r_1 < r_2 < \cdots < r_M$ be a sequence of entry ratios and set

$$A_i = A_X(r_i), B_i = B_X(r_i), 0 \leq i \leq M.$$

Suppose that $A_i = A_X(r_i) > A_X(r_{i-1}), 1 \leq i \leq M$ and $A_M = 1$. Set $\Delta A_i = A_i - A_{i-1}, \Delta B_i = B_i - B_{i-1}$. Note that $\mu\frac{\Delta B_i}{\Delta A_i}, 1 \leq i \leq M$, is the mean value of the untrended loss over the interval $[\mu r_{i-1}, \mu r_i]$. For $1 \leq i \leq M$, set

$$\begin{aligned}
\Delta\widetilde{B}_i &= \Delta A_i \left(\psi\left(\mu\frac{\Delta B_i}{\Delta A_i}\right)\right) \\
\widetilde{B}_i &= \sum_{k=1}^{i} \Delta\widetilde{B}_k.
\end{aligned}$$

Since ψ is order preserving, it is reasonable to assume that $\psi\left(\mu\frac{\Delta B_i}{\Delta A_i}\right)$ is a good estimate of the mean value of the trended losses on the interval

$[\psi(\mu r_{i-1}), \psi(\mu r_i)]$, (the smaller the interval, the more accurate the estimate).

$$\begin{aligned}
\widetilde{B}_M &= \sum_{k=1}^{M} \Delta \widetilde{B}_k = \sum_{k=1}^{M} \Delta A_i \left(\psi\left(\mu \frac{\Delta B_i}{\Delta A_i}\right)\right) \\
&= \sum_{k=1}^{M} \left(G(\psi(\mu r_i) - G(\psi(\mu r_{i-1}))\right) \left(\psi\left(\mu \frac{\Delta B_i}{\Delta A_i}\right)\right) \\
&= \sum_{k=1}^{M} \Pr\left(\psi(\mu r_{i-1}) < Y \leq \psi(\mu r_i)\right) \left(\psi\left(\mu \frac{\Delta B_i}{\Delta A_i}\right)\right) \\
&\approx \sum_{k=1}^{M} \Pr\left(\psi(\mu r_{i-1}) < Y \leq \psi(\mu r_i)\right) E[Y \mid \psi(\mu r_{i-1}) < Y \leq \psi(\mu r_i)] \\
&= E[Y] = \nu
\end{aligned}$$

And we have the estimate $\widetilde{B}_M \approx \nu$. The sequence $\{A_i\}$ can be viewed as the cumulative percentage of cases over the intervals of the trended losses and thus approximates the A column of the entry ratio table of the trended losses. The sequence $\left\{\widetilde{B}_i\right\}$ approximates the cumulative losses for the trended loss cases from the corresponding intervals. So the sequence $\left\{\widetilde{B}_i\right\}$ is proportional to the B column of the entry ratio table of the trended losses. Also, we have observed that the sequence $\{\psi(\mu r_i)\}$ provides the endpoints of the corresponding intervals of the trended losses which have overall mean $= \nu \approx \widetilde{B}_M$. So setting:

$$\widehat{r}_i = \frac{\psi(\mu r_i)}{\widetilde{B}_M}, \; \widehat{A}_i = A_i, \; \widehat{B}_i = \frac{\widetilde{B}_i}{\widetilde{B}_M}, \; 0 \leq i \leq M$$

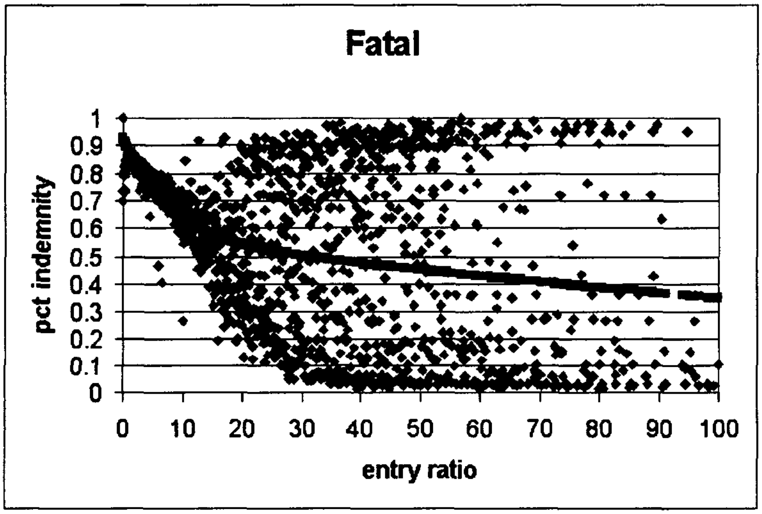
we have approximated the rAB table for the trended losses $rAB_Y \approx \widehat{rAB}$ in the case of differential severity trend. This differential severity trend adjustment to the rAB table is a simple three-step process (1-fix A, 2-estimate B, 3-normalize r and B). In practice, this approximation can yield small negative values for $\delta(r)$ which by Proposition 9 should be set equal to 0.

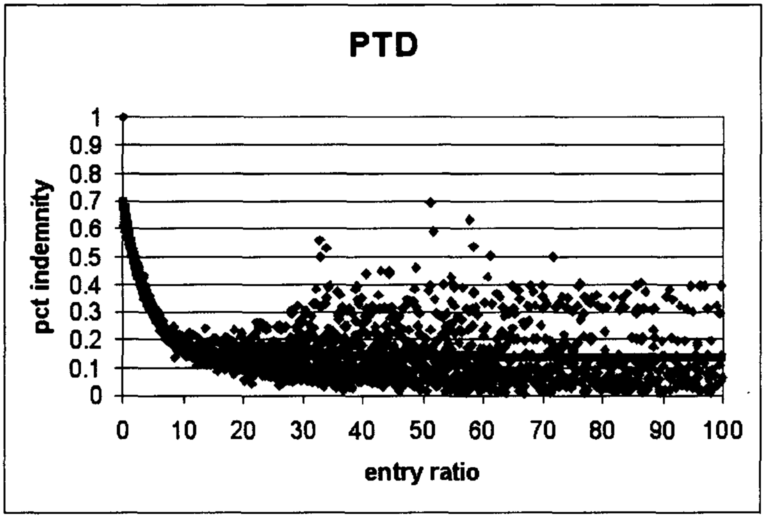
4.1 WC Case Study of Differential Severity Trend

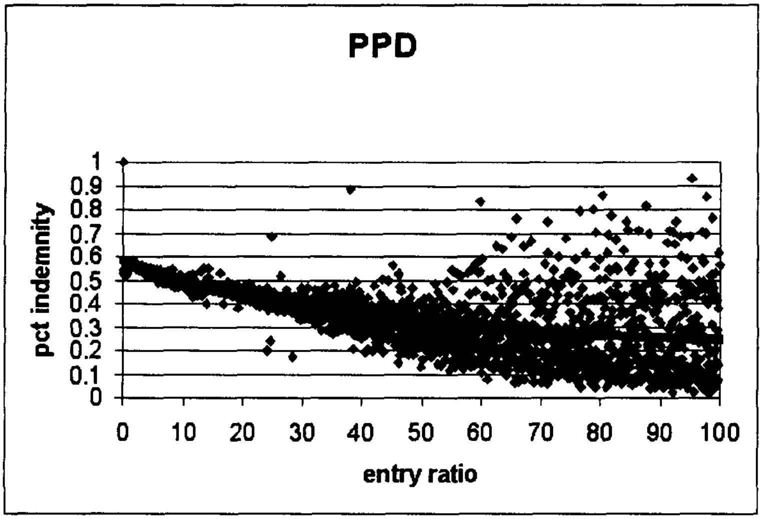
The tables for excess ratios in WC are specific to the five types of WC injury: Fatal, Permanent Total Disability [PTD], Permanent Partial Disability

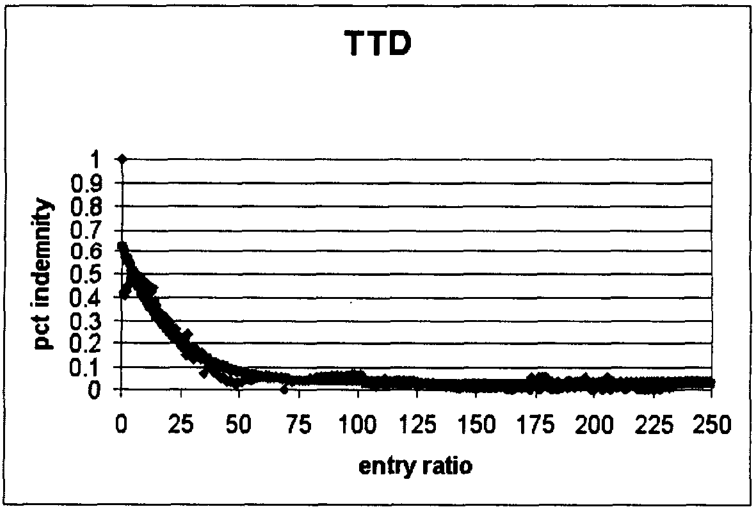
[PPD], Temporary Total Disaility [TTD], and medical only. It is standard to itemize WC losses into medical and indemnity (or wage replacement) components. While indemnity benefits are limited, either implicitly or by statutory maximum aggregates, the medical portion is unlimited and subject to broadly inclusive statutes as regards the medical procedures covered. In any event, it has been noted that as the claim size rises, the percentage of the benefit that goes for medical also rises. This is generally observed within all the injury types (except medical only). A series of charts below provide a more detailed picture of this phenomenon. Combine that observation with the fact that medical losses are subject to greater upward inflationary pressure than wages, and you have a scenario in which to apply the differential severity trend model of the previous section.

In this case study we assume constant annual trend factors of $t_0 = 1.075$ for indemnity and $t_1 = 1.095$ for medical, applicable to all injuries and all loss sizes. Normalized WC loss data by injury type was itemized into medical and indemnity components and used to produce the following charts, by injury type, that show the percentage of the total [=medical + indemnity] loss by entry ratio (the role of the fitted curve will be described later). It is worth noting that the percentages shown in the charts are determined over a common interval width of entry ratio. Since there are typically more claims at lower entry ratios, one consequence is more claims per plotted point at the lower entry ratios, whence the greater spread of the plotted points at the higher entry ratios.


Fatal
pct indemnity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 20 30 40 50 60 70 80 90 100
entry ratio


PTD
pct indemnity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 10 20 30 40 50 60 70 80 90 100
entry ratio





For each injury type = i, a simple curve (akin to a mixed exponential survival curve, and shown on the charts) was fit to the patterns of decreasing indemnity proportion $\pi_i(r)$ by entry ratio r as the loss size increases:

$$\pi_i(r) = a_i \left(b_i e^{\alpha_i r} + c_i e^{\beta_i r} + (1 - b_i - c_i) e^{\gamma_i r} \right)$$

Injury	i	a_i	b_i	c_i	α_i	β_i	γ_i
Fatal	1	0.9280	0.6240	0.3761	−0.0051	−0.1416	−0.4599
PTD	2	0.6928	0.7905	0.2095	−0.2542	−0.0007	−0.4599
PPD	3	0.5811	0.3827	0.6173	0	−0.0281	0
TTD	4	0.6237	0.0397	0.9603	0	−0.0475	−0.4599

We set $h_i(r) = \pi_i(r)t_0 + (1 - \pi_i(r))\,t_1$, then:

$$1 < t_0 < t_1, \frac{d\pi_i}{dr} < 0 \Rightarrow 1 < h_i(r) \text{ and } \frac{dh_i}{dr} = \frac{d\pi_i}{dr}(t_0 - t_1) > 0.$$

and so each injury type other than medical only provides a differential severity trend model.

Letting X_i denote the random variable of losses by injury type and N_{X_i} the corresponding claim counts, the usual formula (readily obtained from Definition 1; see Gillam [1]) for the combined excess ratio over the injury types at attachment A is:

$$XSratio(A) = XSratio(X_1, X_2, X_3, X_4, X_5; A) = \frac{\sum_i N_{X_i}\mu_{X_i}E_{X_i}\left(\frac{A}{\mu_{X_i}}\right)}{\sum_i N_{X_i}\mu_{X_i}}$$

Of course, to accomodate differential severity trend one could produce new rAB tables as detailed above. A simpler alternative is to determine the difference:

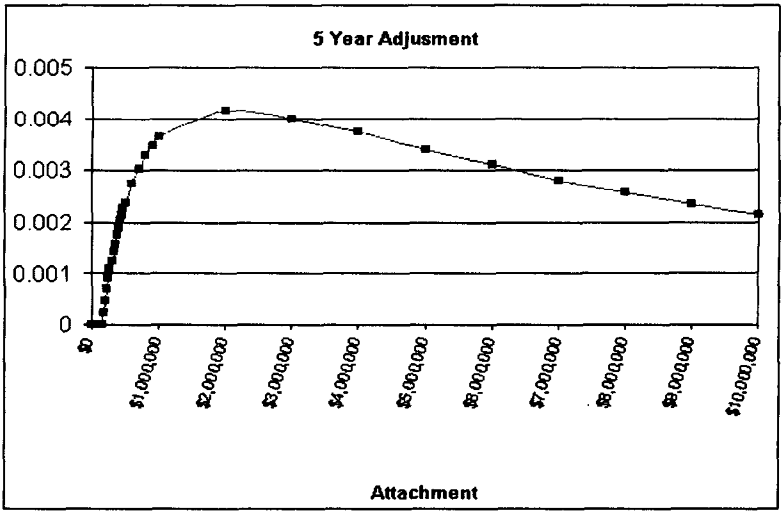
$$\begin{aligned}
& \Delta XSratio(A) \\
= \; & XSratio(Y_1, Y_2, Y_3, Y_4, Y_5; A) - XSratio^*(X_1, X_2, X_3, X_4, X_5; A) \\
= \; & \frac{\sum_i N_{Y_i}\mu_{Y_i}E_{Y_i}\left(\frac{A}{\mu_{Y_i}}\right)}{\sum_i N_{Y_i}\mu_{Y_i}} - \frac{\sum_i N_{Y_i}\mu_{Y_i}E_{X_i}\left(\frac{A}{\mu_{Y_i}}\right)}{\sum_i N_{Y_i}\mu_{Y_i}} \\
= \; & \frac{\sum_i N_{Y_i}\mu_{Y_i}\left(E_{Y_i}\left(\frac{A}{\mu_{Y_i}}\right) - E_{X_i}\left(\frac{A}{\mu_{Y_i}}\right)\right)}{\sum_i N_{Y_i}\mu_{Y_i}} \\
= \; & \frac{\sum_i N_{Y_i}\mu_{Y_i}\delta_{X_iY_i}\left(\frac{A}{\mu_{Y_i}}\right)}{\sum_i N_{Y_i}\mu_{Y_i}}
\end{aligned}$$

expressed in terms of the $\delta_{X_i Y_i}$ and where the $*$ attached to $XSratio^*$ is meant to emphasize that one would consistently use the newer claim counts N_{Y_i} and means μ_{Y_i} in doing the calculation. While in principle you would need updated rAB tables to precisely determine the $\delta_{X_i Y_i}$ terms, if there were a simplified form to approximate that term based on inflation data or other cost trend considerations, this would provide the ability to refine the excess ratio calculation:

$$\begin{aligned} & XSratio(Y_1, Y_2, Y_3, Y_4, Y_5; A) \\ = \; & XSratio^*(X_1, X_2, X_3, X_4, X_5; A) + \Delta XSratio(A) \end{aligned}$$

without recourse to new rAB tables.

The use of entry ratio tables is a very good way to account for inflation when calculating excess ratios. Indeed, even compounded over a five year time interval, the $\Delta XSratio$ adjustment in this case study is very small. The following chart is indicative of what the calculation described here produces. Of course, a bigger difference between medical and indemnity trend or a longer time interval will produce bigger adjustments. Because excess ratios decline with increasing attachment points, as the attachment point increases the adjustment will typically increase as a percentage of the excess ratio.



5 Differential Frequency Trend

Let the function $h(x)$ defined on $[0, \infty)$ be such that $0 \leq h(x) \leq 1$, with h piecewise continuous and non-decreasing on $[0, \infty)$. So as to relate h with the 'untrended' loss variable X, we also assume that there exist $a, b > 0$ such that $h(a) < h(b)$ with h continuous at a and at b and with $f(x) > 0$ for every $x \in (a, b)$. Observe that this clearly forces $a < b$, and so there exist $b_k \in (a, b)$ such that $\lim_{k \to \infty} b_k = b$. But then, since h continuous at b:

$$\begin{aligned} h(a) &< h(b) \Rightarrow h(a) < h(b) = h\left(\lim_{k\to\infty} b_k\right) = \lim_{k\to\infty} h(b_k) \\ &\Rightarrow \text{there exists } M \in \mathbb{N} \text{ such that } h(b_k) > h(a) \text{ for every } k \geq M. \end{aligned}$$

In particular, letting $c = b_M$ we have:

$$\begin{aligned} c &= b_M \in (a, b) \\ &\Rightarrow f(c) > 0, h(c) > h(a) \Rightarrow h(c)f(c) > 0 \Rightarrow 0 < E[h(X)] < E[1] = 1. \end{aligned}$$

We consider the 'trended' loss model defined by the PDF:

$$g(x) = \frac{h(x)f(x)}{E[h(X)]} = \widehat{h}(x)f(x).$$

Think of $h(x)$ as a proportional decline in the incidence rate that decreases with the size of loss x. For the trended loss variable Y, we have:

$$\Pr(Y \leq a) = \int_0^a g(x)dx = \int_0^a \widehat{h}(x)f(x)dx = \Pr(\widehat{h}(X) \leq a).$$

And accordingly, for the differential frequency trend model we take $Y = \widehat{h}(X)$. Also, if h is differentiable (except at perhaps finitely many points), integration by parts gives:

$$\begin{aligned} G(y) &= \int_0^y g(x)dx = \int_0^y \widehat{h}(x)f(x)dx \\ &= \widehat{h}(x)F(x)]_0^y - \int_0^y F(x)\frac{d\widehat{h}}{dx}dx \\ &= \widehat{h}(y)F(y) - \int_0^y F(x)\frac{d\widehat{h}}{dx}dx \\ &\geq \widehat{h}(y)F(y). \end{aligned}$$

For the differential frequency trend model we cannot have $F(x) \geq G(x)$ for all $x \geq 0$, since by the above that would force the contradiction

$$\begin{aligned} G(x) &\geq \widehat{h}(x)F(x) \geq \widehat{h}(x)G(x) \\ \Rightarrow &\ 1 \geq \widehat{h}(x) \text{ for all } x \geq 0 \text{ such that } f(x) > 0 \text{ with } 1 > \widehat{h}(a) \text{ for some } a \geq 0 \text{ such that } f(\\ \Rightarrow &\ 1 < E\left[\widehat{h}(X)\right] = E\left[\frac{h(X)}{E\left[h(X)\right]}\right] = 1 \Rightarrow\Leftarrow \end{aligned}$$

In particular, differential trend models and differential frequency models are disjoint from one another

Remark 10 *The reader should note that unless we make the stronger assumption that h is continuous on $[0, \infty)$, we cannot be assured that this Y is a loss variable, as that term is defined here. The weaker assumption on h is to include the case in which h is a step function. The reader may prefer to demand that h be continuous, in which case some of the arguments can be simplified.*

Proposition 11 *In the case of the differential frequency trend model $g(x) = \widehat{h}(x)f(x)$, as above, $\nu > \mu$.*

Proof. Note that the function $\widehat{h}(x)$ is piecewise continuous and non-decreasing on $[0, \infty)$. We claim that $\widehat{h}(0) < 1$, since otherwise:

$$\widehat{h}(x) \geq 1 \text{ for every } x \geq 0 \Rightarrow g(x) = \widehat{h}(x)f(x) \geq f(x) \text{ for every } x \geq 0.$$

But then $g(x)$ and $f(x)$ are two piecewise continuous funtions on $[0, \infty)$ with the same finite integral $= 1$. So the relation $g(x) \geq f(x)$ entails that $g(x) = f(x)$ except possibly at points of discontinuity of g. So $\widehat{h}(x) = 1$ except for a discrete set of values or where $f(x) = 0$. By our model assumptions, there exist $\alpha, \beta > 0$ such that $h(\alpha) < h(\beta)$ with h continuous at α and at β and with $f(x) > 0$ for every $x \in (\alpha, \beta)$. It follows that $\widehat{h}(x) = 1$ on (α, β), except for perhaps a discrete set of points:

$$\begin{aligned}
&\Rightarrow \text{ there exist } a_i, b_i \in (\alpha, \beta) \text{ such that} \\
\alpha &= \lim_{i\to\infty} a_i, \beta = \lim_{i\to\infty} b_i \text{ and } \widehat{h}(a_i) = \widehat{h}(b_i) = 1 \\
&\Rightarrow \widehat{h}(\alpha) = \widehat{h}\left(\lim_{i\to\infty} a_i\right) = \lim_{i\to\infty} \widehat{h}(a_i) = \lim_{i\to\infty} 1 = 1 \\
&= \lim_{i\to\infty} \widehat{h}(b_i) = \widehat{h}\left(\lim_{i\to\infty} b_i\right) = \widehat{h}(\beta) \\
&\Rightarrow h(\alpha) = \widehat{h}(\alpha)E[h(X)] = \widehat{h}(\beta)E[h(X)] = h(\beta) \\
&\Rightarrow h(\alpha) = h(\beta) > h(\alpha) \quad \Rightarrow\Leftarrow \text{[read "contradiction"]}.
\end{aligned}$$

This contradiction shows that $\widehat{h}(0) < 1$. Similarly, we claim that $\widehat{h}(a) > 1$ for some $a > 0$, since otherwise:

$$\widehat{h}(x) \leq 1 \text{ for all } x \geq 0 \Rightarrow g(x) = \widehat{h}(x)f(x) \leq f(x) \text{ for all } x \geq 0$$

and again $g(x)$ and $f(x)$ are two piecewise continuous funtions on $[0, \infty)$ with the same finite integral. This again entails that they are equal except possibly at points of discontinuity. Then again $\widehat{h}(x) = 1$ except for a discrete set of values or where $f(x) = 0$ and just as before we arrive at a contradiction. So we have

$$\begin{aligned}
\widehat{h}(0) &< 1 < \widehat{h}(a) \\
&\Rightarrow \text{ there exists } b > 0 \text{ such that } \widehat{h}(x) \leq 1 \text{ on } [0, b) \\
\text{and } \widehat{h}(x) &\geq 1 \text{ on } (b, \infty).
\end{aligned}$$

Next we claim that there exists $c > 0$ such that $\widehat{h}(c) \neq 1$ and $f(c) > 0$ since otherwise

$$x > 0, f(x) > 0 \Rightarrow \widehat{h}(x) = 1 \Rightarrow h(x) = E[h(X)]$$

But by our model assumptions, there exist $\alpha, \beta > 0$ such that $h(\alpha) < h(\beta)$

with h continuous at α and at β and with $f(x) > 0$ for every $x \in (\alpha, \beta)$:

$$\begin{aligned} & \Rightarrow \text{ there exists } c \in \left(\alpha, \frac{\alpha+\beta}{2}\right), d \in \left(\frac{\alpha+\beta}{2}, \beta\right) \\ \text{such that } h(c) & \neq h(d),\ f(c) > 0, f(d) > 0 \\ & \Rightarrow E[h(X)] = h(c) \neq h(d) = E[h(X)]. \end{aligned}$$

It follows that there exists $c > 0$ such that $\widehat{h}(c) \neq 1$ and $f(c) > 0$ and we have:

$$\begin{aligned} \nu - \mu &= \int_0^\infty xg(x)dx - \int_0^\infty xf(x)dx = \int_0^\infty x\left(g(x) - f(x)\right)dx \\ &= \int_0^\infty x\left(\widehat{h}(x) - 1\right)f(x)dx \\ &= \int_0^b x\left(\widehat{h}(x) - 1\right)f(x)dx + \int_b^\infty x\left(\widehat{h}(x) - 1\right)f(x)dx \end{aligned}$$

$$\begin{aligned} &> b\int_0^b \left(\widehat{h}(x) - 1\right)f(x)dx + b\int_b^\infty \left(\widehat{h}(x) - 1\right)f(x)dx \\ &= b\int_0^\infty \left(\widehat{h}(x) - 1\right)f(x)dx = b\int_0^\infty \left(g(x) - f(x)\right)dx \\ &= b\left(\int_0^\infty g(x)dx - \int_0^\infty f(x)dx\right) = b(1-1) = 0 \\ &\Rightarrow \nu > \mu \end{aligned}$$

as required. ■

As in the case of differential severity trend in the preceding section, we again are considering a change that increases the mean severity. Suppose we use a fixed entry ratio table to calculate excess ratios. Then for a fixed attachment point A, we have declining entry ratios $\frac{A}{\mu} > \frac{A}{\nu}$ and the lookup into the same entry ratio table leads to excess ratios that increase from $E_X\left(\frac{A}{\mu}\right)$

to $E_X\left(\frac{A}{\nu}\right)$. In the case of differential severity trend, we observed in Proposition 9 of the previous section that the increase is consistently understated. In the case of differential frequency trend, however, we will show that the increase may be either overstated or understated. This may at first seem somewhat counterintuitive for the two "trends" to move the mean upward but the normalized excess ratio tabular amounts in perhaps opposite directions. However, the entry ratio lookup is dominated by the change in the mean. For differential severity trend the overall trend factor consistently understates the impact of trend on the largest loss amounts, which helps explain why the calculation consistently understates the excess ratio. But the case of differential frequency trend is quite different: selectively removing smaller sized losses will have a leveraged upward impact on the overall mean severity while leaving the size of the largest claims unchanged.

With differential frequency trend we have, from the proof of Proposition 11:

$$\begin{aligned} x &\geq b \Rightarrow \nu R_Y(x) - \mu R_X(x) = \int_x^\infty (y-x)\left(g(y) - f(y)\right) dy \\ &= \int_x^\infty (y-x)\left(\widehat{h}(y) - 1\right) f(y) dy \geq 0 \\ &\Rightarrow R_Y(x) \geq \frac{\mu}{\nu} R_X(x). \end{aligned}$$

But again the relationship between $E_Y(r)$ and $E_X(r)$ is more subtle.

In the case that X has a maximum loss $M = M_X < \infty$, since $\widehat{h}(x)$ is non-decreasing on $[0,\infty)$ and there exists $c > 0$ such that $\widehat{h}(c) > 0$ and $f(c) > 0$, and we have $c \leq M$ and $\widehat{h}(d) > 0$ for every $d \geq c$, whence:

$$M_Y = \sup\{x|g(x) > 0\} = \sup\{x|\widehat{h}(x)f(x) > 0\} = \sup\{x|f(x) > 0\} = M.$$

So too must Y have maximum loss M and Proposition 8 assures us that $E_Y(r) \leq E_X(r)$ for large enough r. More precisely, we have:

Proposition 12 *In the case of the differential frequency trend model $g(x) = \widehat{h}(x)f(x)$, as defined above, in which X has a maximum loss $M_X < \infty$, there exists $b > 0$ such that $E_Y(b) < E_X(b)$ and $E_Y(r) \leq E_X(r)$ for all $r \geq b$.*

Before stating a result that deals with the relationship between $E_Y(r)$ and $E_X(r)$ in the case $M_X = \infty$, it is instructive to make a few observations.

Note that since the non-decreasing function h is bounded above by 1, it is reasonable (but not necessary) to have the decline in frequency flatten out for large losses, say in the sense that the derivative $\frac{dh}{dx} \to 0$ as $x \to \infty$. We also observe that:

Proposition 13 *In the case of the differential frequency trend model* $g(x) = \widehat{h}(x)f(x)$, *as above, the limit* $\lim_{x\to\infty} \widehat{h}(x) = \lambda$ *exists and* $\frac{\nu}{\mu} \leq \lambda$.

Proof. Since h is non decreasing and bounded above by 1, existence of the limit is apparent. We evidently have:

$$\widehat{h}(x) \leq \lambda \text{ for all } x \geq 0 \Rightarrow \nu = E[X\widehat{h}(X)] \leq E[X\lambda] = \lambda E[X] = \lambda\mu \Rightarrow \frac{\nu}{\mu} \leq \lambda$$

as required. ■

Proposition 14 *Assume* $M_X = \infty$, *then for any* $\rho > 1$ *for which the limit* $\lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ *exists:*

$$\lim_{x\to\infty} \frac{S(\rho x)}{S(x)} \leq \frac{1}{\rho}.$$

Proof. Note that $M_X = \infty$ is equivalent to $S(x) > 0$ for every $x > 0$ and so $\frac{S(\rho x)}{S(x)}$ is always well defined. Thus the expression $\lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ makes sense and further our assumption is that the limit exists for some $\rho > 1$. Note that the integral $\int_0^\infty S(x)dx = \mu < \infty$. Suppose, by way of contradiction, that $\lim_{x\to\infty} \frac{S(\rho x)}{S(x)} > \frac{1}{\rho}$. Then, using the change of variable $z = \rho x$, we would have:

$$\begin{aligned}
\text{there exists } c &> 0 \text{ such that } \rho S(\rho x) > S(x) \text{ for every } x > c \\
&\Rightarrow \int_c^\infty \rho S(\rho x)dx > \int_c^\infty S(x)dx \\
&\Rightarrow \frac{1}{\rho}\int_c^\infty S(x)dx < \int_c^\infty S(\rho x)dx \\
\int_c^\infty S(\rho x)dx &= \frac{1}{\rho}\int_c^\infty S(\rho x)\rho dx \\
&= \frac{1}{\rho}\int_{\rho c}^\infty S(z)dz \\
\rho &> 1 \Rightarrow \frac{1}{\rho}\int_{\rho c}^\infty S(z)dz < \frac{1}{\rho}\int_c^\infty S(z)dz \\
&\Rightarrow \frac{1}{\rho}\int_c^\infty S(x)dx < \frac{1}{\rho}\int_c^\infty S(z)dz \Rightarrow\Leftarrow
\end{aligned}$$

This contradiction completes the proof. ■

Remark 15 *Appendix A considers the implications of the existence of the limit* $\lim_{x\to\infty}\frac{S(\rho x)}{S(x)}$. *The discussion shows that if you assume that the limit* $\lim_{x\to\infty}\frac{S(\rho x)}{S(x)}$ *exists for all* $\rho > 1$ *and is not identically* 0 *for all* $\rho > 1$, *then the tail behavior is essentially determined up to just a single parameter. More precisely, consider the one-parameter survival function:*

$$T(\beta;x) = \left\{ \begin{array}{cc} 1 & x \leq 1 \\ x^{-\beta} & x > 1 \end{array} \right\}.$$

For $T(\beta;x)$ *such limits exist and are particularly manageable as we clearly have*

$$\rho,\beta,x \geq 1 \Rightarrow \frac{T(\beta;\rho x)}{T(\beta;x)} = \frac{(\rho x)^{-\beta}}{x^{-\beta}} = \rho^{-\beta} = \lim_{y\to\infty}\frac{T(\beta;\rho y)}{T(\beta;y)}.$$

It turns out that for a loss variable X *with* $S = S_X$ *and for which there exist* $\rho_k > 1, k \in \mathbb{N}$ *such that* $\lim_{k\to\infty}\rho_k = 1$ *and* $\lim_{x\to\infty}\frac{S(\rho_k x)}{S(x)}$ *exists for every* $k \in \mathbb{N}$, *then for all* $\rho > 1$:

$$\begin{aligned} \textit{either } \lim_{x\to\infty}\frac{S(\rho x)}{S(x)} &= 0 \\ \textit{or } \lim_{x\to\infty}\frac{S(\rho x)}{S(x)} &= \rho^{-\beta} \textit{ where } \beta = -\ln\left(\lim_{x\to\infty}\frac{S(ex)}{S(x)}\right) \geq 1. \end{aligned}$$

We see that under these assumptions, the conditional probability of survival $\frac{S(y)}{S(x)}$ *for* $y > x$ *and* x *large is asymptotically the same as that of* $T(\beta;x)$ *for some unique* β, *with* $1 \leq \beta \leq \infty$.

Example 16 *For the "thin-tailed" exponential density* $S(x) = e^{-\frac{x}{\theta}}$ *we have, for any constant* $\rho > 1$, *that*

$$\lim_{x\to\infty}\frac{S(\rho x)}{S(x)} = \lim_{x\to\infty}\frac{e^{-\frac{\rho x}{\theta}}}{e^{-\frac{x}{\theta}}} = \lim_{x\to\infty} e^{-\frac{(\rho-1)x}{\theta}} = 0.$$

Example 17 *For the "thicker tailed" Pareto density* $S(x) = \left(\frac{\theta}{\theta+x}\right)^{\alpha}$ *we have, for any constant* $\rho > 1$, *that*

$$\lim_{x\to\infty}\frac{S(\rho x)}{S(x)} = \lim_{x\to\infty}\frac{\left(\frac{\theta}{\theta+\rho x}\right)^{\alpha}}{\left(\frac{\theta}{\theta+x}\right)^{\alpha}} = \lim_{x\to\infty}\left(\frac{\theta+x}{\theta+\rho x}\right)^{\alpha} = \rho^{-\alpha}.$$

Example 18 *This example shows that the inequality in Proposition 14 cannot, in general, be improved. Consider the survival function:*

$$
\begin{aligned}
S(x) &= \frac{e}{(x+e)\left(\ln(x+e)\right)^2} \\
\mu &= \int_0^\infty S(x)dx = \int_e^\infty \frac{e}{u\left(\ln(u)\right)^2}du \text{ where } u = x+e \\
&= e\int_1^\infty \frac{1}{w^2}dw \text{ where } w = \ln(u) \\
&= e\left[-\frac{1}{w}\right]_1^\infty = e < \infty
\end{aligned}
$$

with finite mean. We have, with several applications of L'Hôpital:

$$
\begin{aligned}
\lim_{x\to\infty}\frac{S(\rho x)}{S(x)} &= \lim_{x\to\infty}\left(\frac{e}{(\rho x+e)\left(\ln(\rho x+e)\right)^2}\frac{(x+e)\left(\ln(x+e)\right)^2}{e}\right) \\
&= \lim_{x\to\infty}\left(\left(\frac{x+e}{\rho x+e}\right)\frac{\left(\ln(x+e)\right)^2}{\left(\ln(\rho x+e)\right)^2}\right) \\
&= \lim_{x\to\infty}\left(\frac{x+e}{\rho x+e}\right)\lim_{x\to\infty}\left(\frac{\ln(x+e)}{\ln(\rho x+e)}\right)^2 \\
&= \frac{1}{\rho}\left(\lim_{x\to\infty}\frac{\ln(x+e)}{\ln(\rho x+e)}\right)^2 \\
&= \frac{1}{\rho}\left(\lim_{x\to\infty}\frac{\frac{1}{x+e}}{\frac{\rho}{\rho x+e}}\right)^2 \\
&= \frac{1}{\rho}\left(\lim_{x\to\infty}\frac{\rho x+e}{\rho x+\rho e}\right)^2 \\
&= \frac{1}{\rho}.
\end{aligned}
$$

Example 19 *Define the function:*

$$h(x) = \left\{ \begin{array}{cc} x & 0 \le x < 1 \\ 2x - 1 & 1 \le x < 2 \\ 3 & 2 \le x < 4 \\ 1 + \frac{x}{2} & 4 \le x < 8 \\ \vdots & \vdots \\ k + 1 & 2^{k-1} \le x < 2^k \text{ and } k > 1 \text{ even} \\ k - 2 + \frac{x}{2^{k-2}} & 2^{k-1} \le x < 2^k \text{ and } k > 1 \text{ odd} \end{array} \right\}$$

then the reader can readily verify that h is continuous and non-decreasing with $h(0) = 0$ and $\lim_{x\to\infty} h(x) = \infty$. It follows that $S(x) = e^{-h(x)}$ is a survival function. Let X be a nonnegative random variable with $S = S_X$. The reader can verify the following:

$$\begin{aligned} h(4x) &= h(x) + 2 \text{ for } x > 2 \\ h(2^k) &= \left\{ \begin{array}{cc} k+2 & k > 1 \text{ odd} \\ k+1 & k > 1 \text{ even} \end{array} \right\} \end{aligned}$$

And we find that for $x > 2$:

$$\begin{aligned} \frac{S(4x)}{S(x)} &= \frac{e^{-h(4x)}}{e^{-h(x)}} = e^{h(x)-h(4x)} = e^{-2} \\ &\Rightarrow \lambda(4) = \lim_{x\to\infty} \frac{S(4x)}{S(x)} = \frac{1}{e^2}. \end{aligned}$$

Since $\lambda(4) = \frac{1}{e^2} < \frac{1}{4}$ it is at least possible for this distribution to have a finite mean; and indeed, the reader can readily verify that:

$$\begin{aligned} x &> 2 \Rightarrow h(x) \ge \frac{\ln x}{\ln 2} - 1 \\ &\Rightarrow S(x) \le e x^{-\frac{1}{\ln 2}} \\ &\Rightarrow \int_2^\infty S(x)dx \le \int_2^\infty e x^{-\frac{1}{\ln 2}} dx = e \left(\frac{2^{1-\frac{1}{\ln 2}}}{\frac{1}{\ln 2} - 1} \right) < \infty \\ &\Rightarrow \mu_X = \int_0^\infty S(x)dx < \infty \end{aligned}$$

and we see that X is a loss variable. Observe that:

$$\begin{aligned} h(2^k) &= \left\{ \begin{array}{cc} h(2^{k-1})+2 & k>1 \text{ odd} \\ h(2^{k-1}) & k>1 \text{ even} \end{array} \right\} \\ &\Rightarrow \frac{S(2\cdot 2^{k-1})}{S(2^{k-1})} = \frac{S(2^k)}{S(2^{k-1})} = \frac{e^{-h(2^k)}}{e^{-h(2^{k-1})}} \\ &= e^{h(2^{k-1})-h(2^k)} = \left\{ \begin{array}{cc} e^{-2} & k>1 \text{ odd} \\ 1 & k>1 \text{ even} \end{array} \right\} \\ &\Rightarrow \lim_{x\to\infty} \frac{S(2x)}{S(x)} \text{ fails to exist.} \end{aligned}$$

Finally, observe that should $\lim_{x\to\infty}\frac{S(4x)}{S(x)}$ exist, that is not sufficient to guarantee that $\lim_{x\to\infty}\frac{S(\beta x)}{S(x)}$ exists for $\beta > 4$. Indeed, setting $x_k = \frac{2^k}{5}$ we have:

$$\begin{aligned} h(5x_k) - h(x_k) &= \left\{ \begin{array}{cc} \frac{14}{5} & k>3 \text{ odd} \\ 2 & k>3 \text{ even} \end{array} \right\} \\ &\Rightarrow \frac{S(5x_k)}{S(x_k)} = e^{-h(5x_k)+h(x_k)} = \left\{ \begin{array}{cc} e^{-\frac{14}{5}} & k>3 \text{ odd} \\ e^{-2} & k>3 \text{ even} \end{array} \right\} \\ &\Rightarrow \lim_{x\to\infty} \frac{S(5x)}{S(x)} \text{ fails to exist.} \end{aligned}$$

This example is meant to provide some additional insight into the nature of the assumption made in the very special case considered in the above remark, namely that $\lim_{x\to\infty}\frac{S(\rho x)}{S(x)}$ exists for all $\rho > 1$.

The two limits $\lim_{x\to\infty}\widehat{h}(x) \geq \frac{\nu}{\mu} \geq 1$ (Propositions 11 and 13) and $\lim_{x\to\infty}\frac{S(\frac{\nu}{\mu}x)}{S(x)} = \lim_{x\to\infty}\frac{S(\nu x)}{S(\mu x)} \leq \frac{\mu}{\nu} \leq 1$ play a key role in determining the sign of $\delta(r)$ for large enough entry ratio r, as demonstrated in the following:

Proposition 20 *In the case of the differential frequency trend model $g(x) = \widehat{h}(x)f(x)$, as defined above, assume that $M_X = \infty$, that h is differentiable on $(0,\infty)$ (except at perhaps finitely many points) and that there exists $c > 0$ with $\frac{dh}{dx} = 0$ for all $x \geq c$. Let $\rho = \frac{\nu}{\mu}$ and assume that the limit $\lambda = \lim_{x\to\infty}\frac{S(\rho x)}{S(x)}$ exists. Then*

$$\begin{aligned} \lambda\widehat{h}(c) &> 1 \Rightarrow \text{there exists } b>0 \text{ such that } E_Y(r) > E_X(r) \text{ for all } r \geq b \\ \lambda\widehat{h}(c) &< 1 \Rightarrow \text{there exists } b>0 \text{ such that } E_Y(r) < E_X(r) \text{ for all } r \geq b. \end{aligned}$$

Proof. To compare $E_Y(r)$ and $E_X(r)$ for large entry ratios, we again investigate the derivative of $\delta(r)$:

$$\begin{aligned}\frac{d\delta}{dr} &= G(\nu r) - F(\mu r)\\ &= \int_0^{\nu r} g(x)dx - \int_0^{\mu r} f(x)dx\\ &= \int_{\mu r}^{\nu r} \widehat{h}(x)f(x)dx + \int_0^{\mu r}\left(\widehat{h}(x)-1\right)f(x)dx\end{aligned}$$

Observe that the first integral is always ≥ 0 and converges to 0 as $r \to \infty$ and that the second integral is an increasing function of r for r large enough to force $\widehat{h}(\mu r) > 1$ and the second integral also converges to 0 as $r \to \infty$. Let $r > \frac{c}{\mu}$, our assumptions together with $\frac{dh}{dx} \geq 0$, give us:

$$\begin{aligned}G(\nu r) &= \widehat{h}(\nu r)F(\nu r) - \int_0^{\nu r} F(x)\frac{d\widehat{h}}{dx}dx\\ &= \widehat{h}(c)F(\nu r) - \int_0^{c} F(x)\frac{d\widehat{h}}{dx}dx\\ &= \widehat{h}(c)F(\nu r) - \gamma \text{ for some constant } \gamma \geq 0.\end{aligned}$$

Taking the limit as $r \to \infty$:

$$\begin{aligned}1 &= \widehat{h}(c) - \gamma\\ 1-\widehat{h}(c) &= -\gamma\\ G(\nu r) &= \widehat{h}(c)F(\nu r) - \gamma\\ &= \widehat{h}(c)F(\nu r) + 1 - \widehat{h}(c)\\ &= -\widehat{h}(c)\left(1-F(\nu r)\right) + 1\\ &\Rightarrow \frac{d\delta}{dr} = -\widehat{h}(c)\left(1-F(\nu r)\right) + 1 - F(\mu r)\\ &= -\widehat{h}(c)S(\nu r) + S(\mu r).\end{aligned}$$

Now suppose $\lambda\widehat{h}(c) > 1$:

$$\begin{aligned}
\lambda\widehat{h}(c) &> 1 \Rightarrow \lim_{x\to\infty}\frac{\widehat{h}(c)S(\rho x)}{S(x)} > 1 \\
\text{there exists } b &> 0 \text{ such that } \widehat{h}(c)S(\rho x) > S(x) \text{ for every } x \geq \mu b \\
\rho &= \frac{\nu}{\mu} \Rightarrow \widehat{h}(c)S(\nu r) > S(\mu r) \text{ for every } \mu r \geq \mu b \\
&\Rightarrow \widehat{h}(c)S(\nu r) > S(\mu r) \text{ for every } r \geq b \\
&\Rightarrow -\widehat{h}(c)S(\nu r) < -S(\mu r) \text{ for every } r \geq b \\
&\Rightarrow \frac{d\delta}{dr} = -\widehat{h}(c)S(\nu r) + S(\mu r) < 0 \text{ for every } r \geq b.
\end{aligned}$$

And it follows that $\delta(r)$ is decreasing for $r \geq b$. Since $\delta(r) \to 0$ as $r \to \infty$ it follows that $E_Y(r) - E_X(r) = \delta(r) > 0$ for $r \geq b$. We have established:

$$\begin{aligned}
\lambda\widehat{h}(c) &> 1 \Rightarrow \text{there exists } b > 0 \text{ such that} \\
E_Y(r) - E_X(r) &= \delta(r) > 0 \Rightarrow E_Y(r) > E_X(r) \text{ for all } r \geq b.
\end{aligned}$$

Reversing inequalities in the above argument shows:

$$\begin{aligned}
\lambda\widehat{h}(c) &< 1 \Rightarrow \text{there exists } b > 0 \text{ such that} \\
E_Y(r) - E_X(r) &= \delta(r) < 0 \Rightarrow E_Y(r) < E_X(r) \text{ for all } r \geq b
\end{aligned}$$

completing the proof. ■

An immediate consequence is that distributions with an infinite but comparatively thin tail act like distributions with finite support:

Corollary 21 *In the case of the differential frequency trend model $g(x) = \widehat{h}(x)f(x)$, as defined above, assume that $M_X = \infty$, that h is differentiable on $(0,\infty)$ (except at perhaps finitely many points), that there exists $c > 0$ with $\frac{dh}{dx} = 0$ for $x \geq c$, and further that $\lim_{x\to\infty}\frac{S((\frac{\nu}{\mu})x)}{S(x)} = 0$. Then there exists $b > 0$ such that $E_Y(r) < E_X(r)$ for all $r \geq b$.*

Example 22 *As a general example of a differential freqency trend model we may take $h = F$, then $g(x) = aF(x)f(x)$ for a uniquely determined constant a. But clearly F^2 is itself a distribution function and setting:*

$$\begin{aligned}
G &= F^2 \\
&\Rightarrow \frac{dG}{dx} = 2F\frac{dF}{dx} = 2Ff \text{ is a PDF} \Rightarrow a = 2
\end{aligned}$$

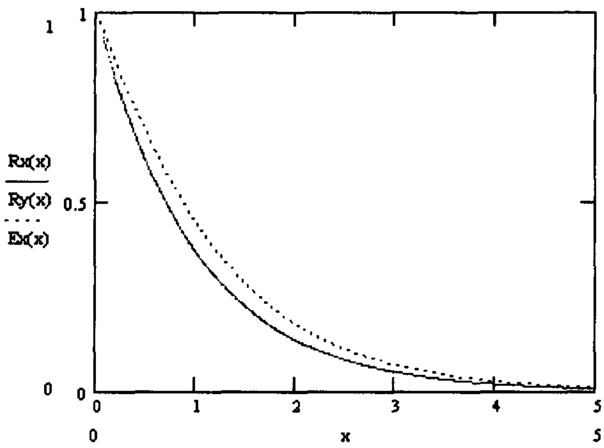
and the increase in the mean is:

$$\begin{aligned}
\nu - \mu &= \int_0^\infty (1 - G(y))dy - \int_0^\infty S(y)dy \\
&= \int_0^\infty (1 - F(y)^2)dy - \int_0^\infty S(y)dy \\
&= \int_0^\infty (1 - F(y))(1 + F(y))dy - \int_0^\infty S(y)dy \\
&= \int_0^\infty S(y)(1 + F(y))dy - \int_0^\infty S(y)dy \\
&= \int_0^\infty S(y)(1 + F(y) - 1)dy \\
&= \int_0^\infty S(y)F(y)dy.
\end{aligned}$$

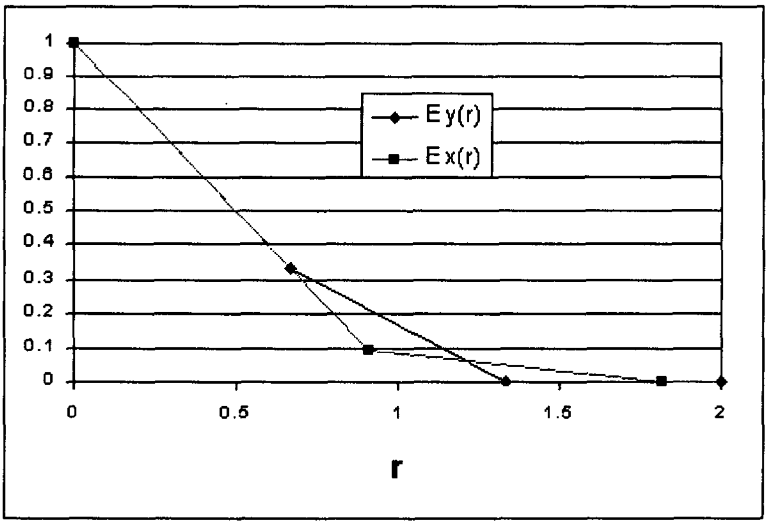
Example 23 *Let X be an exponential denisty with $f(x) = e^{-x}$ and set $h(x) = \frac{x}{x+1}$. Then from numerical integration applied directly to the definitions:*

$$\begin{aligned}
E[h(X)] &= 0.404 \\
\mu &= 1 \\
\nu &= 1.477
\end{aligned}$$

The following graphs the excess ratio functions $R_X(x) = E_X(x), R_Y(x)$, and $E_Y(x)$; from the graph we see that: $E_Y(x) < E_X(x) = R_X(x) < R_Y(x)$.



Example 24 *Consider the case of* 10 *losses per year:* 9 *of amount* 1 *and* 1 *of amount* 2 *and let* X *denote the corresponding random variable. Suppose there is a decline in frequency to a rate of just* 2 *losses per year:* 1 *of amount* 1 *and* 1 *of amount* 2 *with random variable* Y. *The following graphs the excess ratio functions* $E_X(r)$, *and* $E_Y(r)$. *In this case we see that* $E_Y(1) > E_X(1)$ *and* $E_Y(1.5) < E_X(1.5) > 0$.



Example 25 *Consider a Pareto density with survival function* $S(x) = \left(\frac{\theta}{x+\theta}\right)^{\alpha}$ *and a linear frequency decline of the form* $h(x) = Min\left(\frac{x+d}{c+d}, 1\right)$. *We provide the results of a direct evaluation via numerical methods for two cases:*

$$\begin{aligned}
\theta &= 2, \alpha = 5, c = 2, d = 1 \\
\mu &= 0.5, \nu \approx 0.695, \rho = \frac{\nu}{\mu} \approx 1.39 \\
\frac{1}{\lambda} &= \lim_{x\to\infty} \frac{S(x)}{S(\rho x)} \approx 5.16 > 2.04 \approx \widehat{h}(c) \\
E_Y(x) &< E_X(x)
\end{aligned}$$

and:

$$\begin{aligned}
\theta &= 2, \alpha = 5, c = 10, d = 5 \\
\mu &= 0.5, \nu \approx 0.575, \rho = \frac{\nu}{\mu} \approx 1.149 \\
\frac{1}{\lambda} &= \lim_{x\to\infty} \frac{S(x)}{S(\rho x)} \approx 1.978 < 2.728 \approx \widehat{h}(c) \\
E_Y(x) &> E_X(x).
\end{aligned}$$

In both cases, Proposition 17 holds for any $b > 0$. *This gives an instance for which the same untrended loss variable and two functions for h, both of linear frequency decline proportions with the same range of* $\left[\frac{1}{3}, 1\right]$, *can produce opposite sign impacts on the normalized excess ratio function.*

As to the rAB table for this differential frequency trend model, as before let

$$r_0 = 0 < r_1 < r_2 < \cdots < r_M$$

be a sequence of entry ratios and set

$$A_i = A_X(r_i), B_i = B_X(r_i), 0 \le i \le M.$$

Suppose that $A_i = A_X(r_i) > A_X(r_{i-1}), 1 \le i \le M$ and $A_M = 1$. Set $\Delta A_i = A_i - A_{i-1}, \Delta B_i = B_i - B_{i-1}$. Again note that $\mu\frac{\Delta B_i}{\Delta A_i}, 1 \le i \le M$, is the mean value of the untrended loss over the interval $[\mu r_{i-1}, \mu r_i]$, which we assume can be taken as an estimate for the mean of the trended loss. This would hold provided that, within sufficiently narrow entry ratio layers, the removed claims (and whence the retained) are representative of all claims in that layer. This would hold exactly, for example, in case the function h is a step function that is constant on the intervals $[r_{i-1}, r_i)$. For $1 \leq i \leq M$, set

$$\begin{aligned}
\widetilde{A}_i &= \widehat{h}(\mu r_i)A_i,\ \Delta\widetilde{A}_i = \widetilde{A}_i - \widetilde{A}_{i-1} \\
\Delta\widetilde{B}_i &= \Delta\widetilde{A}_i\left(\mu\frac{\Delta B_i}{\Delta A_i}\right) \\
\widetilde{B}_i &= \sum_{k=1}^{i}\Delta\widetilde{B}_k.
\end{aligned}$$

Assuming then that $\mu\frac{\Delta B_i}{\Delta A_i}$ is a good estimate of the mean value of the trended losses on the interval $[\mu r_{i-1}, \mu r_i]$, we have:

$$\begin{aligned}
\frac{\widetilde{B}_M}{\widetilde{A}_M} &= \sum_{k=1}^{M}\frac{\Delta\widetilde{B}_k}{\widetilde{A}_M} = \sum_{k=1}^{M}\frac{\Delta\widetilde{A}_k}{\widetilde{A}_M}\left(\mu\frac{\Delta B_k}{\Delta A_k}\right) \\
&\approx \sum_{k=1}^{M}\Pr\left(\mu r_{k-1} < Y \leq \mu r_k\right)\left(\mu\frac{\Delta B_k}{\Delta A_k}\right) \\
&\approx \sum_{k=1}^{M}\Pr\left(\mu r_{k-1} < Y \leq \mu r_k\right)E[Y\ |\mu r_{k-1} < Y \leq \mu r_k] \\
&= E[Y] = \nu
\end{aligned}$$

and we infer, as before, that $\nu \approx \frac{\widetilde{B}_M}{\widetilde{A}_M}$ and that the two sequences $\left\{\widetilde{A}_i\right\}$ and $\left\{\widetilde{B}_i\right\}$ are nearly equal to the cumulative cases and losses of not necessarily normalized trended losses. So they only need to be rescaled to give the A and B columns of the trended losses. Whence they are very nearly proportional to the A and B columns of the entry ratio for the trended losses (and albeit with different proportionality constants). So setting:

$$\widehat{r}_i = \frac{\mu r_i}{\frac{\widetilde{B}_M}{\widetilde{A}_M}} = \frac{\mu r_i\widetilde{A}_M}{\widetilde{B}_M},\ \widehat{A}_i = \frac{\widetilde{A}_i}{\widetilde{A}_M},\ \widehat{B}_i = \frac{\widetilde{B}_i}{\widetilde{B}_M},\ 0 \leq i \leq M$$

we have approximated the rAB table for the trended losses: $rAB_Y \approx \widehat{rAB}$. Finally, note that this simple three-step differential frequency trend adjustment to the rAB table (adjust A, estimate B, renormalize r, A, and B) can be done quite generally to account for a change in frequency by size of loss and does not formally demand that $\frac{dh}{dx} > 0$ on $(0, \infty)$, although order preserving is needed to justify the calculation.

5.1 WC Case Study of Differential Frequency

The tables for excess ratios in WC are produced by five types of WC injury: Fatal, Permanent Total Disability [PTD], Permanent Partial Disability [PPD], Temporary Total Disaility [TTD], and medical only [MO]. The WC system in the US has seen a persistent decline in claim frequency over the past 10-15 years. The decline is observed within each of the injury types and over the spectrum of US industries. There is no consensus on how long this pattern can persist, or even on its underlying causes. One pattern that has emerged, both in NCCI investigations as well as from studies by the Department of Labor, is that this decline has not been uniform by size of loss. Small WC claims have declined proportionally more than have large WC claims. That is the motivation for this look at how differential frequency trend impacts entry ratio tables.

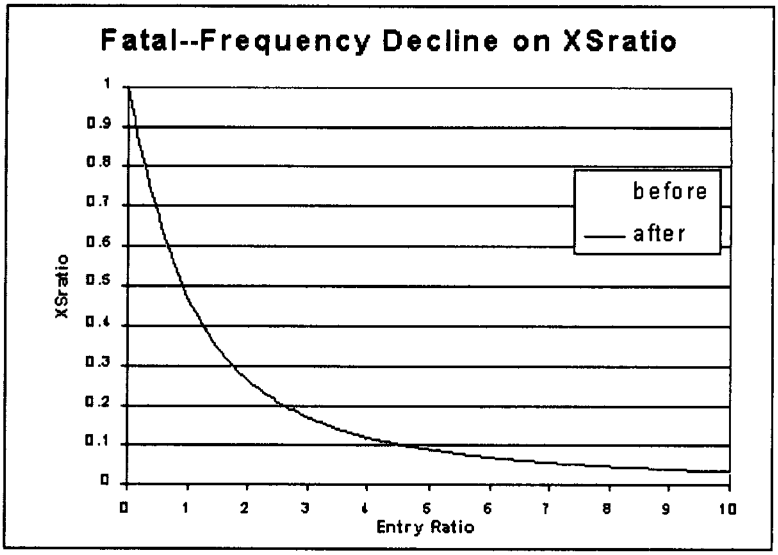
A recent NCCI study produced the following table of percentage changes in claim frequency (per unit of wage-adjusted payroll exposure):

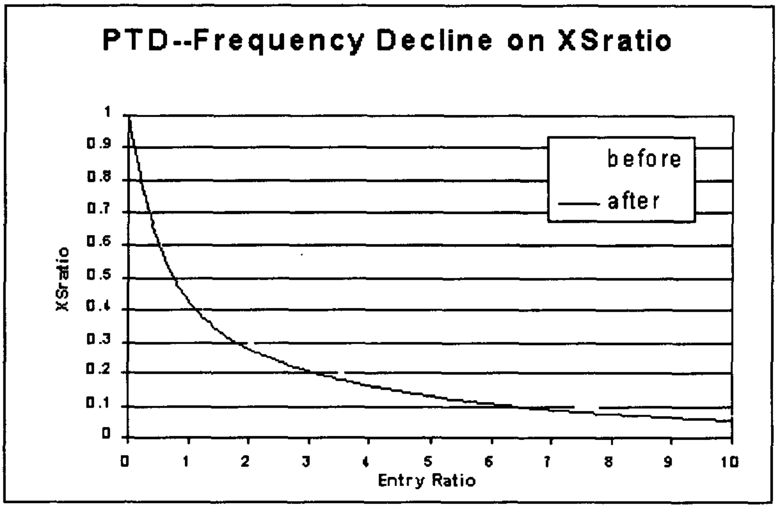
	Fatal	PTD	PPD	TTD	MO
Smallest third of claims	−6.2%	−52.4%	−23.7%	−32.8%	−26.7%
Middle third of claims	−7.9%	−18.5%	−12.8%	−20.4%	−29.9%
Largest third of claims	−10.3%	4.3%	−8.7%	−8.5%	−13.8%

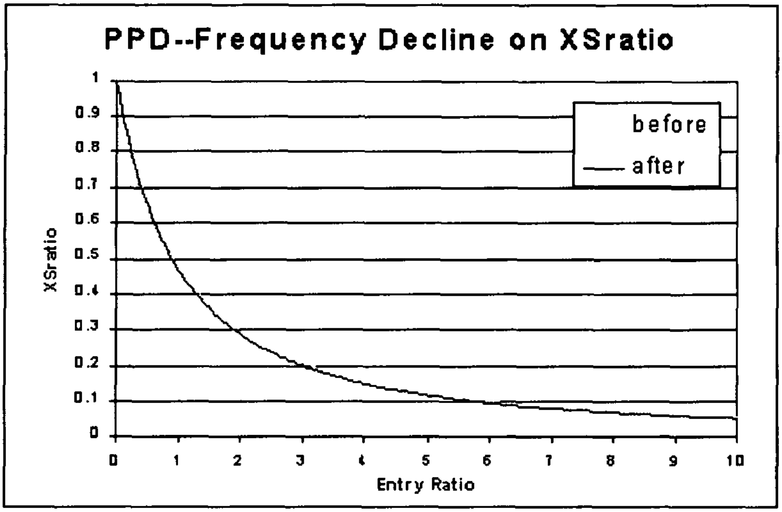
With the exceptions of the fatal and medical only injury types, the table conforms to the by now familiar pattern of a smaller decline in frequency with increasing claim size. These percent changes were used to define a proportional change in frequency function $h_i(r)$ as a step function of entry ratio r for each injury type i. Even a smoothed version of $h_i(r)$ would not likely conform to the differential frequency trend model assumptions for injury types Fatal $[i = 1]$ and Medical Only$[i = 5]$:

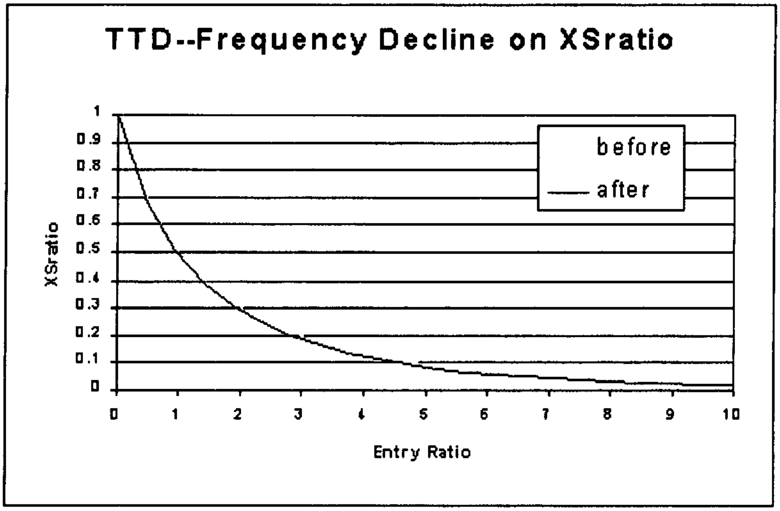
Range of r	$h_1(r)$	$h_2(r)$	$h_3(r)$	$h_4(r)$	$h_5(r)$
$0 \le A(r) < \frac{1}{3}$	0.9382	0.476	0.7628	0.6718	0.7329
$\frac{1}{3} \le A(r) < \frac{2}{3}$	0.9211	0.8151	0.8723	0.7957	0.7014
$\frac{2}{3} \le A(r) \le 1$	0.8967	1.043	0.9134	0.9151	0.8624

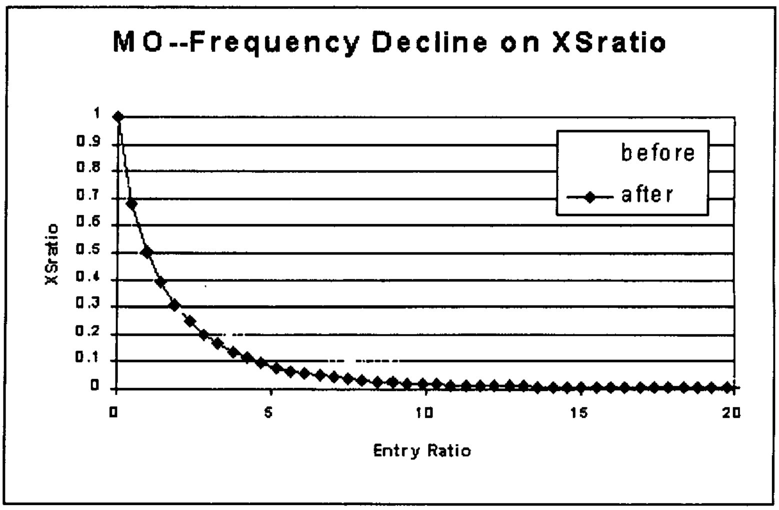
Even though the assumptions of the differential frequency trend model are technically not met in this case study, the discussion still makes it clear how to determine, for each injury type, a trended entry ratio table from the untrended table. The graphs below show the excess ratio functions $E_{X_i}(r)$, and $E_{Y_i}(r)$ by injury type i before (X_i) and after (Y_i) trend. With the exceptions of the fatal and medical only injury types, we again see that $E_Y(r) - E_X(r) \leq 0$. For each injury type except perhaps medical only, the two curves are very close, which indicates that little or no frequency trend adjustment to the rAB table is indicated.




PTD--Frequency Decline on XSratio
XSratio
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
7
8
9
10
Entry Ratio
before
after


PPD--Frequency Decline on XSratio
XSratio
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
7
8
9
10
Entry Ratio
before
after





As in the earlier case study, it is straightforward to combine differential frequency trend impacts by injury type into a combined impact on the normalized excess ratio.

6 Matching the Mean and Median Loss

Suppose we are presented with an entry ratio table rAB_X together with some constant $\varepsilon \neq 0$, we next discuss how to build the entry ratio table rAB_{X^ε}. Here we consider the trended random variable to be $Y = \psi(X) = X^\varepsilon$ where the transformation $\psi(x) = x^\varepsilon$ has $\frac{d\psi}{dx} = \varepsilon x^{\varepsilon-1}$ and is order preserving for $\varepsilon > 0$ and order reversing for $\varepsilon < 0$. Thus, as we did for differential severity trend, we have:

$$\begin{aligned} G(\psi(x)) &= \Pr(Y \le \psi(x)) \\ &= \Pr(\psi(X) \le \psi(x)) = \left\{ \begin{array}{l} \Pr(X \le x) = F(x) \Leftrightarrow \varepsilon > 0 \\ \Pr(X \ge x) = S(x) \Leftrightarrow \varepsilon < 0 \end{array} \right\}. \end{aligned}$$

Let $r_0 = 0 < r_1 < r_2 < \cdots < r_M$ be a sequence of entry ratios and set

$$A_i = A_X(r_i), B_i = B_X(r_i), 0 \le i \le M.$$

As before, suppose that $A_i = A_X(r_i) > A_X(r_{i-1}), 1 \le i \le M$ and $A_M = 1$. Set $\Delta A_i = A_i - A_{i-1}, \Delta B_i = B_i - B_{i-1}$. Note that $\mu\frac{\Delta B_i}{\Delta A_i}, 1 \le i \le M$, is the mean value of the untrended loss over the interval $[\mu r_{i-1}, \mu r_i]$. For $1 \le i \le M$, set

$$\begin{aligned} \Delta\widetilde{B}_i &= \Delta A_i \left(\psi\left(\mu\frac{\Delta B_i}{\Delta A_i}\right)\right) = \Delta A_i \left(\mu\frac{\Delta B_i}{\Delta A_i}\right)^\varepsilon \\ \widetilde{B}_i &= \sum_{k=1}^{i} \Delta\widetilde{B}_k. \end{aligned}$$

Assuming, as usual, that $\left(\mu\frac{\Delta B_i}{\Delta A_i}\right)^\varepsilon$ is a good estimate of the mean value of the trended losses within the interval $\left[\mu^\varepsilon r_{i-1}^\varepsilon, \mu^\varepsilon r_i^\varepsilon\right]$ leads to the familiar estimate $\nu \approx \widetilde{B}_M$ and, as before, the two sequences $\{A_i\}$ and $\left\{\widetilde{B}_i\right\}$ approximate the cumulative claim and loss percentages of the trended losses. A change of scale to normalize the trended losses corresponds to adjusting the two sequences $\{A_i\}$ and $\left\{\widetilde{B}_i\right\}$ by constant factors. So the sequences are very nearly proportional to the A and B columns of the entry ratio for the trended losses. Setting:

$$\widehat{r}_i = \frac{\mu^\varepsilon r_i^\varepsilon}{\widetilde{B}_M},\ \widehat{A}_i = A_i, \text{ and } \widehat{B}_i = \frac{\widetilde{B}_i}{\widetilde{B}_M},\ 0 \le i \le M$$

we have approximated the rAB table for the trended losses: $rAB_Y = rAB_{X^\varepsilon} \approx \widehat{rAB}$.

Now abstract from this and suppose only that you are provided an entry ratio table Θ in the form of three finite increasing sequences of M numbers:

$$\begin{aligned} r_0 &= 0 < r_1 < r_2 < \cdots < r_M \\ A_0 &= 0 < A_1 < A_2 < \cdots < A_M = 1 \\ B_0 &= 0 < B_1 < B_2 < \cdots < B_M = 1 \end{aligned}$$

We will assume that these table values were constructed using some loss variable X and so Θ at least conforms to the properties of an entry ratio table. Given $\varepsilon > 0$ we can formally construct a new entry ratio by mimicking the above and assuming, with no loss of generality, that $\mu_X = 1$. For $1 \le i \le M$, set $\Delta A_i = A_i - A_{i-1}$ and $\Delta B_i = B_i - B_{i-1}$ and define

$$\begin{aligned} \Delta\widetilde{B}_i &= \Delta A_i \left(\frac{\Delta B_i}{\Delta A_i}\right)^{\varepsilon} \\ \widetilde{B}_i &= \sum_{k=1}^{i} \Delta\widetilde{B}_k. \end{aligned}$$

And construct a new table $\widehat{\Theta}$ from the increasing sequences:

$$\widehat{r}_i = \frac{r_i^{\varepsilon}}{\widetilde{B}_M}, \ \widehat{A}_i = A_i, \text{ and } \widehat{B}_i = \frac{\widetilde{B}_i}{\widetilde{B}_M}, \ 0 \le i \le M.$$

The significance of this construction for adapting entry ratio tables to changing conditions will become clear from the following:

Proposition 26 *Let $1 \le x_1 < x_2 < \cdots < x_M$ be an increasing sequence of $M > 1$ numbers. Then for any fixed number w with $0 < w < 1$ and integer k, $1 \le k < M$, there exist $\alpha, \beta > 0$ such that setting $y_i = \alpha {x_i}^{\beta}$ we have:*

$$\frac{1}{M}\sum_{i=1}^{M} y_i = 1 \quad \textit{and} \quad y_k = w$$

Proof. Let $z_i = \frac{x_i}{x_k}, 1 \le i \le M$ and define $\varphi(v) = \frac{1}{M}\sum\limits_{i=1}^{M} z_i^v$ then φ is a continuous function of v and invoking the Intermediate Value Theorem[IVT]:

$$\begin{aligned}\varphi(0) &= 1 \\ z_M &> 1 \Rightarrow \lim_{v\to\infty}\varphi(v) = \infty \\ 1 &< \frac{1}{w} < \infty, \text{IVT} \Rightarrow \text{there exists } \beta > 0 \text{ such that } \varphi(\beta) = \frac{1}{w}\end{aligned}$$

Now set $\alpha = \frac{w}{x_k^\beta}$, then we have:

$$\begin{aligned} y_i &= \alpha {x_i}^\beta = \frac{w}{x_k^\beta}{x_i}^\beta = w z_i^\beta, 1 \le i \le M \\ &\Rightarrow \frac{1}{M}\sum_{i=1}^{M} y_i = \frac{w}{M}\sum_{i=1}^{M} z_i^\beta \\ &= w\varphi(\beta) = \frac{w}{w} = 1 \;\text{ and }\; y_k = w z_k^\beta = w 1^\beta = w \end{aligned}$$

completing the proof. ■

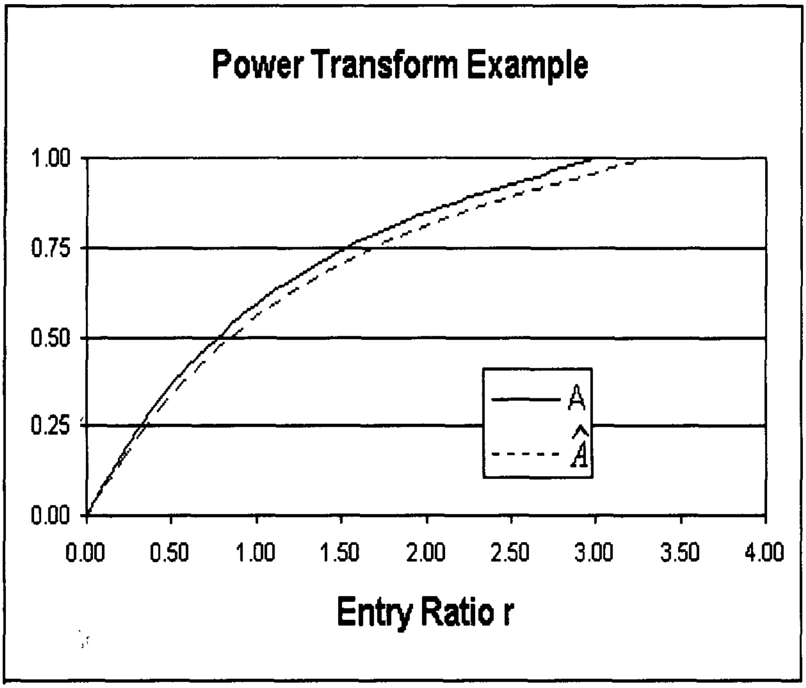
This means that, quite generally, for discrete loss data the power transform $Y = \alpha X^\beta$ enables us not only to normalize to mean 1 but also to simultaneously specify the entry ratio $w(= r)$ of any selected percentile $\frac{k}{M}(= A(r))$. As a very general example, suppose you are provided an rAB table and some loss data with random variable X. Suppose further that you observe a median $= m$ and mean $= \mu$, so the observed entry ratio of the median $= \frac{m}{\mu}$. Now suppose further that in the given rAB table you observe that $A(\frac{m}{\mu})$ is well removed from $\frac{1}{2}$. This suggests to you that the given rAB table may not be suited to the task, say, of looking up excess ratios $R_X(x)$ for the given loss data. Now assume that the given entry ratio table rAB has $A(w) = \frac{1}{2}$ for some $w < 1$—this is not unreasonable since loss distributions typically have median less than the mean. From Proposition 23, there is a power transform $Y = \alpha X^\beta$ whose median has entry ratio equal to w. But this, in turn, suggests that the given entry ratio table rAB may be suitable as an entry ratio table for the transformed losses $Y = \alpha X^\beta$, i.e. $rAB \approx rAB_Y$, inasmuch as the transformed losses have the ratio of mean to median implicit in the table. While a power transform may not be the exact relationship for how losses trend, it is reasonable to assume some structural relationship between the given rAB table and the given losses. By Proposition 5, $R_X(x) = R_{Y^{\frac{1}{\beta}}}(\alpha^{\frac{1}{\beta}} x)$ and we find that all we require to customize the table lookup of excess ratios is an entry ratio table for $Y^{\frac{1}{\beta}}$. But the above discussion provides an algorithm for determining the entry ratio table of a power transform. So let $\widehat{rAB}$ be determined, as above, from the original rAB table under the power transformation $\varepsilon = \frac{1}{\beta}$, then $\widehat{rAB} \approx rAB_{Y^\varepsilon} = rAB_{Y^{\frac{1}{\beta}}}$. This enables us to look up the excess ratio $R_{Y^{\frac{1}{\beta}}}(\alpha^{\frac{1}{\beta}}x)$. Finally, note that all this simplifies to the usual process of looking up the entry ratio of the loss limit, but in the adjusted entry ratio table:

$$\begin{aligned} Y &= \alpha X^\beta \Rightarrow Y^{\frac{1}{\beta}} = \alpha^{\frac{1}{\beta}} X \Rightarrow \mu_{Y^{\frac{1}{\beta}}} = \alpha^{\frac{1}{\beta}} \mu_X \\ &\Rightarrow R_X(x) = R_{Y^{\frac{1}{\beta}}}(\alpha^{\frac{1}{\beta}}x) = E_{Y^{\frac{1}{\beta}}}\left(\frac{\alpha^{\frac{1}{\beta}}x}{\mu_{Y^{\frac{1}{\beta}}}}\right) \\ &= E_{Y^{\frac{1}{\beta}}}\left(\frac{\alpha^{\frac{1}{\beta}}x}{\alpha^{\frac{1}{\beta}}\mu_X}\right) = E_{Y^{\frac{1}{\beta}}}\left(\frac{x}{\mu_X}\right) \\ &\Rightarrow R_X(x) \approx \widehat{E}\left(\frac{x}{\mu_X}\right) \end{aligned}$$

So to summarize, this example illustrates a general technique to deal with the case in which "trend" has impacted the shape of the severity distribution as evidenced by a change in the relationship between the mean and the median loss. In fact, the discussion details how to "trend" the old entry ratio table, rAB, to a new table $\widehat{rAB}$.

The challenge with this approach comes in finding α and β. At first, it would seem to require a calculation involving the complete loss variable X, or at least a very robust and representative claim subsample. And such calculation (the proof of Propostion 23 coupled with a binary search algorithm might prove useful), if doable at all, would suggest that direct calculation of the excess ratio, or even an entirely new rAB table, may be more practical. However, notice that only β is required to construct $\widehat{rAB}$ from rAB and it is a straightforward spreadsheet application to try different values for β until the resulting $\widehat{rAB}$ satisfies $\widehat{A}(w) = \frac{1}{2}$. This approach may well provide a β that works even when $w \geq 1$ and the technique can be applied equally well to other percentiles than the median. Consequently, the technique is both general and constructive.

Example 27 *This example considers an entry ratio table rAB (columns r,A,B) that reflects a loss distribution for which the median is about $\frac{4}{5}^{th}$ of* *the mean. Assume that later data revealed that the entry ratio of the median loss had grown from* 0.8 *to* 0.85. *A power transform with* $\beta = 2$ *is illustrated. Appendix B includes the table and displays a trended entry ratio table* $\widehat{rAB}$ *(columns* $\widehat{r}$, $\widehat{A}$, $\widehat{B}$*) which may better fit the newer data. The following chart shows the corresponding change in the normalized cumulative distribution function, from* $A \rightarrow \widehat{A}$*:*



We just saw how a calculation similar to that of the differential severity trend approach can adapt the rAB table to a power transform $Y = \alpha X^{\beta}$. We conclude this section by describing how the set up of the differential frequency trend calculation can adapt the rAB table to a proportional hazard transform $S_Y = (S_X)^{\alpha}$. In the notation used for differential frequency trend, we have:

$$g(x) = -\frac{dS_Y}{dx} = \alpha S_X(x)^{\alpha-1} f(x) \Rightarrow \widehat{h}(x) = \alpha S_X(x)^{\alpha-1}.$$

Now abstract from this as above and suppose again that you are provided an entry ratio table Θ in the form of three finite increasing sequences of M

numbers:

$$\begin{aligned} r_0 &= 0 < r_1 < r_2 < \cdots < r_M \\ A_0 &= 0 < A_1 < A_2 < \cdots < A_M = 1 \\ B_0 &= 0 < B_1 < B_2 < \cdots < B_M = 1. \end{aligned}$$

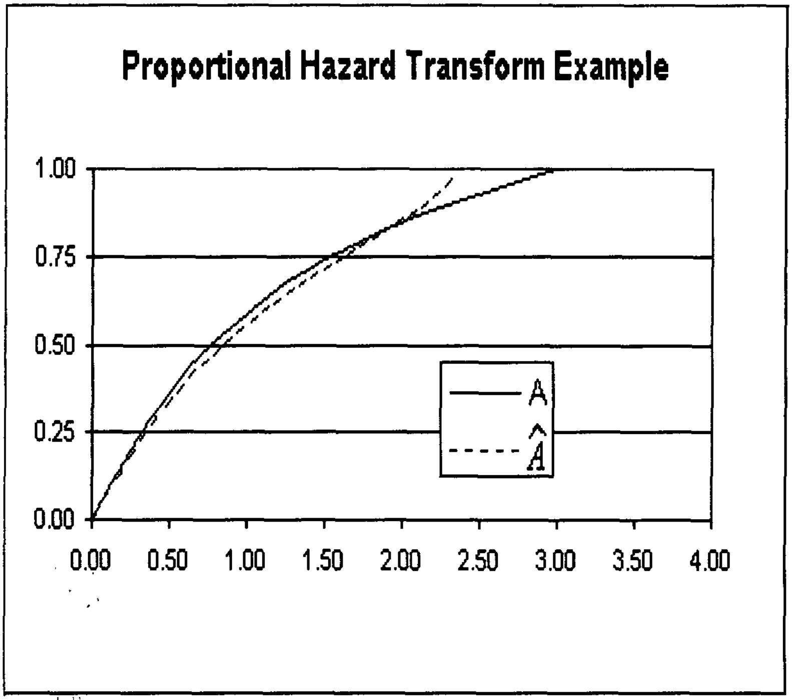
Given $\alpha > 0$ we can formally construct a new entry ratio table by employing the three-step process for the frequency differential trend, again assuming for convenience and with no loss of generality that the mean of the loss variable of the given table is 1. Set $\Delta A_i = A_i - A_{i-1}, \Delta B_i = B_i - B_{i-1}$ and define

$$\begin{aligned} \widetilde{A}_i &= 1 - (1 - A_i)^{\alpha}, \Delta\widetilde{A}_i = \widetilde{A}_i - \widetilde{A}_{i-1},\ 0 \le i \le M \\ \Delta\widetilde{B}_i &= \Delta\widetilde{A}_i \left(\frac{\Delta B_i}{\Delta A_i}\right) \\ \widetilde{B}_i &= \sum_{k=1}^{i} \Delta\widetilde{B}_k, 1 \le i \le M \end{aligned}$$

From which we construct a new table $\widehat{\Theta}$ from the increasing sequences:

$$\widehat{r}_i = \frac{r_i \widetilde{A}_M}{\widetilde{B}_M},\ \widehat{A}_i = \frac{\widetilde{A}_i}{\widetilde{A}_M},\ \widehat{B}_i = \frac{\widetilde{B}_i}{\widetilde{B}_M},\ 0 \le i \le M$$

Example 28 *This example begins with the same entry ratio table rAB as the previous example. A proportional hazard transform $\alpha = \frac{5}{7}$, selected to again adjust the median to an entry ratio of about 0.85–the table is included in Appendix.B. The following chart shows the corresponding change in the normalized cumulative distribution function, from $A \to \widehat{A}$:*



The two examples illustrate the rather different ways in which the power transform (which bears a formal similarity with the differential severity trend set up) and the proportional hazard transform (which bears a formal similarity with the differential frequency trend set up) achieve raising the relativity of the median to the mean loss. The power transform disproportionately increases the larger losses, including increasing the maximum loss amount from 3 to around 3.3, so that proportionally fewer losses above 0.8 are needed for an overall mean = 1. By contrast, the proportional hazard adjustment removes the largest losses, including dropping the maximum loss amount from 3 to about 2.3, forcing the smaller losses to increase in order to maintain an overall mean = 1. Accordingly, it is advisable to consider the impact of trend on the largest losses when selecting a trend adjustment to update an entry ratio table.

It is also worth comparing what the WC case studies suggest in regard to the justification for trending an entry ratio table. Medical inflation has outstripped overall wage growth very consistently and the reasons why are well understood. Also, WC medical coverage is not subject to the statutory limitations imposed on wage-replacement benefits. Finally, in the case of excess ratios, the direction of the change in the tablular values is consistent and readily explained. So in the case of differential severity trend, there is a strong argument to be made that the underlying dynamics are persistent.

The case of differential frequency trend provides a contrast. The decline in WC claim frequency, while persistent over the past decade, is neither readily explained nor well understood. Experts disagree on whether the decline will, or even can, continue. While no one is surprised that medical inflation outstrips wage growth, the observation that the WC frequency decline is greater for smaller claims is a fairly new and a largely unforseen observation. In the case of excess ratios and differential frequency trend, the direction of the change in the tablular values is neither consistent nor straightforward. While the dymanics of differential severity trend are extemely unlikely to reverse, that cannot be said for differential frequency trend.

As with any trend adjustment, there is the concern that missing turning points will result in trend adjustments leading to worse estimates rather than better estimates. This is especially so when the direction of the numerical change is itself problematic. In the case of entry ratio tables, there is a built in correction for short term changes in severity that works very well. And so any "trend" adjustment must be justified over a long time window as improving the estimate. This study suggests that while a fairly strong argument can be made for incorporating the differential severity trend adjustment to WC entry ratio tables, the case is much weaker for differential frequency trend.

7 Conclusion

In the case of a differential severity trend in which large losses trend upward faster (slower) than do smaller losses, the use of an entry ratio table assumes an average trend which corresponds with a severity distribution whose tail is not thickening (thinning) in response to the non-uniform trend. Ideally, the normalized excess ratios from the rAB table should be increased (decreased) to offset this.

In the case of a differential frequency trend in which the frequency of small losses declines faster (slower) than for large losses, the impact of the frequency decline on the mean severity is leveraged. Over the range of attachment points, the use of an untrended entry ratio table may sometimes overstate or sometimes understate the change in the excess ratio.

The two models described here, the differential severity trend and differential frequency trend scenarios, are meant to act independently of one another. Differential severity trend assumes that all trend is due to inflationary movement and none is due to a change in claim emergence. Differential frequency trend holds loss amounts fixed while applying a proportional change in the density. Therefore, it is perhaps not too surprising that while both act to increase the mean severity, they can impact the normalized excess ratio in opposite directions and may offset one another when updating an entry ratio table.

Another very general technique that can be used to accomodate a non-uniform trend is to use a power transformation or a proportional hazard transformation, in lieu of just dividing by the mean loss when performing the lookup into the entry ratio table. The technique provides another way to trend an entry ratio table. More precisely, the ratio between the mean loss and a fixed percentile loss may be observed to change over time. And this calculation gives a way to periodically modify the entry ratio table to accomodate that movement.

References

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

[2] Lee, Y. S., "The Mathematics of Excess of Loss Coverages and Retrospective Rating: A Graphical Approach," *PCAS* LXXV, 1988, pp. 49-64.

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

APPENDIX A

In this appendix we invoke the notation and assumptions of the Differential Frequency Trend (section 4) of the main paper and let X be a loss variable with survival function $S(x)$ for which $M_X = \infty$. We consider the implications of the assumption that the limit $\lambda(\rho) = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists for all $\rho > 1$. Proposition 14 of the paper gives:

Proposition 29 *Let X be a loss variable with $M_X = \infty$ and $S = S_X$, then for any $\rho > 1$ for which the limit $\lambda(\rho) = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists:*

$$\lambda(\rho) \leq \frac{1}{\rho} < 1.$$

Note that when the limit $\lambda(\rho) = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists:

$$\begin{aligned}
\lim_{x\to\infty} \frac{S(\rho^2 x)}{S(x)} &= \lim_{x\to\infty} \frac{S(\rho^2 x)}{S(\rho x)} \frac{S(\rho x)}{S(x)} \\
&= \lim_{x\to\infty} \frac{S(\rho(\rho x))}{S(\rho x)} \lim_{x\to\infty} \frac{S(\rho x)}{S(x)} \\
&= \left(\lim_{x\to\infty} \frac{S(\rho x)}{S(x)} \right)^2 \\
&\Rightarrow \lambda(\rho^2) = \lambda(\rho)^2.
\end{aligned}$$

More generally, we have:

Proposition 30 *Let X be a loss variable with $M_X = \infty$ and $S = S_X$, then for any $m \in \mathbb{N}$, if the limit $\lambda(\rho) = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists then*

$$\lambda(\rho^m) = \lim_{x\to\infty} \frac{S(\rho^m x)}{S(x)} = \lambda(\rho)^m.$$

Proof. The verification is a straightforward induction, the result has been observed to hold for $m = 1, 2$, we have:

$$\begin{aligned}
\lim_{x\to\infty} \frac{S(\rho^{m+1} x)}{S(x)} &= \lim_{x\to\infty} \frac{S(\rho^{m+1} x)}{S(\rho^m x)} \frac{S(\rho^m x)}{S(x)} \\
&= \lim_{x\to\infty} \frac{S(\rho(\rho^m x))}{S(\rho^m x)} \lim_{x\to\infty} \frac{S(\rho^m x)}{S(x)} \\
&= \lambda(\rho)\lambda(\rho^m) \\
&= \lambda(\rho)\lambda(\rho)^m = \lambda(\rho)^{m+1}
\end{aligned}$$

completing the induction and the proof. ■

When such limits all exist, this generalizes to:

Proposition 31 *Let X be a loss variable with $M_X = \infty$ and $S = S_X$, and assume that the limit $\lambda(\rho) = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists for all $\rho > 1$. Then*

$$\lambda(\rho)^{\omega} = \lambda(\rho^{\omega}) \text{ for any positive real number } \omega.$$

Proof. Observe that since the limit $\lambda(\rho^{\frac{1}{n}}) = \lim_{x\to\infty} \frac{S(\rho^{\frac{1}{n}} x)}{S(x)}$ is assumed to exist, we must have:

$$\begin{aligned}
\frac{S(\rho x)}{S(x)} &= \frac{S\left(\rho^{\frac{1}{n}}\left(\rho^{\frac{n-1}{n}}x\right)\right)}{S\left(\rho^{\frac{n-1}{n}}x\right)} \frac{S\left(\rho^{\frac{1}{n}}\left(\rho^{\frac{n-2}{n}}x\right)\right)}{S\left(\rho^{\frac{n-2}{n}}x\right)} \cdots \frac{S\left(\rho^{\frac{1}{n}}\left(\rho^{\frac{n-n}{n}}x\right)\right)}{S\left(\rho^{\frac{n-n}{n}}x\right)} \\
&\Rightarrow \lim_{x\to\infty}\left(\frac{S\left(\rho^{\frac{1}{n}}x\right)}{S(x)}\right)^{n} = \lim_{x\to\infty}\frac{S(\rho x)}{S(x)} = \lambda(\rho) \\
&\Rightarrow \lim_{x\to\infty}\frac{S\left(\rho^{\frac{1}{n}}x\right)}{S(x)} = \lambda(\rho)^{\frac{1}{n}}.
\end{aligned}$$

But then for any positive integers m, n we have:

$$\lambda(\rho^{\frac{m}{n}}) = \lim_{x\to\infty}\frac{S(\rho^{\frac{m}{n}}x)}{S(x)} = \lim_{x\to\infty}\frac{S(\left(\rho^{\frac{1}{n}}\right)^{m}x)}{S(x)} = \left(\lambda(\rho)^{\frac{1}{n}}\right)^{m} = \lambda(\rho)^{\frac{m}{n}}.$$

Whence $\lambda(\rho^{a}) = \lambda(\rho)^{a}$ for any positive rational a. Now let ω be a positive real, then there are sequences of positive rationals:

$$\begin{aligned}
a_k, b_k &\in \mathbb{Q}, k \in \mathbb{N} \\
\text{such that } 0 &< a_1, a_k \le a_{k+1}, b_k \ge b_{k+1} \\
\text{and with } \lim_{k\to\infty} a_k &= \lim_{k\to\infty} b_k = \omega.
\end{aligned}$$

This clearly forces $a_k \le \omega \le b_k$ and since S is a continuous, non-increasing function, we have:

$$\begin{aligned}
a_k &\leq \omega \leq b_k \Rightarrow \rho^{a_k} \leq \rho^{\omega} \leq \rho^{b_k} \\
&\Rightarrow \rho^{a_k}x \leq \rho^{\omega}x \leq \rho^{b_k}x \text{ for all } x > 0 \\
&\Rightarrow S\left(\rho^{a_k}x\right) \geq S\left(\rho^{\omega}x\right) \geq S\left(\rho^{b_k}x\right) \text{ for all } x > 0 \\
&\Rightarrow \frac{S\left(\rho^{a_k}x\right)}{S(x)} \geq \frac{S\left(\rho^{\omega}x\right)}{S(x)} \geq \frac{S\left(\rho^{b_k}x\right)}{S(x)} \text{ for all } x > 0 \\
&\Rightarrow \lim_{x\to\infty} \frac{S\left(\rho^{a_k}x\right)}{S(x)} \geq \lim_{x\to\infty} \frac{S\left(\rho^{\omega}x\right)}{S(x)} \geq \lim_{x\to\infty} \frac{S\left(\rho^{b_k}x\right)}{S(x)} \\
&\Rightarrow \lambda(\rho)^{a_k} = \lambda(\rho^{a_k}) \geq \lambda(\rho^{\omega}) \geq \lambda(\rho^{b_k}) = \lambda(\rho)^{b_k} \\
\lambda(\rho)^{\omega} &= \lambda(\rho)^{\lim_{k\to\infty} a_k} \geq \lambda(\rho^{\omega}) \geq \lambda(\rho)^{\lim_{k\to\infty} b_k} = \lambda(\rho)^{\omega} \\
&\Rightarrow \lambda(\rho)^{\omega} = \lambda(\rho^{\omega})
\end{aligned}$$

and we see that $\lambda(\rho)^{\omega} = \lambda(\rho^{\omega})$ for any positive real number ω, completing the proof. ■

An immediate consequence is:

Corollary 32 *Let X be a loss variable with $M_X = \infty$ and $S = S_X$, and assume that the limit $\lambda(\rho) = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists for all $\rho > 1$ Then*

1. there exists $\rho > 1$ such that $\lambda(\rho) = 0 \Leftrightarrow \lambda(\rho) = 0$ for every $\rho > 1$
2. there exists $\rho > 1$ such that $\lambda(\rho) \neq 0 \Leftrightarrow \lambda(\rho) \neq 0$ for every $\rho > 1$.

Consider the one-parameter survival function:

$$\begin{aligned}
T(x) &= T(\beta; x) = \left\{ \begin{array}{cc} 1 & x \leq 1 \\ x^{-\beta} & x > 1 \end{array} \right\} \\
\rho, \beta, x &\geq 1 \Rightarrow \frac{T(\rho x)}{T(x)} = \frac{(\rho x)^{-\beta}}{x^{-\beta}} = \rho^{-\beta} = \lim_{y\to\infty} \frac{T(\rho y)}{T(y)}
\end{aligned}$$

Note that $T(\beta; x)$ has a finite mean if and only if $\beta > 1$. By convention, we include the (discontinuous) possibility that $\beta = \infty$ by setting $T(\infty; x) = x^{-\infty} = 0$ for $x > 1$.

Proposition 33 *Let X be a loss variable with $M_X = \infty$ and $S = S_X$ and assume that the limit $\lambda(\rho) = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists for all $\rho > 1$. Then*

$$\begin{aligned}
\lambda(\rho) &= \lim_{x\to\infty} \frac{S(\rho x)}{S(x)} = \rho^{-\beta} \\
\textit{for all } \rho &> 1, \textit{ where } \beta = -\ln(\lambda(e)) \geq 1.
\end{aligned}$$

Proof. Consider first the case when there is some $\rho_0 > 1$ such that $\lambda(\rho_0) \neq 0$. Then from Proposition 29 and Corollary 32 we find that $\lambda(e) \in (0, 1)$. Then for any real $\rho > 1$ we have:

$$\begin{aligned} \lambda(\rho) &= \lambda(e^{\ln \rho}) = \lambda(e)^{\ln \rho} = \left(e^{-\beta}\right)^{\ln \rho} = \left(e^{\ln \rho}\right)^{-\beta} = \rho^{-\beta} \\ \text{where } \lambda(e) &= e^{-\beta} \Leftrightarrow \beta = -\ln(\lambda(e)) \end{aligned}$$

and since by Proposition 29:

$$\lambda(e) \leq \frac{1}{e} \Rightarrow e \leq \frac{1}{\lambda(e)} \Rightarrow 1 = \ln(e) \leq \ln\left(\frac{1}{\lambda(e)}\right) = -\ln(\lambda(e)) = \beta$$

the result follows in this case. For the remaining case $\lambda(\rho) = 0$ for all $\rho > 1$ we have from Corollary 32, with minimally abusive notation and our conventions:

$$\begin{aligned} -\ln(\lambda(e)) &= -\ln(0) = \infty \\ &\Rightarrow \lim_{x\to\infty} \frac{S(\rho x)}{S(x)} = 0 = \rho^{-\infty} \text{ for all } \rho > 1 \end{aligned}$$

and the result holds in this case as well. The proof is complete. ■

Corollary 34 *Let X be a loss variable with $M_X = \infty$ and $S = S_X$ and assume that the limit $\lambda(\rho) = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists for all $\rho > 1$ and further that there is some $\rho_0 > 1$ such that $\lambda(\rho_0) \neq 0$. Then*

$$\begin{aligned} \lambda(\rho) &= \lim_{x\to\infty} \frac{S(\rho x)}{S(x)} = \lim_{x\to\infty} \frac{T(\beta; \rho x)}{T(\beta; x)} = \rho^{-\beta} \\ \textit{for all } \rho &> 1, \textit{ where } 1 \leq \beta = -\ln(\lambda(e)) < \infty. \end{aligned}$$

Proposition 35 *Let X be a loss variable with $M_X = \infty$ and $S = S_X$, then the following are equivalent:*

1. $\lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ *exists for all* $\rho > 1$
2. *there exist* $\rho_k > 1, k \in \mathbb{N}$ *such that* $\lim_{k\to\infty} \rho_k = 1$ *and* $\lim_{x\to\infty} \frac{S(\rho_k x)}{S(x)}$ *exists for every* $k \in \mathbb{N}$.

Proof. It is apparent that $1 \Rightarrow 2$. To establish the meaningful direction $2 \Rightarrow 1$, we begin with the claim that:

$\alpha_k > 0$ for all $k \in \mathbb{N}$ and $\lim_{k\to\infty} \alpha_k = 0 \Rightarrow \{m\alpha_k | k, m \in \mathbb{N}\}$ is dense in $[0, \infty)$.

Indeed, given $\epsilon > 0, b \in (0, \infty)$:

$$\lim_{k\to\infty} \alpha_k = 0 \Rightarrow \exists k \in \mathbb{N} \ni 0 < \alpha_k < \frac{\epsilon}{2}$$

and setting

$$\begin{aligned} b_m &= m\alpha_k \Rightarrow b_{m+1} - b_m = \alpha_k < \frac{\epsilon}{2} \\ &\Rightarrow \text{ there exists } m \in \mathbb{N} \text{ such that } b_m \in (b - \epsilon, b + \epsilon) \\ &\Rightarrow m\alpha_k \in (b - \epsilon, b + \epsilon), k, m \in \mathbb{N} \end{aligned}$$

Since this holds for any $\epsilon > 0$, it follows that $\{m\alpha_k | k, m \in \mathbb{N}\}$ is dense in $[0, \infty)$ as claimed. And since the log function $\ln : [1, \infty) \to [0, \infty)$ is bicontinuous and bijective, we see that

$\rho_k > 1$ for all $k \in \mathbb{N}$ and $\lim_{k\to\infty} \rho_k = 1 \Rightarrow \{\rho_k^m | k, m \in \mathbb{N}\}$ is dense in $[1, \infty)$.

Now we have our assumption:

$$\begin{aligned} \text{there exist } \rho_k &> 1, k \in \mathbb{N} \text{ such that} \\ \lim_{k\to\infty} \rho_k &= 1 \text{ and } \lim_{x\to\infty} \frac{S(\rho_k x)}{S(x)} \text{ exists for every } k \in \mathbb{N} \end{aligned}$$

and we select any $\rho > 1$ and seek to prove that this assumption is sufficient to imply that the limit $\lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exits. So assume, by way of contradiction, that $\lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ does not exist. We have, by density:

$$\begin{aligned} \text{there exist } a_k, b_k &\in \{\rho_l^m | l, m \in \mathbb{N}\}, k \in \mathbb{N} \\ \text{such that } 1 &< a_1, a_k \le a_{k+1}, a_k \le \rho \text{ and with } \lim_{k\to\infty} a_k = \rho \\ \text{and such that } b_k &\ge b_{k+1}, b_k \ge \rho \text{ and with } \lim_{k\to\infty} b_k = \rho. \end{aligned}$$

Now S is a continuous, non-increasing function on $[0, \infty)$ and so we have:

$$\begin{aligned} a_k &\le \rho \le b_j \text{ for all } j, k \in \mathbb{N} \\ &\Rightarrow a_k x \le \rho x \le b_j x \text{ for all } x > 0, \text{ and for all } j, k \in \mathbb{N} \\ &\Rightarrow S(a_k x) \ge S(\rho x) \ge S(b_j x) \text{ for all } x > 0, \text{ and for all } j, k \in \mathbb{N} \\ &\Rightarrow \frac{S(a_k x)}{S(x)} \ge \frac{S(\rho x)}{S(x)} \ge \frac{S(b_j x)}{S(x)} \text{ for all } x > 0, \text{ and for all } j, k \in \mathbb{N}. \end{aligned}$$

Consider the two sets:

$$A = \left\{ \lim_{x\to\infty} \frac{S(a_k x)}{S(x)}, k \in \mathbb{N} \right\}$$
$$B = \left\{ \lim_{x\to\infty} \frac{S(b_k x)}{S(x)}, k \in \mathbb{N} \right\}$$

The above inequalities clearly force:

$$\beta \le \alpha \le 1 \quad \text{for all } \alpha \in A \text{ and for all } \beta \in B.$$

Observe that by Proposition 29:

$$\begin{aligned} \lim_{x\to\infty} \frac{S(a_k x)}{S(x)} &\le \frac{1}{a_k} < 1 \\ &\Rightarrow \alpha < 1 \text{ for all } \alpha \in A. \end{aligned}$$

We also have, for any $k \in \mathbb{N}$, that:

$$\begin{aligned} a_k &\le \rho \\ &\Rightarrow a_k x \le \rho x \text{ for all } x > 0 \\ &\Rightarrow S(a_k x) \ge S(\rho x) \text{ for all } x > 0 \\ &\Rightarrow \frac{S(a_k x)}{S(x)} \ge \frac{S(\rho x)}{S(x)} \text{ for all } x > 0 \\ &\Rightarrow \lim_{x\to\infty} \frac{S(a_k x)}{S(x)} \ge \frac{S(\rho x)}{S(x)} \ge 0 \text{ for all } x > 0. \end{aligned}$$

We claim that $\lim_{x\to\infty} \frac{S(a_k x)}{S(x)} > 0$ for every $k \in \mathbb{N}$, since otherwise:

$$\lim_{x\to\infty} \frac{S(a_k x)}{S(x)} = 0 \Rightarrow \text{existence of the limit } \lim_{x\to\infty} \frac{S(\rho x)}{S(x)} = 0 \Rightarrow\Leftarrow .$$

And we established:

$$\begin{aligned} 0 &< \alpha < 1 \text{ for all } \alpha \in A \\ &\Rightarrow A \subset (0,1). \end{aligned}$$

Now set:

$$\alpha = \inf A, \beta = \sup B$$

then clearly $0 \leq \beta \leq \alpha \leq 1$. We claim that:

$$\alpha = \beta.$$

Indeed, suppose, again by way of contradiction, that $\alpha \neq \beta$. Then we would have:

$$\beta < \alpha.$$

Now

$$A \subset (0,1) \Rightarrow \text{ there exists } c \in \{a_k | k \in \mathbb{N}\} \text{ with } 1 > \gamma = \lim_{x\to\infty} \frac{S(cx)}{S(x)} > 0$$

and we have, for any given $\epsilon > 0$:

$$\begin{aligned}
1 &> \gamma > 0 \Rightarrow \exists n \in \mathbb{N} \text{ such that } \gamma^{\frac{1}{n}} > 1 - \epsilon \\
\lim_{k\to\infty} \rho_k &= 1 \Rightarrow \exists m \in \mathbb{N} \text{ such that } \rho_m < c^{\frac{1}{n}} \\
&\Rightarrow \rho_m^n < c \\
&\Rightarrow \rho_m^n x < cx \text{ for all } x > 0 \\
&\Rightarrow S(\rho_m^n x) \geq S(cx) \text{ for all } x > 0 \\
&\Rightarrow \frac{S(\rho_m^n x)}{S(x)} \geq \frac{S(cx)}{S(x)} \text{ for all } x > 0 \\
&\Rightarrow \lim_{x\to\infty} \frac{S(\rho_m^n x)}{S(x)} \geq \lim_{x\to\infty} \frac{S(cx)}{S(x)} = \gamma \\
&\Rightarrow \left(\lim_{x\to\infty} \frac{S(\rho_m x)}{S(x)}\right)^n = \lim_{x\to\infty} \frac{S(\rho_m^n x)}{S(x)} \geq \gamma \\
\text{Proposition 29} &\Rightarrow 1 > \frac{1}{\rho_m} \geq \lim_{x\to\infty} \frac{S(\rho_m x)}{S(x)} \geq \gamma^{\frac{1}{n}} > 1 - \epsilon
\end{aligned}$$

And we have established:

$$\begin{aligned}
\text{For any given } \epsilon &> 0 \text{ there exists } \varphi_1 > 1 \\
&\text{such that the limit } \lim_{x\to\infty} \frac{S(\varphi_1 x)}{S(x)} \text{ exists and is in } (1-\epsilon, 1).
\end{aligned}$$

We next claim that:

$$\text{There exists } \varphi_2 > 1 \text{ such that } \lim_{x\to\infty} \frac{S(\varphi_2 x)}{S(x)} \in (\beta, \alpha).$$

Let $\beta_1 = \frac{\alpha+\beta}{2}$, then clearly $0 < \beta_1 < \alpha$ and $(\beta_1, \alpha) \subseteq (\beta, \alpha)$. Now let $\delta = \ln\alpha - \ln\beta_1 > 0$. Then letting $\epsilon = 1 - e^{-\delta}$ we have $\epsilon > 0$ and so by an earlier claim there exists $\varphi_1 > 1$ such that

$$\begin{aligned} \lim_{x\to\infty}\frac{S(\varphi_1 x)}{S(x)} &\in (1-\epsilon, 1) \\ &\Rightarrow \ln\left(\lim_{x\to\infty}\frac{S(\varphi_1 x)}{S(x)}\right) \in (-\delta, 0). \end{aligned}$$

Set $\eta = -\ln\left(\lim_{x\to\infty}\frac{S(\varphi_1 x)}{S(x)}\right)$, then:

$$\begin{aligned} -\eta &= \ln\left(\lim_{x\to\infty}\frac{S(\varphi_1 x)}{S(x)}\right) \in (-\delta, 0) = (\ln\beta_1 - \ln\alpha, 0), 0 < \eta < \delta \\ \left|\frac{\ln\beta}{-\eta} - \frac{\ln\alpha}{-\eta}\right| &= \frac{\ln\alpha}{\eta} - \frac{\ln\beta}{\eta} = \frac{\delta}{\eta}\left(\frac{\ln\alpha - \ln\beta}{\delta}\right) > \frac{\ln\alpha - \ln\beta}{\delta} = 1 \\ &\Rightarrow \text{there exists } l \in \mathbb{N} \text{ such that } l \in \left(\frac{\ln\alpha}{-\eta}, \frac{\ln\beta_1}{-\eta}\right) \\ &\Rightarrow -\eta l \in (\ln\beta_1, \ln\alpha) \Rightarrow e^{-\eta l} \in (\beta_1, \alpha) \subseteq (\beta, \alpha) \end{aligned}$$

and it follows that, setting $\varphi_2 = \varphi_1^l$ we have:

$$\begin{aligned} e^{-\eta l} &= e^{l\ln\left(\lim_{x\to\infty}\frac{S(\varphi_1 x)}{S(x)}\right)} = \left(e^{\ln\left(\lim_{x\to\infty}\frac{S(\varphi_1 x)}{S(x)}\right)}\right)^l \\ &= \left(\lim_{x\to\infty}\frac{S(\varphi_1 x)}{S(x)}\right)^l = \lim_{x\to\infty}\frac{S(\varphi_1^l x)}{S(x)} = \lim_{x\to\infty}\frac{S(\varphi_2 x)}{S(x)} \\ &\Rightarrow \lim_{x\to\infty}\frac{S(\varphi_2 x)}{S(x)} = e^{-\eta l} \in (\beta, \alpha) \end{aligned}$$

and the claim is established. Recalling how α and β were defined, we have:

$$\begin{aligned} \lim_{x\to\infty}\frac{S(\varphi_2 x)}{S(x)} &\in (\beta, \alpha) = (\sup B, \inf A) \\ &\Rightarrow \lim_{x\to\infty}\frac{S(b_j x)}{S(x)} < \lim_{x\to\infty}\frac{S(\varphi_2 x)}{S(x)} < \lim_{x\to\infty}\frac{S(a_k x)}{S(x)}, \forall j, k \in \mathbb{N} \end{aligned}$$

and we also have that:

$$\begin{aligned}
a_k &> \varphi_2 \Rightarrow \frac{S(a_k x)}{S(x)} \leq \frac{S(\varphi_2 x)}{S(x)} \\
&\Rightarrow \lim_{x\to\infty} \frac{S(a_k x)}{S(x)} \leq \lim_{x\to\infty} \frac{S(\varphi_2 x)}{S(x)} < \alpha = \inf A \Rightarrow\Leftarrow \\
b_j &< \varphi_2 \Rightarrow \frac{S(b_j x)}{S(x)} \geq \frac{S(\varphi_2 x)}{S(x)} \\
&\Rightarrow \lim_{x\to\infty} \frac{S(b_j x)}{S(x)} \geq \lim_{x\to\infty} \frac{S(\varphi_2 x)}{S(x)} > \beta = \sup B \Rightarrow\Leftarrow
\end{aligned}$$

and we are lead to:

$$a_k \leq \varphi_2 \leq b_j \text{ for all } k, j \in \mathbb{N}.$$

But this, in turn, leads to

$$\begin{aligned}
\lim_{k\to\infty} a_k &= \rho = \lim_{k\to\infty} b_k \Rightarrow \varphi_2 = \rho \\
&\Rightarrow \text{existence of the limit } \lim_{x\to\infty} \frac{S(\varphi_2 x)}{S(x)} = \lim_{x\to\infty} \frac{S(\rho x)}{S(x)} \Rightarrow\Leftarrow
\end{aligned}$$

and with this contradiction we have established our claim that that $\alpha = \beta$. Now by the definition of the set A and $\alpha = \inf A$ we find that for any given $\epsilon > 0$:

$$\begin{aligned}
\text{there exists } k_1 &\in \mathbb{N} \text{ such that } \alpha + \frac{\epsilon}{2} > \lim_{x\to\infty} \frac{S(a_{k_1} x)}{S(x)} \\
&\Rightarrow \text{there exists } x_1 > 0 \text{ such that} \\
\alpha + \epsilon &> \frac{S(a_{k_1} x)}{S(x)} \geq \frac{S(\rho x)}{S(x)}, \text{ for every } x \geq x_1 \\
&\Rightarrow \text{there exists } x_1 > 0 \text{ such that} \\
\alpha + \epsilon &> \frac{S(\rho x)}{S(x)}, \text{ for every } x \geq x_1.
\end{aligned}$$

And similarly, by the definition of the set B with $\alpha = \beta = \sup B$, we find that for any given $\epsilon > 0$

$$\begin{aligned}
\text{there exists } k_2 &\in \mathbb{N} \text{ such that } \beta - \frac{\epsilon}{2} < \lim_{x\to\infty} \frac{S(b_{k_2}x)}{S(x)} \\
&\Rightarrow \text{there exists } x_2 > 0 \text{ such that} \\
\beta - \epsilon &< \frac{S(b_{k_2}x)}{S(x)} \leq \frac{S(\rho x)}{S(x)}, \text{ for every } x \geq x_2 \\
&\Rightarrow \text{there exists } x_2 > 0 \text{ such that} \\
\alpha - \epsilon &= \beta - \epsilon < \frac{S(\rho x)}{S(x)}, \text{ for every } x \geq x_2.
\end{aligned}$$

Therefore:

$$\begin{aligned}
\text{given any } \epsilon &> 0, \text{there exists } x_3 > 0 \text{ such that} \\
\left\{\frac{S(\rho x)}{S(x)}, x \geq x_3\right\} &\subseteq (\alpha - \epsilon, \alpha + \epsilon) \\
&\Rightarrow \text{the limit } \lim_{x\to\infty} \frac{S(\rho x)}{S(x)} = \alpha \text{ exists} \Rightarrow\Leftarrow
\end{aligned}$$

and this final contradiction establishes that the limit $\lim_{x\to\infty} \frac{S(\rho x)}{S(x)}$ exists for all $\rho \geq 1$ and completes the proof. ∎

APPENDIX B

Power Transform Example $(\beta = 2)$

r	A	B	ΔA	ΔB	$\tilde{r}$	$\Delta\tilde{B}$	$\tilde{B}$	$\hat{r}$	$\hat{A}$	$\hat{B}$
0	0	0	0	0	0	0	0	0	0	0
0.1	0.082907	0.004145	0.082907	0.004145	0.316228	0.01853862	0.018539	0.110346	0.082907	0.020457
0.2	0.163781	0.016276	0.080874	0.012131	0.447214	0.03132221	0.049861	0.220692	0.163781	0.05502
0.3	0.236026	0.034338	0.072245	0.018061	0.547723	0.03612254	0.085983	0.331039	0.236026	0.094879
0.4	0.303221	0.057856	0.067195	0.023518	0.632456	0.03975295	0.125736	0.441385	0.303221	0.138745
0.5	0.363143	0.084821	0.059922	0.026965	0.707107	0.04019696	0.165933	0.551731	0.363143	0.183101
0.6	0.417715	0.114835	0.054572	0.030015	0.774597	0.04047166	0.206405	0.662077	0.417715	0.22776
0.7	0.467217	0.147012	0.049502	0.032177	0.83666	0.03991004	0.246315	0.772424	0.467217	0.271799
0.8	0.512363	0.180871	0.045146	0.033859	0.894427	0.03909732	0.285412	0.88277	0.512363	0.314942
0.9	0.554109	0.216355	0.041746	0.035484	0.948683	0.03848772	0.3239	0.993116	0.554109	0.357411
1	0.591878	0.252236	0.03777	0.035881	1	0.03681339	0.360713	1.103462	0.591878	0.398034
1.1	0.626693	0.288792	0.034815	0.036555	1.048809	0.03567449	0.396388	1.213809	0.626693	0.437399
1.2	0.658777	0.325688	0.032084	0.036897	1.095445	0.03440628	0.430794	1.324155	0.658777	0.475365
1.3	0.688675	0.363061	0.029898	0.037373	1.140175	0.03342733	0.464222	1.434501	0.688675	0.512251
1.4	0.716234	0.400266	0.027559	0.037204	1.183216	0.03202034	0.496242	1.544847	0.716234	0.547584
1.5	0.741513	0.43692	0.025279	0.036655	1.224745	0.03044021	0.526682	1.655193	0.741513	0.581174
1.6	0.765774	0.474524	0.024261	0.037604	1.264911	0.03020417	0.556886	1.76554	0.765774	0.614503
1.7	0.788177	0.51149	0.022404	0.036966	1.30384	0.02877787	0.585664	1.875886	0.788177	0.646258
1.8	0.809417	0.54866	0.02124	0.037169	1.341641	0.02809741	0.613762	1.986232	0.809417	0.677263
1.9	0.828932	0.584763	0.019515	0.036103	1.378405	0.02654347	0.640305	2.096578	0.828932	0.706552
2	0.847016	0.620025	0.018084	0.035263	1.414214	0.02525227	0.665557	2.206925	0.847016	0.734417
2.1	0.864609	0.656092	0.017594	0.036067	1.449138	0.02519015	0.690747	2.317271	0.864609	0.762214
2.2	0.88129	0.691955	0.01668	0.035863	1.48324	0.02445823	0.715206	2.427617	0.88129	0.789202
2.3	0.897579	0.728605	0.016289	0.03665	1.516575	0.02443335	0.739639	2.537963	0.897579	0.816164
2.4	0.912953	0.764736	0.015375	0.036131	1.549193	0.02356912	0.763208	2.64831	0.912953	0.842171
2.5	0.927842	0.801212	0.014888	0.036476	1.581139	0.02330379	0.786512	2.758656	0.927842	0.867886
2.6	0.942066	0.837484	0.014224	0.036272	1.612452	0.02271451	0.809226	2.869002	0.942066	0.892951
2.7	0.955979	0.874354	0.013913	0.03687	1.643168	0.02264905	0.831875	2.979348	0.955979	0.917943
2.8	0.970255	0.913613	0.014276	0.039259	1.67332	0.0236738	0.855549	3.089694	0.970255	0.944066
2.9	0.98386	0.952387	0.013605	0.038774	1.702939	0.02296765	0.878517	3.200041	0.98386	0.96941
3	1	1	0.01614	0.047613	1.732051	0.02772158	0.906238	3.310387	1	1

Proportional Hazard Transform Example $(\alpha = 5/7)$

r	A	B	ΔA	ΔB	$\tilde{A}$	$\Delta\tilde{A}$	$\Delta\tilde{B}$	$\tilde{B}$	$\hat{r}$	$\hat{A}$	$\hat{B}$
0	0	0	0	0	0	0	0	0	0	0	0
0.1	0.08290724	0.004145	0.082907	0 004145	0.059947	0.05994702	0.00299735	0.002997	0.078635	0.059947	0.002357
0.2	0.16378083	0.016276	0.080874	0.012131	0.119936	0 05998875	0.00899831	0.011996	0.15727	0.119936	0.009433
0 3	0.23602591	0.034338	0.072245	0.018061	0.174942	0.055005747	0.01375144	0.025747	0.235905	0.174942	0.020246
0 4	0.30322066	0 057856	0.067195	0 023518	0.227453	0.052510987	0.01837885	0.044126	0.31454	0.227453	0.034698
0.5	0.36314275	0 084821	0.059922	0.026965	0 275514	0.048061374	0.02162762	0.065754	0.393175	0.275514	0.051705
0.6	0.41771473	0.114835	0.054572	0.030015	0.320421	0.044907259	0.02469899	0.090453	0.47181	0.320421	0 071127
0 7	0.46721704	0.147012	0.049502	0 032177	0.362208	0.041787246	0.02716171	0.117614	0.550445	0.362208	0.092486
0 8	0.51236273	0.180871	0.045146	0 033859	0.401296	0.03908775	0 02931581	0.14693	0.629081	0.401296	0.115539
0.9	0 55410853	0.216355	0.041746	0.035484	0.438371	0 037075025	0.03151377	0.178444	0.707716	0.438371	0.140319
1	0.59187827	0.252236	0 03777	0.035881	0.472779	0.034407898	0.0326875	0.211131	0.786351	0 472779	0.166023
1.1	0.62669301	0.288792	0.034815	0 036555	0.50531	0.032531155	0.03415771	0.245289	0.864986	0.50531	0.192883
1.2	0.65877703	0 325688	0.032084	0 036897	0.536066	0.030756108	0.03536952	0.280659	0.943621	0.536066	0 220696
1 3	0.68867534	0.363061	0.029898	0.037373	0.56548	0.02941379	0 03676724	0.317426	1.022256	0.56548	0.249608
1.4	0.71623406	0.400266	0.027559	0.037204	0.593316	0 027835655	0 03757813	0.355004	1.100891	0.593316	0.279158
1.5	0.74151327	0.43692	0.025279	0.036655	0.619536	0.026220703	0.03802002	0.393024	1.179526	0.619536	0.309055
1.6	0.76577385	0.474524	0.024261	0 037604	0.645399	0 025862853	0.04008742	0.433111	1.258161	0.645399	0.340577
1.7	0.78817739	0.51149	0.022404	0.036966	0 669971	0.024572019	0 04054383	0 473655	1.336796	0.669971	0.372459
1.8	0.80941703	0.54866	0.02124	0 037169	0.693963	0 023991374	0.0419849	0.51564	1.415431	0 693963	0.405474
1.9	0.82893218	0 584763	0.019515	0.036103	0.716689	0.022726515	0.04204405	0.557684	1.494066	0.716689	0 438535
2	0.84701571	0.620025	0 018084	0.035263	0.73842	0.02173056	0.04237459	0.600059	1.572701	0.73842	0.471857
2.1	0.86460927	0 656092	0.017594	0 036067	0.760279	0.02185905	0 04481105	0.64487	1.651336	0.760279	0.507094
2.2	0.88128965	0.691955	0.01668	0.035863	0.781767	0.021488162	0.04619955	0.691069	1.729971	0.781767	0.543423
2.3	0.89757855	0 728605	0.016289	0.03665	0.803602	0.021835308	0.04912944	0.740199	1 808607	0 803602	0.582056
2.4	0.91295335	0 764736	0.015375	0 036131	0.825144	0.021541886	0 05062343	0.790822	1.887242	0.825144	0.621864
2.5	0.92784159	0 801212	0.014888	0 036476	0.847071	0.021926519	0 05371997	0.844542	1.965877	0.847071	0.664106
2.6	0.94206597	0.837484	0.014224	0.036272	0.869268	0.022197493	0.05660361	0.901146	2.044512	0.869268	0 708617
2.7	0.95597917	0.874354	0.013913	0.03687	0.892555	0.023287235	0 06171117	0.962857	2.123147	0.892555	0.757143
2.8	0.97025501	0.913613	0.014276	0.039259	0.918795	0.026239833	0.07215954	1.035017	2.201782	0.918795	0.813886
2.9	0.98385987	0.952387	0.013605	0.038774	0.947527	0.028732009	0.08188622	1 116903	2.280417	0.947527	0.878277
3	1	1	0 01614	0 047613	1	0.052472719	0 15479452	1.271697	2.359052	1	1

The 2004 NCCI Excess Loss Factors

Dan Corro and Greg Engl*

October 17, 2005

1 Introduction

An in-depth review of the NCCI excess loss factors (ELFs) was recently completed and changes were implemented in the 2004 filing season. The most significant change was to incorporate the latest data, but the methodology was thoroughly reviewed and a number of methodological changes were made as well. Among the methodological items considered were:

1. Individual Claim Development
   Our intent here was to follow the method in Gillam and Couret [5] and merely update the parameters. However our treatment of reopened claims is new as is the way we implement individual claim development. This is covered in detail in section 2.

2. Organization of Data
   The prior procedure fit countrywide loss distributions by injury type and then adjusted the means of those distributions to be appropriate for each individual state. We extend this idea to match the first two moments. The prior procedure implicitly gives each state's data a weight proportional to the number of claims in the given state, and thus even the largest states do not get very much weight in the countrywide distributions. We give much more weight to individual states' own data and thus fit state specific loss distributions. For credibility reasons the

*We gratefully acknowledge the creative contributions of the many people involved in this project, including, but not limited to, NCCI staff and NCCI's Retrospective Rating Working Group.

prior loss distributions combined permanent total injuries with major permanent partial injuries, and minor permanent partial injuries with temporary total injuries. We fit fatal, permanent total (PT), permanent partial (PP), temporary total (TT), and medical only distributions separately. In order to do this we use data at third, fourth, and fifth report for fatal and permanent total injuries. Mahler [10] also uses data at third, fourth, and fifth report. For permanent partial, temporary total, and medical only injuries, where there is adequate data, we only use data at fifth report. This is covered in section 3.

3. Fitting Method
   We follow Mahler [10] and rely on the empirical data for the small claims and only fit a distribution to the tail. We fit a mixed exponential distribution to the tail. Keatinge [8] discusses the mixed exponential distribution. Rather than fitting with the traditional maximum likelihood method we choose to fit the excess ratio function of the mixed exponential to the empirical excess ratio function using a least squares approach. This yields an extremely good fit to the data. It should be noted that we do not fit the raw data, but rather the data adjusted to reflect individual claim development as described in section 2. This results in a data set that has already been smoothed significantly and so we were not concerned that the mixed exponential tail might drop off too rapidly. Mahler [10] noted that the excess ratios are not very sensitive to the splice point, i.e. the point where the empirical data ends and the tail fit begins. Thus we preferred to not attach too far out into the tail so that we could have some confidence in the tail probability, i.e. the probability of a claim being greater than the splice point. We generally chose splice points that resulted in a tail probability between 5% and 15%. This is covered in section 4.

4. Treatment of Occurrences
   We put a firmer foundation under the modeling of occurrences by basing it on a collective risk model. In the end we find that the difference between per claim excess ratios and per occurrence excess ratios is almost negligible. This is quite a sharp contrast with the past. Once, per occurrence excess ratios were assumed to be 10% higher than per claim excess ratios. This was later refined by Gillam [4] to the assumption that the cost of the average occurrence was 10% higher than the aver-

age claim. Gillam and Couret [5] then refined this even further to apply by injury type: 3.9% for fatal injuries, 6.6% for permanent total and major permanent partial injuries, and 0% for minor permanent partial and temporary total injuries. Our analysis shows that per occurrence excess ratios are less than .2% more than per claim excess ratios. This is covered in section 5.

In section 6 we discuss updating the loss distributions. The current procedure is to update the loss distributions annually by a scale transformation and to refit the loss distributions based on new data fairly infrequently. The scale transformation assumption is extremely convenient and is discussed by Venter [12]. What is needed is a method to decide when a scale transformation is adequate and when the loss distributions need to be refit. We conclude by reviewing the methodology changes. While the focus of this paper is on methodology, we also take the opportunity to briefly discuss the impact of the changes.

2 Individual Claim Development

When evaluating aggregate loss development it is not necessary to account for the different patterns that individual claims may follow as they mature to closure. In aggregate it does not matter whether ten claims of $100 each all increase by $10 or whether just one claim increases by $100 to produce an ultimate loss of $1,100 and an aggregate loss development factor (LDF) of 1.1. But if you are interested in the excess of $110 per claim, it makes all the difference. Gillam and Couret [5] address the need to replace a single aggregate LDF with a distribution of LDFs in order to account for different possibilities for the ultimate loss of any immature claim. They refer to this as dispersion, and the name has stuck. Here, the term dispersion refers to a way of modelling ultimate losses that replaces each open claim with a loss distribution whose loss amounts correspond to the possibilities expected for that individual claim at closure.

The loss distribution used to determine the ELF should reflect the loss at claim closure. The calculation is done by injury type and uses incurred losses. It must reflect maturity in the incurred loss beyond its reporting maturity fully to closure, including any change in claim status (open/closed) and change in the incurred loss amount. Moreover, it must accommodate the reality that not all claims mature in the same way. Age to age aggregate incurred LDFs are determined from 1^{st} to 5^{th} report by state, injury type, and separately for indemnity and medical losses. The source is Workers Compensation Statistical Plan Data (WCSP), as adjusted for use in class ratemaking. As WCSP reporting ceases at 5^{th} report, 5^{th} to ultimate incurred LDFs, again separately for indemnity and medical losses, are determined from financial call data, typically in concert with the overall rate-level indication.

Individual claim WCSP data by injury type and report is the data source for the claim severity distributions. PP, TT, and medical only claims are included at a 5^{th} report basis. The far less frequent but often much larger Fatal and PT claims are included at 3^{rd}, 4^{th} and 5^{th} report basis. The WCSP data elements captured include state, injury type, report, incurred indemnity loss, incurred medical loss, and claim status. This detailed WCSP loss data is captured into a model for the empirical undeveloped loss distribution. That model consists of a discrete probability space to capture the probability of occurrence of individual claims together with two random variables for the claims' undeveloped medical and indemnity losses as well as four characteristic variables for state, injury type, report, and claim status. Eventually, this is refined into a model for the ultimate loss severity distribution that consists of a probability space together with one random variable for the claims' ultimate loss as well as two characteristic variables for state and injury type.

Because dispersion is exclusively focussed on open claims, without some accommodation, claims reported closed but that later reopen would not be correctly incorporated in the dispersion model. Accordingly, it is advisable to account for reopened claims prior to dispersing losses. The loss amounts considered are the total of the medical and indemnity losses for each claim. The methodology adjusts those loss amounts and probabilities by claim status and injury type, so as to model the impact of reopening claims. The details for the specific calculations used can be found in Appendix A and Appendix C. It is based on the observation that the few closed claims that reopen after a 5^{th} report (0.2%) are not typical, but are on average larger (by a factor of 8) and have a smaller CV (by a factor of 0.4). Appendix A shows quite generally how to calculate the resulting means and variances when a subset of claims have their status changed from closed to open.

The probability, mean, and variance of the three subsets of the loss model:

1. claims reported closed at 5^{th} report

2. claims reported open at 5^{th} report

3. claims that reopen subsequent to a 5^{th} report

completely determine the probability, mean, and variance of the complementary subsets:

1. claims 'truly closed' at 5^{th} report (those reported closed that do not reopen)

2. the complement set of 'truly open' claims.

That is, there is only one possibility for the probability, mean, and variance of the truly open and closed subsets, even though there are multiple possibilities for what particular claims reported closed at 5^{th} later reopen. In fact, those values can be explicitly determined from the formulas derived in Appendix A.

Knowing the probabilities of the truly open and closed subsets, we adjust the loss model by proportionally shifting the probabilities. The probability of each open claim is increased by a constant factor while the probability of each closed claim is correspondingly decreased by another factor. Knowing the mean and variance of the truly open subset lets us adjust the undeveloped combined medical and indemnity loss amounts of the open claims to match the two revised moments for open claims; this is done via a power transformation as described in Appendix C. The closed claim loss amounts are similarly adjusted. The result is a model of empirical undeveloped losses that reflects a trued up claim status as of a 5^{th} report, in the sense that no closed claims will reopen. That model, in turn, provides the input to the dispersion calculation. This approach is a refinement from that of Gillam and Couret [5] who account for the reopening of just a very few closed claims by dispersing all closed claims by just a very little. The idea here is to perform the adjustment prior to dispersion so that it is exactly the set of 'truly closed' claims whose losses are deemed to be at their ultimate cost and it is the complement set of 'truly open' claims that are dispersed.

In the resulting model for the empirical undeveloped loss distribution, the claim status variable is assumed to be correct in the sense that the loss amount for each closed claim is taken to be the known ultimate loss on the claim. Dispersion is applied only to open claims. Accordingly, the LDF applicable to all claims is adjusted to one appropriate for open claims only, and all development occurs on exactly the open claims. For each state, injury type, and report, one average LDF is determined from the medical and indemnity LDFs to apply to the sum of the medical and indemnity incurred losses of each claim. That combined incurred LDF is then modified to apply to just the open claims. More precisely, the relationship used to focus an aggregate LDF onto just the open claims is simply:

$$\begin{aligned} L_c &= \text{Aggregate undeveloped loss for closed claims} \\ L_o &= \text{Aggregate undeveloped loss for open claims} \\ \lambda &= \text{Aggregate LDF applicable to all claims} \\ \widetilde{\lambda} &= \text{Open only LDF} \\ \lambda\left(L_c + L_o\right) &= L_c + \widetilde{\lambda} L_o \Rightarrow \widetilde{\lambda} = \lambda + \left(\lambda - 1\right)\frac{L_c}{L_o} \end{aligned}$$

The adjusted to open only LDFs are determined and applied by state, injury type, and report.

Even though the adjusted LDFs are applied to all open claims independent of loss size, because the proportion of claims that remain open correlates with size of loss, the application of dispersion varies by the size of loss layer. Typically, larger losses are more likely to be open, and this application of development factors will have a greater impact in the higher loss layers. It follows that the application of loss development changes the shape of the severity distribution, making it better reflect the ultimate loss severity distribution.

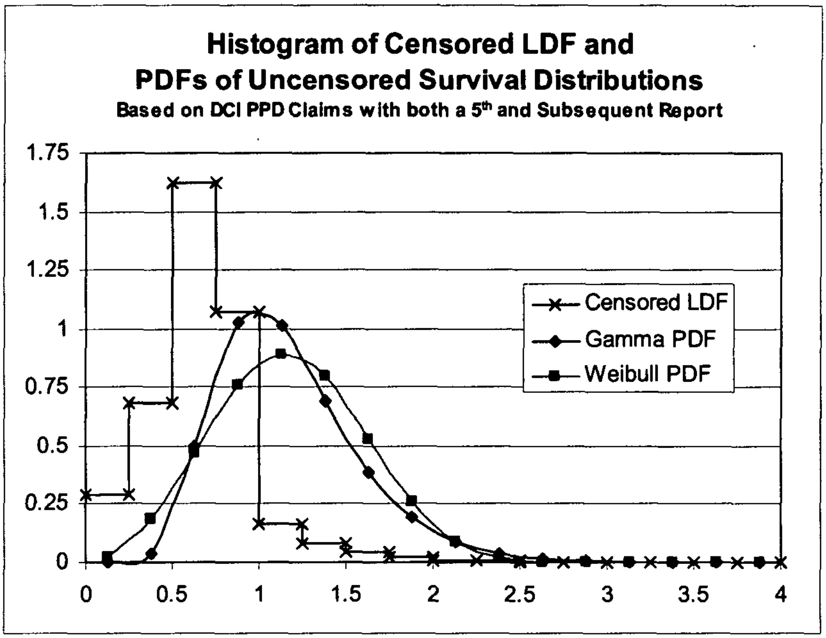
The next step is to apply dispersion to open claims. The technique used to disperse losses is formally equivalent to that used by Gillam and Couret [5]. The technique bears some similarity to kernel density estimation in which an assumed known density function (the kernel) is averaged across the observed data points so as to create a smoothed approximation. More precisely, the idea is to replace each open claim with a distribution of claims that reflect the various possibilities for the loss that is ultimately incurred on that claim. The expected loss at closure is just the applicable to ultimate LDF times the undeveloped loss. The LDF is varied according to an inverse transformed gamma distribution and multiplied by the undeveloped loss to model the possibilities for the ultimate loss.

The NCCI Detailed Claim Information (DCI) database was used to build a data set of observed LDFs beyond a 5^{th} report. We studied DCI claims open at 5^{th} report for which a subsequent DCI report was available. The observed LDF was determined as the ratio of the incurred loss at the latest available report divided by the incurred loss at 5^{th} report. If the claim remained open at that latest report, the observed LDF was considered "right censored." Censored regression of the kind used to study survival was used to fit this data. Open claims were identified as the censored observations, i.e. closed claims were deemed "dead" and open claims "alive" in the survival model. The survival model was used to determine an appropriate form to represent the distribution.

More precisely, the SAS PROC LIFEREG procedure was used to estimate accelerated failure time models from the LDF observations. Letting Y denote the observed LDF, the model was specified by the simplest possible equation $Y = \widetilde{\lambda} + \varepsilon$, where $\widetilde{\lambda}$ represents a constant and ε a variable error term. That is, the model specifies just an intercept term with no covariates at all. That model specification was selected because it corresponds to the application of a constant LDF ($\widetilde{\lambda}$) to open claims. Moreover, the error term of the model corresponds precisely with dispersion, as that term is used here. Consequently, this application of survival analysis is somehat unconventional inasmuch as the issue is not the survival curve or the goodness of fit of the parameter estmate $\widetilde{\lambda}$ that is key. Rather, the interest here is on the distribution of the error term ε. The SAS LIFEREG procedure is well suited to this because not only does it account for censored observations, it also allows for different structural forms to be assumed for the error term ε when estimating accelerated failure time models.

In this application, the estmated parameter for the intercept was not used since the LDF factors by state, injury type and report were taken from ratemaking data. What was of interest is the form and parameters that specify the error distribution. The Weibull, the Lognormal, the Gamma, and the generalized Gamma distribution were considered. In fact, the two-parameter Weibull, two-parameter Gamma, and the two-parameter Lognormal are all special cases of the three-parameter generalized Gamma (the Weibull and Gamma directly via parameter constraint, the Lognormal only asymptotically). The solutions for the generalized gamma implied that its three parameters enabled it to outperform the two parameter distributions. The three-parameter model guided the specification of the functional form and parameter values for the LDF distributions used in the dispersion calculation.

With the eventual goal to calculate excess ratios, it was important to assess whether the error term varies by size of loss. Gillam and Couret [5] assume that the CV of the dispersion distribution does not vary by size of loss. In addition to specifying different structural forms for the error term, models were fit to quintiles of the data, where by a quintile we mean that



the observations were divided into five equal volume groups according to claim size. It was observed that the CV of the error term did not show any significant variation by size of loss. This affirmed the prior assumption of a constant CV, and that assumption was again used in this dispersion calculation.

The LIFEREG procedure outputs the parameters that specify the dispersion pattern, by injury type, that relates a fifth report loss amount with the probable distribution of the incurred cost at "death" of the claim, i.e. at claim closure. Combining that with average LDFs from ratemaking, the uncensored distribution of the ultimate loss severity can be calculated. For any fixed open claim, the uncensored LDF distribution values times the (undeveloped) loss amount corresponds with the probable values for that claim at closure. It follows that the uncensored LDF distribution corresponds to age to ultimate LDFs applicable on a per open claim basis. The above chart illustrates how the survival model anticipates rightward movement of the reported empirical losses and fills out the right hand tail.

Because the mean LDF was already known, our primary focus was on the CV. This follows the approach of Gillam and Couret [5], whose decision to use a two-parameter gamma distribution for the reciprocal of the LDF was also followed. The use of the gamma to model the reciprocal amounts to the use of an inverse gamma for the LDF. That choice was reaffirmed by the DCI data and is illustrated somewhat in the above chart. We actually used a three-parameter inverse transformed gamma distribution, as the survival model suggested that would yield a better representation of the LDF distribution. The first two parameters, denoted α, τ in Klugman, et. al. [9] determine the CV of the distribution, which varies by report and injury type as indicated in the following table:

Injury	Report	α	τ	CV
Fatal & PT	3	5.7134	0.8	0.7
Fatal & PT	4	6.8664	0.8	0.6
Fatal & PT	5	8.7775	0.8	0.5
PP	5	8.7775	0.8	0.5
TT	5	12	3	0.1
Med Only	5	12	3	0.1

The third parameter, denoted β in Klugman, et. al. [9], determines the mean LDF and was directly solved for to make that mean equal the age to ultimate aggregate open claim LDF by state, report, and injury type. Even though open TT and Med only claims are not assumed to develop in aggregate (mean LDF = 1), the open TT and Med only claims are dispersed, but with a small CV.

Gillam and Couret [5] used a CV of 0.9 for the LDF on open claims; that selection was dictated to some degree by the need to account for potential unobserved large losses. The current ratemaking methodology makes separate provision for very large losses. This, in turn, enables this ELF revision to rely less on judgment and more on empirical data. The empirical data suggested the lower CVs used for the LDF distributions. All else equal, lowering the CV lowers the ELF at the largest attachment points. Much sensitivity analysis was done to assess the impact of this change in the assumed CV. It was determined that the selection did not represent an unreasonable reduction in the ELFs.

As is typical with kernel density models, Gillam and Couret [5] used a closed form integration formula to implement dispersion. However, in order to be able to perform the downstream data adjustments (in particular, adjusting to state conditions as discussed in the next section), we instead used the device of representing each open claim by 173 variants. The variants are determined by multiplying the undeveloped loss amount by 173 different LDFs. The variant LDFs have mean equal to the applicable overall LDF (as applicable to open claims only) and a CV of 0.5 for 5^{th} report Fatal, PT, and PP claims. The mean LDF applicable for medical only and TT cases is 1, as those cases are assumed not to develop in aggregate beyond a 5^{th} report. So even open medical only and TT claims are dispersed, albeit so as not to change the aggregate loss (and with a smaller CV of 0.1 for the LDF distribution). The choice of 173 points was done to enable the calculation to better capture the tail. Very small and very large LDFs are included in the model (corresponding to the 0.000001^{st} and 99.999999^{th} percentile of the inverse transformed gamma) albeit with a correspondingly very small weight (about 0.000001) being assigned to such variants. Dispersion does not change the contribution of any claim to the aggregate developed loss. It was determined that the use of 173 points provided a very close approximation to the continuous form. Additional details on that calculation can be found in Appendix B.

To summarize, the dispersion calculation starts with a finite probability space of claims together with a random variable giving the undeveloped claim values. Then both the probability measure and the random variable are adjusted to account for reopened claims. That gives a modified probability space of claims. Replacing each open claim with a distribution of 173 expected loss amounts at closure yields a developed dispersed probability space of claims with a random variable giving the ultimate claim value. This is done for each injury type and for all NCCI states. The next section describes how those random variables are adjusted to state specific conditions so as to yield the empirical distributions used in fitting the data to severity distributions.

3 Organization of Data

The idea of estimating excess ratios by injury type goes back at least to Uhthoff [11] and has been used as well by Harwayne [6], Gillam [4], and Gillam and Couret [5]. While we follow this approach as well, it should be noted that alternatives have recently been identified by Brooks [2] and Mahler [10].

Owing to the relatively few fatal and permanent total claims it is desirable to combine data across states. Differences between states preclude doing this without adjustment however. Gillam [4] addressed this by grouping states according to benefit structure. For an interesting recent approach incorporating benefit structure see Gleeson [7]. With the current dominance of medical costs this approach is less satisfactory. In the prior approach, Gillam and Couret [5] addressed the problem "by dividing each claim by the average cost per case for the appropriate state-injury-type combination." We refer to this data adjustment technique as mean normalization. This results in a countrywide database with mean of 1. Loss distributions were then fit to this normalized database. The countrywide loss distributions are then adjusted via a scale transformation (see Venter [12]) to be appropriate for each particular state. Thus the data for different states is adjusted to have the same mean. A natural variant of this would be median normalization, the thought being that the median might be more stable than the mean. A natural extension is to try and match more than one moment. We considered five data adjustment techniques altogether:

1. Mean Normalization
   As mentioned above, for a given injury type, each claim in state i, denoted by x_i (here x_i denotes the incurred loss on a claim from state i developed to ultimate), is transformed by $x_i \rightarrow x_i/\mu_i$, where μ_i denotes the mean of the x_i. The normalized claims for all states are now combined into a countrywide database. To get a database appropriate for state j, each normalized claim is then scaled up by the mean in state j, i.e. $x_i/\mu_i \rightarrow \mu_j \cdot x_i/\mu_i$

2. Median Normalization
   This is analogous to mean normalization, but claims are now normalized by the median rather than the mean.

3. Logarithmic Standardization
   A natural generalization of mean normalization would be to standardize claims, $x_i \rightarrow \frac{x_i - \mu_i}{\sigma_i}$. To avoid negative claim values when transforming the standardized database to a particular state we standardize the logged losses, $\log x_i \rightarrow \frac{\log x_i - \mu_i}{\sigma_i}$, where now μ_i, σ_i denote the mean and standard deviation of the logged losses. This results in a standardized countrywide database, which can then be adjusted to a given state j by $\frac{\log x_i - \mu_i}{\sigma_i} \rightarrow \sigma_j \cdot \frac{\log x_i - \mu_i}{\sigma_i} + \mu_j$. Appendix C discusses this in more detail.

4. Generalized Standardization
   This is analogous to logarithmic standardization except that instead of the mean and variance, percentiles can be used. For example, instead of the mean we could use the median and instead of the standard deviation we could use the 85^{th} percentile minus the median.

5. Power Transform
   Lastly, we considered a power transform, $x_i \to ax_i^b$, where the values of a and b are chosen so that the transformed values have the mean and variance of state j. That this is possible is shown in Appendix C. Thus for each state i there is a different power transform that takes the unadjusted state i claims and adjusts them to what they would be in state j, in the sense that the transformed claims from state i match the mean and variance in state j. Combining all of the adjusted claims results in an expanded state j specific database. Notice that the unadjusted state j claims appear in the expanded state j database and so the expanded state j database is indeed an expansion of the state j data. It should also be noted that the power transform generalizes both mean normalization and logarithmic standardization and the moments are matched in dollar space rather than in log space. This is discussed in more detail in Appendix C.

Extensive performance testing was conducted to decide which data adjustment techniques to use. The idea was to postulate realistic loss distributions for the states, based on realistic parameters, simulate data from the postulated loss distributions and see which techniques best recovered the postulated distributions. Initial tests showed that median normalization and generalized standardization performed poorly and so further tests concentrated on the remaining techniques. Based on our performance tests we chose to use logarithmic standardization for Fatal and Permanent Total (PT) claims and the power transform for Permanent Partial (PP), Temporary Total (TT), and Medical Only claims. It seemed that when there were only a limited number of claims and the difference in CVs between states was large the exponent in the power transform could occasionally be quite large, leading the power transform to underperform logarithmic standardization.

Gillam and Couret [5] call modeling PT and PP claims separately the "common sense approach." Owing to the scarcity of PT claims they have in the past been combined with Major PP claims. Due to our improved data adjustment techniques we are able to separate PT from PP. We also used data at 3^{rd}, 4^{th}, and 5^{th} report for Fatal and PT claims because of their relative scarcity, whereas we only used data at 5^{th} report for the other injury types.

In the prior approach, each state's weight in the countrywide database was proportional to the number of claims it contributed to the countrywide total. This seems implicitly like assigning a state's data a credibility of n/N, where n is the number of claims in the state and N is the countrywide total. Further, this implicit credibility did not vary by injury type. This makes sense when there is only one countrywide database. We however, use a different database for each state and give each state's data a weight of $\sqrt{n/N}$ in the state specific database, where n is the number of claims in the state and N is a standard based on actuarial judgment. Our view was that most states would have enough data to fit loss distributions for Medical Only, but that no state would have enough claims to fit a Fatal loss distribution and only the largest states would have enough PT claims. We thought it reasonable that three quarters of the states would have enough Medical Only claims, half of them would have enough TT claims and about a quarter of them would have enough PP claims. With this in mind, we chose N, the standard for full pooling weight, to be $2,000$ for Fatal claims, $1,500$ for PT claims, $7,000$ for PP claims, $8,500$ for TT claims and $20,000$ for Medical Only claims. It is intuitively sensible that the standard for Medical Only should be higher than for PT because excess ratios are driven by large claims and most PT claims are large whereas most Medical Only claims are typically small.

4 Fitting

Traditionally a parametric loss distribution would be fit to the entire data set by maximum likelihood. The first problem with this approach is that distributions which fit the tail well may not fit the small claims so well and thus there is a trade-off between fitting the tail well and fitting the small claims well. The need for a fitted loss distribution is really only in the tail as the number of small claims is quite large. Mahler [10] has recently used the empirical distribution for small claims and spliced a fitted loss distribution onto the tail. This is the approach we follow as well and we describe it in detail in Appendix E. Fitting the tail alone is of course much easier and the fits are much better than they have been in the past. The second problem with the traditional approach is that maximizing the likelihood function is somewhat indirect. While maximum likelihood fits typically result in loss distributions with excess ratio functions that do fit the data well, there is no intrinsic interest in the likelihood function itself. The primary objective is a loss distribution whose excess ratio function fits the data well and so instead of maximum likelihood we use least squares to fit the excess ratio function directly. Appendix D gives some general facts about excess ratio functions. In particular, Proposition 12 shows that a distribution is determined by its excess ratio function and so there is no loss of information in working with excess ratio functions rather than densities or distribution functions.

Mahler [10] uses a Pareto-exponential mixture to fit the tail. We use two to four term mixed exponentials. The mixed exponential distribution is described by Keatinge [8]. All things being equal, the mixed exponential is a thinner tailed distribution than has been used in the past. It has moments of all orders, whereas some loss distributions in use do not even have finite variances. However, the loss data used to fit the mixed exponential is driven by the inverse transformed gamma distribution of LDFs, as described in section 2, and the inverse transformed gamma is not a thin tailed distribution. This prevents the tail of the fitted loss distribution from being too thin. The mixed exponential also has an increasing mean residual life, and this is quite typical of Workers Compensation claim data. Fat tailed distributions may make sense in the presence of catastrophic loss potential, but recently NCCI has made a separate CAT filing so the new ELFs are for the first time explicitly non-CAT. From a geometrical perspective, the density function over the tail region should be decreasing and have no inflection points, as occurs where the first derivative of the density function is negative and its second derivative is positive. The mixed exponential class of distributions has alternating sign derivatives of all orders. And conversely any distribution with alternating sign derivatives of all orders can be approximated by a mixed exponential to within any desired degree of accuracy. Functions with this alternating derivative property are called completely monotone and this characterization of them follows from a theorem by Bernstein. (See Feller [3].) We initially considered using other distributions besides the mixed exponential, but the mixed exponential fits were so good that it was not necessary to consider other distributions further.

Mahler [10] noted that the excess ratios are not very sensitive to the splice point, i.e. the point where the empirical data ends and the tail fit begins. We found that to be the case as well. We were concerned with large losses being under represented in the data. Thus we preferred to not attach too far out into the tail so that we could have some confidence in the tail probability, i.e. the probability of a claim being greater than the splice point. So we generally chose splice points that resulted in a tail probability between 5% and 15%. While this gave us some confidence in the tail probability, we were still concerned about claims in the \$10 million to \$50 million range being under represented in the data. (Claims larger than \$50 million would be accounted for in the separate CAT filing.) The new excess ratios are based on one to three years of data, depending on the injury type, but the largest WC claims and events occur with return periods exceeding three years. WC catastrophe modeling indicates that claims and occurrences in the \$10 million to \$50 million range are underrepresented in the data used to fit the new curves. Because of this, we included an additional provision for individual claims and occurrences between \$10 million and \$50 million. This new provision is broadly grounded in the results of several WC catastrophe models, and known large WC occurrences. Previous excess ratio curves included a provision for anti-selection of 0.005, which has been eliminated in the new curves. The new provision, per-claim or per-occurrence, is .003 up to \$10 million, 0 for \$50 million or greater, and declines linearly from .003 to 0 between \$10 million and \$50 million. Thus the final adjusted excess ratio is 0.997 times the excess ratio before this adjustment, plus this adjustment. That is, if L is the loss limit and $R(L)$ is the unadjusted per claim or per occurrence excess ratio, then the adjusted excess ratio is given by

$$R'(L) = \begin{cases} .997R(L) + .003 & \text{if } L \leq \$10M \\ .997R(L) - \frac{.003}{\$40M}L + .00375 & \text{if } \$10M < L < \$50M \\ .997R(L) & \text{if } L \geq \$50M \end{cases}.$$

5 Modelling Occurrences

Data is typically collected on a per claim basis. This makes it a challenge to produce per occurrence excess ratios. The first attempt to address this was to merely increase the per claim excess ratios by 10% to account for occurrences. For low attachment points this could lead to excess ratios greater than 1. Gillam [4] improved this approach by assuming only that the average occurrence cost 10% more than the average claim. This affects the entry ratio used to compute the excess ratio. Gillam and Couret [5] then refined this approach still further by breaking down the 10% by injury type: 3.9% for fatal injuries, 6.6% for permanent total and major permanent injuries, and 0% for minor permanent partial and temporary total injuries. These approaches, while reasonable, rely heavily on actuarial judgment.

The first attempt to base per occurrence excess ratios more solidly on per occurrence data was by Mahler [10], who attempted to group claims into occurrences based on hazard group, accident date, and policy number. NCCI has a CAT[1] code which identifies claims in multiple claim occurrences. Singleton claims (occurrences with only one claim) have a CAT code of 00, all claims in the first multi-claim occurrence would have a CAT code of 01, claims in the second multi-claim occurrence would have a CAT code of 02, etc. Unfortunately there were several problems with the CAT code:

1. missing CAT codes
   For singleton claims it is permissible to report a blank field for the CAT code. This would then be converted to a 00. However there was no way of knowing whether a blank field was deliberately reported as a blank or inadvertently omitted.

2. orphans
   There were claims observed with nonzero CAT codes, but with no other claims with the same CAT code. One carrier, for example, appeared to have numbered the claims in a multiple claim occurrence sequentially.

3. variance in injury dates
   Claims were observed with the same CAT code, but with different injury dates. In one case the injury dates were 14 months apart.

4. grouping of CAT claims
   It is permissible to group small med only claims in reporting. This is not permissible however in the case of CAT claims. Nevertheless there was some evidence of grouped reporting for CAT claims.

Further complicating things was the fact that even with optimal reporting, multiple claim occurrences appear to be extremely rare. Based on an examination of data from carriers known to report their data well, it would appear that .2% is a reasonable estimate of the portion of all claims that

[1]Here a catastrophe is merely an occurrence with more than one claim. The term 'catastrophe' in this context has no implications as to the size of the occurrence.

occur as part of multi-claim occurrences. Based on the above problems, we decided not to try and build a per occurrence data base, but rather to use a collective risk model. From the per claim loss distributions we could easily get an overall per claim severity distribution. We estimated the frequency distribution for multiple claim occurrences from carriers thought to have recorded the CAT code correctly. The mean number of claims in a multiple claim occurrence is about 3, but most multiple claim occurrences consist of two claims.

Unfortunately the severity distribution of claims in multiple claim occurrences seemed to be different from the severity distribution of singleton claims. First, the mix of injury types in multiple claim occurrences was more severe than in singleton claims. Second, even when fixing an injury type, claims occurring as part of a multiple claim occurrence were more severe. We chose to address this issue by assuming that the severity distribution of claims in multiple claim occurrences differed from the distribution of singletons only by a scale transformation. This assumption goes at least as far back as Venter [12].

More formally, let X_i be the random variable giving the cost of a singleton claim of injury type i and let F_{X_i} be the distribution function of X_i. If S is the random variable giving the overall cost of a singleton occurrence then $F_S = \sum w_i F_{X_i}$, where w_i is the probability that a singleton claim is of injury type i. That is, the per claim severity distribution is a mixture of the injury type distributions. If Y_i is the random variable giving the cost of a claim of injury type i in a multiple claim occurrence then we assume that Y_i differs from X_i by a scale transform, i.e. $Y_i = a_i X_i$ for some constant a_i. If Z is the random variable giving the overall cost of a claim in a multiple claim occurrence then $F_Z = \sum w'_i F_{Y_i}$, where w'_i is the probability that a claim in a multiple claim occurrence is of injury type i. Then $M = Z_1 + \cdots + Z_N$ is the cost of a multiple claim occurrence, where N is the random variable giving the number of claims in a multiple claim occurrence and the Z_i are iid random variables with the same distribution as Z. Finally, the per occurrence severity distribution is given by $F = rF_S + (1 - r)F_M$, where r is the probability that an occurrence consists of a single claim.

Because r is so close to 1 there is very little difference between per claim and per occurrence loss distributions. Per occurrence excess ratios are no more than .2% more than per claim excess ratios. This is a sharp contrast with the prior approaches.

6 Updating

Overall excess ratios are computed as a weighted average of the injury type excess ratios. Let $R(L)$ be the overall excess ratio at a loss limit of L, and let $R_i(r)$ be the excess ratio for injury type i at an entry ratio of r, then

$$R(L) = \sum_i w_i R_i(L/\mu_i),$$

where w_i is the percentage of losses of type i and μ_i is the mean loss of type i. The injury type weights, w_i, and average costs per case, μ_i, are updated annually, but the injury type excess ratio functions, R_i, are updated only infrequently. The idea is that the shape of the loss distributions changes much more slowly than the scale. The annual update thus involves adjusting the mix of injury types and adjusting the loss distributions by a scale transformation. Updating via a scale transformation is extremely convenient and is discussed by Venter [12].

The key question is how to determine when a simple scale transformation update is adequate and when the loss distributions need to be refit. If X is the random variable corresponding to last year's loss distribution and Y is the random variable corresponding to this year's loss distribution, then the scale transformation updating assumption is that there is some constant, c, such that Y and cX have the same distribution. Then the normalized distribution, Y/μ_Y has the same distribution as $cX/c\mu_X = X/\mu_X$ and thus $\mathrm{Var}(Y/\mu_Y) = \mathrm{Var}(X/\mu_X) = \sigma_X^2/\mu_X^2 = CV_X^2$. So if successive year's loss distributions really did differ only by a scale transform then the CV would remain constant over time. Thus monitoring the CV over time might give a criterion for when it is necessary to update the underlying loss distributions and not just the injury type weights and average costs per case.

Since the injury type loss distributions are normalized to have mean 1, applying a uniform trend factor would have no impact. Thus the losses used for fitting are typically not trended to a future effective date. This is extremely convenient in that it does not require us to decide in advance when the loss distributions need to be updated. However, if the trend is not uniform, then it could result in a change in the shape of the loss distributions. This could for instance happen if there was a persistent difference in medical and indemnity trends and the percentage of loss due to medical costs varied by claim size, as it typically does, even after controlling for injury type. How significant this phenomenon is remains an open question. It is in some sense limited as medical trends cannot exceed inflation forever without the medical sector consuming an unacceptably large fraction of GDP. Nevertheless, this does suggest that monitoring the difference in cumulative medical and indemnity trends might provide a guide as to when the shape of the loss distributions needs to be updated.

7 Conclusion

With the present revision we have implemented several changes to the methodology as summarized in the table below. We retained the general approach to dispersion of individual claim development due to Gillam and Couret [5], using an inverse transformed gamma for the distribution of LDFs, but lowering the CV from .9 to .5. Instead of fitting a loss distribution to all of the claims, we followed Mahler [10] and fit only the tail, using the empirical distribution for the small claims. For the tail we used a mixed exponential as compared to the prior transformed betas fit to the entire distribution. Instead of combining PT with Major PP claims, we fit PT and PP claims separately, using data at 3^{rd}, 4^{th}, and 5^{th} report for Fatal and PT claims. The prior approach used only data at 5^{th} report. To adjust the data from one state to be comparable with another state we used logarithimic standardization for Fatal and PT claims and power transforms for PP, TT, and Med Only. The prior approach was to use mean normalization for all injury types. We then fit state specific loss distributions rather than the countywide ones used before. Finally, to go from per claim data to per occurrence ELFs we used a collective risk model of occurrences. This contrasts sharply with prior approaches based on estimates of how much the mean occurrence cost exceeded the mean claim cost. The prior approach implicitly assumed a 3.9% load for Fatal claims, a 6.6% load for PT/Major PP claims, and a 0% load for TT and Med Only claims.

	new approach	**prior approach**
dispersion	CV = .5	CV = .9
fitting	fit tail only	fit whole distribution
form of distribution	empirical/mixed exponential	transformed beta
injury types	PT, PP separate	PT, Major PP combined
data	3^{rd}, 4^{th}, 5^{th} report for F, PT	5^{th} report
data adjustment	logarithmic standardization, power transform	mean normalization
applicability of distributions	state specific	countrywide
per occurrence	collective risk	3.9% F, 6.6% PT/Maj PP

While the changes made to the ELF methodology were significant, they were more evolutionary than revolutionary. Nevertheless, the new ELFs are quite a bit lower than the old ones at the larger limits in many states. We examined carefully the impact of the change in the dispersion CV and the use of mixed exponential rather than transformed beta distributions. Had we used a dispersion CV of 0.9 rather than 0.5, the ELFs would have been higher than the new ones. But at the higher limits, where the decrease was most pronounced, ELFs based on a CV of 0.9 would still be much closer to the new ELFs than the old. We also refit the old transformed beta distributions to the new data and found that even with the old distributional forms, fit to the entire distribution, the result is a much thinner tail than in the distributions underlying the old ELFs. We thus concluded that changes in the empirical loss distributions underlying the prior and the revised ELFs are what drive the reduction in ELFs. The prior review of ELFs relied on data that preceded the decline of WC claim frequency that so dominated WC experience in the 1990s, and beyond. There are solid theoretical reasons to suggest that this is just the sort of dynamic that can significantly change the shape of the loss distributions in a fashion that may not be captured by scale adjustments and as such require the development of new ELFs.

APPENDIX

A Adjusting for Reopened Claims

This appendix details some calculations referenced in section 2 on developing individual claims, in particular on the treatment of reopened claims. We consider a set of observed individual claims grouped by their open/closed claim status and determine how the first two moments of the open and closed subsets change when some claims are 'reopened,' i.e. when some claims are reclassified from the closed to the open subset. The discussion applies quite generally to show how the first two moments are impacted by a change in a characteristic, like claim status, to a selected subset of observations. The mean and variance of a finite set of observed values have natural generalizations to vector valued observations. It is convenient to express the findings as they apply in a multi-dimensional context, even though the specific application in this paper requires only the one-dimensional case.

Suppose we have a finite set of claims C and that a vector $x_c \in \mathbb{R}^n$ is associated with each $c \in C$. Suppose each $c \in C$ is also assigned a probability of occurrence $\omega_c > 0$ For any nonempty subset $A \subseteq C$, we make the following definitions

$$\begin{aligned}
\text{Probability of the set } A &= |A|_\omega = \sum_{a \in A} \omega_a \\
\text{Mean of } A &= \mu_A = \frac{1}{|A|_\omega} \sum_{a \in A} \omega_a x_a \in \mathbb{R}^n \\
\text{Variance of } A &= \sigma_A^2 = \frac{1}{|A|_\omega} \sum_{a \in A} \omega_a \|x_a - \mu_A\|^2 \geq 0
\end{aligned}$$

and we make the usual convention that for the empty set $|\phi|_\omega = \sigma_\phi^2 = 0$ and $\mu_\phi = \overrightarrow{0}$ is the 0-vector.

Observe that the mean is a vector and the variance a scalar and that for $n = 1$ this defines the mean and variance associated with the probability density function $f(a) = \frac{\omega_a}{|A|_\omega}$ on A when we view the subset A as a probability space in its own right. A natural WC application of multi-dimensionality is the case $n = 2$ in which the first coordinate measures the indemnity loss amount and the second component the medical loss of a claim $c \in C$. Note that we have the usual relationship between the mean, the variance and the second moment:

$$\begin{aligned}
\sigma_A^2 &= \frac{1}{|A|_\omega}\sum_{a\in A}\omega_a \left\|x_a-\mu_A\right\|^2 = \frac{1}{|A|_\omega}\sum_{a\in A}\omega_a (x_a-\mu_A)\cdot(x_a-\mu_A) \\
&= \frac{1}{|A|_\omega}\sum_{a\in A}\omega_a (x_a\cdot x_a - 2\mu_A\cdot x_a + \mu_A\cdot\mu_A) \\
&= \frac{1}{|A|_\omega}\sum_{a\in A}\omega_a \left\|x_a\right\|^2 - 2\left(\mu_A\cdot\frac{1}{|A|_\omega}\sum_{a\in A}\omega_a x_a\right) + \frac{|A|_\omega}{|A|_\omega}\left\|\mu_A\right\|^2 \\
&= \frac{1}{|A|_\omega}\sum_{a\in A}\omega_a \left\|x_a\right\|^2 - 2(\mu_A\cdot\mu_A) + \left\|\mu_A\right\|^2 \\
&= \frac{1}{|A|_\omega}\sum_{a\in A}\omega_a \left\|x_a\right\|^2 - 2\left\|\mu_A\right\|^2 + \left\|\mu_A\right\|^2 = \frac{1}{|A|_\omega}\sum_{a\in A}\omega_a \left\|x_a\right\|^2 - \left\|\mu_A\right\|^2
\end{aligned}$$

And thus

$$\left\|\mu_A\right\|^2 + \sigma_A^2 = \frac{1}{|A|_\omega}\sum_{a\in A}\omega_a \left\|x_a\right\|^2.$$

There are the evident relationships with the union and intersection of subsets $A, B \subseteq C$; for the mean we have:

$$\begin{aligned}
\mu_{A\cup B} &= \frac{1}{|A\cup B|_\omega}\sum_{c\in A\cup B}\omega_c x_c = \frac{1}{|A\cup B|_\omega}\left(\sum_{a\in A}\omega_a x_a + \sum_{b\in B}\omega_b x_b - \sum_{c\in A\cap B}\omega_c x_c\right) \\
&= \frac{1}{|A\cup B|_\omega}\left(|A|_\omega\mu_A + |B|_\omega\mu_B - |A\cap B|_\omega\mu_{A\cap B}\right)
\end{aligned}$$

And thus

$$\mu_{A\cup B} + \frac{|A\cap B|_\omega}{|A\cup B|_\omega}\mu_{A\cap B} = \frac{|A|_\omega}{|A\cup B|_\omega}\mu_A + \frac{|B|_\omega}{|A\cup B|_\omega}\mu_B.$$

and similarly for the variance:

$$\begin{aligned}
|A\cup B|_\omega\left(\left\|\mu_{A\cup B}\right\|^2 + \sigma_{A\cup B}^2\right) &= \sum_{c\in A\cup B}\omega_c\left\|x_c\right\|^2 \\
&= \sum_{a\in A}\omega_a\left\|x_a\right\|^2 + \sum_{b\in B}\omega_b\left\|x_b\right\|^2 - \sum_{c\in A\cap B}\omega_c\left\|x_c\right\|^2 \\
&= |A|_\omega\left(\left\|\mu_A\right\|^2+\sigma_A^2\right) + |B|_\omega\left(\left\|\mu_B\right\|^2+\sigma_B^2\right) \\
&\quad - |A\cap B|_\omega\left(\left\|\mu_{A\cap B}\right\|^2+\sigma_{A\cap B}^2\right)
\end{aligned}$$

And thus

$$\begin{aligned}\sigma^2_{A\cup B} + \frac{|A\cap B|_\omega}{|A\cup B|_\omega}\sigma^2_{A\cap B} &= \frac{|A|_\omega}{|A\cup B|_\omega}\sigma^2_A + \frac{|B|_\omega}{|A\cup B|_\omega}\sigma^2_B \\ &+\frac{1}{|A\cup B|_\omega}\left(|A|_\omega \|\mu_A\|^2 + |B|_\omega \|\mu_B\|^2 - |A\cap B|_\omega \|\mu_{A\cap B}\|^2\right) \\ &- \|\mu_{A\cup B}\|^2\end{aligned}$$

We are especially interested in the case when C is a disjoint union, so we make the assumption:

$$C = A\cup B \qquad A\cap B = \phi \qquad A \neq \phi$$

Think of the decomposition as reflecting a two-valued claim status, like open and closed. The goal is to determine how the mean and variance change after "moving" a subset D from A to B. The example of this paper is when the claim decomposition reflects claim closure status as of a 5^{th} report, (A = closed and B=open) and D is a set of closed claims that reopen after a 5^{th} report.

In this case of a disjoint union, it is especially easy to express μ_C and σ^2_C in terms of the corresponding statistics for A and B. From the above formula for the mean of a union:

$$\begin{aligned}\mu_C &= \mu_{A\cup B} + \overrightarrow{0} = \mu_{A\cup B} + \frac{|A\cap B|_\omega}{|A\cup B|_\omega}\mu_{A\cap B} \\ &= \frac{|A|_\omega}{|A\cup B|_\omega}\mu_A + \frac{|B|_\omega}{|A\cup B|_\omega}\mu_B. \\ &= w\mu_A + (1-w)\mu_B \quad \text{where } w = \frac{|A|_\omega}{|C|_\omega} \in (0,1].\end{aligned}$$

The second moments are similarly weighted averages, with the same subset weights w and $1-w$. From what we just saw for the mean of a disjoint union combined with the above formula for the variance of a union:

$$\begin{aligned}\sigma^2_C &= \sigma^2_{A\cup B} + 0 = \sigma^2_{A\cup B} + \frac{|A\cap B|_\omega}{|A\cup B|_\omega}\sigma^2_{A\cap B} \\ &= w\sigma^2_A + (1-w)\sigma^2_B + w\|\mu_A\|^2 + (1-w)\|\mu_B\|^2 - \|w\mu_A + (1-w)\mu_B\|^2 \\ &= w\sigma^2_A + (1-w)\sigma^2_B + w\|\mu_A\|^2 + (1-w)\|\mu_B\|^2 \\ &\quad -w^2\|\mu_A\|^2 - 2w(1-w)\mu_A\cdot\mu_B - (1-w)^2\|\mu_B\|^2 \\ &= w\sigma^2_A + (1-w)\sigma^2_B + w(1-w)\left(\|\mu_A\|^2 - 2\mu_A\cdot\mu_B + \|\mu_B\|^2\right) \\ &= w\sigma^2_A + (1-w)\sigma^2_B + w(1-w)\|\mu_A - \mu_B\|^2\end{aligned}$$

This expresses the variance of a disjoint union in terms of the means and variances of the subsets.

Notice that these formulas for μ_C and σ_C^2 show how the mean and variance of the subset A are constrained by those of the superset C. For the remainder of this appendix we assume $\sigma_C > 0$ and so we have:

$$\sigma_C^2 = w\sigma_A^2 + (1-w)\,\sigma_B^2 + w\,(1-w)\,\|\mu_A - \mu_B\|^2 \geq w\sigma_A^2 \Rightarrow w\left(\frac{\sigma_A}{\sigma_C}\right)^2 \leq 1$$

Observe that assigning the difference vector δ and scalar ratio r as:

$$\delta = \mu_A - \mu_C \qquad\qquad r = \frac{\sigma_A}{\sigma_C}$$

then we also have:

$$\begin{aligned}
\mu_C &= w\,(\mu_C + \delta) + (1-w)\,\mu_B \Rightarrow \mu_B = \mu_C - \frac{w\delta}{1-w} \\
&\Rightarrow \mu_A - \mu_B = \mu_C + \delta - \left(\mu_C - \frac{w\delta}{1-w}\right) = \frac{\delta\,(1-w) + w\delta}{1-w} = \frac{\delta}{1-w}
\end{aligned}$$

But then:

$$\begin{aligned}
\sigma_C^2 &= w\sigma_A^2 + (1-w)\,\sigma_B^2 + w\,(1-w)\left\|\frac{\delta}{1-w}\right\|^2 \geq w\sigma_A^2 + \frac{w\,\|\delta\|^2}{1-w} = wr^2\sigma_C^2 + \frac{w\,\|\delta\|^2}{1-w} \\
&\Rightarrow \left(1 - wr^2\right)\sigma_C^2 \geq \frac{w\,\|\delta\|^2}{1-w} \Rightarrow r \leq \sqrt{\frac{1}{w}} \text{ and } \|\delta\| \leq \sigma_C\sqrt{\frac{(1-w)\,(1-wr^2)}{w}}
\end{aligned}$$

and we see how, for any nonempty subset A, the mean difference vector δ is constrained by the probability allocation together with the deviation ratio r and the standard deviation of C.

Now suppose we have "local information" on how the proper subset $D \subset A$ fits within A, captured in the two numbers p, r and the difference vector δ:

$$\begin{aligned}
p &= \frac{|D|_\omega}{|A|_\omega} \\
r\sigma_A &= \sigma_D \\
\delta &= \mu_D - \mu_A
\end{aligned}$$

in which we specify that $r = 1$ should $\sigma_A = 0$. From what we've just seen, applying the above to any nonempty subset $D \subset A$, the following two inequalities must hold:

$$r \leq \sqrt{\frac{1}{p}} \qquad \|\delta\| \leq \sigma_A \sqrt{\frac{(1-p)(1-pr^2)}{p}}$$

Define the sets:

$$\begin{aligned} \widehat{A} &= A \backslash D = \{a \in A | a \notin D\} \\ \widehat{B} &= B \cup D \\ &\Rightarrow C = \widehat{A} \cup \widehat{B} \qquad \widehat{A} \cap \widehat{B} = \phi \qquad \widehat{A} \neq \phi \neq \widehat{B} \end{aligned}$$

In terms of the above open/closed claim example, this second decomposition represents the "truly closed" verses the "truly open" claims, as of a 5^{th} report.

With transparent notation, we seek to determine the subset probability and the moments $\widehat{w}$, $\mu_{\widehat{A}}$, $\mu_{\widehat{B}}$, $\sigma_{\widehat{A}}$, $\sigma_{\widehat{B}}$ in terms of the original subset probability and moments w, μ_A, μ_B, σ_A, σ_B together with the local information p, r and δ. The calculations only require some persistence:

$$\begin{aligned} p &= \frac{|D|_\omega}{|A|_\omega} \Rightarrow |D|_\omega = p\,|A|_\omega \Rightarrow \left|\widehat{A}\right|_\omega = |A|_\omega - |D|_\omega = |A|_\omega - p\,|A|_\omega = (1-p)\,|A \\ &\Rightarrow \quad \widehat{w} = \frac{\left|\widehat{A}\right|_\omega}{|C|_\omega} = \frac{\left|\widehat{A}\right|_\omega}{|A|_\omega}\frac{|A|_\omega}{|C|_\omega} = (1-p)\,w \end{aligned}$$

Continuing in turn, we have:

$$\begin{aligned} \mu_A &= p\mu_D + (1-p)\,\mu_{\widehat{A}} = p\,(\mu_A + \delta) + (1-p)\,\mu_{\widehat{A}} \\ &\Rightarrow (1-p)\,\mu_{\widehat{A}} = \mu_A - p\mu_A - p\delta = (1-p)\,\mu_A - p\delta \\ &\Rightarrow \mu_{\widehat{A}} = \mu_A - \left(\frac{p}{1-p}\right)\delta. \end{aligned}$$

And since we now know $\widehat{w}$ and $\mu_{\widehat{A}}$, we determine $\mu_{\widehat{B}}$ from:

$$\mu_C = \widehat{w}\mu_{\widehat{A}} + (1-\widehat{w})\,\mu_{\widehat{B}} \Rightarrow \mu_{\widehat{B}} = \frac{\mu_C - \widehat{w}\mu_{\widehat{A}}}{1-\widehat{w}}$$

And we get $\sigma_{\widehat{A}}$ from:

$$\begin{aligned} \sigma_A^2 &= p\sigma_D^2 + (1-p)\,\sigma_{\widehat{A}}^2 + p\,(1-p)\left\|\mu_{\widehat{A}} - \mu_D\right\|^2 \\ &\Rightarrow \sigma_{\widehat{A}}^2 = \frac{\sigma_A^2 - p\sigma_D^2}{1-p} - p\left\|\mu_{\widehat{A}} - \mu_D\right\|^2 \end{aligned}$$

And finally, we can obtain $\sigma_{\widehat{B}}$ from:

$$\begin{aligned}\sigma_C^2 &= \widehat{w}\sigma_{\widehat{A}}^2 + (1-\widehat{w})\,\sigma_{\widehat{B}}^2 + \widehat{w}\,(1-\widehat{w})\left\|\mu_{\widehat{A}} - \mu_{\widehat{B}}\right\|^2 \\ \Rightarrow\quad \sigma_{\widehat{B}}^2 &= \frac{\sigma_C^2 - \widehat{w}\sigma_{\widehat{A}}^2}{1-\widehat{w}} - \widehat{w}\left\|\mu_{\widehat{A}} - \mu_{\widehat{B}}\right\|^2\end{aligned}$$

The requisite formulas for the adjusted moments and subset probabilities are summarized in the following proposition:

Proposition 1 *Let $C = A \cup B$ be a decomposition of C into mutually exclusive subsets, as above, and suppose D is a proper subset of A and set*

$$\begin{aligned} w &= \frac{|A|_\omega}{|C|_\omega} \\ p &= \frac{|D|_\omega}{|A|_\omega} \\ \delta &= \mu_D - \mu_A. \end{aligned}$$

Then for the alternative decomposition $C = \widehat{A} \cup \widehat{B}$ where

$$\begin{aligned} \widehat{A} &= A \backslash D = \{a \in A | a \notin D\} \\ \widehat{B} &= B \cup D \\ \widehat{w} &= \frac{\left|\widehat{A}\right|_\omega}{|C|_\omega} \end{aligned}$$

we have:

$$\begin{aligned} \phi &= \widehat{A} \cap \widehat{B} \\ \widehat{A} &\neq \phi \neq \widehat{B} \\ \widehat{w} &= (1-p)\,w \\ \mu_{\widehat{A}} &= \mu_A - \left(\frac{p}{1-p}\right)\delta \\ \mu_{\widehat{B}} &= \frac{\mu_C - \widehat{w}\mu_{\widehat{A}}}{1-\widehat{w}} \\ \sigma_{\widehat{A}}^2 &= \frac{\sigma_A^2 - p\sigma_D^2}{1-p} - p\left\|\mu_{\widehat{A}} - \mu_D\right\|^2 \\ \sigma_{\widehat{B}}^2 &= \frac{\sigma_C^2 - \widehat{w}\sigma_{\widehat{A}}^2}{1-\widehat{w}} - \widehat{w}\left\|\mu_{\widehat{A}} - \mu_{\widehat{B}}\right\|^2 \end{aligned}$$

Proof. Clear from the above. ■

It is straightforward to generalize the formulas that express the mean and variance of a disjoint union of two sets to apply to partitions of more than two sets. The formula for the mean is immediate:

$$C = \bigcup_{i=1}^{m} A_i \qquad A_i \cap A_j = \phi \text{ for } i \neq j \qquad w_i = \frac{|A_i|_\omega}{|C|_\omega} > 0$$

$$\begin{aligned} \mu_C &= \frac{1}{|C|_\omega} \sum_{c \in C} \omega_c x_c = \frac{1}{|C|_\omega} \sum_{i=1}^{m} \sum_{a \in A_i} \omega_a x_a = \frac{1}{|C|_\omega} \sum_{i=1}^{m} \frac{|A_i|_\omega}{|A_i|_\omega} \sum_{a \in A_i} \omega_a x_a \\ &= \frac{1}{|C|_\omega} \sum_{i=1}^{m} |A_i|_\omega \, \mu_{A_i} = \sum_{i=1}^{m} w_i \mu_{A_i} \end{aligned}$$

and for the variance we first consider the expression for the second moment:

$$\begin{aligned} \|\mu_C\|^2 + \sigma_C^2 &= \frac{1}{|C|_\omega} \sum_{c \in C} \omega_c \|x_c\|^2 = \frac{1}{|C|_\omega} \sum_{i=1}^{m} \sum_{a \in A_i} \omega_a \|x_a\|^2 \\ &= \frac{1}{|C|_\omega} \sum_{i=1}^{m} |A_i|_\omega \left(\|\mu_{A_i}\|^2 + \sigma_{A_i}^2 \right) \\ &= \sum_{i=1}^{m} w_i \left(\|\mu_{A_i}\|^2 + \sigma_{A_i}^2 \right) = \sum_{i=1}^{m} w_i \|\mu_{A_i}\|^2 + \sum_{i=1}^{m} w_i \sigma_{A_i}^2 \end{aligned}$$

and we find that:

$$\begin{aligned}
\sigma_C^2 &= \sum_{i=1}^{m} w_i \left\|\mu_{A_i}\right\|^2 + \sum_{i=1}^{m} w_i \sigma_{A_i}^2 - \|\mu_C\|^2 \\
&= \sum_{i=1}^{m} w_i \left(\mu_{A_i} \cdot \mu_{A_i}\right) + \sum_{i=1}^{m} w_i \sigma_{A_i}^2 - \left(\sum_{i=1}^{m} w_i \mu_{A_i}\right) \cdot \left(\sum_{i=1}^{m} w_i \mu_{A_i}\right) \\
&= \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i=1}^{m} w_i \left(\mu_{A_i} \cdot \mu_{A_i}\right) - \left(\sum_{i=1}^{m} w_i^2 \left(\mu_{A_i} \cdot \mu_{A_i}\right) + 2\sum_{i<j} w_i w_j \left(\mu_{A_i} \cdot \mu_{A_j}\right)\right) \\
&= \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i=1}^{m} \left(w_i - w_i^2\right) \left(\mu_{A_i} \cdot \mu_{A_i}\right) - 2\sum_{i<j} w_i w_j \left(\mu_{A_i} \cdot \mu_{A_j}\right) \\
&= \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i=1}^{m} w_i \left(1 - w_i\right) \left(\mu_{A_i} \cdot \mu_{A_i}\right) - 2\sum_{i<j} w_i w_j \left(\mu_{A_i} \cdot \mu_{A_j}\right) \\
&= \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i=1}^{m} w_i \left(\sum_{j\neq i} w_j\right) \left(\mu_{A_i} \cdot \mu_{A_i}\right) - 2\sum_{i<j} w_i w_j \left(\mu_{A_i} \cdot \mu_{A_j}\right) \\
&= \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i<j} w_i w_j \left(\mu_{A_i} \cdot \mu_{A_i} + \mu_{A_j} \cdot \mu_{A_j}\right) - 2\sum_{i<j} w_i w_j \left(\mu_{A_i} \cdot \mu_{A_j}\right) \\
&= \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i<j} w_i w_j \left(\mu_{A_i} \cdot \mu_{A_i} + \mu_{A_j} \cdot \mu_{A_j} - 2\mu_{A_i} \cdot \mu_{A_j}\right) \\
&= \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i<j} w_i w_j \left(\mu_{A_i} - \mu_{A_j}\right) \cdot \left(\mu_{A_i} - \mu_{A_j}\right) \\
&= \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i<j} w_i w_j \left\|\mu_{A_i} - \mu_{A_j}\right\|^2
\end{aligned}$$

and the generalization of the formula for the variance of a partition is:

$$\sigma_C^2 = \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sum_{i<j} w_i w_j \left\|\mu_{A_i} - \mu_{A_j}\right\|^2.$$

Consider the special case of the set of m mean vectors $M = \left\{\mu_{A_i}\right\}$ expressed as a disjoint union of singleton subsets in which the vector μ_{A_i} is assigned the probability w_i. Then the formula gives:

$$\begin{aligned}\sigma_M^2 &= \sum_{i=1}^{m} w_i \sigma^2_{\{\mu_{A_i}\}} + \sum_{i<j} w_i w_j \left\| \mu_{A_i} - \mu_{A_j} \right\|^2 \\ &= \sum_{i=1}^{m} w_i \,(0) + \sum_{i<j} w_i w_j \left\| \mu_{A_i} - \mu_{A_j} \right\|^2 \\ &= \sum_{i<j} w_i w_j \left\| \mu_{A_i} - \mu_{A_j} \right\|^2\end{aligned}$$

But this is just the second term in the earlier expression for σ_C^2 and we find that

$$\sigma_C^2 = \sum_{i=1}^{m} w_i \sigma_{A_i}^2 + \sigma_M^2$$

which generalizes the usual decomposition of the variance into the sum of the within and the between variance. This has application to cluster analysis, where it affords a useful geometrical interpretation. In cluster analysis it is common to work with vectors so as to capture the influence of multiple data fields. So as above assume each claim $c \in C$ is assigned a vector of values that captures information about the claim that we seek to organize into a classification scheme. Viewing the m subsets $A_i \subseteq C$ as defining clusters of vectors, the set of m mean vectors $M = \{\mu_{A_i}\}$ is the set of 'centroids' of those clusters. The goal of cluster analysis is to separate the data into like clusters, but there is both a local and a global perspective to that classification problem: selecting like data in each cluster (minimize the within clusters variance) and separating the clusters (maximize the between clusters variance). The above shows that the two are one and the same when the Euclidean metric is used to measure the distance between observations. Indeed, decreasing the within clusters variance is the same as increasing the between centroids variance, as the two sum to the constant σ_C^2.

B Discrete Individual Claim Development

We want to populate the tails of the LDF distribution so that the dispersion model contemplates a claim developing quite dramatically. Accordingly, we seek a finite set of probabilities

$$0 < p_1 < p_2 < \cdots < p_n < 1$$

that cover $(0,1)$ with an emphasis on populating the right and left hand tails near 0 and 1. We are confronted with a practical working limit of no more than 200 points. We have also observed that 100 equally spaced points will result in the dispersion reflecting too confined a range, about 1/3 to 3-fold for the full range of dispersion. To cover a wider range, we use 171 non-uniform probabilities, and focus on the tails. Then treating the probabilities p_i as defining percentiles, we determine the corresponding percentile values u_i from a gamma distribution. That finite sequence $\{u_i\}$ of values is the starting point to capture a gamma density. But this representation is then refined, replacing the percentiles with the means over the 172 intervals $[0, u_1), [u_1, u_2), \ldots, [u_{170}, u_{171})$ and $[u_{171}, \infty)$. The new sequence of values, again denoted as $\{u_i\}$, is an optimized discrete approximation to a gamma. It is "weighted" in the sense that mean value u_i has associated with it the frequency weight v_i, where

$$v_1 = p_1, v_2 = p_2 - p_1, ..., v_{171} = p_{171} - p_{170}, v_{172} = 1 - p_{171}.$$

The interval width provides the weight assigned to the corresponding percentile value and is selected to be at most $\frac{1}{100}$ so that the usual "percentiles" are "covered." By definition, inverting and transforming those observations produces a discrete approximation to values from an inverse transformed gamma distribution. These are the candidates for the set of loss development factors used for dispersion. Parameters were selected so as to achieve a target mean LDF as well as a target CV for the LDFs. In order to assure the correct mean, one more observation is added, forcing the weighted mean of the sequence $\{u_i|1 \leq i \leq 173\}$ to be exactly the appropriate open claim only LDF. There is the concern that if that final observation is allotted too little weight, it will have the potential for becoming an outlier. So the added observation has weight $\frac{1}{100}$ and the other weights are adjusted by a factor of $\frac{99}{100}$, making the 173 weights $\{v_i|1 \leq i \leq 173\}$ again total to 1. From this construction, it is expected that the $\{u_i|1 \leq i \leq 173\}$ will exhibit a slightly smaller variance than the theoretical inverse transformed gamma, and that is indeed observed to be the case in the calculations. For example, when targeting a CV of 0.500, the model yielded a CV of 0.495.

This discussion does not describe the (comparatively minor) adjustment for reopened claims. The reopened claim adjustment is achieved by first using the results of Appendix A to determine means and variances after reclassifying some closed claims as open, and then matching two moments using the power transform as detailed in Appendix C.2. In this way, the "truly open" claims are dispersed.

We now fill in the details of the algorithm used to build the dispersion model. The first step is to specify the α and τ parameters, by injury type and report, for the inverse transformed gamma. The parameters were selected from an analysis of LDF distributions as presented in section 2. The parameterizations follow that of the Appendix of Klugman et al. [9].

Recall that the α and τ parameters determine the CV and once they are set, the θ parameter dictates the mean.

The next step is to build a discrete approximation to a Gamma distribution with parameters α and θ=1. This is captured in two finite sequences, u and v. The u sequence captures the values while the v sequence stores the corresponding probability of occurrence "weights." We identify the "percentile" u-value of the distribution function associated with the following list of probabilities p_i, $1 \leq i \leq 171$:

$$\begin{aligned}
p_0 &= 0 \\
p_i &= p_{i-1} + 10^{-6} && 1 \leq i \leq 10 \\
p_i &= p_{i-1} + 10^{-5} && 11 \leq i \leq 19 \\
p_i &= p_{i-1} + 10^{-4} && 20 \leq i \leq 28 \\
p_i &= p_{i-1} + 10^{-3} && 29 \leq i \leq 37 \\
p_i &= p_{i-1} + 10^{-2} && 38 \leq i \leq 86 \\
p_{86+i} &= 1 - p_{86-i} && 1 \leq i \leq 85.
\end{aligned}$$

These probabilities were selected to give greater granularity to the right and left tails. This corresponds to 171 finite intervals: $[u_0 = 0, u_1), \ldots, [u_i, u_{i+1})$ for $0 \leq i \leq 170$ and the right hand tail interval $[u_{171}, \infty)$. We let $\Gamma(\alpha; u)$ denote the incomplete gamma function as formally defined in the Appendix of Klugman et al. [9], where that function is also noted to be the distribution function of a gamma distribution with parameters α and $\theta = 1$ (and for the transformed gamma with parameters α , $\theta = 1$, and $\tau = 1$). A binary search routine is used to associate the value u_i with the probability p_i, finding u_i that satisfies:

$$|\Gamma(\alpha; u_i) - p_i| < 0.00000000001 \quad 1 \leq i \leq 171.$$

The first difference of the p_i gives the frequency probability v_i of an observation falling within the interval $[u_{i-1}, u_i)$, i.e. between percentile p_{i-1} and

p_i. The mean value over each of the 172 intervals is readily determined from the observation that given $f(\alpha, 1; x) = \frac{x^\alpha e^{-x}}{x\Gamma(\alpha)}$, we get

$$xf(\alpha, 1; x) = \frac{x^{\alpha+1}e^{-x}}{x\Gamma(\alpha)} = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \cdot \frac{x^{\alpha+1}e^{-x}}{x\Gamma(\alpha+1)} = \alpha\frac{x^{\alpha+1}e^{-x}}{x\Gamma(\alpha+1)} = \alpha f(\alpha{+}1, 1; x)$$

and thus

$$\int_0^z xf(\alpha, 1; x)dx = \int_0^z \alpha f(\alpha+1, 1; x)dx = \alpha\Gamma(\alpha+1; z).$$

This lets us specify the sequence u of length 172 whose components are the mean value of the inverse transformed gamma over the 172 intervals. The sequence v, also of length 172, with components equaling the corresponding frequency provides a sequence of weights to apply to the corresponding LDF values captured in the sequence u.

Denote the applicable development factor for open claims as $\widetilde{\lambda}$. The next step in building the dispersion model is to specify a sequence Λ of length 173 whose component values (properly weighted) are distributed as an inverse transformed gamma distribution of mean $\widetilde{\lambda}$ (and CV determined from the corresponding α and τ parameters). The formula for the expectation of an inverse transformed gamma random variable, X, allows us to calculate the θ parameter:

$$E[X] = \frac{\theta \cdot \Gamma(\alpha - \frac{1}{\tau})}{\Gamma(\alpha)} \Rightarrow \theta = \widetilde{\lambda}\left(\frac{\Gamma(\alpha)}{\Gamma(\alpha - \frac{1}{\tau})}\right).$$

The dispersion model uses the inverse transformed gamma as the LDF distribution. By definition, a distribution is inverse transformed gamma exactly if, when transformed and inverted, it conforms to a gamma distribution like that approximated by the sequences u and v of discrete values and weights, respectively. Following the parametrization of the Appendix of Klugman et al. [9], to make the finite sequence Λ contain values distributed as the inverse transform gamma, we just use the equivalence:

$$\left(\frac{\theta}{\Lambda_i}\right)^\tau = u_i \Leftrightarrow \frac{\Lambda_i}{\theta} = u_i^{-\frac{1}{\tau}} = \frac{1}{u_i^{\frac{1}{\tau}}} \Leftrightarrow \Lambda_i = \frac{\theta}{u_i^{\frac{1}{\tau}}}.$$

Since we are using a discrete approximation, and to assure we do get the correct expected developed loss, we augment the Λ sequence by an additional value in order to force the weighted mean to be $\widetilde{\lambda}$. More precisely we set $v_{173} = \frac{1}{100}$, rescale the other weights by setting $v_j = \frac{99}{100}v_j$ for $1 \leq j \leq 172$, and set $\Lambda_{173} = 100\left(\widetilde{\lambda} - \sum_{j=1}^{172} v_j\Lambda_j\right)$, which assures that:

$$\sum_{j=1}^{173} v_j = 1 \qquad \sum_{j=1}^{173} v_j\Lambda_j = \widetilde{\lambda}.$$

Having the Λ and v sequences in hand, completing the dispersion loss severity model is then very straightforward. Individual claim data are captured from WCSP data into observations that include state, injury type, claim status, a weight w, and a loss amount l, as described in section 2. Closed and open claims are separated into two subsets of observations, L_c and L_o respectively. Then for each open claim of weight w and undeveloped loss amount equal to l in L_o, 173 "dispersed" observations are captured into the data set $\widetilde{L_o}$ using the sequences Λ and v to assign the observations with weights equal to the product $w \times v_i$ and developed loss amounts equal to the product $\Lambda_i \times l$, $1 \leq i \leq 173$. Losses in $\widetilde{L_o}$ are adjusted to be at least \$1. Finally, forming the union $L_c \cup \widetilde{L_o}$ of two sets, each consisting of observations of individual claim data at closure, results in the dispersion model for ultimate claim severity.

C Data Adjustment Techniques

Let $x_{i1}, x_{i2}, \ldots, x_{in_i}$ be the incurred loss amounts on the claims (of a given injury type) in state i and let $\mu_i = \frac{1}{n_i}\sum_j x_{ij}$ be the sample mean. Under mean normalization we divide each claim amount by the state sample mean to get x_{ij}/μ_i. Pooling all the mean normalized claims for all states gives us a countrywide mean normalized database, $\{x_{ij}/\mu_i\}$. This database has mean 1 of course. If we fix a state k and multiply each mean normalized claim amount in the countrywide database by μ_k we get a database, $\{\mu_k x_{ij}/\mu_i\}$, that has mean μ_k. This database augments the claims in state k with out of state claims that have been adjusted to the state k level. We now generalize this simple idea to the case of standardization as well as the power transform.

C.1 Logarithmic Standardization

A natural way to generalize mean normalization would be to standardize claims, i.e. to subtract the state sample mean from every claim and divide by the standard deviation, $x_i \to \frac{x_i - \mu_i}{\sigma_i}$. Pooling all of the standardized claims would result in a countrywide standardized database with mean 0 and standard deviation 1. Then for a given state k, we might multiply each standardized claim by σ_k and and add μ_k to get a database, $\left\{\sigma_k \frac{x_i - \mu_i}{\sigma_i} + \mu_k\right\}$, appropriate for state k. Unfortunately, this can result in negative claim amounts so we prefer to work with logged losses and standardize them by mapping $\log x_i \to \frac{\log x_i - \mu_i}{\sigma_i}$, where now μ_i, σ_i denote the sample mean and standard deviation of the logged losses. This results in a standardized database of logged losses, $\left\{\frac{\log x_i - \mu_i}{\sigma_i}\right\}$. To get a database appropriate for a given state k it is natural to multiply each standardized logged loss by σ_k, add μ_k, and then exponentiate to get a database, $\left\{\exp(\sigma_k \frac{\log x_i - \mu_i}{\sigma_i} + \mu_k)\right\}$. The linear transformation, $\frac{\log x_i - \mu_i}{\sigma_i} \to \sigma_k \frac{\log x_i - \mu_i}{\sigma_i} + \mu_k$, results in a database that matches the mean and variance of the logged losses in state k, but upon exponentiation we lose this property. That is, the database, $\left\{\exp(\sigma_k \frac{\log x_i - \mu_i}{\sigma_i} + \mu_k)\right\}$, may not have the same mean and variance as the claims in state k. However, under reasonable conditions we can find μ, σ such that the database, $\left\{\exp(\sigma \frac{\log x_i - \mu_i}{\sigma_i} + \mu)\right\}$, will have the mean and variance in state k. We proceed now to establish this. We begin with a lemma.

Lemma 2 *Let $x_1, x_2, \ldots, x_n$ be a finite sequence of real numbers, not all equal, and let $\varphi : (0, \infty) \to \mathbb{R}$ by*

$$\varphi(t) = \frac{\left(\sum_{i=1}^{n} t^{x_i}\right)^2}{n \sum_{i=1}^{n} t^{2x_i}},$$

then

1. $\varphi(1) = 1$
2. φ is strictly increasing on $(0, 1)$ and strictly decreasing on $(1, \infty)$

3. $\lim_{t\to\infty}\varphi(t) = k/n$, where k is the number of i such that $x_i = \max\{x_j | 1 \le j \le n\}$.

Proof. We have $\varphi(1) = \frac{n^2}{n\cdot n} = 1$, thus proving item 1.

To prove item 2, first note that

$$\begin{aligned}
\frac{d\varphi}{dt} &= \frac{\left(\sum_{i=1}^n t^{2x_i}\right)\left(2\sum_{i=1}^n t^{x_i}\right)\left(\sum_{i=1}^n x_i t^{x_i-1}\right) - \left(\sum_{i=1}^n t^{x_i}\right)^2\left(\sum_{i=1}^n 2x_i t^{2x_i-1}\right)}{n\left(\sum_{i=1}^n t^{2x_i}\right)^2} \\
&= \frac{2\sum_{i=1}^n t^{x_i}}{n\left(\sum_{i=1}^n t^{2x_i}\right)^2}\left[\left(\sum_{i=1}^n t^{2x_i}\right)\left(\sum_{i=1}^n x_i t^{x_i-1}\right) - \left(\sum_{i=1}^n t^{x_i}\right)\left(\sum_{i=1}^n x_i t^{2x_i-1}\right)\right]
\end{aligned}$$

As the term $2\sum_{i=1}^n t^{x_i}/n\left(\sum_{i=1}^n t^{2x_i}\right)^2$ is positive, we note that, after relabelling indices for convenience, $\frac{d\varphi}{dt}$ has the same sign as

$$\begin{aligned}
\gamma(t) &= \left(\sum_{i=1}^n t^{2x_i}\right)\left(\sum_{j=1}^n x_j t^{x_j-1}\right) - \left(\sum_{j=1}^n t^{x_j}\right)\left(\sum_{i=1}^n x_i t^{2x_i-1}\right) \\
&= \sum_{1\le i,j\le n} x_j t^{2x_i+x_j-1} - \sum_{1\le i,j\le n} x_i t^{2x_i+x_j-1} \\
&= \sum_{1\le i,j\le n} (x_j - x_i) t^{2x_i+x_j-1} \\
&= \sum_{1\le i<j\le n} (x_j - x_i) t^{2x_i+x_j-1} + \sum_{1\le j<i\le n} (x_j - x_i) t^{2x_i+x_j-1} \\
&= \sum_{1\le i<j\le n} (x_j - x_i) t^{x_i} t^{x_i+x_j-1} + \sum_{1\le j<i\le n} (x_j - x_i) t^{x_i} t^{x_i+x_j-1} \\
&= \sum_{1\le i<j\le n} (x_j - x_i) t^{x_i} t^{x_i+x_j-1} + \sum_{1\le i<j\le n} (x_i - x_j) t^{x_j} t^{x_i+x_j-1} \\
&= \sum_{1\le i<j\le n} (x_j - x_i) t^{x_i} t^{x_i+x_j-1} - \sum_{1\le i<j\le n} (x_j - x_i) t^{x_j} t^{x_i+x_j-1} \\
&= \sum_{1\le i<j\le n} (x_j - x_i)(t^{x_i} - t^{x_j}) t^{x_i+x_j-1}.
\end{aligned}$$

Observe that for $t < 1$, the differences $x_j - x_i$ and $t^{x_i} - t^{x_j}$, not all of which are 0, have the same sign, which implies that $\gamma(t) > 0$. Similarly, for $t > 1$, those differences have opposite signs, hence $\gamma(t) < 0$, thus proving item 2.

To prove item 3, we sort and relabel the x_i as necessary so that $x_1 = \cdots = x_k = \max\{x_i\}$ and $x_i < x_1$ for $i > k$. We then find that

$$\varphi(t) = \frac{\left(\sum_{i=1}^{n} t^{x_i}\right)^2}{n\sum_{i=1}^{n} t^{2x_i}} = \frac{t^{2x_1}\left(\sum_{i=1}^{k} 1 + \sum_{i=k+1}^{n} t^{x_i - x_1}\right)^2}{nt^{2x_1}\left(\sum_{i=1}^{k} 1 + \sum_{i=k+1}^{n} t^{2(x_i - x_1)}\right)} = \frac{\left(k + \sum_{i=k+1}^{n} t^{x_i - x_1}\right)^2}{n\left(k + \sum_{i=k+1}^{n} t^{2(x_i - x_1)}\right)}.$$

Since $x_i - x_1 < 0$ for $i > k$ it follows that

$$\lim_{t\to\infty} \varphi(t) = \lim_{t\to\infty} \frac{\left(k + \sum_{i=k+1}^{n} t^{x_i - x_1}\right)^2}{n\left(k + \sum_{i=k+1}^{n} t^{2(x_i - x_1)}\right)} = \frac{\left(k + \sum_{i=k+1}^{n} 0\right)^2}{n\left(k + \sum_{i=k+1}^{n} 0\right)} = \frac{k^2}{nk} = \frac{k}{n}$$

as claimed. This completes the proof of item 3 and the lemma. ∎

Now interpreting $x_1, x_2, \ldots, x_n$ to be the standardized logged losses this lemma allows us to prove the following proposition which shows that standardization of logged losses, followed by a linear transformation and re-exponentiation does what we want under reasonable conditions.

Proposition 3 *Let $x_1, x_2, \ldots, x_n$ be a finite sequence of real numbers, not all equal, and let k be the number of i such that $x_i = \max\{x_j | 1 \leq j \leq n\}$. Then for any pair of positive real numbers, μ, σ, such that $\mu^2/(\mu^2+\sigma^2) > k/n$, there exists a unique pair of real numbers, m, s, with $s > 0$ such that the finite sequence, $e^{m+sx_1}, e^{m+sx_2}, \ldots, e^{m+sx_n}$, has mean μ and standard deviation σ. More precisely, if $y_i = e^{m+sx_i}$, then*

$$\mu = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \textit{and} \quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2.$$

Proof. From the lemma, there exists a unique $t > 1$ with $\varphi(t) = \frac{\mu^2}{\mu^2+\sigma^2}$. Observe that since $\sum_{i=1}^{n} t^{x_i} > 0$ we can define

$$s = \ln t > 0 \quad \text{and} \quad m = \ln\left(\frac{n\mu}{\sum_{i=1}^{n} t^{x_i}}\right).$$

Then setting $y_i = e^{m+sx_i}$ for $1 \leq i \leq n$, we have

$$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{e^m}{n}\sum_{i=1}^{n}(e^s)^{x_i} = \frac{e^m}{n}\sum_{i=1}^{n} t^{x_i} = \left(\frac{n\mu}{n\sum_{i=1}^{n} t^{x_i}}\right)\sum_{i=1}^{n} t^{x_i} = \mu.$$

We also have

$$\begin{aligned}\frac{n\mu^2}{\sum_{i=1}^{n} y_i^2} &= \frac{n\mu^2}{\sum_{i=1}^{n}(e^m t^{x_i})^2} = \frac{(n\mu)^2}{ne^{2m}\sum_{i=1}^{n} t^{2x_i}} \\ &= \frac{\left(e^m\sum_{i=1}^{n} t^{x_i}\right)^2}{ne^{2m}\sum_{i=1}^{n} t^{2x_i}} = \frac{\left(\sum_{i=1}^{n} t^{x_i}\right)^2}{n\sum_{i=1}^{n} t^{2x_i}} \\ &= \varphi(t) = \frac{\mu^2}{\mu^2+\sigma^2}\end{aligned}$$

which implies $\frac{1}{n}\sum_{i=1}^{n} y_i^2 = \mu^2 + \sigma^2$ and thus

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}(y_i-\mu)^2 &= \frac{1}{n}\sum_{i=1}^{n}(y_i^2 - 2\mu y_i + \mu^2) \\ &= \frac{1}{n}\sum_{i=1}^{n} y_i^2 - 2\mu\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) + \mu^2 \\ &= \mu^2 + \sigma^2 - 2\mu\mu + \mu^2 \\ &= \sigma^2.\end{aligned}$$

To prove uniqueness, let $\hat{m}, \hat{s}$ be another such pair, and set $\hat{y}_i = e^{\hat{m}+\hat{s}x_i}$ for $1 \leq i \leq n$. It follows that

$$\frac{\mu^2}{\mu^2+\sigma^2} = \frac{\left(\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i\right)^2}{\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i^2} = \frac{\left(e^{\hat{m}}\sum_{i=1}^{n}(e^{\hat{s}})^{x_i}\right)^2}{ne^{2\hat{m}}\sum_{i=1}^{n}(e^{\hat{s}})^{2x_i}} = \frac{\left(\sum_{i=1}^{n}(e^{\hat{s}})^{x_i}\right)^2}{n\sum_{i=1}^{n}(e^{\hat{s}})^{2x_i}} = \varphi(e^{\hat{s}}).$$

Since $\hat{s} > 0$ implies $e^{\hat{s}} > 1$, it follows that $e^{\hat{s}} = t = e^s$ and thus $\hat{s} = s$. Finally, we have

$$\hat{y}_i = e^{\hat{m}+\hat{s}x_i} = e^{\hat{m}+sx_i} = e^{\hat{m}-m+m+sx_i} = e^{\hat{m}-m}e^{m+sx_i} = e^{\hat{m}-m}y_i$$

for $1 \leq i \leq n$, which implies $\mu = \frac{1}{n}\sum_{i=1}^{n}\hat{y}_i = \frac{e^{\hat{m}-m}}{n}\sum_{i=1}^{n} y_i = e^{\hat{m}-m}\mu$. Since $\mu \neq 0$, it follows that $\hat{m} = m$ and the proof is complete. ∎

It is possible to generalize the previous result from a finite sample, $x_1, x_2, \ldots, x_r$ to a distribution with finite support. The argument mirrors that for the discrete case. As before, we begin with a lemma.

Lemma 4 *Let f be a continuous probability density on the finite interval, $[a,b]$, and let $\varphi : (0,\infty) \to \mathbb{R}$ by*

$$\varphi(t) = \frac{\left(\int\limits_a^b t^x f(x)dx\right)^2}{\int\limits_a^b t^{2x} f(x)dx},$$

then

1. $\varphi(1) = 1$
2. φ *is strictly increasing on* $(0,1)$ *and strictly decreasing on* $(1,\infty)$
3. $\lim_{t\to\infty} \varphi(t) = 0$.

Proof. We have $\varphi(1) = \frac{1^2}{1} = 1$, thus proving item 1.

To prove item 2, first note that

$$\begin{aligned}
\frac{d\varphi}{dt} &= \left[\left(\int_a^b t^{2x} f(x)dx\right)\left(2\int_a^b t^x f(x)dx\right)\left(\frac{d}{dt}\int_a^b t^x f(x)dx\right)\right. \\
&\quad \left. - \left(\int_a^b t^x f(x)dx\right)^2 \left(\frac{d}{dt}\int_a^b t^{2x} f(x)dx\right)\right] / \left(\int_a^b t^{2x} f(x)dx\right)^2 \\
&= \left[\left(\int_a^b t^{2x} f(x)dx\right)\left(2\int_a^b t^x f(x)dx\right)\left(\int_a^b x t^{x-1} f(x)dx\right)\right. \\
&\quad \left. - \left(\int_a^b t^x f(x)dx\right)^2 \left(\int_a^b 2x t^{2x-1} f(x)dx\right)\right] / \left(\int_a^b t^{2x} f(x)dx\right)^2 \\
&= \frac{2\int_a^b t^x f(x)dx}{\left(\int_a^b t^{2x} f(x)dx\right)^2}\left[\left(\int_a^b t^{2x} f(x)dx\right)\left(\int_a^b x t^{x-1} f(x)dx\right)\right. \\
&\quad \left. - \left(\int_a^b t^x f(x)dx\right)\left(\int_a^b x t^{2x-1} f(x)dx\right)\right]
\end{aligned}$$

As the term $2\int_a^b t^x f(x)dx / \left(\int_a^b t^{2x} f(x)dx\right)^2$ is positive, we note that $\frac{d\varphi}{dt}$ has the same sign as

$$\gamma(t) = \left(\int_a^b t^{2x} f(x)dx\right)\left(\int_a^b x t^{x-1} f(x)dx\right) - \left(\int_a^b t^x f(x)dx\right)\left(\int_a^b x t^{2x-1} f(x)dx\right).$$

Relabelling dummy variables for convenience, we get

$$\begin{aligned}
\gamma(t) &= \left(\int_a^b t^{2x} f(x)dx\right)\left(\int_a^b y t^{y-1} f(y)dy\right) - \left(\int_a^b t^{y} f(y)dy\right)\left(\int_a^b x t^{2x-1} f(x)dx\right) \\
&= \int_a^b \int_a^b y t^{2x+y-1} f(x)f(y)dxdy - \int_a^b \int_a^b x t^{2x+y-1} f(x)f(y)dxdy \\
&= \int_a^b \int_a^b (y-x) t^{2x+y-1} f(x)f(y)dxdy \\
&= \int_a^b \int_a^y (y-x) t^{2x+y-1} f(x)f(y)dxdy + \int_a^b \int_a^x (y-x) t^{2x+y-1} f(x)f(y)dydx \\
&= \int_a^b \int_a^y (y-x) t^{2x+y-1} f(x)f(y)dxdy - \int_a^b \int_a^y (y-x) t^{2y+x-1} f(x)f(y)dxdy \\
&= \int_a^b \int_a^y (y-x)(t^{2x+y-1} - t^{2y+x-1}) f(x)f(y)dxdy \\
&= \int_a^b \int_a^y (y-x)(t^{x} - t^{y}) t^{x+y-1} f(x)f(y)dxdy.
\end{aligned}$$

Observe that for $t < 1$, the differences $y - x$ and $t^x - t^y$ have the same sign, which implies that $\gamma(t) > 0$. Similarly, for $t > 1$, those differences have opposite signs, hence $\gamma(t) < 0$, thus proving item 2.

To prove item 3, first consider the case when $f(x) > 0$ for all $x \in [a, b]$. Since f is continuous on $[a, b]$, it is uniformly continuous on $[a, b]$. Thus, for any $\epsilon > 0$, there is a partition

$$[a, b] = \bigcup_{i=1}^{n} [a_i, b_i] \text{ with } a = a_1, a_i < b_i = a_{i+1}, b_n = b$$

such that

$$x_1, x_2 \in [a_i, b_i] \Longrightarrow |f(x_1) - f(x_2)| \le \epsilon.$$

Let $\alpha = \min\{f(x) | x \in [a, b]\} > 0$ and let $\alpha_i = \min\{f(x) | x \in [a_i, b_i]\}$, then $\{f(x) | x \in [a_i, b_i]\} \subseteq [\alpha_i, \alpha_i + \epsilon]$. We claim that

$$\lim_{t\to\infty} \frac{\int_{a_i}^{b_i} t^x dx \int_{a_j}^{b_j} t^x dx}{\int_a^b t^{2x} dx} = 0.$$

To see this assume that $b_i \le a_j$, then

$$\begin{aligned}
\frac{\int_{a_i}^{b_i} t^x dx \int_{a_j}^{b_j} t^x dx}{\int_a^b t^{2x} dx} &= \frac{\frac{t^x}{\ln t}\Big|_{a_i}^{b_i} \frac{t^x}{\ln t}\Big|_{a_j}^{b_j}}{\frac{t^{2x}}{2\ln t}\Big|_a^b} \\
&= \frac{2}{\ln t}\left(\frac{\left(t^{b_i} - t^{a_i}\right)\left(t^{b_j} - t^{a_j}\right)}{\left(t^{2b} - t^{2a}\right)}\right) \\
&= \frac{2}{\ln t}\left(\frac{\left(t^{b_i - b_j} - t^{a_i - b_j}\right)\left(1 - t^{a_j - b_j}\right)}{\left(t^{2(b-b_j)} - t^{2(a-b_j)}\right)}\right).
\end{aligned}$$

Thus

$$\lim_{t\to\infty} \frac{\int_{a_i}^{b_i} t^x dx \int_{a_j}^{b_j} t^x dx}{\int_a^b t^{2x} dx} = \lim_{t\to\infty} \frac{2}{\ln t}\left(\frac{(0-0)(1-0)}{\left(t^{2(b-b_j)} - 0\right)}\right) = 0$$

as claimed.

From what we've just claimed, for every i and j, there exists a $t_{i,j}$ such that for all $t > t_{i,j}$ we have

$$\frac{\int_{a_i}^{b_i} t^x dx \int_{a_j}^{b_j} t^x dx}{\int_a^b t^{2x} dx} \le \frac{\alpha\epsilon}{n^2(\alpha_i + \epsilon)(\alpha_j + \epsilon)}.$$

Then for all $t > t_{i,j}$ we have

$$\begin{aligned}
\frac{\int_{a_i}^{b_i} t^x f(x) dx \int_{a_j}^{b_j} t^x f(x) dx}{\int_a^b t^{2x} f(x) dx} &\le \frac{(\alpha_i + \epsilon)(\alpha_j + \epsilon)\int_{a_i}^{b_i} t^x dx \int_{a_j}^{b_j} t^x dx}{\alpha \int_a^b t^{2x} dx} \\
&\le \frac{(\alpha_i + \epsilon)(\alpha_j + \epsilon)}{\alpha} \frac{\alpha\epsilon}{n^2(\alpha_i + \epsilon)(\alpha_j + \epsilon)} = \frac{\epsilon}{n^2}.
\end{aligned}$$

Thus for $t > \max\{t_{i,j}\}$, it follows that

$$\begin{aligned}
\varphi(t) &= \frac{\left(\int_a^b t^x f(x) dx\right)^2}{\int_a^b t^{2x} f(x) dx} = \frac{\left(\sum_{i=1}^n \int_{a_i}^{b_i} t^x f(x) dx\right)^2}{\int_a^b t^{2x} f(x) dx} \\
&= \sum_{1\le i,j\le n} \frac{\int_{a_i}^{b_i} t^x f(x) dx \int_{a_j}^{b_j} t^x f(x) dx}{\int_a^b t^{2x} f(x) dx} \le \sum_{1\le i,j\le n} \frac{\epsilon}{n^2} = \epsilon.
\end{aligned}$$

Thus, $\lim_{t\to\infty} \varphi(t) = 0$ in the case when $f(x) > 0$ for all $x \in [a, b]$. Finally, if we set $g(x) = \frac{f(x)+1}{b-a+1}$ then g is a positive, continuous probability density function on $[a, b]$ and we have

$$
\begin{aligned}
0 &= (b-a+1)\lim_{t\to\infty} \frac{\left(\int_a^b t^x g(x)dx\right)^2}{\int_a^b t^{2x} g(x)dx} \\
&= (b-a+1)\lim_{t\to\infty} \frac{\left(\int_a^b t^x (\frac{f(x)+1}{b-a+1})dx\right)^2}{\int_a^b t^{2x} (\frac{f(x)+1}{b-a+1})dx} \\
&= \lim_{t\to\infty} \frac{\left(\int_a^b t^x (f(x)+1)dx\right)^2}{\int_a^b t^{2x} (f(x)+1)dx} = \lim_{t\to\infty} \frac{\left(\int_a^b t^x f(x)dx + \int_a^b t^x dx\right)^2}{\int_a^b t^{2x} f(x)dx + \int_a^b t^{2x} dx} \\
&= \lim_{t\to\infty} \frac{\left(\int_a^b t^x f(x)dx + \left(\frac{t^b - t^a}{\ln t}\right)\right)^2}{\int_a^b t^{2x} f(x)dx + \left(\frac{t^{2b}-t^{2a}}{\ln t}\right)} \\
&= \lim_{t\to\infty} \frac{\left(\int_a^b t^{x-b} f(x)dx + \left(\frac{1-t^{(a-b)}}{\ln t}\right)\right)^2}{\int_a^b t^{2(x-b)} f(x)dx + \left(\frac{1-t^{2(a-b)}}{\ln t}\right)} \\
&= \lim_{t\to\infty} \frac{\left(\int_a^b t^{x-b} f(x)dx\right)^2}{\int_a^b t^{2(x-b)} f(x)dx} = \lim_{t\to\infty} \frac{\left(\int_a^b t^x f(x)dx\right)^2}{\int_a^b t^{2x} f(x)dx} \\
&= \lim_{t\to\infty} \varphi(t).
\end{aligned}
$$

This completes the proof of the lemma. ∎

Now interpreting f to be the density of the standardized logged losses the lemma allows us to prove the analog of Proposition 3 in the continuous case, namely that there is a linear transformation of the standardized logged losses such that after re-exponentiation we get the desired mean and variance.

Proposition 5 *Let f be a continuous probability density on the finite interval, $[a,b]$. Then for any pair of positive real numbers, μ, σ, there exists a unique pair of real numbers, m, s, with $s > 0$ such that*

$$g(y) = \frac{1}{sy} f\left(\frac{\ln y - m}{s}\right)$$

is a continuous probability density on $[e^{m+as}, e^{m+bs}]$ with mean μ and standard deviation σ.

Proof. From the lemma, there exists a unique $t > 1$ with $\varphi(t) = \frac{\mu^2}{\mu^2+\sigma^2}$. Observe that $\varphi(t) > 0$ implies $\int_a^b t^x f(x)dx > 0$, thus we can define

$$s = \ln t > 0 \quad \text{and} \quad m = \ln\left(\frac{\mu}{\int_a^b t^x f(x)dx}\right).$$

Let $c = e^{m+as}$ and $d = e^{m+bs}$. We also introduce the change of variable $x = \frac{\ln y - m}{s} \Leftrightarrow y = e^{m+sx}$, hence $\frac{dy}{dx} = ys$, which implies $dy = ysdx$. Then

$$\int_c^d g(y)dy = \int_a^b \frac{1}{sy} f(x)ysdx = \int_a^b f(x)dx = 1.$$

Further, we have

$$\begin{aligned}
\int_c^d yg(y)dy &= \int_a^b y\frac{1}{sy} f(x)ysdx = \int_a^b yf(x)dx \\
&= \int_a^b e^{m+sx} f(x)dx = e^m \int_a^b (e^s)^x f(x)dx \\
&= e^m \int_a^b t^x f(x)dx = \left(\frac{\mu}{\int_a^b t^x f(x)dx}\right) \int_a^b t^x f(x)dx = \mu.
\end{aligned}$$

Since f is continuous, g is continuous as well and we have shown that g is a continuous probability density function on $[c,d] = [e^{m+as}, e^{m+bs}]$ with mean

μ. As in the discrete case, we note that

$$\begin{aligned}
\frac{\mu^2}{\int_c^d y^2 g(y)dy} &= \frac{\left(\int_c^d y g(y)dy\right)^2}{\int_a^b (e^{m+sx})^2 f(x)dx} = \frac{\left(\int_a^b e^{m+sx} f(x)dx\right)^2}{\int_a^b (e^m t^x)^2 f(x)dx} \\
&= \frac{\left(e^m \int_a^b t^x f(x)dx\right)^2}{e^{2m} \int_a^b t^{2x} f(x)dx} = \frac{\left(\int_a^b t^x f(x)dx\right)^2}{\int_a^b t^{2x} f(x)dx} \\
&= \varphi(t) = \frac{\mu^2}{\mu^2+\sigma^2},
\end{aligned}$$

which implies $\int_c^d y^2 g(y)dy = \mu^2 + \sigma^2$. Thus

$$\begin{aligned}
\int_c^d (y-\mu)^2 g(y)dy &= \int_c^d y^2 g(y)dy - 2\mu \int_c^d y g(y)dy + \mu^2 \int_c^d g(y)dy \\
&= \mu^2 + \sigma^2 - 2\mu^2 + \mu^2 \\
&= \sigma^2.
\end{aligned}$$

To prove uniqueness, let $\hat{m}, \hat{s}$ be another such pair, and let

$$\hat{g}(y) = \frac{1}{\hat{s}y} f\left(\frac{\ln y - \hat{m}}{\hat{s}}\right) \text{ for } y \in \left[\hat{c}, \hat{d}\right] = \left[e^{\hat{m}+a\hat{s}}, e^{\hat{m}+b\hat{s}}\right].$$

From a similar change of variable as above, it follows that

$$\begin{aligned}
\frac{\mu^2}{\mu^2+\sigma^2} &= \frac{\left(\int_{\hat{c}}^{\hat{d}} y \hat{g}(y)dy\right)^2}{\int_{\hat{c}}^{\hat{d}} y^2 \hat{g}(y)dy} = \frac{\left(\int_a^b e^{\hat{m}+\hat{s}x} f(x)dx\right)^2}{\int_a^b e^{2(\hat{m}+\hat{s}x)} f(x)dx} \\
&= \frac{\left(e^{\hat{m}} \int_a^b e^{\hat{s}x} f(x)dx\right)^2}{e^{2\hat{m}} \int_a^b e^{2\hat{s}x} f(x)dx} = \frac{\left(\int_a^b \left(e^{\hat{s}}\right)^x f(x)dx\right)^2}{\int_a^b \left(e^{\hat{s}}\right)^{2x} f(x)dx} = \varphi(e^{\hat{s}}).
\end{aligned}$$

Since $\hat{s} > 0$, it follows that $e^{\hat{s}} > 1$, implying that $e^{\hat{s}} = t = e^s$ and thus $\hat{s} = s$. Finally, we have

$$\mu = \int_a^b e^{\hat{m}+\hat{s}x} f(x)dx = \int_a^b e^{\hat{m}+sx} f(x)dx = e^{\hat{m}-m} \int_a^b e^{m+sx} f(x)dx = e^{\hat{m}-m}\mu.$$

Since $\mu > 0$, it follows that $\hat{m} = m$, and the proof is complete. ∎

Remark 6 *What holds for the continuous case with infinite support is not so straightforward. For example, letting* $f(x) = \frac{e^{-\frac{x}{\theta}}}{\theta}$ *be the exponential density on* $[0, \infty)$*, the interested reader can readily verify that the condition* $0 < \frac{\sigma}{\mu} < \sqrt{1+\sqrt{2}}$ *is both necessary and sufficient for the existence of positive numbers* m *and* s *as in the proposition. More generally, if* $f(x)$ *is a probability density on* $[a, \infty)$ *with moment generating function* $M_X(t) = \int_a^\infty e^{tx} f(x)dx$ *and we are given a target mean* μ *and standard deviation* σ*, then one suggestion is to first try to determine* s *by solving the following implicit equation for the target coefficient of variation:*

$$\frac{\sigma}{\mu} = \frac{\sqrt{M_X(2s) - M_X(s)^2}}{M_X(s)}$$

and, if successful, determine m *from:*

$$m = \ln\left(\frac{\mu}{M_X(s)}\right).$$

C.2 The Power Transform

A more subtle way to transform claims is with a power transform, $x \rightarrow ax^b$. With $a = 1/\mu$ and $b = 1$ we can see that the power transform generalizes mean normalization. With logarithmic standardization we first log the data, then standardize, and then re-exponentiate: $x \rightarrow \log x \rightarrow \frac{\log x - \mu}{\sigma} \rightarrow \exp(\frac{\log x - \mu}{\sigma})$. But $\exp(\frac{\log x - \mu}{\sigma}) = e^{-\mu/\sigma} x^{1/\sigma}$ and so the power transform generalizes logarithmic standardization as well. Thus the power transform could potentially outperform both mean normalization and logarithmic standardization. In addition, with the power transform there is no need to log the losses and then re-exponentiate. The moments are matched in dollar space rather than in log space. The idea is to choose a and b so that the transformed losses from one state match the mean and variance of the losses from another state. In this way we can use the out of state losses to build an expanded database for each state. We now prove, under reasonable conditions, that it is possible to choose a and b in the power transform so that the transformed losses from one state do indeed match the mean and variance of another state.

Proposition 7 *Let $x_1, x_2, \ldots, x_n$ be a finite sequence of positive real numbers, not all equal and and let k be the number of i such that $x_i = \max\{x_j | 1 \leq j \leq n\}$. Then given $\mu > 0$ and $\gamma \in \left[0, \sqrt{\frac{n-k}{k}}\right)$, there exist unique constants $a > 0$ and $b \geq 0$ such that the database $\{ax_i^b\}$ has mean μ and CV γ.*

Proof. If $\gamma = 0$ then we must take $b = 0$ and $a = \mu$, and the result holds. So assume $\gamma > 0$. Let $\sigma = \gamma\mu > 0$ and set

$$z_i = \ln x_i \text{ for } 1 \leq j \leq n.$$

Then clearly k is the number of i such that $z_i = \max\{z_j | 1 \leq j \leq n\}$. We have:

$$\begin{aligned} \gamma &\in \left(0, \sqrt{\frac{n-k}{k}}\right) \\ &\Leftrightarrow \frac{\sigma^2}{\mu^2} < \frac{n}{k} - 1 \\ &\Leftrightarrow \frac{\mu^2 + \sigma^2}{\mu^2} < \frac{n}{k} \\ &\Leftrightarrow \mu^2/(\mu^2 + \sigma^2) > k/n \end{aligned}$$

and so by Proposition 3 there is a unique pair of real numbers m, s with $s > 0$ such that the finite sequence, $e^{m+sz_1}, e^{m+sz_2}, \ldots, e^{m+sz_n}$, has mean μ and standard deviation σ. Letting $a = e^m$ and $b = s$ we have:

$$e^{m+sz_i} = e^m \left(e^{z_i}\right)^s = e^m \left(e^{\ln x_i}\right)^s = ax_i^b \text{ for } 1 \leq j \leq n$$

and the existence of the constants a and b is proved. Uniqueness of $a = e^m$ and $b = s$ follows from the uniqueness of m and s, and the proof is complete. ■

D Excess Ratio Functions

We collect here some facts about excess ratio functions. We show how to recover the distribution function from the excess ratio function, give a characterization of excess ratio functions, and discuss the mixed exponential case. We start with some basic definitions and results.

Definition 8 *A random variable X is a loss variable if it is nonnegative valued, has finite nonzero mean, and has a density f that is continuous when restricted to $[0, +\infty)$. We denote by F the distribution function of X. The survival function of X is $S = 1 - F$. The excess ratio function of X is given by $R(r) = \int_r^\infty (x - r) f(x) dx / E[X]$ for $r \geq 0$. We denote by $\widehat{F}$ the function given by $\widehat{F}(r) = \int_0^r x f(x) dx / E[X]$. We use subscripts on F, $\widehat{F}$, S, and R when necessary to indicate dependence on X.*

The following proposition expresses the excess ratio function in terms of F and $\widehat{F}$.

Proposition 9 *Let X be a loss variable with mean μ, then*

$$R(r) = 1 - \widehat{F}(r) - \frac{r}{\mu}[1 - F(r)].$$

Proof. From the definition of $R(r)$ we have

$$\begin{aligned} R(r) &= \frac{1}{\mu} \int_r^\infty (x - r) f(x) dx \\ &= \frac{1}{\mu} \left[\int_r^\infty x f(x) dx - r \int_r^\infty f(x) dx \right] \\ &= \frac{1}{\mu} \left[\mu - \int_0^r x f(x) dx - r S(r) \right] \\ &= 1 - \frac{1}{\mu} \int_0^r x f(x) dx - \frac{r}{\mu} S(r) \\ &= 1 - \widehat{F}(r) - \frac{r}{\mu}[1 - F(r)]. \end{aligned}$$

■

It is well known (see, for example, Billingsley [1], page 282) that the mean of a nonnegative random variable, X, can be expressed in terms of its survival function as $E[X] = \int_0^\infty S(x) dx$. It is easy to see that a similar result also holds for excess ratio functions.

Proposition 10 *Let X be a loss variable with survival function S and excess ratio function R, then*

$$R(r) = \frac{\int_r^\infty S(x) dx}{\int_0^\infty S(x) dx}.$$

Proof. Let X have density f, then noting that $S'(x) = -f(x)$ and using integration by parts, we have

$$\begin{aligned}\int_r^\infty S(x)dx &= xS(x)|_r^\infty + \int_r^\infty xf(x)dx \\ &= -rS(r) + \int_r^\infty xf(x)dx \\ &= -r\int_r^\infty f(x)dx + \int_r^\infty xf(x)dx \\ &= \int_r^\infty (x-r)f(x)dx,\end{aligned}$$

where the second equality follows as $xS(x) = x\int_x^\infty f(y)dy \leq \int_x^\infty yf(y)dy \to 0$ as $x \to \infty$ since X has finite mean. Thus $R(r) = \int_r^\infty (x-r)f(x)dx/E[X] = \int_r^\infty S(x)dx/\int_0^\infty S(x)dx$. ■

Survival functions and excess ratio functions share several elementary properties given in the next proposition.

Proposition 11 *If g is a survival function of a loss variable or an excess ratio function then*

1. $g(0) = 1$ *(and $g(x) = 1$ for $x < 0$ if g is a survival function)*
2. *g is non increasing*
3. $\lim_{x\to\infty} g(x) = 0$

The following proposition shows how to recover the distribution function from the excess ratio function. Thus the excess ratio function characterizes a loss distribution and so there is no loss of information in considering excess ratio functions rather than densities or distribution functions.

Proposition 12 *Let X be a loss variable with survival function S, and excess ratio function R, then $\frac{d}{dr}R(r) = -S(r)/E[X]$. Further, if we set $g(x) = S(x)/E[X]$ then $R(r) = \int_r^\infty g(x)dx$.*

Proof. For the first assertion, we have

$$\begin{aligned}\frac{d}{dr}R(r) &= \frac{1}{E(X)}\frac{d}{dr}\int_r^\infty (x-r)f(x)dx \\ &= \frac{1}{E(X)}\left[\frac{d}{dr}\int_r^\infty xf(x)dx - \frac{d}{dr}r\int_r^\infty f(x)dx\right].\end{aligned}$$

By the Fundamental Theorem of Calculus, we have $\frac{d}{dr}\int_r^\infty xf(x)dx = -rf(r)$ and $\frac{d}{dr}\int_r^\infty f(x) = -f(r)$, thus

$$\frac{d}{dr}R(r) = \frac{1}{E(X)}\left[-rf(r) + rf(r) - \int_r^\infty f(x)dx\right] = \frac{-S(r)}{E(X)}$$

For the second assertion, by Proposition 10 we have

$$R(r) = \int_r^\infty S(x)dx \bigg/ \int_0^\infty S(x)dx = \int_r^\infty \frac{S(x)}{E(X)}dx = \int_r^\infty g(x)dx$$

■

This proposition also shows that the excess ratio function of a loss variable X is also the survival function of another random variable with density $S(x)/E[X]$. We next give characterizations of survival functions and excess ratio functions.

Proposition 13 *Let* $g : [0, +\infty) \to \mathbb{R}$ *be differentiable with* g' *continuous,* $g(0) = 1$*, and* $\lim_{x\to\infty} g(x) = 0$*, and let*

$$\widetilde{g}(x) = \begin{cases} 1 & \textit{if } x < 0 \\ g(x) & \textit{if } x \geq 0 \end{cases},$$

then $\widetilde{g}$ *is the survival function of some nonnegative random variable* X *with density,* f*, that is continuous when restricted to* $[0, +\infty)$ *if and only if* $g' \leq 0$.

Proof. Suppose $\widetilde{g} = S_X$ for some nonnegative random variable X with density, f, that is continuous when restricted to $[0, +\infty)$. Then for $x \geq 0$

$$g(x) = \widetilde{g}(x) = S_X(x) = \int_x^\infty f(y)dy,$$

and so by the Fundamental Theorem of Calculus $g'(x) = -f(x) \leq 0$.

Conversely, suppose $g' \leq 0$ and define

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ -g'(x) & \text{if } x \geq 0 \end{cases},$$

then f restricted to $[0, +\infty)$ is continuous and

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{\infty} -g'(x)dx = -g(x)|_0^\infty = -\lim_{x\to\infty} g(x) + g(0) = -0 + 1 = 1,$$

so f is a probability density function of some nonnegative random variable X. For $x < 0$ we have $S_X(x) = 1 = \widetilde{g}(x)$ and for $x \geq 0$

$$S_X(x) = \int_{x}^{\infty} f(y)dy = \int_{x}^{\infty} -g'(y)dy = -g(y)|_x^\infty = -\lim_{y\to\infty} g(y) + g(x) = g(x) = \widetilde{g}(x).$$

Thus $\widetilde{g} = S_X$. ■

Proposition 14 *Let $g : [0, +\infty) \to \mathbb{R}$ be twice differentiable with g'' continuous, $g(0) = 1$, and $\lim_{x\to\infty} g(x) = 0$, then g is the excess ratio function of some loss variable if and only if $g' \leq 0$ and $g'' \geq 0$.*

Proof. Suppose $g = R_X$ for some loss variable X with density f, survival function S, and mean μ. Then by Proposition 12,

$$g' = -S/\mu \leq 0 \text{ and } g'' = -S'/\mu = f/\mu \geq 0.$$

Conversely, suppose $g' \leq 0$ and $g'' \geq 0$. Since $g'' \geq 0$ we know that g' is non decreasing. So if $g'(0) = 0$ then $g'(x) = 0$ for all x as $g' \leq 0$. This would imply that g is constant and so $g(x) = g(0) = 1$ for all x, but this contradicts our hypothesis that $\lim_{x\to\infty} g(x) = 0$. Thus we must have $g'(0) < 0$. Observe also that

$$\int_0^\infty |g'(x)|\, dx = -\int_0^\infty g'(x)dx = -g(x)|_0^\infty = -0 + g(0) = 1,$$

and so $\lim_{x\to\infty} g'(x) = 0$. If we let

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ -\frac{1}{g'(0)} g''(x) & \text{if } x \geq 0 \end{cases}$$

then $f \geq 0$ and f is continuous when restricted to $[0, +\infty)$. Further

$$\int_0^\infty f(x)dx = \left(-\frac{1}{g'(0)}\right) g'(x)|_0^\infty = \left(-\frac{1}{g'(0)}\right) \lim_{x\to\infty} g'(x) + \frac{g'(0)}{g'(0)} = -0+1 = 1.$$

Thus f is a density function for some nonnegative random variable X. Since for $t \geq 0$

$$S_X(t) = \int_t^\infty f(x)dx = \left(-\frac{1}{g'(0)}\right) g'(x)|_t^\infty = \left(-\frac{1}{g'(0)}\right) \lim_{x\to\infty} g'(x) + \frac{g'(t)}{g'(0)} = \frac{g'(t)}{g'(0)},$$

it follows that

$$\begin{aligned} E(X) &= \int_0^\infty S_X(t)dt = \int_0^\infty \frac{g'(t)}{g'(0)}dt = \left(\frac{1}{g'(0)}\right) g(x)|_0^\infty \\ &= \left(\frac{1}{g'(0)}\right) \lim_{x\to\infty} g(x) - \frac{g(0)}{g'(0)} = 0 - \frac{1}{g'(0)} = -\frac{1}{g'(0)}. \end{aligned}$$

Thus, $0 < E(X) = -\frac{1}{g'(0)} < \infty$ and so X is a loss variable. Finally, by Proposition 10 we have

$$R_X(t) = \frac{\int_t^\infty S_X(x)dx}{\int_0^\infty S_X(x)dx} = \frac{\int_t^\infty g'(x)dx}{\int_0^\infty g'(x)dx} = \frac{g(x)|_t^\infty}{g(x)|_0^\infty} = \frac{0 - g(t)}{0 - g(0)} = g(t).$$

■

We can now characterize excess ratio functions in terms of survival functions.

Proposition 15 *Excess ratio functions are exactly the restrictions to $[0, +\infty)$ of survival functions of nonnegative random variables with densities that when restricted to $[0, +\infty)$ have nonpositive, continuous derivatives.*

Proof. Let $g = R_X$ be an excess ratio function of a loss variable X. Then by Proposition 14, $g' \leq 0$ and $g'' \geq 0$. Proposition 13 then implies there is a nonnegative random variable Y such that

$$S_Y(x) = \begin{cases} 1 & \text{if } x < 0 \\ g(x) & \text{if } x \geq 0 \end{cases},$$

and Y has a density function, f, that is continuous when restricted to $[0, +\infty)$. For $x \geq 0$ we have $g(x) = \int_x^\infty f(y)dy$ and so $g'(x) = -f(x)$, which implies that $f' = -g'' \leq 0$ and f' is continuous as g'' is continuous.

Conversely, let X be a nonnegative random variable with a density function, f, that when restricted to $[0, +\infty)$ has a continuous derivative and $f' \leq 0$. Let $g : [0, +\infty) \to \mathbb{R}$ by

$$g(x) = \int_x^\infty f(y)dy.$$

Then $g' = -f \leq 0$, which implies that $g'' = -f' \geq 0$. Then by Proposition 14, g is the excess ratio function of some loss variable. ■

In the exponential case things are particularly simple as the next proposition shows.

Proposition 16 *Let* $f(x) = \frac{1}{m}e^{-x/m}$ *be an exponential density, then* $R(x) = S(x) = e^{-x/m}$*. That is, for an exponential distribution the excess ratio function is the same as the survival function.*

Proof. This follows directly from applying Definition 8 and using integration by parts. ■

For finite mixtures we have the following proposition.

Proposition 17 *Let* $f_1, f_2, \ldots, f_n$ *be densities with corresponding excess ratio functions* $R_1, R_2, \ldots, R_n$ *and means* $\mu_1, \mu_2, \ldots, \mu_n$*. Then given weights* $w_i \in (0, 1)$ *with* $\sum w_i = 1$*, the mixed density* $f = w_1 f_1 + w_2 f_2 + \cdots + w_n f_n$ *has excess ratio function*

$$R = \widehat{w}_1 R_1 + \widehat{w}_2 R_2 + \cdots + \widehat{w}_n R_n,$$

where $\widehat{w}_i = w_i \mu_i / \mu$ *and* μ *is the mean of the mixed distribution.*

Proof. From the definition of the excess ratio function, we have

$$\begin{aligned} R(r) &= \frac{1}{\mu}\int_r^\infty (x - r) f(x) dx \\ &= \frac{1}{\mu}\int_r^\infty (x - r)\left[\sum\nolimits_{i=1}^n w_i f_i(x)\right] dx \\ &= \sum\nolimits_{i=1}^n \frac{w_i}{\mu}\int_r^\infty (x - r) f_i(x) dx \\ &= \sum\nolimits_{i=1}^n \left(\frac{w_i \mu_i}{\mu}\right)\frac{1}{\mu_i}\int_r^\infty (x - r) f_i(x) dx \\ &= \sum\nolimits_{i=1}^n \widehat{w}_i R_i(r). \end{aligned}$$

■

Corollary 18 *If $f(x) = \sum_{i=1}^{n} w_i \frac{1}{m_i} e^{-x/m_i}$ is a finite mixed exponential density, then its excess ratio function is given by*

$$R(x) = \frac{\sum w_i m_i e^{-x/m_i}}{\sum w_i m_i}.$$

E Splicing Loss Distributions

We start with a loss variable, X (see Definition 8). The interpretation is that this represents the empirical losses. We then choose a point $l > 0$, such that $\Pr(X > l) > 0$ and $\Pr(X = l) = 0$. The point l, is called the splice point because we want to rely on X for claims less than l, but we want to splice on a distribution for claims larger than l. We let $Y = X - l$ conditional on $X \geq l$. That is, we truncate and shift X. More formally, if $X : \Omega \to [0, +\infty)$ then let $\Omega_0 = \{\omega \in \Omega | X(\omega) \geq l\}$ and define $Y : \Omega_0 \to [0, +\infty)$ by $Y(\omega) = X(\omega) - l$. The following proposition expresses the survival function, the density, and expected value of Y in terms of X.

Proposition 19 *Let X be a loss variable and let l be the splice point as above, then*

1. $S_Y(r - l) = \frac{1 - F_X(r)}{1 - F_X(l)}$ for $r \geq l$,
2. $f_Y(r - l) = \frac{f_X(r)}{1 - F_X(l)}$ for $r \geq l$, and
3. $E[Y] = \frac{E[X] R_X(l)}{1 - F_X(l)}$.

Proof. To prove item 1 we note first that $\Pr(X \geq l) = \Pr(X > l)$, then for $r \geq l$ we have

$$\begin{aligned} S_Y(r - l) &= \Pr(Y > r - l) = \Pr(X - l > r - l | X \geq l) \\ &= \Pr(X > r | X \geq l) = \frac{S_X(r)}{S_X(l)} = \frac{1 - F_X(r)}{1 - F_X(l)}. \end{aligned}$$

For item 2 we note that

$$F_Y(r - l) = 1 - S_Y(r - l) = 1 - \frac{1 - F_X(r)}{1 - F_X(l)} = \frac{F_X(r) - F_X(l)}{1 - F_X(l)}.$$

Then

$$f_Y(r-l) = \frac{d}{dr}\left[\frac{F_X(r)-F_X(l)}{1-F_X(l)}\right] = \frac{f_X(r)}{1-F_X(l)}.$$

For item 3 we have,

$$\begin{aligned} E[Y] &= \int_0^\infty y f_Y(y)dy \\ &= \int_l^\infty (r-l) f_Y(r-l)dr \\ &= \int_l^\infty (r-l)\left(\frac{f_X(r)}{1-F_X(l)}\right)dr \\ &= \frac{E[X]}{1-F_X(l)} \frac{\int_l^\infty (r-l) f_X(r)dr}{E[X]} \\ &= \frac{E[X]R_X(l)}{1-F_X(l)}, \end{aligned}$$

completing the proof. ■

We want to fit an excess ratio function (see Definition 8), R_0, from a mixed exponential distribution to R_Y. More precisely, we want to replace the empirical loss variable, X, with a loss variable $\widetilde{X}$ such that if $\widetilde{Y} = \widetilde{X} - l$ conditional on $\widetilde{X} \geq l$ then

1. $f_{\widetilde{X}}(x) = f_X(x)$ for $x \leq l$
2. $R_{\widetilde{Y}} = R_0$.

We now derive the distribution function, the probability density function, and the excess ratio function of the spliced distribution $\widetilde{X}$.

Proposition 20 *The distribution function of the spliced random variable $\widetilde{X}$ is given by*

$$F_{\widetilde{X}}(r) = \begin{cases} F_X(r) & \text{if } r \leq l \\ 1-[1-F_X(l)]S_{\widetilde{Y}}(r-l) & \text{if } r > l \end{cases}.$$

Proof. For $r \leq l$, we have $f_{\widetilde{X}}(x) = f_X(x)$. Thus

$$F_{\widetilde{X}}(r) = \int_0^r f_{\widetilde{X}}(x)dx = \int_0^r f_X(x)dx = F_X(r).$$

From this and Proposition 19, we see that for $r > l$

$$S_{\tilde{Y}}(r-l) = \frac{1 - F_{\tilde{X}}(r)}{1 - F_{\tilde{X}}(l)} = \frac{1 - F_{\tilde{X}}(r)}{1 - F_X(l)}$$

and therefore

$$F_{\tilde{X}}(x) = 1 - [1 - F_X(l)]S_{\tilde{Y}}(r-l).$$

■

This allows us to determine the distribution function of $\widetilde{X}$ since we know the empirical distribution F_X and our assumption that $R_{\widetilde{Y}} = R_0$ determines the distribution of $\widetilde{Y}$ as well by Proposition 12. We have thus shown that the following two conditions

1. $f_{\widetilde{X}}(x) = f_X(x)$ for $x \leq l$
2. $R_{\widetilde{Y}} = R_0$,

uniquely determine a random variable $\widetilde{X}$. What we have not shown is that the above two conditions are consistent, i.e. that there exists a random variable, $\widetilde{X}$, that satisfies them. We do this by working with the density of $\widetilde{X}$ and show that $\widetilde{X}$ is a loss variable as well.

Proposition 21 *The density of $\widetilde{X}$ is given by*

$$f_{\widetilde{X}}(r) = \begin{cases} f_X(r) & \text{if } r \leq l \\ [1 - F_X(l)]f_{\widetilde{Y}}(r-l) & \text{if } r > l \end{cases}.$$

and this defines a valid density function of a loss variable with mean given by

$$E[\widetilde{X}] = E[X]\widehat{F}_X(l) + S_X(l)\left(E[\widetilde{Y}] + l\right).$$

Proof. Item 1 of the definition of $\tilde{X}$ ensures that $f_{\tilde{X}}(r) = f_X(r)$ for $r \leq l$. For $r > l$, we have from Proposition 20 that $F_{\widetilde{X}}(r) = 1 - [1 - F_X(l)]S_{\widetilde{Y}}(r-l)$ and thus

$$f_{\tilde{X}}(r) = [1 - F_X(l)]f_{\widetilde{Y}}(r-l).$$

It remains to show that $f_{\tilde{X}}$ is a valid density function. To show this, we compute

$$\begin{aligned}\int_0^\infty f_{\tilde{X}}(r)dr &= \int_0^l f_X(r)dr + [1 - F_X(l)]\int_l^\infty f_{\tilde{Y}}(r-l)dr \\ &= \int_0^l f_X(r)dr + [1 - F_X(l)]\int_0^\infty f_{\tilde{Y}}(r)dr \\ &= F_X(l) + [1 - F_X(l)] = 1.\end{aligned}$$

From $f_{\tilde{X}}$ we can compute the mean of $\widetilde{X}$.

$$\begin{aligned}E[\widetilde{X}] &= \int_0^\infty x f_{\tilde{X}}(x)dx = \int_0^l x f_X(x)dx + S_X(l)\int_l^\infty x f_{\tilde{Y}}(x-l)dx \\ &= \int_0^l x f_X(x)dx + S_X(l)\int_0^\infty (x+l) f_{\tilde{Y}}(x)dx \\ &= \int_0^l x f_X(x)dx + S_X(l)\left(\int_0^\infty x f_{\tilde{Y}}(x)dx + l\right) \\ &= E[X]\widehat{F}_X(l) + S_X(l)\left(E[\widetilde{Y}] + l\right).\end{aligned}$$

This shows that $\widetilde{X}$ is a loss variable because by assumption $\widetilde{Y}$ is. ■

Now we turn to the excess ratio function of $\widetilde{X}$.

Proposition 22 *The excess ratio function of the spliced random variable $\widetilde{X}$ is given by*

$$R_{\tilde{X}}(r) = \begin{cases} 1 - \frac{E[X]}{E[\tilde{X}]}[1 - R_X(r)] & \text{if } r \le l \\ R_{\tilde{X}}(l)R_{\tilde{Y}}(r-l) & \text{if } r > l \end{cases}.$$

Proof. Using Definition 8 we first note that for $r \le l$ we have

$$\hat{F}_{\tilde{X}}(r) = \int_0^r x f_{\tilde{X}}(x)dx / E[\tilde{X}] = \frac{E[X]}{E[\tilde{X}]}\int_0^r x f_X(x)dx/E[X] = \frac{E[X]}{E[\tilde{X}]}\hat{F}_X(r).$$

Then using this relation and Propositions 9 and 20 we have for $r \leq l$

$$\begin{aligned} R_{\tilde{X}}(r) &= 1 - \hat{F}_{\tilde{X}}(r) - \frac{r}{E[\tilde{X}]}\left[1 - F_{\tilde{X}}(r)\right] \\ &= 1 - \frac{E[X]}{E[\tilde{X}]}\hat{F}_X(r) - \frac{r}{E[\tilde{X}]}\left[1 - F_X(r)\right] \\ &= 1 - \frac{E[X]}{E[\tilde{X}]}\left[\hat{F}_X(r) + \frac{r}{E[X]}(1 - F_X(r))\right] \\ &= 1 - \frac{E[X]}{E[\tilde{X}]}[1 - R_X(r)]. \end{aligned}$$

Now for $r > l$, using Propositions 19, 20, and 21 we get

$$\begin{aligned} R_{\tilde{X}}(r) &= \frac{1}{E[\tilde{X}]}\int_r^\infty (x-r) f_{\tilde{X}}(x)dx \\ &= \frac{1}{E[\tilde{X}]}\left[\int_r^\infty (x-r)[1 - F_X(l)] f_{\tilde{Y}}(x-l)dx\right] \\ &= \frac{1}{E[\tilde{X}]}\left[\int_{r-l}^\infty (x-(r-l)) S_X(l) f_{\tilde{Y}}(x)dx\right] \\ &= \frac{S_X(l)E[\tilde{Y}]}{E[\tilde{X}]}\left[\frac{\int_{r-l}^\infty (x-(r-l)) f_{\tilde{Y}}(x)dx}{E[\tilde{Y}]}\right] \\ &= \frac{S_{\tilde{X}}(l)}{E[\tilde{X}]}\frac{E[\tilde{X}]R_{\tilde{X}}(l)}{1 - F_{\tilde{X}}(l)} R_{\tilde{Y}}(r-l) \\ &= R_{\tilde{X}}(l) R_{\tilde{Y}}(r-l). \end{aligned}$$

■

We would typically start with a distribution X that has mean 1 and so we would naturally normalize $\widetilde{X}$ and work with $\widetilde{X}/\widetilde{\mu}$ where $\widetilde{\mu} = E[\widetilde{X}]$. We use the following slightly more general proposition.

Proposition 23 *Let X be a random variable with density f_X and distribution function F_X, and let $\alpha > 0$, then*

1. $f_{X/\alpha}(x) = \alpha f_X(\alpha x)$
2. $F_{X/\alpha}(x) = F_X(\alpha x)$

Proof. We note that

$$\begin{aligned} F_{X/\alpha}(x) &= \Pr(X/\alpha \le x) = \Pr(X \le \alpha x) = F_X(\alpha x) \\ &= \int_{-\infty}^{\alpha x} f(y)dy = \int_{-\infty}^{x} \alpha f(\alpha y)dy. \end{aligned}$$

Thus $f_{X/\alpha}(x) = \alpha f_X(\alpha x)$ and $F_{X/\alpha}(x) = F_X(\alpha x)$. ■

From this and Proposition 20 we get the following.

Proposition 24 *The distribution function of the normalized spliced random variable $\widetilde{X}/\widetilde{\mu}$, where $\widetilde{\mu} = E[\widetilde{X}]$, is given by*

$$F_{\widetilde{X}/\widetilde{\mu}}(r) = \begin{cases} F_X(\widetilde{\mu}r) & \text{if } r \le l/\widetilde{\mu} \\ 1 - [1 - F_X(l)]S_{\widetilde{Y}}(\widetilde{\mu}r - l) & \text{if } r > l/\widetilde{\mu} \end{cases}.$$

We can similarly recast Proposition 22.

Proposition 25 *Let $E[X] = 1$, then the excess ratio function of the normalized spliced random variable $\widetilde{X}/\widetilde{\mu}$,where $\widetilde{\mu} = E[\widetilde{X}]$, is given by*

$$R_{\widetilde{X}/\widetilde{\mu}}(r) = \begin{cases} 1 - \frac{1}{\widetilde{\mu}}[1 - R_X(\widetilde{\mu}r)] & \text{if } r \le l/\widetilde{\mu} \\ R_{\widetilde{X}}(l)R_{\widetilde{Y}}(\widetilde{\mu}r - l) & \text{if } r > l/\widetilde{\mu} \end{cases}.$$

Proof. By Proposition 23 and the change of variables $y = \tilde{\mu}x$, we have

$$\begin{aligned} R_{\widetilde{X}/\widetilde{\mu}}(r) &= \frac{\int_r^\infty (x - r) f_{\widetilde{X}/\widetilde{\mu}}(x)dx}{E[\widetilde{X}/\widetilde{\mu}]} = \frac{\int_r^\infty (\tilde{\mu}x - \tilde{\mu}r) f_{\tilde{X}}(\tilde{\mu}x)dx}{E[\tilde{X}]/\tilde{\mu}} \\ &= \frac{\frac{1}{\tilde{\mu}}\int_{\tilde{\mu}r}^\infty (y - \tilde{\mu}r) f_{\tilde{X}}(y)dy}{E[\tilde{X}]/\tilde{\mu}} = R_{\tilde{X}}(\tilde{\mu}r). \end{aligned}$$

Then by application of Proposition 22 we have

$$R_{\tilde{X}/\tilde{\mu}}(r) = R_{\tilde{X}}(\tilde{\mu}r) = \begin{cases} 1 - \frac{1}{E[\tilde{X}]}[1 - R_X(\tilde{\mu}r)] & \text{if } r \le l/\tilde{\mu} \\ R_{\tilde{X}}(l)R_{\tilde{Y}}(\tilde{\mu}r - l) & \text{if } r > l/\tilde{\mu} \end{cases}$$

■

In our case we fit a mixed exponential to the tail of the empirical random variable X. More precisely, we assume that $\widetilde{Y}$ is a mixed exponential. That is, using the parameterization in Klugman, et. al. [9] (see page 43 on mixture models as well), we assume

$$f_{\widetilde{Y}}(x) = \sum_i w_i \frac{1}{m_i} e^{-x/m_i}$$

and thus

$$F_{\widetilde{Y}}(x) = \sum_i w_i(1 - e^{-x/m_i}) = 1 - \sum_i w_i e^{-x/m_i}.$$

Then by Corollary 18,

$$R_{\widetilde{Y}}(r) = \frac{\sum w_i m_i e^{-r/m_i}}{\sum w_i m_i}.$$

We now state Propositions 24 and 25 in the mixed exponential case.

Proposition 26 *If $\widetilde{Y}$ has a mixed exponential distribution as above then the distribution function of the normalized spliced random variable $\widetilde{X}/\widetilde{\mu}$ is given by*

$$F_{\widetilde{X}/\widetilde{\mu}}(r) = \begin{cases} F_X(\widetilde{\mu}r) & \text{if } r \le l/\widetilde{\mu} \\ F_{\widetilde{X}/\widetilde{\mu}}(l/\widetilde{\mu}) + \left[1 - F_{\widetilde{X}/\widetilde{\mu}}(l/\widetilde{\mu})\right]\left[1 - \sum w_i e^{-(\widetilde{\mu}r - l)/m_i}\right] & \text{if } r > l/\widetilde{\mu} \end{cases}.$$

Proof. From Proposition 24 for $r > l/\widetilde{\mu}$ we get

$$\begin{aligned} F_{\widetilde{X}/\widetilde{\mu}}(r) &= 1 - [1 - F_X(l)]S_{\widetilde{Y}}(\widetilde{\mu}r - l) \\ &= 1 - [1 - F_X(l)][1 - F_{\widetilde{Y}}(\widetilde{\mu}r - l)] \\ &= 1 - [1 - F_X(l) - F_{\widetilde{Y}}(\widetilde{\mu}r - l) + F_X(l)F_{\widetilde{Y}}(\widetilde{\mu}r - l)] \\ &= F_X(l) + F_{\widetilde{Y}}(\widetilde{\mu}r - l) - F_X(l)F_{\widetilde{Y}}(\widetilde{\mu}r - l) \\ &= F_X(l) + [1 - F_X(l)]\, F_{\widetilde{Y}}(\widetilde{\mu}r - l) \\ &= F_{\widetilde{X}}(l) + \left[1 - F_{\widetilde{X}}(l)\right] F_{\widetilde{Y}}(\widetilde{\mu}r - l) \\ &= F_{\widetilde{X}/\widetilde{\mu}}(l/\widetilde{\mu}) + \left[1 - F_{\widetilde{X}/\widetilde{\mu}}(l/\widetilde{\mu})\right]\left[1 - \sum w_i e^{-(\widetilde{\mu}r - l)/m_i}\right] \end{aligned}$$

■

Proposition 27 *If $\widetilde{Y}$ has a mixed exponential distribution as above and $E[X] = 1$, then the excess ratio function of the normalized spliced random variable $\widetilde{X}/\widetilde{\mu}$ is given by*

$$R_{\widetilde{X}/\widetilde{\mu}}(r) = \begin{cases} 1 - \frac{1}{\widetilde{\mu}}[1 - R_X(\widetilde{\mu}r)] & \text{if } r \le l/\widetilde{\mu} \\ R_{\widetilde{X}}(l) \frac{\sum w_i m_i e^{-(\widetilde{\mu}r - l)/m_i}}{\sum w_i m_i} & \text{if } r > l/\widetilde{\mu} \end{cases}.$$

References

[1] Billingsley, Patrick, *Probability and Measure*, 2nd edition, Wiley, 1986.

[2] Brooks, Ward, "The 2001 California Retrospective Rating Plan–Developing Losses to Ultimate," unpublished manuscript.

[3] Feller, William, *An Introduction to Probability Theory and Its Applications,* Volume II, Wiley, 1966.

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

[5] Gillam, W.R., and Couret, J.R., "Retrospective Rating: 1997 Excess Loss Factors," *PCAS* LXXXIV, 1997, pp. 450–481.

[6] Harwayne, Frank, "Accident Limitations for Retrospective Rating," *PCAS* LXIII, 1976, pp. 1–31.

[7] Gleeson, Owen M., "Ratemaking for Excess Workers Compensation," *CAS Forum*, Winter 2001, pp. 61-114.

[8] Keatinge, C.L., "Modeling Losses with the Mixed Exponential Distribution," *PCAS* LXXXVI, 1999, pp. 654–698.

[9] Klugman, Stuart A., Panjer, Harry H., Willmot, Gordon E., *Loss Models: From Data to Decisions,* 2nd edition, Wiley, 2004.

[10] Mahler, H.C., "Workers Compensation Excess Ratios: An Alternative Method of Estimation," *PCAS* LXXXV, 1998, pp. 132–156.

[11] Uhthoff, Dunbar R., "Excess Loss Ratios via Loss Distributions," *PCAS* XXXVII, 1950, pp. 82–93.

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

Insurance Capital as a Shared Asset

Donald Mango

Abstract
Merton and Perold (1993) offered a framework for determining risk capital in a financial firm based on the cost of the implicit guarantee the firm provides to its subsidiaries to make up any operating shortfall. Merton and Perold assume the price of such guarantees is observable from the market at large. For an insurer, this may not be a realistic assumption. This paper proposes an insurance-specific framework for determining the cost of those parental guarantees, and utilizing that cost in pricing and portfolio mix evaluation. An insurer's capital is treated as a shared asset, with the insurance contracts in the portfolio having simultaneous rights to access potentially all that shared capital. By granting underwriting capacity, an insurer's management team is implicitly issuing a set of options to draw upon the common capital pool—similar in structure to letters of credit (LOC), except they are not loans but grants. The paper will (i) discuss the valuation of parental guarantees, beginning with Merton and Perold; (ii) treat insurer capital as a shared asset and explore the dual nature of insurer capital usage; (iii) offer a method for determining insurer capital usage cost; and (iv) demonstrate how this method could be used for product pricing and portfolio mix evaluation using economic value added concepts.

Keywords: Merton-Perold, capital allocation, capital consumption, economic value added.

1. VALUATION OF PARENTAL GUARANTEES

Merton and Perold (1993) (M-P) define risk capital as the amount required to guarantee payment of an asset or liability. In their first section, they present three related examples of a financial firm, Mortgage Bank or "MB," making a risky one-year bridge loan of $100M, financed by the issuing of a note to a note holder ("NH"). The only risk in any of these cases is the possible default of repayment by the bridge loan recipient ("BL"). They posit three outcome scenarios:

- Anticipated (A): bridge loan is repaid with interest of 20% at maturity in one year—e.g., for a loan of $100M, the repayment would be $120M;
- Disaster (D): amount repaid at maturity is only half that of Anticipated;
- Catastrophe (C): amount repaid at maturity is zero.

They discuss three cases, which differ mainly by which party bears the ultimate cost of any default. Under their *Case 1*, the note holder wishes to purchase a default-free note. The note holder is insulated from the default risk of BL by MB's purchase of "note insurance" from a third-party guarantor. The free market cost of this is *assumed to be* $5M. It is the cost of this guarantee that M-P considers to be risk capital. Merton and Perold never discuss the determination of that $5M price tag. They assume it to be a given figure, observable from the market.

Valuation of the Insurer Parental Guarantee
Similar to MB, every insurance contract in an insurer's portfolio receives a parental guarantee: should it be unable to pay for its own claims, a contract can draw upon the available funds of the company. Philbrick and Painter (2001) (P-P) elucidate:

First published in *ASTIN Bulletin*, Volume 35, No. 2. Used by permission.

> "When an insurance company writes a policy, a premium is received. A portion of this policy can be viewed as the loss component. When a particular policy incurs a loss, the company can look to three places to pay the loss. The first place is the loss component (together with the investment income earned) of the policy itself. In many cases, this will not be sufficient to pay the loss. The second source is unused loss components of other policies. In most cases, these two sources will be sufficient to pay the losses. In some years, it will not, and the company will have to look to a third source, the surplus, to pay the losses." (p. 124)

The only market from which an insurer might be able to observe the value of a guarantee is the reinsurance market. However, this market is limited, with relatively low numbers of participants, and a great diversity among products. A reinsurance valuation exercise is similar to that for over-the-counter (OTC) derivatives, in that it requires proprietary (as opposed to public) information, as well as a specific valuation methodology. Also, reinsurers, the liability holders of last resort, do not have the luxury of market prices for the guarantees they offer their portfolio segments. This suggests that, at a minimum, reinsurers must have an internal valuation framework of their own. It is argued here that an insurer must value the guarantee it provides to its portfolio, either explicitly or implicitly.

This paper proposes one such insurance-specific valuation methodology for the insurer parental guarantee. It is based upon the following premises:

- An insurer's capital is a shared asset, with all insurance contracts in the portfolio having simultaneous rights to access potentially all that shared capital.
- The impacts on an insurer from underwriting a contract are (i) the occupation of some of its finite underwriting capacity over a period of time (as determined by required capital calculations), and (ii) the extension of a guarantee by the firm to the contract holder to fulfill legitimate claims. These impacts represent distinct types of usage of the insurer's capital.
- Each distinct capital usage type will result in a unique charge: a capacity occupation cost and a capital consumption cost. The capacity occupation cost is an upfront cost which is a function of premium and expected reserve balances. Thus it can be treated as a fixed cost. The capital consumption cost is a variable cost depending on the amount of shortfall, which is scenario-specific. Therefore, its expected value over all possible contract outcome scenarios is used.
- The sum of these two costs will be called the capital usage cost, and will be treated as an expense in the contract pricing evaluation. The contribution to the insurer of a contract is therefore not a return on capital, like the ratio of expected profit to allocated capital, but rather the profit less the capital usage cost.
- The recommended decision metric then becomes economic value added or EVA[1], a means of risk-adjusting return by subtracting the opportunity cost of capital.

The paper will proceed by framing insurer capital as a shared asset, exploring the dual nature of insurer capital usage, proposing a method for determining insurer capital usage cost, and demonstrating how this method could be used for product pricing and portfolio mix evaluation using economic value added concepts.

---

[1] EVA is a registered trademark of Stern Stewart & Co. See www.sternstewart.com.

2. INSURER CAPITAL IS A SHARED ASSET

Shared assets or resources are entities conjointly owned by a community or group, for the use of their members. Shared assets can be scarce and essential public entities (e.g., reservoirs, fisheries, national forests), or desirable private entities (e.g., hotels, golf courses, beach houses). The access to and use of the assets is controlled and regulated by their owners; this control and regulation is essential to preserve the asset for future use. Examples of controls include usage rules (standards of care), limitations on the number of users (e.g., occupancy limits in a restaurant, swimmer limits at a pool), limitations on duration of usage (e.g., campsites at national parks), and limitations on amount of consumption (e.g., tons of fish taken from a fishery). It is particularly critical with essential assets that over-use by some members not compromise the future viability of the asset for the entire group. This *aggregation risk* is a common characteristic of shared asset usage, since shared assets typically have more members who could potentially use the asset than the asset can safely bear. Owners cannot count on individual users taking steps to preserve the asset. These users have their own incentives, and due to limited perspective and information, cannot see the implications of their individual actions upon the larger whole.

Consumptive and Non-Consumptive Uses

Shared assets are typically used in one of two manners, what is termed consumptive or non-consumptive[2]. *Consumptive* use involves the transfer of a portion or share of the asset from the communal asset to the member. Examples of consumptive use include water from a reservoir, livestock grazing on common pasture, or logging from national forests.

Non-consumptive use differs from consumptive use in several fundamental ways:

- Non-consumptive use involves temporary, limited transfer of control.
- Non-consumptive use is intended to be non-depletive—proper use of the asset leaves it intact for subsequent users.
- Non-consumptive use has a time element. Users occupy or borrow the asset for a period of time, then return it to the owner's control.

Examples of non-consumptive use include boating on a reservoir, hiking in a national forest, playing on a golf course, or renting a car or hotel room. The main aggregation concern from non-consumptive use relates to either capacity limitations or insufficient maintenance. Capacity limitation examples include caps on the number of water ski boats allowed on a lake, the number of campsites at national parks, or the number of available tee times at a golf course.

Shared assets are typically used in only one of the two manners. However, some shared assets can be used in either a consumptive or non-consumptive manner, depending on the situation. A good example is the renting of a hotel room. The intended use of the hotel

---

[2] These terms are used extensively in areas such as water and wildlife management. See Appendix C of the United States' Environmental Protection Agency's "Interim Economic Guidance for Water Quality Standards," www.epa.gov/waterscience/econ/appendc.html. Another good reference is the "Addis Ababa Principles and Guidelines for the Sustainable Use of Biodiversity," Convention on Migratory Species, www.cms.int/bodies/COP/cop8/documents/meeting_docs/en/Inf_15_AAPG_Sustainable_Use.pdf.

room is benign occupancy—the guest stays in the room, leaves it intact, and after cleaning the room is ready for subsequent rental. However, if a guest leaves the water running and floods their floor, or falls asleep with a lit cigarette and burns down a wing of the hotel, their use has become *consumptive*, because the capacity itself has been destroyed. The hotelier must rebuild the damaged rooms (invest additional capital) before the rooms can again be rented.

Insurer Capacity

Insurers sell promises to pay claims, so legitimate counterparty standing (i.e., claims paying rating) is vital. Other factors enter into a rating decision, but a key variable is the *capital adequacy ratio* (CAR). Different rating agencies use different approaches, but essentially CAR is the ratio of actual capital to required capital. Typically the rating agency formulas generate required capital from three broad sources: premiums, reserves, and assets. Current year underwriting activity will generate required *premium* capital. As that business ages, reserves will be established, which will generate required *reserve* capital. As those reserves run off, the amount of generated required reserve capital diminishes, eventually disappearing once the reserve balance reaches zero.

There are usually minimum CAR levels associated with each rating level. Thus, if an insurer has a desired rating, a given amount of actual capital corresponds to a maximum amount of rating agency required capital. This means required premium capital is an excellent proxy for *underwriting capacity*. It represents an externally imposed constraint on the amount of new business that can be written. Since total required capital consists of portions attributable to premium, reserves and assets, the maximum required premium capital is also a function of the amount of required reserve capital.

In summary, an insurer's actual capital *creates* underwriting capacity, and underwriting activity (either past or present) *ties up or occupies* potentially available underwriting capacity.

Consumptive and Non-Consumptive Use of Insurer Capital

Per the rating agency required capital formula, the presence of either premium balances (representing current year underwriting) or reserve balances (representing previous years' underwriting) results in required capital being calculated. This temporarily reduces the amount of underwriting capacity available for other underwriting uses. Being temporary, it is similar to capacity occupancy, a *non-consumptive use* of the shared asset.

Capital consumption occurs when reserves are increased. This involves a transfer of funds from the capital account to the reserve account, and eventually out of the firm as claims payments. P-P also introduced this concept:

> "The entire surplus is available to every policy to pay losses in excess of the aggregate loss component. Some policies are more likely to create this need than others are, even if the expected loss portions are equal. Roughly speaking, for policies with similar expected losses, we would expect the policies with a large variability of possible results to require more contributions from surplus to pay the losses. We can envision an insurance company instituting a charge for the access to the surplus. This charge should depend, not just on the likelihood that surplus might be needed, but on the amount of such a surplus call." (p. 124)

The two distinct impacts of underwriting an insurance portfolio on the insurer in total are therefore:

(i) Certain occupation of underwriting capacity for a period of time, and
(ii) Possible consumption of capital.

This "bi-polar" capital usage is structurally similar to a line of credit (LOC) as issued by banks. The dual impacts on a bank of issuing a LOC are:

(i) Certain occupation of capacity to issue LOC's, for the term of the LOC, and
(ii) Possible loan to the LOC holder.

Banks receive income for the issuance of LOC's in two ways:

(i) An access fee (i.e., option fee) for the right to draw upon the credit line, and
(ii) Loan payback with interest.

This dual form of payments for the dual nature of usage will be adapted for the unique characteristics of insurance.

3. THE COST OF USING INSURER CAPITAL

The cost of the insurer's parental guarantee therefore has two pieces: (i) a Capacity Occupation Cost, similar to the LOC access fee, and (ii) a Capital Call Cost, similar to the payback costs of accessing an LOC, but adjusted for the facts that the call is not for a loan but for a permanent transfer, and that the call destroys future underwriting capacity.

(i) Capacity Occupation Cost
The capacity occupation cost is an opportunity cost, compensating the firm for preclusion of other opportunities. It can be thought of as a minimum risk-adjusted hurdle rate. It will be the product of an opportunity cost rate and the amount of required capital generated over the active life of the contract. In continuous time, the formula would be:

$$\int_{t=0}^{T} RC_t \cdot r_{Opp} \cdot dt \,, \qquad (3.1)$$

where

- r_{Opp} is the "instantaneous" opportunity cost of capacity (similar to the force of interest); and
- RC_t is the required capital amount for the segment or contract at each point in time t, with t going from 0 (contract inception) to T (final resolution of all payments).

Rating agency required capital formulas are a discrete approximation of the continuous time situation:

$$\left\{\sum_{i=1}^{T} RC_i\right\} * r_{Opp} \qquad (3.2)$$

RC_i is the required capital for time period i. For $i=1$, it would be a function of premium; for all subsequent periods, it will be a function of reserves.

(ii) Capital Call Cost

Let v be the random variable representing the present value at inception of all insurance cash flows associated with an insurance contract—premium, expenses and loss payments (but not required capital). For simplicity assume $p(v)$ is the discrete distribution with possible outcomes v_i, $i = 1$ to n.

Let $f(x)$ be the capital call cost function that charges for a particular capital call. We will assume that a capital call is necessary when the present value of insurance flows v_i falls below zero. The magnitude of capital call for outcome v_i would be $-\min(0, v_i)$, which will be denoted s_i for shortfall of outcome i, a non-negative number. The cost of a capital call for outcome i will be denoted $f(s_i)$. The expected cost of capital calls over all outcomes would be:

$$\sum_{i=1}^{n} p_i * f(s_i) \qquad (3.4)$$

The form of $f(s_i)$ can be determined in part based on an understanding that a capital call *destroys future underwriting capacity*. Therefore, any call cost function should at least equal the amount of the call (payback of the capital grant). It should also compensate for lost opportunity cost. In this case, the destroyed capacity would need to be replenished by some means (e.g., recoupment from the product line's future returns, or capital infusion from parent). Whatever the source, the lost capacity could cost the firm the equivalent of *m years* of "capacity downtime," what one might call an *inconvenience* premium. Such an understanding leads to one possible means for determining the capital call cost function $f(s_i)$:

$$f(s_i) = \left(1 + m * r_{Opp}\right) \qquad (3.5)$$

The determination of m could be based on the volatility of a product's pricing cycles—that is, the likelihood that temporary capital impairment would lead to missed opportunity to write business at higher price levels.

Economic Value Added (EVA)

EVA, a registered trademark of Stern Stewart & Co., is a powerful metric used in financial analysis. The formula for EVA is:

EVA = NPV Return – Opportunity Cost of Capital

EVA is typically expressed as an amount. An activity with a positive EVA is said to "add value," while one with a negative EVA "destroys value."

EVA is simple to calculate using the shared asset framework:

EVA = NPV Return – Opportunity Cost of Capital *Usage*

EVA balances both risk and reward, and will be used as the key decision variable in the application examples to follow.

4. Application in Reinsurance Contract Evaluation

This section will demonstrate the application of this approach to two reinsurance contracts. Both examples use the following key parameters:

- $r_{Opp} = 10\%$
- $m = 5$
- $f(s) = 5*10\% = 50\%$

High Layer Property Excess of Loss Contract
Consider a high-layer contract, with a 2% chance of incurring a loss (i.e., 1 in 50 years). When a loss occurs, it is assumed to be a full limit loss. Example 1 shows the details:

Example 1
Property Catastrophe Contract

			Comments
(1) Premium	$	500,000	= 5% Net Rate on Line
(2) Limit	$	10,000,000	
Capacity Occupation Cost			
(3) Required Capital Factor		*50.0%*	Rating Agency
(4) Required Capital	$	250,000	= (3) * (1)
(5) Opportunity Cost for Capacity		*10.0%*	r_{Opp}
(6) Capacity Occupation Cost	$	25,000	= (4) * (5)
Capital Call Cost			
(7) Probability		2.0%	
(8) Loss	$	10,000,000	Full limit loss
(9) Capital Call Amount	$	9,500,000	= (8) - (1)
(10) Capital Call Cost Function		*50.0%*	= 5 * r_{Opp}
(11) Capital Call Charge	$	4,750,000	= (10) * (9)
(12) Expected Capital Call Cost	$	95,000	= (11) * (7)
EVA			
(13) Expected NPV	$	300,000	= (1) - (7) * (8)
(14) Expected Capital Usage Cost	$	120,000	= (6) + (12)
(15) EVA	$	180,000	= (13) - (14)

Example 1a shows the premium for a zero EVA:

Example 1a
Property Catastrophe Contract @ Zero EVA

				Comments
(1)	Premium	$	312,500	= 5% Net Rate on Line
(2)	Limit	$	10,000,000	
	Capacity Occupation Cost			
(3)	Required Capital Factor		*50.0%*	Rating Agency
(4)	Required Capital	$	156,250	= (3) * (1)
(5)	Opportunity Cost for Capacity		*10.0%*	r_{Opp}
(6)	Capacity Occupation Cost	$	15,625	= (4) * (5)
	Capital Call Cost			
(7)	Probability		2.0%	
(8)	Loss	$	10,000,000	Full limit loss
(9)	Capital Call Amount	$	9,687,500	= (8) - (1)
(10)	Capital Call Cost Function		*50.0%*	$= 5 * r_{Opp}$
(11)	Capital Call Charge	$	4,843,750	= (10) * (9)
(12)	Expected Capital Call Cost	$	96,875	= (11) * (7)
	EVA			
(13)	Expected NPV	$	112,500	= (1) - (7) * (8)
(14)	Expected Capital Usage Cost	$	112,500	= (6) + (12)
(15)	EVA	$	-	= (13) - (14)

Since this is a short payment tail line, there are no required capital charges for reserves, and discounting is ignored for simplicity. The two pieces of the capital usage cost are calculated separately. The EVA formula is straightforward, being NPV minus capital usage cost.

Longer Tail Excess of Loss Contract

Now consider a high-layer excess of loss contract on a liability product, with the same probability of loss, severity profile, limit, and premium, but a five-year payout. Example 2 shows the calculation details.

Example 2
Longer Tail Excess of Loss Contract

			Comments
(1) Premium	$	500,000	= 5% Net Rate on Line
(2) Limit	$	10,000,000	
Capacity Occupation Cost			
(3) Required Capital Factor - Premium		*50.0%*	Rating Agency
(3a) Required Capital Factor - Reserves		*35.0%*	Rating Agency
(3b) Reserve Amount	$	156,705	
(3c) Reserve Duration		5.00	Years
(4) Required Capital	$	524,234	= (3) * (1) + (3a) * (3b) * (3c)
(5) Opportunity Cost for Capacity		*10.0%*	r_{Opp}
(6) Capacity Occupation Cost	$	52,423	= (4) * (5)
Capital Call Cost			
(7) Probability		2.0%	
(8) Loss (NPV @ 5%)	$	7,835,262	Full limit loss, discounted
(9) Capital Call Amount	$	7,335,262	= (8) - (1)
(10) Capital Call Cost Function		*50.0%*	$= 5 * r_{Opp}$
(11) Capital Call Charge	$	3,667,631	= (10) * (9)
(12) Expected Capital Call Cost	$	73,353	= (11) * (7)
EVA			
(13) Expected NPV	$	343,295	= (1) - (7) * (8)
(14) Expected Capital Usage Cost	$	125,776	= (6) + (12)
(15) EVA	$	217,519	= (13) - (14)

The major differences between Examples 1 and 2:

- The Capacity Occupation Cost now includes charges for required capital needs over time on reserves. This increases the capacity occupation fee from $25,000 to $52,423.
- The loss payment has been discounted at 5% (the assumed default-free rate) for five years (assumed payment delay). This reduces the expected capital call cost from $95,000 to $73,353.
- The total capital usage cost stayed about the same, changing from $120,000 to $125,776.
- The EVA increased from $180,000 to $217,519. However, this is partly due to the premium being held constant at $500,000. The market price for the longer payment tail would likely have factored in the loss discounting.

Example 2a shows the liability contract premium that would give zero EVA:

Example 2a Longer Tail Excess of Loss Contract @ Zero EVA			Comments
(1) Premium	$	273,418	= 5% Net Rate on Line
(2) Limit	$	10,000,000	
Capacity Occupation Cost			
(3) Required Capital Factor - Premium		*50.0%*	Rating Agency
(3a) Required Capital Factor - Reserves		*35.0%*	Rating Agency
(3b) Reserve Amount	$	156,705	
(3c) Reserve Duration		5.00	Years
(4) Required Capital	$	410,943	= (3) * (1) + (3a) * (3b) * (3c)
(5) Opportunity Cost for Capacity		*10.0%*	r_{Opp}
(6) Capacity Occupation Cost	$	41,094	= (4) * (5)
Capital Call Cost			
(7) Probability		2.0%	
(8) Loss (NPV @ 5%)	$	7,835,262	Full limit loss, discounted
(9) Capital Call Amount	$	7,561,844	= (8) - (1)
(10) Capital Call Cost Function		*50.0%*	$= 5 * r_{Opp}$*
(11) Capital Call Charge	$	3,780,922	= (10) * (9)
(12) Expected Capital Call Cost	$	75,618	= (11) * (7)
EVA			
(13) Expected NPV	$	116,713	= (1) - (7) * (8)
(14) Expected Capital Usage Cost	$	116,713	= (6) + (12)
(15) EVA	$	0	= (13) - (14)

5. Application in Portfolio Mix Evaluation

This section will describe a Portfolio Mix Evaluation model based on the proposed approach. A simple example will be used to demonstrate the concepts. It will follow four steps:

1. Loss Distributions
2. Deviations from Mean
3. Capital Usage Cost Calculation
4. Evaluation Metrics

1. Loss Distributions

The model has three lines of business (abbreviated "LOB"), each with losses distributed Log-Normal, with expected value of $1,000,000, but different variances reflected by different *sigma* parameters. The parameters are shown here:

1) Loss Distributions	*LOB 1*	*LOB 2*	*LOB 3*	*TOTAL*
Log Normal Mu	13.771	13.691	13.571	
Log Normal Sigma	30.0%	50.0%	70.0%	
Expected Loss	1,000,000	1,000,000	1,000,000	*3,000,000*
Profit Margin	10.0%	10.0%	10.0%	
Premium	1,111,111	1,111,111	1,111,111	*3,333,333*
Return $	111,111	111,111	111,111	*333,333*

The model uses 100 independent scenarios drawn from these distributions, each of which is stored on its own row in the spreadsheet. Premium is assumed equal to expected losses plus a profit margin (expressed as a percentage of premium). Expenses are ignored.

2. Deviations from Mean

For simplicity, this model ignores discounting. The capital calls are therefore assumed to happen under those scenarios where a segment's losses are higher than expected. Section 2 of the model subtracts scenario loss from expected loss by segment. This amount is called *"deviation from mean,"* denoted d_{ij} for scenario i and segment j.

3. Capital Usage Cost Calculation

This table summarizes the major inputs for Capital Usage Cost.

3) Capital Usage Cost Inputs		LOB 1	LOB 2	LOB 3
Rating Agency Required Premium Capital Charge		40.0%	40.0%	40.0%
Opportunity Cost	10.0%			
m Years of Lost Opportunity	3.00			
Capital Call Cost Factor	30.0%			
Rating Agency Required Premium Capital		444,444	444,444	444,444

Here are detailed descriptions of each element:

- The rating agency required premium capital formula is a factor (40%) multiplied by premium.
- $r_{Opp} = 10\%$.
- $m = 3$ years.
- *Capital Call Cost Factor* $f = 3*10\% = 30\%$.

The Capital Usage Cost is calculated in the following steps:

- For scenario *i*, *portfolio* shortfall $s_i = -\min(0, d_i)$.
- For scenario *i*, *portfolio* capital call cost $c_i = f \cdot s_i$.
- Allocate c_i back to segment using the *RMK algorithm*. The RMK algorithm is a conditional risk allocation method developed by Ruhm, Mango and Kreps (2004)[3].
- For scenario *i*, segment *j* shortfall $s_{ij} = -\min(0, d_{ij})$.
- For scenario *i*, the sum of segment shortfalls $s_{i.} = \sum_{j=1}^{3}\left[-\min(0, d_{ij})\right]$.

[3] It is conceptually similar to concepts in Buhlmann, "An Economic Premium Principle," *ASTIN Bulletin* 11 (1980), p. 52-60. Ruhm and Mango (2003), and Kreps (2004), independently derived the approach, known as "RMK" for short. Kreps derived it under the name "riskiness leverage models"; Ruhm and Mango derived it under the name "Risk Charge Based on Conditional Probability." The method begins at the aggregate or portfolio level for evaluating risk, and allocates the total portfolio risk charge by each component's contribution to total portfolio risk. The result is an internally consistent allocation of diversification benefits.

- For scenario *i*, segment *j* share of the portfolio capital call cost $= c_{ij} = \frac{c_i \cdot s_{ij}}{s_{i.}}$.
- For segment *j*, expected capital call cost $c_{.j} = \sum_{i=1}^{n} p_i * c_{ij}$.

4. Evaluation Metrics

This table summarizes the major evaluation metrics:

4) Portfolio Evaluation Metrics	LOB 1	LOB 2	LOB 3	TOTAL
Premium	$ 1,111,111	$ 1,111,111	$ 1,111,111	$ 3,333,333
Required Capital	$ 444,444	$ 444,444	$ 444,444	$ 1,333,333
Return	$ 111,111	$ 111,111	$ 111,111	$ 333,333
(a) Expected Capital Usage Cost $	$ 60,099	$ 76,802	$ 109,481	$ 246,382
(b) Capital Usage Cost as % of Capital	13.5%	17.3%	24.6%	18.5%
(c) Occupation Cost	10.0%	10.0%	10.0%	10.0%
(d) Capital Call Cost	3.5%	7.3%	14.6%	8.5%
(e) EVA $	$ 51,012	$ 34,309	$ 1,630	$ 86,951
(f) Prob of Exceeding Required Capital	8.0%	15.0%	23.0%	9.7%

Premium, Required Capital, and Return are all inputs. Elements (a) – (f) will be discussed in detail:

- (a) Expected Capital Usage Cost $ = expected capital call cost c_j + capacity occupation cost.
  - Capacity Occupation Cost = Rating Agency Required Premium Capital * Opportunity Cost. The values are the same for each segment (line of business or "LOB"):
    - *Rating Agency Required Premium Capital = $444,444*
    - *Opportunity Cost = 10%*
    - *Capacity Occupation Cost = $444,444*10% = $44,444*
- (b) Capital Usage Cost as % of Capital = (a) divided by Rating Agency Required Premium Capital. Items (c) and (d) split this value into its two components:
  - (c) Occupation Cost = Opportunity Cost
  - (d) Capital Call Cost = (b) – (c)
  - The average value for the entire portfolio is 18.5%. This is the figure that would be calibrated to company cost of capital.
- (e) EVA $ = Expected Return minus (a)
- (f) Prob of Exceeding Required Capital = percentage of scenarios where shortfall was larger in magnitude than the required premium capital. This is one indicator as to how much "risk-sensitivity" underlies the capital factors. For example, if the capital factors were derived from a method based on a constant probability of default by segment—e.g., 5%—then this value would be 5% for every LOB.

Each LOB used the same required capital factor (40%), yet the variances (i.e., the riskiness) were markedly different. The method has corrected for this by indicating different *capital usage costs*:

- LOB 1 (low variance): *13.5%*

- LOB 2 (medium variance): *17.3%*
- LOB 3 (high variance): *24.6%*

This represents a true implementation of *RAROC*—risk-adjusted return on capital. As an alternative, we could use the model to calculate *RORAC*—return on risk-adjusted capital. We do this by varying the required capital factors until all three lines have the same return of 18.5%. The output is:

4) Portfolio Evaluation Metrics - RORAC								
		LOB 1		***LOB 2***		***LOB 3***		***TOTAL***
Premium	$	1,111,111	$	1,111,111	$	1,111,111	$	3,333,333
Required Capital	$	184,633	$	381,639	$	767,061	$	1,333,333
Return	$	111,111	$	111,111	$	111,111	$	333,333
(a) Expected Capital Usage Cost $	$	34,118	$	70,522	$	141,743	$	246,382
(b) Capital Usage Cost as % of Capital		18.5%		18.5%		18.5%		**18.5%**
(c) Occupation Cost		10.0%		10.0%		10.0%		**10.0%**
(d) Capital Call Cost		8.5%		8.5%		8.5%		**8.5%**
(e) EVA $	$	76,993	$	40,589	$	(30,632)	$	86,951
(f) Prob of Exceeding Required Capital		25.0%		16.0%		16.0%		9.7%

As compared with a constant 40% capital charge under RAROC, the resulting RORAC capital charges are:

3) Capital Usage Cost Inputs			
	LOB 1	***LOB 2***	***LOB 3***
Rating Agency Required Premium Capital Charge	**16.6%**	**34.3%**	**69.0%**

With this much higher capital charge, the EVA for LOB 3 becomes negative. This is because the product of its required capital and return is higher than in the base case.

All three LOB show positive EVA at these price levels. This table shows the premiums required to bring all three LOB to zero EVA using RAROC:

4) Portfolio Evaluation Metrics								
		LOB 1		***LOB 2***		***LOB 3***		***TOTAL***
Premium	$	1,057,974	$	1,075,373	$	1,109,414	$	3,242,760
Required Capital	$	423,189	$	430,149	$	443,765	$	1,297,104
Return	$	57,974	$	75,373	$	109,414	$	242,760
(a) Expected Capital Usage Cost $	$	57,973	$	75,373	$	109,413	$	242,759
(b) Capital Usage Cost as % of Capital		13.7%		17.5%		24.7%		**18.7%**
(c) Occupation Cost		10.0%		10.0%		10.0%		**10.0%**
(d) Capital Call Cost		3.7%		7.5%		14.7%		**8.7%**
(e) EVA $	$	-	$	-	$	-	$	-
(f) Prob of Exceeding Required Capital		9.0%		15.0%		23.0%		9.7%

The profit margins required to achieve this are:

	LOB 1	***LOB 2***	***LOB 3***
Profit Margin	**5.5%**	**7.0%**	**9.9%**

These might be thought of as *risk-based pricing benchmarks*, all calibrated to a zero-EVA level.

If we assume the starting point 10% profit margins are *given* from the market, we might seek the *portfolio mix* that maximizes total EVA, subject to a maximum rating agency required premium capital amount. The results of such a search (using Excel Solver) are shown here:

4) Portfolio Evaluation Metrics								
		LOB 1		***LOB 2***		***LOB 3***		***TOTAL***
Premium	$	1,953,704	$	870,370	$	509,259	$	3,333,333
Required Capital	$	781,481	$	348,148	$	203,704	$	1,333,333
Return	$	195,370	$	87,037	$	50,926	$	333,333
(a) Expected Capital Usage Cost $	$	117,709	$	58,927	$	43,396	$	220,033
(b) Capital Usage Cost as % of Capital		15.1%		16.9%		21.3%		16.5%
(c) Occupation Cost		10.0%		10.0%		10.0%		10.0%
(d) Capital Call Cost		5.1%		6.9%		11.3%		6.5%
(e) EVA $	$	77,661	$	28,110	$	7,530	$	113,300
(f) Prob of Exceeding Required Capital		8.0%		15.0%		23.0%		5.8%

The resulting EVA—$113,300—is far higher than the base case EVA of $86,951.

6. CONCLUSIONS

This paper introduces a method for assessing the cost of capital usage based on a shared asset view of insurer's capital. The shared asset view eliminates the need for allocation of capital, and is far more grounded in insurer realities. The method also shows promise for use with a portfolio risk model to evaluate portfolio mixes.

REFERENCES

Kreps, Rodney (2004), Riskiness Leverage Models, to appear in the *Proceedings of the CAS*.

Merton, Robert and Perold, Andre (1993), "Theory of Risk Capital in Financial Firms," *Journal of Applied Corporate Finance*, Volume 6, Number 3, Fall 1993, 16-32.

Philbrick, Stephen W., and Painter, Robert W. (2001), DFA Insurance Company Case Study, Part II: Capital Adequacy and Capital Allocation, *CAS 2001 Spring Forum*, DFA Call Paper Program, 99-152.

Ruhm, David, and Mango, Donald, (2003), A Risk Charge Calculation Based on Conditional Probability, presented at the *2003 Thomas P. Bowles Jr. Symposium*, available online at www.casact.org/coneduc/specsem/sp2003/papers/ruhm-mango.doc.

Ruhm, David, Mango, Donald, and Kreps, Rodney (2004), Applications of the Ruhm-Mango-Kreps Conditional Risk Charge Algorithm, to appear in *ASTIN Bulletin*.

Schnapp, Frederic (2004), The Cost of Conditional Risk Financing, *CAS 2004 Winter Forum*, Ratemaking Call Paper Program, 121-152.

Author's address: GE Insurance Solutions, 6 Hunt Lane, Gladstone, NJ, 07934, USA. Telephone: +1-908-234-1756; Fax: +1-908-234-1690; E-mail: don.mango@ge.com.

Discussion of Insurance Capital as a Shared Asset

Robert A. Bear, FCAS, MAAA, FCA, CPCU

Abstract

In his 2005 ASTIN paper (reprinted in the CAS 2006 Fall Forum), Donald Mango's ground-breaking work [1] in developing the concepts of insurance capital as a shared asset and Economic Value Added (EVA) are discussed with special emphasis on the purpose and calculation of the important Capital Call Costs. The EVA approach permits one to charge for risk (capital usage) and measure profitability at any desired level of definition while satisfying the key additivity property for risk charges without needing to allocate capital. Test examples are discussed that illustrate the impact on profitability of rate changes, changes in the distributions of premium written by line of business, inaccurate pricing due to parameter and model risk, correlation between lines of business, alternative reinsurance programs, and alternative selections for the Capital Call Cost function which is central to the EVA approach.

For those who prefer to measure returns as a percentage of invested capital, a Risk Return on Capital model (RROC) is suggested as an alternative way to integrate desirable properties of the EVA approach and the return on risk adjusted capital (RORAC) approach based upon riskiness leverage models. This method measures returns that are a reward for exposing capital to risk of loss after reflecting the cost of required rating agency capital.

Keywords. Capital allocation, cost of capital, enterprise risk management, return on equity, RMK algorithm, risk load.

1. INTRODUCTION

Actuaries frequently allocate capital to line of business or individual risk in an effort to calculate risk loads or evaluate profitability by calculating a risk adjusted return in the form of a return on equity (ROE) metric. Concerns have been expressed about ROE methods [7], especially the fact that the value inherent in the unallocated surplus is ignored (the entire surplus supports each and every risk). In his 2005 ASTIN paper on "Insurance Capital as a Shared Asset" [1], Donald Mango has introduced a method that eliminates the need for allocation of capital which he believes is more grounded in insurer realities.

2. SUMMARY WITH COMMENTS

Donald Mango treats insurance capital as a shared asset, with the insurance contracts having simultaneous rights to access potentially all of that shared capital. Shared assets can be scarce and essential public entities (e.g., reservoirs, fisheries, national forests), or desirable private entities (e.g., hotels, golf courses, beach houses). The access to and use of the assets is controlled and regulated by their owners; this control and regulation is essential to preserve the asset for future use. The aggregation risk is a common characteristic of shared asset usage, since shared assets typically have more members who could potentially use the asset than the asset can safely bear [1].

Mr. Mango differentiates between consumptive and non-consumptive use of an asset. A consumptive use involves the transfer of a portion or share of the asset from the communal asset to an individual, such as in the reservoir water usage and fishery examples. Non-consumptive use involves temporary, limited transfer of control which is intended to be non-depletive in that it is left intact for subsequent users. Examples of non-consumptive use include boating on a reservoir, playing on a golf course or renting a hotel room [1].

While shared assets are typically used in only one of the two manners, some shared assets can be used in either a consumptive or non-consumptive manner, depending on the situation. Mr. Mango gives the example of renting a hotel room. While the intended use is benign occupancy (non-consumptive), there is the risk that a guest may fall asleep with a lit cigarette and burn down a wing of the hotel (clearly consumptive) [1].

Mr. Mango notes that rating agencies use different approaches in establishing ratings, but the key variable is the capital adequacy ratio (CAR) which is the ratio of actual capital to required capital. Typically the rating agency formulas generate required capital from three sources: premiums, reserves, and assets. Current year underwriting activity will generate required premium capital. As that premium ages, reserves will be established that will generate required reserve capital. As the reserves are run off, the amount of required reserve capital will diminish and eventually reach zero when all claims are settled. As there are usually minimum CAR levels associated with each rating level, Mr. Mango points out that a given amount of actual capital corresponds to a maximum amount of rating agency required capital. Given reserve levels, this implies a limit to premium capital and thus to how much business can be written. Mr. Mango summarizes by stating than an insurer's actual capital creates underwriting capacity, while underwriting activity (either past or present) uses up underwriting capacity [1].

Mr. Mango notes that the generation of required capital, whether by premiums or reserves, temporarily reduces the amount of capacity available for other underwriting. Being temporary, it is similar to capacity occupancy, a non-consumptive use of the shared asset. Capacity consumption occurs when reserves must be increased beyond planned levels. Mr. Mango points out that this involves a transfer of funds from the capital account to the reserve account, and eventually out of the firm. Mr. Mango summarizes by stating that the two distinct impacts of underwriting an insurance portfolio are as follows [1]:

(1) Certain occupation of underwriting capacity for a period of time.

(2) Possible consumption of capital.

He notes that this "bi-polar" capital usage is structurally similar to a bank issuing a letter of credit (LOC). The dual impacts of a bank issuing a LOC are as follows [1]:

(1) Certain occupation of capacity to issue LOC's, for the term of the LOC.

(2) Possible loan to the LOC holder.

Mr. Mango notes that banks receive income for the issuance of LOC's in two ways [1]:

(1) An access fee (i.e., option fee) for the right to draw upon the credit line.

(2) Loan payback with interest.

Mr. Mango notes that every insurance contract in an insurer's portfolio receives a parental guarantee: Should it be unable to pay for its own claims, the contract can draw upon the available funds of the company. He states that the cost of this guarantee has two pieces [1]:

(1) A Capacity Occupation Cost, similar to the LOC access fee.

(2) A Capital Call Cost, similar to the payback costs of accessing an LOC, but adjusted for the facts that the call is not for a loan but for a permanent transfer, and that the call destroys future underwriting capacity.

Mr. Mango states that a capacity occupation cost is an opportunity cost, and thinks of it as a minimum risk adjusted hurdle rate. He computes it as the product of an opportunity cost rate and the amount of required rating agency capital generated over the active life of the contract. However, he does not explicitly credit interest on supporting surplus in his formula or in his examples, but usually interprets the opportunity cost of capital as a spread above investment returns on capital. In the examples discussed below, I show that this can be a significant factor. I think it reasonable to credit the mean interest earned over all simulations on required rating agency capital using a risk free rate, as we are already recognizing the opportunity cost of earmarking this capital to support the business written.

Mr. Mango also develops a formula for computing capital call costs which are his true risk loads, and defines the expected capital usage cost to be the sum of the capacity occupation cost and the expected capital call cost. He defines his key decision metric Economic Value Added (EVA) to be the NPV Return less the expected cost of capital [1]:

EVA = NPV Return – Capacity Occupation Cost – Capital Call Cost

Mr. Mango calculates capital call costs using the following algorithm:

(1) For each iteration (loss scenario) in the simulation, calculate the deviation of the loss for each segment (line of business or individual risk) from the expected loss. If the deviation from the mean is positive, there is no capital call and therefore no capital call cost. If the deviation from the mean is negative, the capital call cost equals the product of the magnitude of the deviation and the Capital Call Cost Factor. Calculate each segment's share of the portfolio capital call cost as the ratio of the segment cost to the total of these costs across all segments.

(2) Use the same procedure to calculate the portfolio capital call cost that was used to calculate segment capital call costs.

(3) Multiply the portfolio capital call cost by the segment shares calculated in (1) to calculate each segment's share of the capital call cost for that scenario.

(4) Each segment's expected capital call cost is the average of (3) over all scenarios.

The allocation procedure in the above algorithm was developed jointly by Mr. Mango, Mr. Rodney Kreps and Mr. David Ruhm [6]. It is a conditional risk allocation method which has become known as the RMK algorithm. Mr. Mango points out that the method extends risk valuation from the aggregate portfolio level down to the segments that comprise the portfolio, reflecting each segment's contribution to the total portfolio risk. The result is an internally consistent allocation of diversification benefits for which risk charges (costs of capital) are additive in any combination.

Mr. Mango notes that any capital cost function should at least equal the amount of the call (payback of the capital grant). It should also compensate for lost opportunity cost (inability to write as much business for several years until capital is replenished). Thus, Mr. Mango suggests the following form for the Capital Call Cost Factor: $1+n*r_{Opp}$.

He suggests that the determination of n could be based on the volatility of a product's pricing cycles (i.e., the likelihood that temporary capital impairment would lead to missed opportunities to write business at higher price levels). The opportunity cost of capacity r_{Opp} selected by Mr. Mango in his examples for the computation of the Capital Call Cost Factor is the same opportunity cost rate used to calculate the Capacity Occupation Cost.

Hence, if n=4 and r_{Opp} = 25%, then the Capital Call Cost Factor is 200%.

If pricing is accurate, this reviewer would theoretically expect capital grants in some years to be offset by redundancies in other years, averaging to the plan loss ratio which would equal the true Expected Loss Ratio (ELR). Hence, this reviewer believes the purpose of the capital call cost is to compensate for lost profits while capital is being replenished. Pricing errors or excessively competitive behavior may lead to market dislocations that permit risk loads of a magnitude that would be viewed by many as "payback," but this would appear in this methodology as a very healthy EVA.

Thus we have an asymmetric dynamic, where the additional capacity from upside scenarios rarely compensates for the lost capacity of downside scenarios. This is particularly true after the occurrence of extreme events, when pricing can become excessive for a limited period of time. Thus, capital call costs are intended to compensate for these missed opportunities.

Seminar notes from the 2005 Seminar on Reinsurance session on "Risk Load, Profitability Measures, and Enterprise Risk Management" may be downloaded from the CAS web site and illustrate the flexibility which this approach permits management in quantifying risk preferences. In Mr. Mango's seminar notes entitled "Insurance Capital as a Shared Asset – Theory and Practice," he points out that rating agency required capital can provide a convenient means to introduce a tail penalty. Rating agency required capital can be calculated at any level of detail, and so an additional charge can be assessed for exceeding allocated rating agency capital (this would be analogous to burning down a wing of the hotel in our illustrative example). In computing the Capital Call Cost, he assesses a moderate charge for damage within a segment's allocation (drawdown on allocated capital), and a much more severe charge for damage beyond a segment's allocation (drawdown of other segments' capital).

Assuming that correlations between segments are estimated with reasonable accuracy, it appears to this reviewer that this two step approach has the advantage of discouraging company threatening accumulations of risk, which is the central goal for an enterprise risk management system. For those willing to allocate capital as an intermediate step in allocating the cost of capital ([2], [4]), the Tail Value at Risk and Semi-Variance metrics [2] would also serve this function.

3. COMPARISON TO OTHER APPROACHES

This reviewer compared the EVA approach to the return on risk adjusted capital (RORAC) approach based upon riskiness leverage models [2] and to a modified RORAC approach which shall be referred to as a risk return on capital (RROC) model. RORAC based upon riskiness leverage models does not reflect rating agency capital requirements, particularly the requirement to hold capital to support reserves until all claims are settled. This is especially important for long tailed Casualty lines. A mean rating agency capital is computed by averaging rating agency required capital from the simulation (capital needed to support premium writings is added to the net present value, NPV, of the capital needed to support reserves on each iteration of the simulation). The mean rental cost of rating agency capital is calculated by multiplying the mean rating agency capital by the selected rental fee, which is an opportunity cost of capacity. Expected underwriting return is computed by adding the mean NPV of interest on reserves and interest on rating agency capital to expected underwriting return (profit and overhead). The expected underwriting return after rental cost of capital is computed by subtracting the mean rental cost of rating agency capital.

In my comparisons of EVA with RORAC and RROC, risk capital is a selected multiple of Excess Tail Value at Risk (XTVAR). XTVAR is defined to be the average value of $X-\mu$ when $X > x_q$, where the quantile x_q is the value of x where the cumulative distribution of X is q. Capital is allocated to line of business based upon Co-Excess Tail Values at Risk (Co-XTVAR) [4]. The same desirable properties hold for TVAR and co-TVAR as well as XTVAR and co-XTVAR [2], [3]:

(1) They can allocate risk down to any desired level of definition.

(2) They satisfy the additivity property (risk load or capital allocated to components of the portfolio sum to the total risk load or capital need for the portfolio).

(3) They are coherent measures of risk. Unlike Value at Risk, they satisfy the subadditivity axiom (the risk of a combination of exposures should not exceed the sum of the risks of the components) [5].

Mr. Venter notes that if capital is set by XTVAR, it would cover average losses in excess of expected losses for those years where the portfolio losses X exceed the q^{th} quantile x_q. It is assumed that expected losses have been fully reflected in pricing and in loss reserves. The capital allocated by co-XTVAR to a line would be the line's average losses above its mean losses in those same adverse years. Mr. Venter notes that there should be some probability level q for which XTVAR or a multiple of it makes sense as a capital standard [4].

RROC is computed as the ratio of expected underwriting return after rental cost of capital to allocated risk capital. RROC represents the expected return for exposing capital to risk of loss, as the cost of benign rental of capital has already been reflected [3]. (It is assumed that expense items like overhead and taxes, as well as returns from any capital excess the rating agency required capital or from riskier investments that would require additional rating agency capital, would be handled at the corporate planning level.)

RROC is analogous to the Capital Call Cost in the EVA approach, here expressed as a return on capital rather than applied as a cost. In his discussion of Tail Value at Risk, Mr. Venter has noted that co-XTVAR may not allocate capital to a line of business that didn't contribute significantly to adverse outcomes [4]. In such a situation, the traditional RORAC calculation may show the line to be highly profitable, whereas RROC may show that the line is unprofitable because it did not cover the mean rental cost of rating agency capital [3].

In the EVA approach, risk preferences are reflected in the function selected and parameterized in computing the Capital Call Cost. In the RORAC and RROC approaches, risk preferences are specified in the selection of the statistic used to measure risk [2], [3]. In practice, the RORAC and RROC approaches would be parameterized to allocate the total capital of the company, which would be maintained to at least cover rating agency capital required for its desired rating. All three approaches utilize the RMK algorithm for allocating risk (measured as a Capital Call Cost in EVA and as risk capital in RORAC and RROC) to line of business [1], [2], [3].

These models were tested and results summarized in the tables below. Table 1 summarizes the test examples, while Table 2 compares simulation results. In the base case, Example 2, all lines are uncorrelated and no reinsurance is purchased. Equal amounts of premium are written in the three lines, and pricing is accurate with the plan loss ratio equaling the true Expected Loss Ratio (ELR) of 80% for each line. Aggregate losses are assumed to be modeled accurately by lognormal distributions with coefficients of variation of 80%, 20% and 40% for lines of business (LOB) 1-3, respectively.

Payout Patterns were generated based upon an exponential settlement lag distribution with mean lags to settlement of one year, five years and ten years for lines of business (LOB) 1-3, respectively. Thus, the payout patterns for LOB 1-3 can be characterized as Fast, Average, and Slow, respectively. Interest is credited on supporting surplus using risk free rates for bonds of duration equal to the average settlement lag in each line of business. In this example, interest rates of 3%, 4% and 5% for LOB 1-3, respectively, were assumed.

These are the same rates that are used to calculate Net Present Value (NPV) reserves and the NPV Reserves Capital component of Required Rating Agency Capital. For simplicity, interest rates and payment patterns are assumed to be deterministic.

Required Rating Agency Capital is computed based upon rating agency premium and reserves capital charge factors assumed appropriate for the Company's desired rating. Somewhat smaller factors were selected for the reinsurance line (LOB 4) under the assumption that the Company would not receive full credit for ceded premium and reserves because a charge for potential uncollectibility would be applied. Capital needed to support reserves for a calendar year is the product of the reserves factors and the previous year-end reserves. Capital needed to support reserves must be calculated for all future calendar years until reserves run off. Required capital to support reserves is the NPV of these capital amounts. Required Rating Agency Capital is computed by adding the products of the plan premiums and the premium capital charge factors to the required capital to support reserves.

For both RORAC and RROC models, capital needed to support the portfolio risk is calculated as 200% of XTVAR. That is, the Company wants twice the capital needed to support average 1 in 50 year or worse deviations from plan. Capital needed to support the portfolio risk is allocated to line of business based upon Co-XTVAR.

Interest is credited on supporting surplus for Example 2, but not for Example 1. In the base example, Example 2, profitability is satisfactory overall, but inadequate for LOB 1 and redundant for LOB 2 and LOB 3. Comparison of Example 1 and 2 test results demonstrates that not crediting interest on supporting surplus can have a significant impact on all three profitability measures.

In Example 3, the margins are adjusted to reflect results in the base case. The ELR's for LOB 1-3 are 60%, 88%, and 85%, respectively. The test results show that overall profitability has increased significantly and is now marginally adequate even for LOB 1 assuming the implied rate change can be achieved. Note that EVA was negative for LOB 1 in the base Example 2, but is now positive with the improved rate adequacy. A negative EVA implies that the line should not be written unless the company is required to do so for regulatory reasons or it is necessary to support other lines with positive EVA (e.g., package policies). The required rating agency capital increases slightly from the base case, but the capital needed to support the portfolio under the ROE measures (RROC and RORAC) decreases by over 22% compared to the base case.

In Example 4, premiums written by line are adjusted to reflect the base example results. Premium written in LOB 1 is reduced by $250,000, while premium written in LOB 2 and in

LOB 3 are both increased by $125,000. The portfolio profitability increases significantly as a result, but remains inadequate for LOB 1. The required rating agency capital increases modestly from the base case, but the capital needed to support the portfolio under the ROE measures decreases by nearly 18% compared to the base case.

In Example 5, a new version of the company's catastrophe model is released right after the renewal season is over. The revised model implies a reduction in the ELR for LOB 1 from 80% to 70%. The test results show that EVA improves dramatically for LOB 1 (EVA is now positive) and for the entire portfolio. The ROE measures (RROC and RORAC) improve significantly for LOB 1 and the entire portfolio. Required rating agency capital is not significantly different compared to the base case, while the capital needed to support the portfolio under the ROE measures decreases by 15%.

In Example 6, a Supreme Court decision declared recent tort reforms to be unconstitutional. The ELR for LOB 3 is revised from 80% to 100%. The EVA deteriorates dramatically for LOB 3 and for the entire portfolio. Similarly, the ROE measures deteriorate dramatically for LOB 3, while deteriorating significantly for the entire portfolio. Because LOB 3 is a long tailed line, RROC declines much more dramatically than RORAC because the mean rental cost of rating agency capital has gone up significantly due to the increased reserves that must be held for a long period of time. In the base case, LOB 3 was viewed as highly profitable by all three measures. In Example 6, LOB 3 is viewed as unprofitable by the EVA approach, marginally profitable by the RROC approach, and highly profitable by the RORAC approach. The required rating agency capital increases by over 9% from the base case, while the capital needed to support the portfolio under the ROE measures increases by over 8% compared to the base case.

Both Examples 5 and 6 demonstrate that inaccurate pricing due to parameter and model risk can significantly impact profitability estimates when those errors are discovered.

In Example 7, LOB 1 and LOB 2 losses are 50% correlated, while losses for both lines are uncorrelated with LOB 3 losses. The EVA deteriorates significantly for LOB 1, LOB 2, and for the entire portfolio. For the ROE measures (RROC and RORAC), profitability has decreased dramatically for LOB 2 because LOB 2 losses now contribute more significantly to adverse scenarios created by LOB 1. Required rating agency capital is not significantly different compared to the base case, while the capital required to support the portfolio under the ROE approaches has increased by 6.5%.

In Example 8, a stop loss reinsurance treaty is purchased for LOB 1 covering a 30% excess 90% loss ratio layer for a 10% rate. The test results show that this program modestly improves all three profitability measures. The required rating agency capital decreases slightly from the base case, while the capital needed to support the net portfolio under the ROE measures decreases by 8.5%.

In Example 9, a 40% quota share is purchased for LOB 1 with commissions just covering variable costs. The test results show that this program had a major positive impact on all three profitability measures. The required rating agency capital decreases by nearly 6% from the base case, while the capital needed to support the net portfolio under the ROE measures decreases by over 35%.

On a technical note, when a reinsurance program is in place for a particular line of business and is invoked by a loss scenario, the average capital call cost factor for the line of business (ratio of the computed capital call charge to the deviation of the simulated loss from the mean) is applied to the deviation of the simulated reinsurance loss from the mean reinsured loss. This generates a credit capital call cost in the reinsurance line which reduces the average capital call cost for the line of business when combined with the reinsurance line.

In Examples 1-9, EVA is computed using the default assumption that the consumption fee for capital less than the required rating agency capital is 50% of the consumption fee for common capital. In Examples 10 and 11, alternative Capital Call Cost functions are parameterized and tested. In Exhibit 10, it is assumed that the consumption fee for capital less than the required rating agency capital is equal to the fee for capital consumed in excess of rating agency capital. In Exhibit 11, it is assumed that the consumption fee for capital less than the required rating agency capital is 25% of the consumption fee for common capital. Otherwise, Exhibits 10 and 11 are identical to Exhibit 9. EVA is dramatically lower in Example 10 compared to Example 9, while it is significantly improved in Example 11. These examples illustrate the importance of the selected Capital Call Cost function to the EVA approach. (The ROE measures differed slightly between Examples 9-11 due to random variation between simulations of 100,000 iterations.) Details of Examples 1-11 may be reviewed in Exhibits 1-11, respectively.

4. CONCLUSIONS

Donald Mango's very innovative work in developing the concepts of insurance capital as a shared asset and Economic Value Added contribute significantly to understanding the ways capital supports an insurance enterprise and must be financed. The EVA approach permits one to charge for risk (capital usage) and measure profitability at any desired level of definition while satisfying the key additivity property for risk charges without needing to allocate capital. The test examples demonstrate that it can be used to measure the impact on profitability of rate changes, changes in the distributions of premium written by line of business, inaccurate pricing due to parameter and model risk, correlation between lines of business, and alternative reinsurance programs. Results for alternative Capital Call Cost functions can be compared using these kinds of test examples.

For those who prefer to measure returns as a percentage of invested capital, a Risk Return on Capital model is suggested as an alternative way to integrate desirable properties of the EVA approach and the return on risk adjusted capital approach based upon riskiness leverage models. This method measures returns on capital after reflecting the mean rental cost of rating agency capital. Thus, returns that are a reward for exposing capital to risk of loss are measured after reflecting the cost of carrying capital to support premium written and loss reserves.

Table 1: Summary of Assumptions Underlying Examples

Example	*Exhibit*	*Key Assumptions*
1	*1*	*Same as base example, Example 2, except interest is not credited on surplus.*
2	*2*	*Base example: Write equal amounts of premium in three lines of business. Pricing is accurate, as the Plan Loss Ratios equal the true ELR's. The ELR's are equal to 80% for all three lines. Aggregate losses are assumed to be modeled accurately by lognormal distributions with coefficients of variation of 80%, 20% and 40% for LOB 1-3, respectively. LOB 1-3 losses are uncorrelated. Interest is credited on supporting surplus.*
3	*3*	*Same as base example, except adjust Margins by line to reflect results. ELR's for LOB 1-3 are 60%, 88% and 85%, respectively.*
4	*4*	*Same as base example, except adjust premiums by line to reflect results. Write $0.250m less in LOB 1, and write $0.125m more in LOB 2 and in LOB 3.*
5	*5*	*Base example, where pricing model is updated after renewal. Right after renewal season, a new version of the company's cat model is released which implies a reduction in the ELR for LOB 1 to 70%. The ELR's for LOB 2 and LOB 3 remain at 80%. The Plan Loss Ratios based upon Price Monitoring are all equal to 80%.*
6	*6*	*Base example, where new information is available after renewal. Right after renewal season, a Supreme Court decision declared recent tort reforms to be unconstitutional. The ELR for LOB 3 is revised to 100%, while the ELR's for LOB 1 and LOB 2 remain at 80%. The Plan Loss Ratios based upon Price Monitoring are all equal to 80%.*
7	*7*	*Same as base example, except that LOB 1 and LOB 2 losses are 50% correlated.*
8	*8*	*Same as the base example, except a 30% xs 90% loss ratio Stop Loss Reinsurance program is purchased for LOB 1 at a 10% rate.*

9	*9*	*Same as base example, except a 40% Quota Share is purchased for LOB 1 with commission just covering variable costs.* *The Consumption Fee for Capital Less than Allocation is 120%, while the Consumption Fee for Common Capital (excess allocation) is 240%.* *These same capital call charge factors have been applied in Examples 1-8.*
10	*10*	*Same assumptions as in Example 9, with the exception of capital call factors.* *The Consumption Fee for Capital Less than Allocation and the Consumption Fee for Common Capital (excess allocation) are both set to 180%.*
11	*11*	*Same assumptions as in Example 9, with the exception of capital call factors.* *The Consumption Fee for Capital Less than Allocation is 100%, while the Consumption Fee for Common Capital (excess allocation) is 400%.*

Table 2: Comparison of Results for Test Examples

	Returns on Risk Adjusted Capital		***Risk Returns on Capital***		***Economic Value Added***	
	Gross	***Net***	***Gross***	***Net***	***Gross***	***Net***
Example	***RORAC***	***RORAC***	***RROC***	***RROC***	***EVA***	***EVA***
1	**11.43%**	**11.43%**	**5.30%**	**5.30%**	**(19,077)**	**(19,077)**
2	**14.60%**	**14.60%**	**7.95%**	**7.95%**	**170,541**	**170,541**
3	**20.18%**	**20.18%**	**12.20%**	**12.20%**	**337,106**	**337,106**
4	**17.91%**	**17.91%**	**10.17%**	**10.17%**	**239,886**	**239,886**
5	**18.68%**	**18.68%**	**11.39%**	**11.39%**	**386,023**	**386,023**
6	**11.78%**	**11.78%**	**4.92%**	**4.92%**	**(187,275)**	**(187,275)**
7	**13.94%**	**13.94%**	**7.47%**	**7.47%**	**133,870**	**133,870**
8	**14.72%**	**15.06%**	**8.03%**	**8.14%**	**170,631**	**185,141**
9	**14.71%**	**20.03%**	**8.04%**	**11.48%**	**170,871**	**235,927**
10	**14.63%**	**19.91%**	**7.97%**	**11.40%**	**(27,654)**	**87,025**
11	**14.69%**	**19.91%**	**8.02%**	**11.41%**	**233,126**	**283,519**

Acknowledgment

The author wishes to thank Donald Mango and the CAS Forum reviewer, Windrie Wong, for their very helpful review and comments. Any errors are the responsibility of the author.

Supplementary Material

Seminar notes from 2005 Seminar on Reinsurance on "Risk Load, Profitability Measures, and Enterprise

Risk Management," which may be downloaded from the CAS web site.

5. REFERENCES

[1] Mango, Donald, "Insurance Capital as a Shared Asset," ASTIN Bulletin, Vol. 35, No. 2, 2005, p. 471-486. Reprinted in the CAS 2006 Fall Forum.

[2] Kreps, Rodney (2005), "Riskiness Leverage Models," CAS Spring 2005 Meeting.

[3] Bear, Robert A. (2005), Discussion of "Riskiness Leverage Models," accepted for publication in 2005 CAS Proceedings.

[4] Venter, Gary G., "Capital Allocation Survey with Commentary," *The 2003 Bowles Symposium, North American Actuarial Journal, April 2004, p. 96.*

[5] Kaye, Paul, "A Guide to Risk Measurement, Capital Allocation and Related Decision Support Issues," *Casualty Actuarial Society 2005 Discussion Paper Program.*

[6] Ruhm, David, Mango, Donald, and Kreps, Rodney (2005), "A General Additive Method for Portfolio Risk Analysis," submitted to ASTIN Bulletin.

[7] McClenahan, Charles L., "Risk Theory and Profit Loads- Remarks," CAS 1990 Spring Forum, 145-162.

Abbreviations and notations

CAR, Capital Adequacy Ratio
ELR, Expected Loss Ratio
EVA, Economic Value Added
Co-TVAR, Co-Tail Value at Risk
Co-XTVAR, Co-Excess Tail Value at Risk
LOB, Line of Business
LOC, letter of credit
RMK algorithm, a conditional risk allocation method
ROE, Return on Equity
RORAC, Return on Risk Adjusted Capital
RROC, Risk Return on Capital After Rental Cost of Ca
TVAR, Tail Value at Risk
VAR, Value at Risk
XTVAR, Excess Tail Value at Risk

Biography of the Author

Robert Bear is currently a Consulting Actuary, Reinsurance Consultant and Arbitrator in the firm he has established, RAB Actuarial Solutions, LLC. He previously served as Senior Vice President and Chief Actuary of PXRE Group. The author began his career at Insurance Services Office and subsequently served as an actuarial manager at Prudential Reinsurance, Signet Star Reinsurance and SCOR Reinsurance Company.

The author's service to the actuarial profession has included terms as Chairperson of the RAA Actuarial Committee and as President of Casualty Actuaries in Reinsurance. He has earned MS degrees in both theoretical and applied mathematics, as well as in economic systems. He is currently serving on the CAS Committee on the Theory of Risk and the Committee on Dynamic Risk Modeling.

The author previously co-authored "Pricing the Impact of Adjustable Features and Loss Sharing Provisions of Reinsurance Treaties" (1990 CAS Proceedings), which won the 1991 Woodward-Fondiller prize. He also authored a discussion of the Pinto-Gogol paper on "An Analysis of Excess Loss Development" (1992 CAS Proceedings) and a discussion of Rodney Kreps' paper on "Riskiness Leverage Models" which will be published in the 2005 CAS Proceedings.

Email: rabsolutions@gmail.com

Web Site: www.rabsolutions.net

Exhibit 1

Page 1

Example 1 Comparing EVA with Returns on Capital (RROC and RORAC) where Interest is Not Credited on Surplus

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus:* *No*

Pricing is accurate, as the Plan Loss Ratio equals the Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines.

All three lines are uncorrelated and no reinsurance is purchased. Correlation Between LOB 1 and LOB 2 Lc *0.0%*

Note: Alternative EVA measures and RARAC are computed before taxes, overhead, and returns on non-allocated capital or attributable to assumption of investment risk.

	Fast Pay	*Average Pay*	*Slow Pay*		
1) Loss Generator	*LOB 1*	*LOB 2*	*LOB 3*	*NET TOTAL*	*GROSS TOTAL*
1A) Expected Loss: Copy and Paste-Special from LOB 4 of (3K).	*1,000,000*	*1,000,000*	*1,000,000*	3,000,000	3,000,000
1B) Coefficient of Variation of Assumed Lognormal Loss Distribution	*80.0%*	*20.0%*	*40.0%*		
1C) Standard Deviation	*800,000*	*200,000*	*400,000*		
1D) Profit and Overhead Margin (includes Brokerage on Reinsurance)	9.0%	8.0%	7.0%	8.0%	8.0%
1E) Variable Expense Ratio	11.0%	12.0%	13.0%	*12.0%*	*12.0%*
1F) Plan Premium	1,250,000	1,250,000	1,250,000	3,750,000	3,750,000
1G) Expected Loss Ratio = (1A)/(1F)	80.0%	80.0%	80.0%	*80.0%*	*80.0%*
1H) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	87,500	300,000	300,000
1I) Plan Loss Ratio	80.0%	80.0%	80.0%	*80.0%*	*80.0%*
1J) Plan Expected Loss	1,000,000	1,000,000	1,000,000	3,000,000	3,000,000
1K) Pricing Error = ((1J)-(1A))/(1A)	0.0%	0.0%	0.0%	0.0%	0.0%
2) Capital Usage Calculation	*LOB 1*	*LOB 2*	*LOB 3*	*NET TOTAL*	*GROSS TOTAL*
2A) Required Capital Charge on Premium	40.0%	40.0%	40.0%	*40.0%*	*40.0%*
2B) Required Capital Charge on Reserves	25.0%	25.0%	25.0%	*25.0%*	*25.0%*
2C) Rental Fee	10.0%				
2D) Consumption Fee for Capital Less than Allocation	*120.0%*	12.00			
2E) Consumption Fee for Common Capital (excess allocation)	*240.0%*	24.00			
2F) Required Premium Capital =(1F)(2A)*	500,000	500,000	500,000	1,500,000	1,500,000
2G) Simulated Required NPV Reserves Capital = (2B)(NPV Future Reserves)*					
2H) Simulated Total Required Rating Agency Capital = (2F)+(2G)					
3) Annual Simulation - Calculation of Capital Call Costs and XTVAR	*LOB 1*	*LOB 2*	*LOB 3*	*NET TOTAL*	*GROSS TOTAL*
3A) Simulated Losses					
3B) Deviations From Plan = (1J)-(3A)					
3C) Segment Level Capital Usage Charges (Capital Call Costs)					
3D) Net Portfolio Capital Usage Cost with RMK Algorithm					
3E) Gross Portfolio Capital Usage Cost with RMK Algorithm					
3F) Deviation from Plan at 2nd Percentile: Copy and Paste-Special from (3M)					
3G) Deviation from Plan when Exceed 1 in 50 Year Result					
3H) Flag to Count Number of Simulations in Excess of 1 in 50 Year Result					
3I) Contribution to Gross 1 in 50 Year Result					
3J) Contribution to Net 1 in 50 Year Result					
Loss Simulation Statistics		Number of Simulations:		100,000	
3K) Expected Loss	1,000,011	1,000,000	999,996	3,000,007	3,000,007
3L) Standard Deviation	800,185	200,004	399,962	916,520	916,520
3M) Percentiles of Deviations from Plan (Negatives are Values at Risk)					
0.1 Percentile (1 in 1000)	(5,866,794)	(809,359)	(2,055,270)	(6,034,577)	(6,034,577)
1st Percentile (1 in 100)	(3,010,869)	(554,457)	(1,275,198)	(3,153,170)	(3,153,170)
2nd Percentile (1 in 50)	(2,311,103)	(472,772)	(1,048,373)	(2,460,701)	(2,460,701)
10.0 Percentile (1 in 10)	(923,344)	(263,898)	(521,231)	(1,091,084)	(1,091,084)
50th Percentile (1 in 2)	219,120	19,417	71,517	174,654	174,654
90th Percentile	682,951	239,216	433,302	919,994	919,994

Exhibit 1

Page 2

Example Comparing EVA with Returns on Capital (RROC and RORAC) where Interest is Not Credited on Surplus

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus:* *No*

Pricing is accurate, as the Plan Loss Ratio equals the Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines.

All three lines are uncorrelated and no reinsurance is purchased.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	87,500	300,000
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	
4D) Mean Net Present Value of Interest Earned on Reserves	27,485	163,602	327,516	518,602
4E) Mean Rating Agency Capital	729,013	1,522,318	2,137,091	4,388,422
4F) Mean Interest Earned on Rating Agency Capital	-	-	-	-
4G) Mean Rental Cost of Rating Agency Capital ((Mean of (2H)) x (2C))	72,901	152,232	213,709	438,842
4H) Gross Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3E)))	340,098	187,529	310,052	837,679
4I) Gross Economic Value Added (GEVA) = (4B)+(4D)+(4F)-(4H)	(200,114)	76,073	104,963	(19,077)
4J) Gross Capital Cost Percentage = (4H)/(4E)	46.7%	12.3%	14.5%	19.1%
4K) Net Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3D)))	340,098	187,529	310,052	
4L) Net Economic Value Added (NEVA) = (4B)+(4D)+(4F)-(4K)	(200,114)	76,073	104,963	
4M Net Capital Cost Percentage = (4K)/(4E)	46.7%	12.3%	14.5%	
4N) Change in EVA Due to Reinsurance = NEVA - GEVA	-			

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
5A) Average Deviation from Plan When Exceed 1 in 50 Year Result (XTVAR)	(3,425,698)	(587,974)	(1,380,969)	(3,575,724)
5B) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,523,075	98,481	538,412	7,159,968
5C) Mean Interest Earned on Rating Agency Capital = (4F)	-	-	-	-
5D) Mean Rental Cost of Rating Agency Capital (4G)	72,901	152,232	213,709	438,842
5E) Expected Underwriting Return After Rental Cost of Capital = (4B)+(4D)+(5C)-(5D)	67,083	111,371	201,306	379,760
5F) Gross Risk Return on Capital = GRROC = (5E)/(5B)	1.03%	113.09%	37.39%	5.30%
5G) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,523,075	98,481	538,412	
5H) Net Risk Return on Capital = NRROC = (5E)/(5G)	1.03%	113.09%	37.39%	
5I) Change in Return Due to Reinsurance = (5E for LOB 4)	-			
5J) Change in Allocated Capital = (5G)-(5B)	-			

6) Returns on Risk Adjusted Capital (RORAC)	LOB 1	LOB 2	LOB 3	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,523,075	98,481	538,412	7,159,968
6B) Interest Earned on Gross Allocated Capital = (4C)x(6A)	-	-	-	-
6C) Gross Expected Total Underwriting Return = (4B)+(4D)+(6B)	139,985	263,602	415,016	818,602
6D) Gross Return on Risk Adjusted Capital = GRORAC = (6C)/(6A)	2.15%	267.67%	77.08%	11.43%
6E) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,523,075	98,481	538,412	
6F) Interest Earned on Net Allocated Capital = (4C)x(6E)	-	-	-	
6G) Net Expected Total Underwriting Return = (4B)+(4D)+(6F)	139,985	263,602	415,016	
6H) Net Return on Risk Adjusted Capital = NRORAC = (6G)/(6E)	2.15%	267.67%	77.08%	
6I) Change in Return Due to Reinsurance = (6G - Net Total) - (6C - Gross Total)	-			
6J) Change in Allocated Capital = (6E - Net Total) - (6A - Gross Total)	-			

Exhibit 2

Base Example 2 Comparing EVA with Returns on Capital (RROC and RORAC) where Interest is Credited on Surplus

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus:* *Yes*

Pricing is accurate, as the Plan Loss Ratio equals the Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines.

All three lines are uncorrelated and no reinsurance is purchased.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	87,500	300,000
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	
4D) Mean Net Present Value of Interest Earned on Reserves	27,485	163,602	327,516	518,602
4E) Mean Rating Agency Capital	729,013	1,522,318	2,137,091	4,388,422
4F) Mean Interest Earned on Rating Agency Capital	21,870	60,893	106,855	189,618
4G) Mean Rental Cost of Rating Agency Capital ((Mean of (2H)) x (2C))	72,901	152,232	213,709	438,842
4H) Gross Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3E)))	340,098	187,529	310,052	837,679
4I) Gross Economic Value Added (GEVA) = (4B)+(4D)+(4F)-(4H)	(178,243)	136,966	211,818	170,541
4J) Gross Capital Cost Percentage = (4H)/(4E)	46.7%	12.3%	14.5%	19.1%
4K) Net Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3D)))	340,098	187,529	310,052	
4L) Net Economic Value Added (NEVA) = (4B)+(4D)+(4F)-(4K)	(178,243)	136,966	211,818	
4M Net Capital Cost Percentage = (4K)/(4E)	46.7%	12.3%	14.5%	
4N) Change in EVA Due to Reinsurance = NEVA - GEVA	-			

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
5A) Average Deviation from Plan When Exceed 1 in 50 Year Result (XTVAR)	(3,425,698)	(587,974)	(1,380,969)	(3,575,724)
5B) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,523,075	98,481	538,412	7,159,968
5C) Mean Interest Earned on Rating Agency Capital = (4F)	21,870	60,893	106,855	189,618
5D) Mean Rental Cost of Rating Agency Capital (4G)	72,901	152,232	213,709	438,842
5E) Expected Underwriting Return After Rental Cost of Capital = (4B)+(4D)+(5C)-(5D)	88,954	172,263	308,161	569,378
5F) Gross Risk Return on Capital = GRROC = (5E)/(5B)	1.36%	174.92%	57.24%	7.95%
5G) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,523,075	98,481	538,412	
5H) Net Risk Return on Capital = NRROC = (5E)/(5G)	1.36%	174.92%	57.24%	
5I) Change in Return Due to Reinsurance = (5E for LOB 4)	-			
5J) Change in Allocated Capital = (5G)-(5B)	-			

6) Returns on Risk Adjusted Capital (RORAC)

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,523,075	98,481	538,412	7,159,968
6B) Interest Earned on Gross Allocated Capital = (4C)x(6A)	195,692	3,939	26,921	226,552
6C) Gross Expected Total Underwriting Return = (4B)+(4D)+(6B)	335,677	267,542	441,936	1,045,155
6D) Gross Return on Risk Adjusted Capital = GRORAC = (6C)/(6A)	5.15%	271.67%	82.08%	14.60%
6E) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,523,075	98,481	538,412	
6F) Interest Earned on Net Allocated Capital = (4C)x(6E)	195,692	3,939	26,921	
6G) Net Expected Total Underwriting Return = (4B)+(4D)+(6F)	335,677	267,542	441,936	
6H) Net Return on Risk Adjusted Capital = NRORAC = (6G)/(6E)	5.15%	271.67%	82.08%	
6I) Change in Return Due to Reinsurance = (6G - Net Total) - (6C - Gross Total)	-			
6J) Change in Allocated Capital = (6E - Net Total) - (6A - Gross Total)	-			

Exhibit 3

Modified Base Example 3 Comparing EVA with Returns on Capital (RROC and RORAC) where Adjust Margins

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting S/h Yes*

Pricing is accurate, as the Plan Loss Ratio equals the true ELR for all three lines. Adjust Margins by line to reflect results of Example 2.

The ELR's for LOB 1, LOB 2, and LOB 3 are now 60%, 88% and 85%, respectively. *All three lines are uncorrelated and no reinsurance is purchased.*

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	362,500	-	25,000	387,500
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	
4D) Mean Net Present Value of Interest Earned on Reserves	20,613	179,963	347,987	548,562
4E) Mean Rating Agency Capital	671,754	1,624,552	2,239,416	4,535,722
4F) Mean Interest Earned on Rating Agency Capital	20,153	64,982	111,971	197,106
4G) Mean Rental Cost of Rating Agency Capital ((Mean of (2H)) x (2C))	67,175	162,455	223,942	453,572
4H) Gross Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3E)))	252,757	205,404	337,901	796,062
4I) Gross Economic Value Added (GEVA) = (4B)+(4D)+(4F)-(4H)	150,509	39,541	147,057	337,106
4J) Gross Capital Cost Percentage = (4H)/(4E)	37.6%	12.6%	15.1%	17.6%
4K) Net Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3D)))	252,757	205,404	337,901	
4L) Net Economic Value Added (NEVA) = (4B)+(4D)+(4F)-(4K)	150,509	39,541	147,057	
4M Net Capital Cost Percentage = (4K)/(4E)	37.6%	12.6%	15.1%	
4N) Change in EVA Due to Reinsurance = NEVA - GEVA	-			

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
5A) Average Deviation from Plan When Exceed 1 in 50 Year Result (XTVAR)	(2,566,035)	(646,459)	(1,468,083)	(2,784,762)
5B) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	4,450,243	149,438	970,231	5,569,913
5C) Mean Interest Earned on Rating Agency Capital = (4F)	20,153	64,982	111,971	197,106
5D) Mean Rental Cost of Rating Agency Capital (4G)	67,175	162,455	223,942	453,572
5E) Expected Underwriting Return After Rental Cost of Capital = (4B)+(4D)+(5C)-(5D)	336,090	82,490	261,016	679,596
5F) Gross Risk Return on Capital = GRROC = (5E)/(5B)	7.55%	55.20%	26.90%	12.20%
5G) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	4,450,243	149,438	970,231	
5H) Net Risk Return on Capital = NRROC = (5E)/(5G)	7.55%	55.20%	26.90%	
5I) Change in Return Due to Reinsurance = (5E for LOB 4)	-			
5J) Change in Allocated Capital = (5G)-(5B)	-			

6) Returns on Risk Adjusted Capital (RORAC)

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	4,450,243	149,438	970,231	5,569,913
6B) Interest Earned on Gross Allocated Capital = (4C)x(6A)	133,507	5,978	48,512	187,996
6C) Gross Expected Total Underwriting Return = (4B)+(4D)+(6B)	516,620	185,940	421,498	1,124,059
6D) Gross Return on Risk Adjusted Capital = GRORAC = (6C)/(6A)	11.61%	124.43%	43.44%	20.18%
6E) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	4,450,243	149,438	970,231	
6F) Interest Earned on Net Allocated Capital = (4C)x(6E)	133,507	5,978	48,512	
6G) Net Expected Total Underwriting Return = (4B)+(4D)+(6F)	516,620	185,940	421,498	
6H) Net Return on Risk Adjusted Capital = NRORAC = (6G/(6E)	11.61%	124.43%	43.44%	
6I) Change in Return Due to Reinsurance = (6G - Net Total) - (6C - Gross Total)	-			
6J) Change in Allocated Capital = (6E - Net Total) - (6A - Gross Total)	-			

Exhibit 4

Modified Base Example 4 Comparing EVA with Returns on Capital where Adjust Premiums by Line

Key Assumptions: Write $0.250m less in LOB 1, and write $0.125m more in LOB 2 and in LOB 3. *Interest Credited on Supporting Surplus: Yes*

Pricing is accurate, as the Plan Loss Ratio equals the true Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines,.

All three lines are uncorrelated and no reinsurance is purchased.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
4A) Plan Premium	1,000,000	1,375,000	1,375,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	90,000	110,000	96,250	296,250
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	
4D) Mean Net Present Value of Interest Earned on Reserves	21,988	179,962	360,267	562,217
4E) Mean Rating Agency Capital	583,215	1,674,549	2,350,798	4,608,562
4F) Mean Interest Earned on Rating Agency Capital	17,496	66,982	117,540	202,018
4G) Mean Rental Cost of Rating Agency Capital ((Mean of (2H)) x (2C))	58,322	167,455	235,080	460,856
4H) Gross Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3E)))	259,101	209,513	351,987	820,600
4I) Gross Economic Value Added (GEVA) = (4B)+(4D)+(4F)-(4H)	(129,616)	147,432	222,070	239,886
4J) Gross Capital Cost Percentage = (4H)/(4E)	44.4%	12.5%	15.0%	17.8%
4K) Net Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3D)))	259,101	209,513	351,987	
4L) Net Economic Value Added (NEVA) = (4B)+(4D)+(4F)-(4K)	(129,616)	147,432	222,070	
4M Net Capital Cost Percentage = (4K)/(4E)	44.4%	12.5%	15.0%	
4N) Change in EVA Due to Reinsurance = NEVA - GEVA	-			

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
5A) Average Deviation from Plan When Exceed 1 in 50 Year Result (XTVAR)	(2,739,812)	(646,227)	(1,519,256)	(2,944,172)
5B) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	4,802,202	153,679	938,356	5,894,237
5C) Mean Interest Earned on Rating Agency Capital = (4F)	17,496	66,982	117,540	202,018
5D) Mean Rental Cost of Rating Agency Capital (4G)	58,322	167,455	235,080	460,856
5E) Expected Underwriting Return After Rental Cost of Capital = (4B)+(4D)+(5C)-(5D)	71,163	189,489	338,977	599,630
5F) Gross Risk Return on Capital = GRROC = (5E)/(5B)	1.48%	123.30%	36.12%	10.17%
5G) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	4,802,202	153,679	938,356	
5H) Net Risk Return on Capital = NRROC = (5E)/(5G)	1.48%	123.30%	36.12%	
5I) Change in Return Due to Reinsurance = (5E for LOB 4)	-			
5J) Change in Allocated Capital = (5G)-(5B)	-			

6) Returns on Risk Adjusted Capital (RORAC)	LOB 1	LOB 2	LOB 3	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	4,802,202	153,679	938,356	5,894,237
6B) Interest Earned on Gross Allocated Capital = (4C)x(6A)	144,066	6,147	46,918	197,131
6C) Gross Expected Total Underwriting Return = (4B)+(4D)+(6B)	256,054	296,110	503,435	1,055,599
6D) Gross Return on Risk Adjusted Capital = GRORAC = (6C)/(6A)	5.33%	192.68%	53.65%	17.91%
6E) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	4,802,202	153,679	938,356	
6F) Interest Earned on Net Allocated Capital = (4C)x(6E)	144,066	6,147	46,918	
6G) Net Expected Total Underwriting Return = (4B)+(4D)+(6F)	256,054	296,110	503,435	
6H) Net Return on Risk Adjusted Capital = NRORAC = (6G)/(6E)	5.33%	192.68%	53.65%	
6I) Change in Return Due to Reinsurance = (6G - Net Total) - (6C - Gross Total)	-			
6J) Change in Allocated Capital = (6E - Net Total) - (6A - Gross Total)	-			

Exhibit 5

Modified Base Example 5 Comparing EVA with Returns on Capital where Update ELR for LOB 1

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus:* *Yes*

Right after renewal season, a new version of company's cat model is released which implies a 10% reduction in the ELR for LOB 1.

The original plan loss ratio for LOB 1 was 80%, but the estimated ELR has been revised to 70%. All lines are uncorrelated and no reinsurance is purchased.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	237,500	100,000	87,500	425,000
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	
4D) Mean Net Present Value of Interest Earned on Reserves	24,048	163,602	327,517	515,168
4E) Mean Rating Agency Capital	700,381	1,522,318	2,137,097	4,359,796
4F) Mean Interest Earned on Rating Agency Capital	21,011	60,893	106,855	188,759
4G) Mean Rental Cost of Rating Agency Capital ((Mean of (2H)) x (2C))	70,038	152,232	213,710	435,980
4H) Gross Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3E)))	259,685	182,851	300,367	742,903
4I) Gross Economic Value Added (GEVA) = (4B)+(4D)+(4F)-(4H)	22,875	141,644	221,504	386,023
4J) Gross Capital Cost Percentage = (4H)/(4E)	37.1%	12.0%	14.1%	17.0%
4K) Net Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3D)))	259,685	182,851	300,367	
4L) Net Economic Value Added (NEVA) = (4B)+(4D)+(4F)-(4K)	22,875	141,644	221,504	
4M Net Capital Cost Percentage = (4K)/(4E)	37.1%	12.0%	14.1%	
4N) Change in EVA Due to Reinsurance = NEVA - GEVA	-			

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
5A) Average Deviation from Plan When Exceed 1 in 50 Year Result (XTVAR)	(2,871,920)	(587,447)	(1,380,805)	(3,038,640)
5B) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	5,359,487	93,120	630,998	6,083,606
5C) Mean Interest Earned on Rating Agency Capital = (4F)	21,011	60,893	106,855	188,759
5D) Mean Rental Cost of Rating Agency Capital (4G)	70,038	152,232	213,710	435,980
5E) Expected Underwriting Return After Rental Cost of Capital = (4B)+(4D)+(5C)-(5D)	212,522	172,263	308,162	692,947
5F) Gross Risk Return on Capital = GRROC = (5E)/(5B)	3.97%	184.99%	48.84%	11.39%
5G) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	5,359,487	93,120	630,998	
5H) Net Risk Return on Capital = NRROC = (5E)/(5G)	3.97%	184.99%	48.84%	
5I) Change in Return Due to Reinsurance = (5E for LOB 4)	-			
5J) Change in Allocated Capital = (5G)-(5B)	-			

6) Returns on Risk Adjusted Capital (RORAC)

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	5,359,487	93,120	630,998	6,083,606
6B) Interest Earned on Gross Allocated Capital = (4C)x(6A)	160,785	3,725	31,550	196,059
6C) Gross Expected Total Underwriting Return = (4B)+(4D)+(6B)	422,333	267,327	446,567	1,136,227
6D) Gross Return on Risk Adjusted Capital = GRORAC = (6C)/(6A)	7.88%	287.08%	70.77%	18.68%
6E) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	5,359,487	93,120	630,998	
6F) Interest Earned on Net Allocated Capital = (4C)x(6E)	160,785	3,725	31,550	
6G) Net Expected Total Underwriting Return = (4B)+(4D)+(6F)	422,333	267,327	446,567	
6H) Net Return on Risk Adjusted Capital = NRORAC = (6G)/(6E)	7.88%	287.08%	70.77%	
6I) Change in Return Due to Reinsurance = (6G - Net Total) - (6C - Gross Total)	-			
6J) Change in Allocated Capital = (6E - Net Total) - (6A - Gross Total)	-			

Exhibit 6

Modified Base Example 6 Comparing EVA with Returns on Capital where Update ELR for LOB 3

Same as base case, but after renewal season a Supreme Court decision declared recent tort reforms to be unconstitutional.
This decision implies a 20% increase in the ELR for LOB 3. The original plan loss ratio for LOB 3 was 80%, but the estimated ELR has been revised to 100%.
All three lines are uncorrelated and no reinsurance is purchased.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	(162,500)	50,000
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	
4D) Mean Net Present Value of Interest Earned on Reserves	27,485	163,603	409,398	600,486
4E) Mean Rating Agency Capital	729,016	1,522,321	2,546,383	4,797,719
4F) Mean Interest Earned on Rating Agency Capital	21,870	60,893	127,319	210,082
4G) Mean Rental Cost of Rating Agency Capital ((Mean of (2H)) x (2C))	72,902	152,232	254,638	479,772
4H) Gross Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3E)))	368,134	198,455	481,255	1,047,844
4I) Gross Economic Value Added (GEVA) = (4B)+(4D)+(4F)-(4H)	(206,279)	126,041	(107,038)	(187,275)
4J) Gross Capital Cost Percentage = (4H)/(4E)	50.5%	13.0%	18.9%	21.8%
4K) Net Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3D)))	368,134	198,455	481,255	
4L) Net Economic Value Added (NEVA) = (4B)+(4D)+(4F)-(4K)	(206,279)	126,041	(107,038)	
4M Net Capital Cost Percentage = (4K)/(4E)	50.5%	13.0%	18.9%	
4N) Change in EVA Due to Reinsurance = NEVA - GEVA	-			

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
5A) Average Deviation from Plan When Exceed 1 in 50 Year Result (XTVAR)	(3,425,534)	(588,031)	(1,977,400)	(3,871,434)
5B) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,218,516	99,077	1,429,846	7,747,439
5C) Mean Interest Earned on Rating Agency Capital = (4F)	21,870	60,893	127,319	210,082
5D) Mean Rental Cost of Rating Agency Capital (4G)	72,902	152,232	254,638	479,772
5E) Expected Underwriting Return After Rental Cost of Capital = (4B)+(4D)+(5C)-(5D)	88,954	172,264	119,579	380,796
5F) Gross Risk Return on Capital = GRROC = (5E)/(5B)	1.43%	173.87%	8.36%	4.92%
5G) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,218,516	99,077	1,429,846	
5H) Net Risk Return on Capital = NRROC = (5E)/(5G)	1.43%	173.87%	8.36%	
5I) Change in Return Due to Reinsurance = (5E for LOB 4)	-			
5J) Change in Allocated Capital = (5G)-(5B)	-			

6) Returns on Risk Adjusted Capital (RORAC)	LOB 1	LOB 2	LOB 3	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,218,516	99,077	1,429,846	7,747,439
6B) Interest Earned on Gross Allocated Capital = (4C)x(6A)	186,555	3,963	71,492	262,011
6C) Gross Expected Total Underwriting Return = (4B)+(4D)+(6B)	326,540	267,566	318,391	912,497
6D) Gross Return on Risk Adjusted Capital = GRORAC = (6C)/(6A)	5.25%	270.06%	22.27%	11.78%
6E) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,218,516	99,077	1,429,846	
6F) Interest Earned on Net Allocated Capital = (4C)x(6E)	186,555	3,963	71,492	
6G) Net Expected Total Underwriting Return = (4B)+(4D)+(6F)	326,540	267,566	318,391	
6H) Net Return on Risk Adjusted Capital = NRORAC = (6G/(6E)	5.25%	270.06%	22.27%	
6I) Change in Return Due to Reinsurance = (6G - Net Total) - (6C - Gross Total)	-			
6J) Change in Allocated Capital = (6E - Net Total) - (6A - Gross Total)	-			

Exhibit 7

Modified Base Example 7 Comparing EVA with Returns on Capital where LOB 1 and LOB 2 are 50% Correlated

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus:* *Yes*

Pricing is accurate, as the Plan Loss Ratio equals the Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines.

Lines 1 and 2 losses are 50% correlated but uncorrelated with line 3. No reinsurance is purchased.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	87,500	300,000
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	
4D) Mean Net Present Value of Interest Earned on Reserves	27,484	163,602	327,518	518,605
4E) Mean Rating Agency Capital	729,012	1,522,317	2,137,104	4,388,433
4F) Mean Interest Earned on Rating Agency Capital	21,870	60,893	106,855	189,618
4G) Mean Rental Cost of Rating Agency Capital ((Mean of (2H)) x (2C))	72,901	152,232	213,710	438,843
4H) Gross Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3E)))	359,046	204,285	311,022	874,353
4I) Gross Economic Value Added (GEVA) = (4B)+(4D)+(4F)-(4H)	(197,191)	120,210	210,851	133,870
4J) Gross Capital Cost Percentage = (4H)/(4E)	49.3%	13.4%	14.6%	19.9%
4K) Net Expected Cost of Capital - Rental and Usage ((4G) + (Mean of (3D)))	359,046	204,285	311,022	
4L) Net Economic Value Added (NEVA) = (4B)+(4D)+(4F)-(4K)	(197,191)	120,210	210,851	
4M Net Capital Cost Percentage = (4K)/(4E)	49.3%	13.4%	14.6%	
4N) Change in EVA Due to Reinsurance = NEVA - GEVA	-			

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	GROSS TOTAL
5A) Average Deviation from Plan When Exceed 1 in 50 Year Result (XTVAR)	(3,422,804)	(587,438)	(1,382,036)	(3,812,609)
5B) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,547,208	607,181	471,515	7,625,903
5C) Mean Interest Earned on Rating Agency Capital = (4F)	21,870	60,893	106,855	189,618
5D) Mean Rental Cost of Rating Agency Capital (4G)	72,901	152,232	213,710	438,843
5E) Expected Underwriting Return After Rental Cost of Capital = (4B)+(4D)+(5C)-(5D)	88,954	172,263	308,163	569,380
5F) Gross Risk Return on Capital = GRROC = (5E)/(5B)	1.36%	28.37%	65.36%	7.47%
5G) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,547,208	607,181	471,515	
5H) Net Risk Return on Capital = NRROC = (5E)/(5G)	1.36%	28.37%	65.36%	
5I) Change in Return Due to Reinsurance = (5E for LOB 4)	-			
5J) Change in Allocated Capital = (5G)-(5B)	-			

6) Returns on Risk Adjusted Capital (RORAC)	LOB 1	LOB 2	LOB 3	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,547,208	607,181	471,515	7,625,903
6B) Interest Earned on Gross Allocated Capital = (4C)x(6A)	196,416	24,287	23,576	244,279
6C) Gross Expected Total Underwriting Return = (4B)+(4D)+(6B)	336,401	287,889	438,594	1,062,884
6D) Gross Return on Risk Adjusted Capital = GRORAC = (6C)/(6A)	5.14%	47.41%	93.02%	13.94%
6E) Net Risk Capital K% of XTVAR, Allocated to Line Based Upon Co-XTVAR's	6,547,208	607,181	471,515	
6F) Interest Earned on Net Allocated Capital = (4C)x(6E)	196,416	24,287	23,576	
6G) Net Expected Total Underwriting Return = (4B)+(4D)+(6F)	336,401	287,889	438,594	
6H) Net Return on Risk Adjusted Capital = NRORAC = (6G)/(6E)	5.14%	47.41%	93.02%	
6I) Change in Return Due to Reinsurance = (6G - Net Total) - (6C - Gross Total)	-			
6J) Change in Allocated Capital = (6E - Net Total) - (6A - Gross Total)	-			

Exhibit 8

Stop Loss Reinsurance Example 8 Comparing EVA with Returns on Capital (RROC and RORAC)

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus: Yes*

Pricing is accurate, as the Plan Loss Ratio equals the Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines.

A 30% xs 90% LR Stop Loss reinsurance program is purchased for LOB 1 for a 10% rate. *All three lines are uncorrelated.*

Refer to Exhibits 1-7 for detailed descriptions of items below.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	*LOB 2*	*LOB 3*	*LOB 4*	*NET TOTAL*	*GROSS TOTAL*
4A) Plan Premium	1,250,000	1,250,000	1,250,000	(125,000)	3,625,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	87,500	(37,555)	262,445	300,000
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	3.0%		
4D) Mean Net Present Value of Interest Earned on Reserves	27,484	163,602	327,518	(3,890)	514,714	518,604
4E) Mean Rating Agency Capital	729,006	1,522,318	2,137,102	(69,680)	4,318,746	4,388,426
4F) Mean Interest Earned on Rating Agency Capital	21,870	60,893	106,855	(2,090)	187,528	189,618
4G) Mean Rental Cost of Rating Agency Capital	72,901	152,232	213,710	(6,968)	431,875	438,843
4H) Gross Expected Cost of Capital - Rental and Usage	340,294	187,406	309,891			837,590
4I) Gross Economic Value Added (GEVA)	(178,440)	137,090	211,982			170,631
4J) Gross Capital Cost Percentage	46.7%	12.3%	14.5%			19.1%
4K) Net Expected Cost of Capital - Rental and Usage	324,725	191,436	320,905	(57,522)	779,545	
4L) Net Economic Value Added (NEVA)	(162,871)	133,059	200,968	13,987	185,141	
4M Net Capital Cost Percentage	44.5%	12.6%	15.0%	82.6%	18.1%	
4N) Change in EVA Due to Reinsurance	14,510					

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	*LOB 2*	*LOB 3*	*LOB 4*	*NET TOTAL*	*GROSS TOTAL*
5A) Average 1 in 50 Year Deviation from Plan (XTVAR)	(3,421,737)	(587,394)	(1,380,739)	287,561	(3,273,740)	(3,543,084)
5B) Gross Risk Capital K% of XTVAR	6,441,898	83,285	561,977			7,087,161
5C) Mean Interest Earned on Rating Agency Capital	21,870	60,893	106,855	(2,090)	187,528	189,618
5D) Mean Rental Cost of Rating Agency Capital	72,901	152,232	213,710	(6,968)	431,875	438,843
5E) Expected Underwriting Return After Rental Cost of Capital	88,953	172,263	308,163	(36,567)	532,812	569,379
5F) Gross Risk Return on Capital = GRROC	1.38%	206.83%	54.84%			8.03%
5G) Net Risk Capital K% of XTVAR	6,425,757	85,226	599,625	(562,210)	6,548,397	
5H) Net Risk Return on Capital = NRROC	1.38%	202.13%	51.39%	6.50%	8.14%	
5I) Change in Return Due to Reinsurance	(36,567)					
5J) Change in Allocated Capital)	(538,763)	*5K) Cost of Additional XTVAR Capital = (5I)/(5J)*				6.8%

6) Returns on Risk Adjusted Capital (RORAC)	*LOB 1*	*LOB 2*	*LOB 3*	*LOB 4*	*NET TOTAL*	*GROSS TOTAL*
6A) Gross Risk Capital K% of XTVAR	6,441,898	83,285	561,977			7,087,161
6B) Interest Earned on Gross Allocated Capital	193,257	3,331	28,099			224,687
6C) Gross Expected Total Underwriting Return	333,241	266,934	443,117			1,043,291
6D) Gross Return on Risk Adjusted Capital	5.17%	320.50%	78.85%			14.72%
6E) Net Risk Capital K% of XTVAR	6,425,757	85,226	599,625	(562,210)	6,548,397	
6F) Interest Earned on Net Allocated Capital	192,773	3,409	29,981	(16,866)	209,297	
6G) Net Expected Total Underwriting Return	332,756	267,011	444,999	(58,311)	986,456	
6H) Net Return on Risk Adjusted Capital	5.18%	313.30%	74.21%	10.37%	15.06%	
6I) Change in Return Due to Reinsurance	(56,835)					
6J) Change in Allocated Capital	(538,763)	*6K) Cost of Additional XTVAR Capital = (6I)/(6J)*				10.5%

Exhibit 9

Quota Share Reinsurance Example 9 Comparing EVA with Returns on Capital (RROC and RORAC)

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus:* *Yes*
Pricing is accurate, as the Plan Loss Ratio equals the Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines.
A 40% Quota Share is purchased for LOB 1 with commission just covering variable costs. *All three lines are uncorrelated.*

Refer to Exhibits 1-7 for detailed descriptions of items below.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	(500,000)	3,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	87,500	(45,000)	255,000	300,000
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	3.0%		
4D) Mean Net Present Value of Interest Earned on Reserves	27,484	163,602	327,517	(10,994)	507,610	518,603
4E) Mean Rating Agency Capital	729,007	1,522,317	2,137,100	(248,282)	4,140,142	4,388,424
4F) Mean Interest Earned on Rating Agency Capital	21,870	60,893	106,855	(7,448)	182,169	189,618
4G) Mean Rental Cost of Rating Agency Capital	72,901	152,232	213,710	(24,828)	414,014	438,842
4H) Gross Expected Cost of Capital - Rental and Usage	340,069	187,586	309,695			837,351
4I) Gross Economic Value Added (GEVA)	(178,215)	136,908	212,177			170,871
4J) Gross Capital Cost Percentage	46.6%	12.3%	14.5%			19.1%
4K) Net Expected Cost of Capital - Rental and Usage	313,366	191,567	324,933	(121,014)	708,852	
4L) Net Economic Value Added (NEVA)	(151,512)	132,928	196,939	57,572	235,927	
4M Net Capital Cost Percentage	43.0%	12.6%	15.2%	48.7%	17.1%	
4N) Change in EVA Due to Reinsurance	65,057					

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
5A) Average 1 in 50 Year Deviation from Plan (XTVAR)	(3,422,444)	(587,552)	(1,381,531)	1,368,533	(2,310,833)	(3,542,615)
5B) Gross Risk Capital K% of XTVAR	6,490,236	93,743	502,173			7,086,151
5C) Mean Interest Earned on Rating Agency Capital	21,870	60,893	106,855	(7,448)	182,169	189,618
5D) Mean Rental Cost of Rating Agency Capital	72,901	152,232	213,710	(24,828)	414,014	438,842
5E) Expected Underwriting Return After Rental Cost of Capital	88,953	172,263	308,162	(38,614)	530,765	569,379
5F) Gross Risk Return on Capital = GRROC	1.37%	183.76%	61.37%			8.04%
5G) Net Risk Capital K% of XTVAR	5,527,702	173,740	1,131,906	(2,211,081)	4,622,267	
5H) Net Risk Return on Capital = NRROC	1.61%	99.15%	27.23%	1.75%	11.48%	
5I) Change in Return Due to Reinsurance	(38,614)					
5J) Change in Allocated Capital)	(2,463,885)	*5K) Cost of Additional XTVAR Capital = (5I)/(5J)*				1.6%

6) Returns on Risk Adjusted Capital (RORAC)	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR	6,490,236	93,743	502,173			7,086,151
6B) Interest Earned on Gross Allocated Capital	194,707	3,750	25,109			223,565
6C) Gross Expected Total Underwriting Return	334,691	267,352	440,126			1,042,169
6D) Gross Return on Risk Adjusted Capital	5.16%	285.20%	87.64%			14.71%
6E) Net Risk Capital K% of XTVAR	5,527,702	173,740	1,131,906	(2,211,081)	4,622,267	
6F) Interest Earned on Net Allocated Capital	165,831	6,950	56,595	(66,332)	163,044	
6G) Net Expected Total Underwriting Return	305,815	270,552	471,613	(122,326)	925,653	
6H) Net Return on Risk Adjusted Capital	5.53%	155.72%	41.67%	5.53%	20.03%	
6I) Change in Return Due to Reinsurance	(116,515)					
6J) Change in Allocated Capital	(2,463,885)	*6K) Cost of Additional XTVAR Capital = (6I)/(6J)*				4.7%

Exhibit 10

Quota Share Reinsurance Example 10 Comparing Alternative Parameterization of EVA with Returns on Capital

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus:* *Yes*
Pricing is accurate, as the Plan Loss Ratio equals the Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines.
A 40% Quota Share is purchased for LOB 1 with commission just covering variable costs. *All three lines are uncorrelated.*
The Consumption Fee for Capital Less than Allocation is assumed to be the same as the Consumption Fee for Common Capital.

Refer to Exhibits 1-7 for detailed descriptions of items below.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	(500,000)	3,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	87,500	(45,000)	255,000	300,000
4C) Interest Rate Assumed	3.0%	4.0%	5.0%	3.0%		
4D) Mean Net Present Value of Interest Earned on Reserves	27,490	163,602	327,516	(10,996)	507,612	518,608
4E) Mean Rating Agency Capital	729,057	1,522,317	2,137,091	(248,298)	4,140,167	4,388,465
4F) Mean Interest Earned on Rating Agency Capital	21,872	60,893	106,855	(7,449)	182,170	189,619
4G) Mean Rental Cost of Rating Agency Capital	72,906	152,232	213,709	(24,830)	414,017	438,846
4H) Gross Expected Cost of Capital - Rental and Usage	466,132	207,661	362,088			1,035,881
4I) Gross Economic Value Added (GEVA)	(304,271)	116,834	159,783			(27,654)
4J) Gross Capital Cost Percentage	63.9%	13.6%	16.9%			23.6%
4K) Net Expected Cost of Capital - Rental and Usage	425,503	213,480	384,642	(165,869)	857,757	
4L) Net Economic Value Added (NEVA)	(263,641)	111,015	137,228	102,424	87,025	
4M Net Capital Cost Percentage	58.4%	14.0%	18.0%	66.8%	20.7%	
4N) Change in EVA Due to Reinsurance	114,679					

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
5A) Average 1 in 50 Year Deviation from Plan (XTVAR)	(3,434,006)	(587,944)	(1,381,452)	1,373,154	(2,326,769)	(3,569,458)
5B) Gross Risk Capital K% of XTVAR	6,509,573	84,590	547,325			7,141,487
5C) Mean Interest Earned on Rating Agency Capital	21,872	60,893	106,855	(7,449)	182,170	189,619
5D) Mean Rental Cost of Rating Agency Capital	72,906	152,232	213,709	(24,830)	414,017	438,846
5E) Expected Underwriting Return After Rental Cost of Capital	88,956	172,263	308,161	(38,615)	530,765	569,380
5F) Gross Risk Return on Capital = GRROC	1.37%	203.65%	56.30%			7.97%
5G) Net Risk Capital K% of XTVAR	5,600,881	151,545	1,141,744	(2,240,352)	4,653,818	
5H) Net Risk Return on Capital = NRROC	1.59%	113.67%	26.99%	1.72%	11.40%	
5I) Change in Return Due to Reinsurance	(38,615)					
5J) Change in Allocated Capital)	(2,487,669)	*5K) Cost of Additional XTVAR Capital = (5I)/(5J)*				+1.6%

6) Returns on Risk Adjusted Capital (RORAC)	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR	6,509,573	84,590	547,325			7,141,487
6B) Interest Earned on Gross Allocated Capital	195,287	3,384	27,366			226,037
6C) Gross Expected Total Underwriting Return	335,277	266,986	442,382			1,044,645
6D) Gross Return on Risk Adjusted Capital	5.15%	315.62%	80.83%			14.63%
6E) Net Risk Capital K% of XTVAR	5,600,881	151,545	1,141,744	(2,240,352)	4,653,818	
6F) Interest Earned on Net Allocated Capital	168,026	6,062	57,087	(67,211)	163,965	
6G) Net Expected Total Underwriting Return	308,016	269,664	472,103	(123,206)	926,577	
6H) Net Return on Risk Adjusted Capital	5.50%	177.94%	41.35%	5.50%	19.91%	
6I) Change in Return Due to Reinsurance	(118,068)					
6J) Change in Allocated Capital	(2,487,669)	*6K) Cost of Additional XTVAR Capital = (6I)/(6J)*				4.7%

Exhibit 11

Quota Share Reinsurance Example 11 Comparing Alternative Parameterization of EVA with Returns on Capital

Key Assumptions: Write equal amounts of premium in three lines of business. *Interest Credited on Supporting Surplus:* *Yes*

Pricing is accurate, as the Plan Loss Ratio equals the Expected Loss Ratio (ELR) for all three lines. The ELR's are equal to 80% for all three lines

A 40% Quota Share is purchased for LOB 1 with commission just covering variable costs. *All three lines are uncorrelated.*

The Consumption Fee for Capital Less than Allocation is assumed to be 25% of the Consumption Fee for Common Capital.

Refer to Exhibits 1-7 for detailed descriptions of items below.

4) Economic Value Added (EVA) where Usage Charges Are Computed Using Two Step Formula

	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
4A) Plan Premium	1,250,000	1,250,000	1,250,000	(500,000)	3,250,000	3,750,000
4B) Expected Underwriting Return (Profit & Overhead)	112,500	100,000	87,500	(45,000)	255,000	300,000
4C) Interest Rate Assumed	3 0%	4.0%	5.0%	3.0%		
4D) Mean Net Present Value of Interest Earned on Reserves	27,484	163,602	327,518	(10,994)	507,610	518,604
4E) Mean Rating Agency Capital	729,007	1,522,317	2,137,103	(248,282)	4,140,145	4,388,427
4F) Mean Interest Earned on Rating Agency Capital	21,870	60,893	106,855	(7,448)	182,170	189,618
4G) Mean Rental Cost of Rating Agency Capital	72,901	152,232	213,710	(24,828)	414,014	438,843
4H) Gross Expected Cost of Capital - Rental and Usage	302,783	180,569	291,744			775,096
4I) Gross Economic Value Added (GEVA)	(140,929)	143,926	230,129			233,126
4J) Gross Capital Cost Percentage	41.5%	11.9%	13.7%			17.7%
4K) Net Expected Cost of Capital - Rental and Usage	280,343	184,040	304,683	(107,805)	661,261	
4L) Net Economic Value Added (NEVA)	(118,489)	140,455	217,191	44,363	283,519	
4M Net Capital Cost Percentage	38.5%	12.1%	14.3%	43.4%	16.0%	
4N) Change in EVA Due to Reinsurance	50,393					

5) Risk Returns on Capital (RROC) After Rental Cost of Capital

Risk Capital Standard (Multiple K of XTVAR): 200%

	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
5A) Average 1 in 50 Year Deviation from Plan (XTVAR)	(3,422,935)	(587,455)	(1,381,379)	1,368,729	(2,325,477)	(3,548,909)
5B) Gross Risk Capital K% of XTVAR	6,475,419	82,920	541,609			7,099,947
5C) Mean Interest Earned on Rating Agency Capital	21,870	60,893	106,855	(7,448)	182,170	189,618
5D) Mean Rental Cost of Rating Agency Capital	72,901	152,232	213,710	(24,828)	414,014	438,843
5E) Expected Underwriting Return After Rental Cost of Capital	88,953	172,263	308,163	(38,614)	530,766	569,379
5F) Gross Risk Return on Capital = GRROC	1.37%	207.75%	56.90%			8.02%
5G) Net Risk Capital K% of XTVAR	5,603,182	157,190	1,133,669	(2,241,273)	4,652,768	
5H) Net Risk Return on Capital = NRROC	1.59%	109.59%	27.18%	1.72%	11.41%	
5I) Change in Return Due to Reinsurance	(38,614)					
5J) Change in Allocated Capital)	(2,447,179)	5K) Cost of Additional XTVAR Capital = (5I)/(5J)				1.6%

6) Returns on Risk Adjusted Capital (RORAC)	LOB 1	LOB 2	LOB 3	LOB 4	NET TOTAL	GROSS TOTAL
6A) Gross Risk Capital K% of XTVAR	6,475,419	82,920	541,609			7,099,947
6B) Interest Earned on Gross Allocated Capital	194,263	3,317	27,080			224,660
6C) Gross Expected Total Underwriting Return	334,246	266,919	442,098			1,043,264
6D) Gross Return on Risk Adjusted Capital	5.16%	321.90%	81.63%			14 69%
6E) Net Risk Capital K% of XTVAR	5,603,182	157,190	1,133,669	(2,241,273)	4,652,768	
6F) Interest Earned on Net Allocated Capital	168,095	6,288	56,683	(67,238)	163,828	
6G) Net Expected Total Underwriting Return	308,079	269,890	471,701	(123,232)	926,439	
6H) Net Return on Risk Adjusted Capital	5.50%	171.70%	41.61%	5.50%	19.91%	
6I) Change in Return Due to Reinsurance	(116,825)					
6J) Change in Allocated Capital	(2,447,179)	6K) Cost of Additional XTVAR Capital = (6I)/(6J)				4.8%