Casualty Actuarial Society E-Forum, Summer 2010

The Summer 2010 CAS *E-Forum*

The Summer 2010 Edition of the CAS *E-Forum* is a cooperative effort between the Committee for the CAS *E-Forum* and various other CAS committees.

2010 Summer *E-Forum*

Table of Contents

A Structural Simulation Model for Measuring General Insurance Risk
Timothy Gault, Len Llaguno, FCAS, MAAA, and Stephen Lowe, FCAS, MAAA, CERA............ 1-57

Credibility for Experience Rating, A Minimum Variance Approach
Lawrence F. Marcus, FCAS, MAAA.. 1-18

E-Forum Committee

Glenn M. Walker, *Chairperson*

Mark A. Florenz
Karl Goring
Dennis L. Lange
Elizabeth A. Smith, *Staff Liaison*
John Sopkowicz
Zongli Sun
Windrie Wong
Yingjie Zhang

For information on submitting a paper to the *E-Forum*, visit http://www.casact.org/pubs/forum/.

A Structural Simulation Model for Measuring General Insurance Risk

Timothy Gault
Len Llaguno, FCAS, MAAA
Stephen Lowe, FCAS, MAAA, CERA
Towers Watson

Abstract

Motivation. As use of economic capital models expands, the need for a robust approach to the measurement of reserving and pricing risk becomes increasingly important. The paper describes a stochastic simulation model developed by the authors that has some attractive advantages over other published approaches to risk measurement. A particular issue is the need to measure reserving and pricing risk over a one-year time horizon; the model does both one-year and run-off risk measurement.

Method. The Structural Model separates overall insurance risk into systematic and non-systematic risk elements, using an Economic Scenario Generator to simulate the former and the Practical method to simulate the latter. In addition, it combines stochastic chain ladder projections of past years with a stochastic ARMA loss ratio model for recent and future years, facilitating the measurement of reserving and pricing risk in an integrated way.

Results. The Structural Model offers several benefits, described in the paper and illustrated using an empirical dataset. Illustrative validation results are also presented. These include some useful ideas about validation that may be applied to other approaches, as well.

Conclusions. The Structural Model is a practical approach to measuring reserve and pricing risk in an integrated way, over either a one-year or run-off risk horizon. Its ability to separate systematic economic risks from general claim misestimation risk is particularly relevant, given the concerns about a resurgence of inflation. It can be successfully validated using historical data on past reserve and pricing errors.

Availability. No software is being made available with the paper.

Keywords

Economic capital, stochastic reserving, financial modeling, inflation risk, economic scenario generator, ARMA model

---

1. INTRODUCTION

The measurement of General Insurance[1] risk is becoming increasingly important, as insurers work to create stronger linkages between portfolio risk, capital utilization and value creation. When combined with a security standard, a General Insurance risk profile can be used to calculate the economic capital required to support a given class of business. It is also a necessary input to models used in the measurement of the Fair Value of claim liabilities and other market-consistent liability measurement schemes. The same risk profile can also be used to assess the adequacy of returns in relationship to the risk, and to set risk-based technical price margins.

This paper describes a method for measuring General Insurance risk, using a structural stochastic simulation approach. It builds upon work by many previous authors, particularly Butsic [1]; Kelly [7]; and Hodes, Feldblum, and Blumsohn [4].

We describe the method as ***structural*** because it seeks to separate the overall risk into (a) systematic[2] elements that are driven by key socio-economic factors such as inflation and unemployment, and (b) non-systematic elements that are more truly random in nature (although there may be a dependency structure to the randomness). Application of the method requires that the user specify the structure of the risk, in contrast to other simulation approaches which typically assume that risk can be represented by a series of independent random variables within a stationary stochastic process. The method seems particularly apropos given the heightened concerns about a resurgence of inflation.

After this introductory section, the next section provides a formal definition of General Insurance risk and articulates the desired characteristics of the model that we sought to develop. We believe the structural model performs well against these desired characteristics.

Section 3 provides a description of the structural simulation model, using a "peel the onion" approach; we start with a basic overview, and then gradually drill down to provide successive detail on the parameterization and operation of the model. The structural model requires specification of many parameters, giving the user a great deal of flexibility. The price one pays for this flexibility is

[1] In the U.S., General Insurance is usually referred to as "Property & Casualty Insurance." In other parts of the world it is referred to as "Non-Life Insurance." Because our model was developed and tested using data from a multinational insurer, we have chosen to adopt their more global terminology.
[2] In this context *systematic risk* is a general term reflecting socio-economic risks that affect all claims; it should not be taken as limited to *equity-systematic* risk as defined in the Capital Asset Pricing Model.

the need for care in setting the parameters, and the need for validation via comparison to historical experience. We conclude this section with a discussion of alternative approaches to validation.

Finally, in Section 4 we provide some test results from the model using publicly available data from a large personal lines insurer. We compare results from the structural model to comparable results for other methods; and also show some validation results, comparing model results to historical reserve errors and pricing volatility.

2. MEASURING GENERAL INSURANCE RISK

Before describing the model itself, we begin with a definition and discussion of the problem we are trying to solve.

2.1 The Nature of General Insurance Risk

The emergence of General Insurance claim payments is subject to a number of underlying stochastic processes relating to the incidence of insured events, the recognition and reporting of claims, the adjusting of the claim, estimates of the value of the claim by the adjuster, and the claim settlement process. These underlying stochastic processes reflect both randomness and systematic variation due to changing socio-economic conditions that may influence the numbers and costs of all claims. As a result, forecasts of future General Insurance claim payments are uncertain.

> *Definition: General Insurance risk is the risk of adverse development in the cost[3] of unpaid claims (net of any offsetting premium or tax benefits) from their current actuarial central estimate.*

Often this risk is represented by a conveniently tractable probability distribution, such as a Lognormal or Pareto. In other instances it is represented by an empirical distribution, for example produced as the output of a hurricane or earthquake simulation model.

[3] We will define "cost" more fully below.

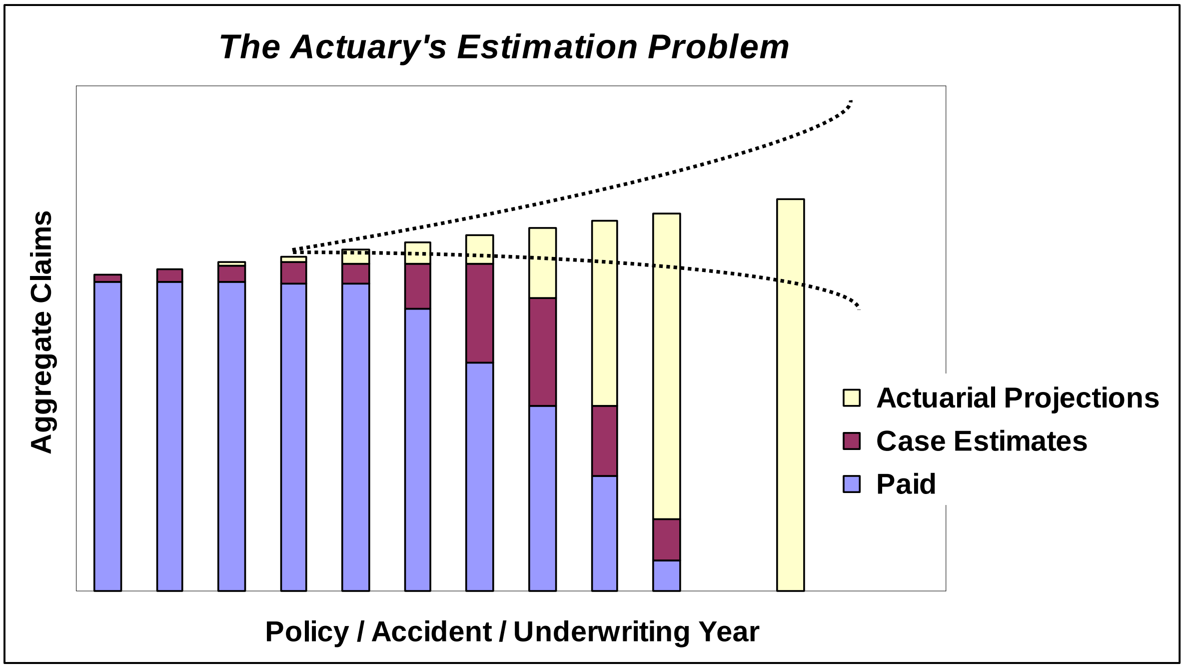
For an insurer who has been in business for many years, the total insurance risk associated with a particular class of business will include the risk associated with all unpaid claims on past business, and also the risk associated with potential claims on future business about to be written. Actuarial central estimates for both types of claim liabilities will be based on the actuary's interpretation of historical experience. While central estimates of the ultimate claims for older accident[4] years will be relatively certain (because many of the claims are paid, and case reserves are relatively reliable), central estimates for more recent accident years will reflect increasing levels of uncertainty (as successively fewer claims will be reported and paid). The actuary's estimation problem is depicted schematically in Chart 2.1, with the central estimates for each accident year represented by a bar and the increasing level of uncertainty represented by the dotted-line "funnel of doubt."

A key point in this schematic is that the "reserving risk" and "pricing risk" components of insurance risk are interrelated. The central estimates on past business are used to construct the central estimate claim costs for pricing of future business. If the central estimates of the former are too low, the central estimates of the latter are likely to be too low as well. The extent of dependency will depend on the characteristics of the class of business. For many classes, we would expect the two risk components to be highly correlated.

Note that under our definition not all variation in underwriting results, or all types of reserve inadequacies, would be classified as insurance risk. For example if a company knowingly followed market prices down in a softening market, any economic losses sustained would be expected, not a variance from expected. Similarly, if management chose to hold claim reserves below the central actuarial estimate, then any resulting adverse development would be expected, not a variance from expected. These types of risks are more appropriately classified as operational risks, rather than insurance risks, as they relate to failures in underwriting and financial controls.

[4] For ease of exposition, the term accident year is used throughout this paper. The concepts apply equally in a policy year, underwriting year, or other cohort-based configuration of claim data.

Chart 2.1: Actuarial Estimation of Claim Liabilities Creates Insurance Risk



While our definition of General Insurance risk and the chart above provide a simple and intuitive depiction of the uncertainty, our framework needs to be refined to account for varying risk horizons. An analogy may be helpful to understand the risk horizon issue. When we discuss market or credit risk in the context of assets, we have to specify the holding period over which we want to measure the risk. For example, the default probability of an investment grade corporate bond is relatively low over a one-year holding period, but can be substantially higher over a longer period. It is highest when the bond is held to maturity.

In an analogous way, insurance risk depends on the period over which it is measured. This period is referred to generically as the risk horizon. The two horizons of greatest interest are the run-off risk horizon (i.e., to "maturity" when all claims are settled) and the one-year risk horizon. Fundamentally, these two risk horizons ask the actuary two different, but related questions:

- In the context of a run-off risk horizon: ***What is the potential adverse variation in the ultimate cost[5] of claim liabilities from the current actuarial central estimate?***
- In the context of a one-year risk horizon: ***What is the potential adverse change in the actuarial central estimate of ultimate claim costs that could occur, with the benefit of one additional year of actual claim emergence, other relevant information, and changes in circumstances that affect the valuation?***

In the context of the run-off risk horizon, the cost of the claim liabilities is measured as the present value of the claim payments and the associated claim handling expenses, discounted using a risk-free yield curve. However, in the context of a one-year risk horizon, the cost of the claim liabilities is measured as the theoretical price at which they could be transferred to a willing third party. In addition to the expected present value of the future claim payments, the price would include a margin for the cost of the economic capital required to support the risk of the liabilities. Returning to our bond analogy, the credit risk associated with a bond over its lifetime is simply the risk of non-payment of any of its coupons or principal; however, the credit risk associated with a bond over a single year will depend on credit-driven movements in the price of the bond, including changes in the probabilities of non-payment of the remaining coupons and principal and the change in the market's price of the remaining credit risk. Concepts and issues relating to alternative risk horizons are discussed more fully in a forthcoming paper by Jing, Lowe and Morin [6].

The one-year risk horizon introduces a new complexity to the problem, because it requires us to go beyond estimating the degree to which our estimates are uncertain. To measure risk over a one-year horizon we must also model how the uncertainty resolves over time — and in particular how much uncertainty resolves during the upcoming year.

[5] Technically, the risk relates to adverse movements in claim costs, net of any offsetting premium or tax benefits. Premium benefits would relate to the additional premiums due on loss sensitive insurance products. Tax benefits would relate the additional tax deductions associated with the adverse development; these would be limited by recoverability considerations. Consideration of these adjustments is outside of the scope of our paper.

Mathematically, we can express the two questions above as follows. Let:

$$C_{a,d}, \quad a = 1,\ldots,n, \qquad d = 1,\ldots,m, \tag{2.1}$$

represent the actual incremental paid claims on the a th accident year in the d th development period. Since they are the result of stochastic processes relating to claim occurrence, reporting and settlement, it is customary to treat all the $C_{a,d}$ as random variables.

We define time $t = 0$ to be the starting point of the first accident year. At time $t = n$ we will then typically have a triangular array of actual values of $C_{a,d}$, for all $C_{a,d}$ where $a + d \le n + 1$, as illustrated below for the case where $n = m = 4$.

$$\begin{bmatrix} C_{1,1} & C_{1,2} & C_{1,3} & C_{1,4} \\ C_{2,1} & C_{2,2} & C_{2,3} & . \\ C_{3,1} & C_{3,2} & . & . \\ C_{4,1} & . & . & . \\ . & . & . & . \end{bmatrix}$$

Note that the array depicted above includes an extra, empty row at the bottom. This row represents the incepting accident year, on which no claims will have occurred or been paid at time $t = n$. Inclusion of this row in the array allows us to consider pricing risk as well as reserving risk, in an integrated way.[6]

These actual $C_{a,d}$ are a sample realization from the underlying stochastic processes; they are an incomplete realization, which will be completed when the balance of the array is filled in. The *reserving* process involves developing a central estimate for all of the missing future $C_{a,d}$, except for the latest row. These latter values are developed via the *pricing* process. The reserving process typically involves a projection across the rows, applying some variant of the chain ladder method to the observed values for each accident year. Since there are not yet any observed values for the incepting accident year, the pricing process typically involves some form of projection down the columns. Insurance risk stems fundamentally from the stochastic nature of the claim payments,

[6] For simplicity we will assume that all policies are annual and they all incept on the first day of the year, such that the accident year and the policy year are identical. In actual application, one may need to consider how to treat the unexpired risk at the end of each year.

which manifests itself as (a) the inability to observe more than a single random sample of the past claim payments, and (b) the randomness of the future claim payments.

The present value of the actual outstanding claim liabilities for an accident year at time t are given by[7]:

$$L_a^{(t)} = \sum_{d=t+2-a}^{m} (v_{d+a-t-1}) \times (C_{a,d}) \tag{2.1a}$$

where $v_{d+a-t-1}$ is the discount factor applicable to payments made $d+a-t-1$ periods after t[8]. For example, $L_{n+1}^{(n)}$ are the actual outstanding claim liabilities for the incepting accident year at time $t=n$, before any claim payments have been made.

The aggregate present value of the total outstanding claim liabilities are given by:

$$L^{(t)} = \sum_{a=1}^{n+1} L_a^{(t)} \tag{2.1b}$$

Let $\hat{C}_{a,d}^{(t)}$ represent the actuarial central estimate of the expected incremental paid claims on accident year a in development period d, based on information at time t. Then the present value of the estimated outstanding claim liabilities for an accident year at time t_1, based on information at time t_2, are given by:

$$\hat{L}_a^{(t_1,t_2)} = \sum_{d=t_1+2-a}^{m} (v_{d+a-t_1-1}) \times (\hat{C}_{a,d}^{(t_2)}) \tag{2.2a}$$

The aggregate actuarial central estimate of the outstanding claim liabilities are given by:

$$\hat{L}^{(t_1,t_2)} = \sum_{a=1}^{n+1} \hat{L}_a^{(t_1,t_2)} \tag{2.2b}$$

The normal case in equations (2.2a) and (2.2b) would be for $t_1 = t_2 = n$. (When $t_1 = t_2$ we will simplify the notation to a single superscript parameter.) The introduction of the additional time variable t_2 allows us to generalize the liability estimates to reflect varying degrees of hindsight, in which the $\hat{C}_{a,d}^{(n)}$ estimated payments are successively replaced by $C_{a,d}$ actual payments, and the remaining $\hat{C}_{a,d}^{(n)}$ estimates are replaced by revised estimates $\hat{C}_{a,d}^{(n+1)}$, $\hat{C}_{a,d}^{(n+2)}$, etc.

---

[7] The use of parentheses around the parameter *t* is to distinguish it from an exponent in the formulae.

[8] The nominal value of the liabilities is, of course, found by setting the present value factors to one.

As was noted at the outset of this section, under the run-off risk horizon, risk is measured from the potential variation of the ultimate cost of outstanding claim liabilities from the current actuarial central estimate. From the definitions above, we can see that General Insurance run-off risk at time $t = n$ can be measured using the probability distribution of the error quantity:

$$e_r = \hat{L}^{(n)} - L^{(n)} \qquad (2.3a)$$

$$e_r = \sum_{a=1}^{n+1} \sum_{d=t_1+2-a}^{m} (v_{d+a-t_1-1}) \times (\hat{C}_{a,d}^{(t_2)} - C_{a,d}) \qquad (2.3b)$$

In contrast, under the one-year risk horizon, risk is measured from the potential changes in the market-consistent value of the outstanding claim liabilities over a single year. Since the one-year risk horizon postulates the cost of the potential transfer of the outstanding liabilities to a third party at the end of the year, it is necessary to focus on valuations rather than just on net present values. For simplicity we define a market-consistent valuation by the addition of a risk margin to the calculated present value. This approach has been suggested by the CRO Forum in their recommendations for Solvency II implementation, based on cost-of-capital considerations.

$$\hat{V}^{(t_1,t_2)} = \sum_{a=1}^{n+1} \sum_{d=t_1+2-a}^{m} (v_{d+a-t_1-1}) \times [(\hat{C}_{a,d}^{(t_2)})] + M^{(t_1,t_2)} \qquad (2.4)$$

Employing the definition above, General Insurance risk can be measured at time $t = n$ using the probability distribution of the error quantity:

$$e_1 = \hat{V}^{(n.n)} - V^{(n,n+1)} \qquad (2.5a)$$

$$e_1 = \hat{V}^{(n,n)} - v_1 (\sum_{a=1}^{n+1} C_{a,t_1+2-a} + \hat{V}^{(n+1,n+1)}) \qquad (2.5b)$$

Where the C_{a,t_1+2-a} are the diagonal of actual claim payments in the next calendar year and v_1 is the discount factor for 1 year at the risk-free rate.

In the one-year risk horizon, risk measurement is concerned both with the movement in the actuarial central estimate over the course of one year, and also with any movement in the risk margin during the same period. At its current state of development, the structural simulation model focuses solely on the movement of the actuarial central estimate. The authors hope to incorporate movements in the risk margin in a subsequent version.

2.2 Desired Characteristics of the Risk Model

While many papers have been written proposing stochastic models for measuring reserving risk, the majority of those papers do not address how the models can be extended to include pricing risk. We sought to develop a model that would measure the totality of insurance risk, incorporating both reserving and pricing components, and explicitly recognize that the two components are interrelated.

General Insurance risk can be measured in two ways.

- *Hindsight Testing*, in which one measures historical claim estimation errors by comparing actuarial central estimates made in the past to the actual subsequent claim emergence. (For example, see Jing, Lebens, and Lowe [5]).
- *Stochastic Modeling*, in which one postulates a model of the stochastic process and measures the risk by applying the model with parameters derived from historical claim experience. Stochastic models can be analytic (for example, regression-based as in Murphy [10]), or simulation-based (for example, Bootstrap, as outlined by England and Verrall [3]).

The principal advantage of hindsight testing is that it is non-parametric; in other words, it does not rely on the assumption of any specific underlying model. Risk measures based on hindsight testing therefore reflect the total risk present in the underlying stochastic processes (even if we can't specify them). Its principal disadvantage is that it requires an extensive history of past estimates and associated run-off claim data. In contrast, stochastic modeling requires less historical information; the data requirements are usually the same as those necessary for estimating claim liabilities, typically the current paid or incurred loss development triangle. The principal disadvantage of stochastic methods is that they are model dependent, requiring estimation of parameter and model error. If one is to use the stochastic modeling approach, it therefore seems important to perform validation testing of the approach. Unfortunately, some of the validation testing of published stochastic models indicates that they perform relatively poorly in many real-world situations[9]. We sought to develop a stochastic model that performed well in validation tests.

As was mentioned earlier, the stochastic processes underlying General Insurance claim emergence include systematic as well as non-systematic (i.e., random) elements. Both general price

---

[9] See, for example, the ROC/GIRO Working Party Paper [12], presenting test results suggesting that the Mack method may understate risk.

inflation and sector-specific cost inflation (for example, traumatic medical care or auto damage repair) affect the ultimate cost of claims. Levels of economic activity, unemployment rates, capital market and credit market conditions all may also affect claim frequency and severity. Historical claim development triangles will reflect the changing state of these systematic variables. In contrast, many of the published stochastic methods assume that the historical loss development is a realization of a stationary random process. We sought to develop a stochastic model that accommodated both systematic and non-systematic components of risk. Such an approach has the advantage of offering much more explanatory value, for example by allowing the user to isolate the impact of inflation risk. This seems particularly relevant in the current economic environment.

To the extent that claims from different classes of business are affected by the same socio-economic variables, the model will provide insight into the correlation of insurance risk across classes of business. Also, since these socio-economic variables may influence asset behavior (at least in the longer term), they provide a linkage between risks on the asset and liability sides of the balance sheet.

Another shortcoming of many of the published stochastic models is that they rely on a particular projection method (typically the chain ladder), usually applied to paid claim development data to avoid the problems of downward development present in reported claim development data. While such models are reasonably well-behaved when applied to classes of business with stable short-tail claim development, several problems arise when they are applied to classes with more volatile long-tail development. First, they often "blow up" when applied to these datasets, indicating extremely high uncertainty. Second, since they usually assume that the development factors from one maturity to the next are independent, they often mistake volatility in the timing of payments for volatility in the ultimate amounts of the payment. More fundamentally, they are really measuring the uncertainty associated with a paid chain ladder projection; and for volatile long-tail lines most actuaries would agree that a paid projection is not an appropriate estimation method, due to the high leverage in the compound development factors. For these classes, actuaries would be more likely to utilize the reported chain ladder method, or a Bornhuetter-Ferguson approach, or even an expected loss ratio approach as the basis of their estimates of the ultimate claims — especially for the more recent accident years. We sought to develop a stochastic model that more faithfully reflected the actual approach to estimation that would take place for these classes.

Finally, we sought to develop a stochastic method that would measure risk over a one-year risk horizon in addition to a run-off risk horizon. The need to look at risk over a one-year horizon is becoming increasingly important in the context of economic capital. Unfortunately, virtually all of the published stochastic models have focused entirely on the run-off risk horizon question. To our knowledge, only Wacek [13] and Wuthrich, Merz and Lysenko [14] have addressed the issue of modeling the one-year risk horizon.

In summary, we sought to develop a stochastic model that met the following needs.

- It would measure reserving and pricing risk in an integrated way, capturing the relationship between the two.
- It would perform well in validation tests.
- It would incorporate both systematic and non-systematic sources of risk.
- It would facilitate the measurement of correlation between classes of business.
- It would reflect actuarial estimation methods actually employed for each class of business.
- It would be capable of measuring General Insurance risk in the context of either a one-year or a run-off risk horizon.

3. THE STRUCTURAL SIMULATION MODEL

In this section, we describe the mechanics of the structural simulation model. We describe both the general concept and the current implementation. The latter reflects where we have traveled thus far, while the former reflects where we ultimately want to go with the development and implementation of the model.

3.1 Overview of the Structural Model

In the structural simulation, claim emergence and development is assumed to be decomposable into the following two components:

- An underlying stationary claim emergence pattern that is the culmination of underlying stochastic claim processes subject to random noise.
- One or more socio-economic factors which distort the stationary emergence pattern, "stretching" or "shrinking" it in some way. Each of these socio-economic factors may itself entail a stochastic process subject to random noise.

Mathematically, we can describe this generically as:

$$C_{a,d} = f(\overline{C}_d, Z) + \varepsilon_{a,d} \tag{2.6}$$

Where $\overline{C}_d$ is the expected payment generated from the underlying stationary emergence, Z is the set of socio-economic factors that modify the payment, and $\varepsilon_{a,d}$ is the residual random error from the stochastic processes. Note that in this general form Z can include past, as well as current values of any socio-economic variables. For example, if calendar year inflation affects development year claim payments, then Z would need to include a vector of calendar year inflation rates for the entire development period.

Since the socio-economic factors affect all claims, they represent the systematic aspect of insurance risk, while the error term $\varepsilon_{a,d}$ represents the non-systematic risk. In the balance of the paper, we will sometimes refer to the socio-economic variables as systematic risk variables.

Typically, the socio-economic factors will be represented by a system of stochastic equations that describe the behavior of each variable over time and how each variable interacts with other variables. Because the equations describe a dynamic system with structure, they are sometimes referred to as a structural model. Hence, we describe our overall model as a structural simulation.

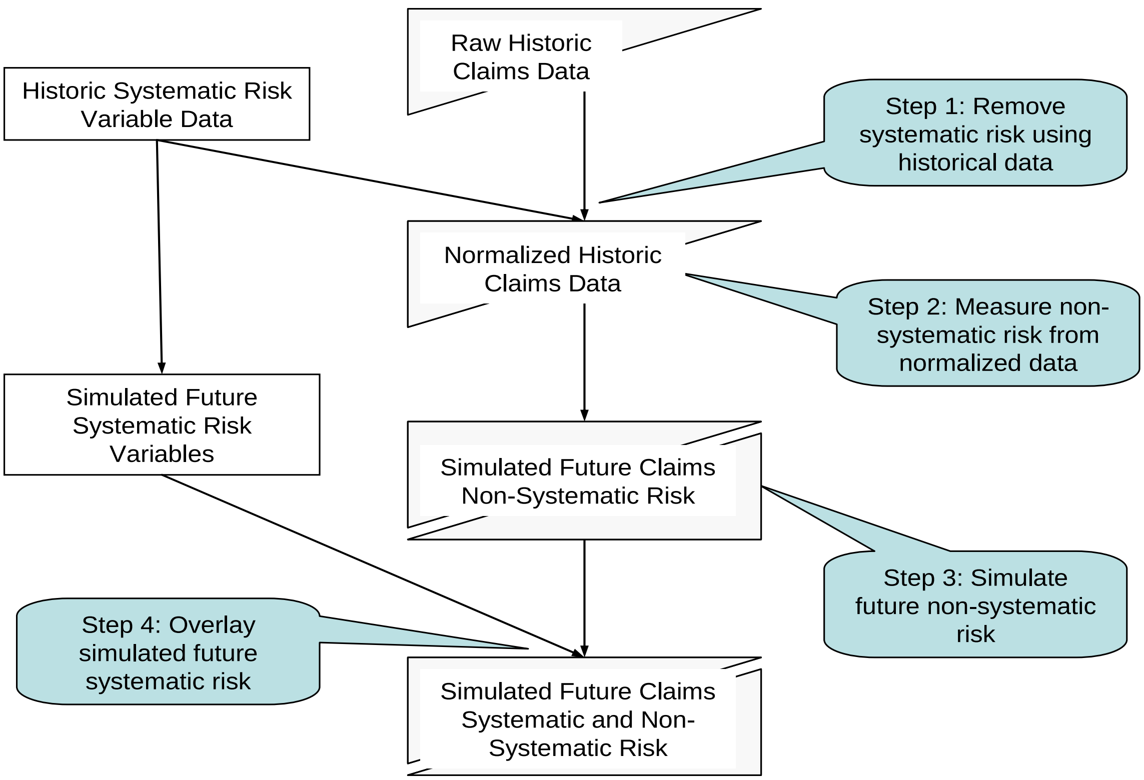
To apply the model, one must determine the relevant Z for the class of business, and the functional form of f. Methods for doing this are beyond the scope of this paper. Taylor [11], Zehnwirth [15], and Christofides [2] have suggested methodologies that may represent workable approaches. In addition, some of the approaches we have taken reflect as-yet unpublished research by the authors' firm.

As a matter of practicality, rather than removing the effects of Z completely from the claim development, it is often simpler to substitute constant "steady-state" expected values of the systematic risk variables Z for the actual varying historical values. Then $\overline{C}_d$ would reflect the stationary emergence under the chosen steady-state values of Z, and the "distortions" would reflect variation of Z from those chosen constant values. The chosen steady-state values can either be normative long-term historical means or current expectations of long-term future means.

Chart 3.1 presents a high-level schematic of the key steps in applying the structural simulation model. As can be seen in Chart 3.1, the model proceeds in four distinct steps. The first two steps entail parameterization of the model; the last two entail running it.

- The first step is to remove systematic risk elements from the historical data, by substituting the effect of constant normative systematic variables for the varying effect of the historical values of those variables.
- Once the systematic risk elements have been normalized out of the historical data, the triangle should reflect only non-systematic noise emanating from a stationary stochastic process. Stationarity can be tested at this point to assure that it is present within a reasonable tolerance. The second step is to measure the amount of non-systematic noise in the normalized historical data and select non-systematic risk parameters for the simulation.

Chart 3.1: Overview of Structural Simulation Procedure



- The third step is to simulate future non-systematic risk for purposes of applying to future development, using the parameters developed in the second step. For each trial, we simulate future development factors and other projection calculations based on the observed historical volatility of the normalized stationary data set.
- The fourth and final step is to overlay future systematic risk. To do this, we use an Economic Scenario Generator ("ESG"), to generate plausible paths of future values for each systematic risk variable. For each trial, we adjust the simulated future development for the effect of the difference between the normative values of systematic variables and the simulated values for that specific trial.

The output from the model are simulated distributions of e_r and e_1, the run-off and one-year claim estimation errors, respectively. Estimation errors can be calculated on a nominal basis, i.e., by setting the discount factors, $(v_{d+a-t-1})$ equal to one. Alternatively, they can be calculated on a present value basis, typically using discount factors based on the risk-free yield curve.

Each of the steps outlined above will be discussed in greater detail in subsequent sections, but before jumping directly to the full detail, we present a simplified example to help the reader understand the basics of the process. The same four steps outlined above are presented in the example. Note that the example ignores a number of technical parameterization issues, which are covered in the subsequent sections.

3.2 A Simplified Example

In Step 1, we begin the model parameterization effort by removing the systematic risk from the historical development triangle. For the purposes of our simple example, we will assume that the only systematic risk is monetary inflation, as measured by the U.S. Consumer Price Index (CPI). Further, we will assume that inflation fully accrues up to the time of the claim payment. Removing the systematic inflation risk requires that we de-cumulate the cumulative claim payment triangle into an incremental triangle, adjust the incremental claim payments to reflect stationary inflation conditions, and re-accumulate the adjusted incremental claim payment triangle to produce a normalized cumulative claim payment triangle.

In our example in Chart 3.2, the actual historical incremental claim payment amounts are 'inflated' by the actual historical CPI index to 2007 levels, and then 'deflated' using an inflation index reflecting constant 2.5% inflation. The net adjustment factors are shown in the upper left table of Chart 3.2. For example, the actual claim payments for the first evaluation period for accident year 2005 are $755. Since these payments occurred in calendar year 2005, they would be inflated to 2007 levels by dividing them by .964; then they would be deflated to reflect constant inflation by multiplying them by .952. The net of these factors is .9874, producing the adjusted paid claims of $745.

In this case, the best choice for the constant inflation rate would be the expected future inflation rate, so that the central value of the resulting claim projections would explicitly reflect current inflation expectations.

Chart 3.2: Step 1 — Removing Systematic Risk

Factors to adjust payments from actual to expected inflation

Calendar Year	Actual Inflation Index	Expected Inflation Index	Net Adjust. Factor
2004	0.948	0.929	0.9797
2005	0.964	0.952	0.9874
2006	0.981	0.976	0.9941
2007	1.000	1.000	1.0000

Actual cumulative paid claim triangle

Accident Year	Earned Premium	Evaluation Date 12	24	36	48
2004	2,243	650	1,049	1,257	1,406
2005	2,340	755	1,218	1,477	
2006	3,288	881	1,483		
2007	2,937	950			
2008	3,000				

Normalized cumulative paid claim triangle

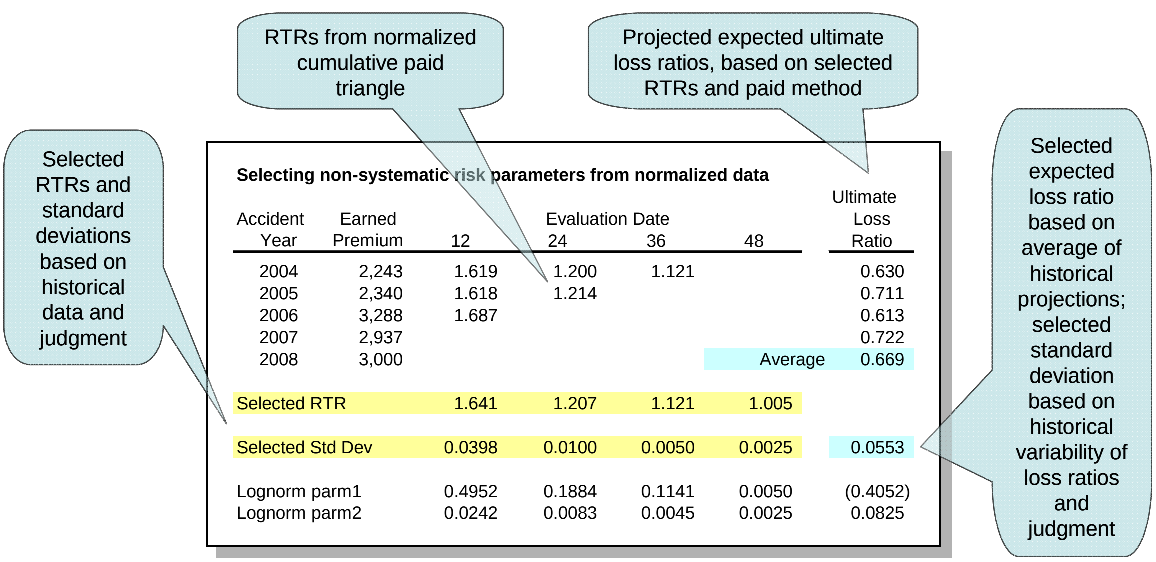
Accident Year	Earned Premium	Evaluation Date 12	24	36	48
2004	2,243	637	1,031	1,237	1,387
2005	2,340	745	1,206	1,465	
2006	3,288	876	1,478		
2007	2,937	950			
2008	3,000				

The historic claim data is normalized by removing historic inflation disturbances from the incremental payments in each calendar year (i.e., by diagonal) and substituting constant inflation expectations

The remaining noise in the development pattern implied by the normalized cumulative loss triangle is attributable solely to non-systematic factors. The normalized triangle can be tested to confirm that the remaining noise is reasonably random, indicating that the normalized triangle reflects a stationary development process.

In Step 2, we develop the parameters necessary to simulate future non-systematic risk. Chart 3.3 below displays the parameterization process. For this simple example, we will assume that (a) the future paid claims on all accident years follows a stochastic paid chain ladder process; and (b) the ultimate losses for the current accident year follow a lognormal process around a temporally stationary expected loss ratio. Ignoring sampling error, the parameters for the stochastic chain ladder process are based on the historical normalized report-to-report development factors (RTRs). Since the observed normalized RTRs are random draws from the underlying lognormal distribution, we can calculate the lognormal parameters from the means and standard deviations of the observed RTRs at each development maturity. The standard deviations for the last two development periods are judgmentally selected, as is the mean for the last development period.

Chart 3.3: Step 2 — Selecting Non-Systematic Risk Parameters



Selecting non-systematic risk parameters from normalized data

Accident Year	Earned Premium	12	Evaluation Date 24	36	48	Ultimate Loss Ratio
2004	2,243	1.619	1.200	1.121		0.630
2005	2,340	1.618	1.214			0.711
2006	3,288	1.687				0.613
2007	2,937					0.722
2008	3,000				Average	0.669
Selected RTR		1.641	1.207	1.121	1.005	
Selected Std Dev		0.0398	0.0100	0.0050	0.0025	0.0553
Lognorm parm1		0.4952	0.1884	0.1141	0.0050	(0.4052)
Lognorm parm2		0.0242	0.0083	0.0045	0.0025	0.0825

The selected RTRs can be interpreted as the expected development at the expected inflation rate. The standard deviations can be interpreted as the variability around the selected RTRs that is due to non-systematic risks associated with the underlying stochastic claim development process. For example, from 12 to 24 months we expect cumulative paid claims to develop upward by 64.1%; the actual development in any year will vary lognormally from that expected value, with a standard deviation of 3.98%.

We also project the expected ultimate loss ratios for the four past years, and measure the mean and the standard deviation of the loss ratios across the four years. This gives us the parameters for the stochastic lognormal accident year loss ratio process, applicable to the 2008 accident year. In our example, the average loss ratio over the past four years is .669. Again, ignoring sampling error, we will assume that .669 is the expected loss ratio for all accident years, and that the standard deviation of the past four years around that average is the expected standard deviation for the 2008 accident year loss ratio due to non-systematic risk.

Based on our selected parameters, the expected unpaid claims on the prior accident year are $1,898. The expected losses on the current accident year are $2,007.

Having parameterized the model, we are now ready to begin the simulation process. In Step 3 we simulate the future emergence of claims using the paid chain ladder model, and the expected loss ratio model for the current year. This process is displayed in Chart 3.4.

We start with the current levels of paid losses (from the actual triangle, the first diagonal in the lower table). We then simulate the future emergence of the past accident years via random draws of the lognormal RTRs to complete the triangle and obtain ultimate losses. For example, the simulated RTR of 1.122 for the 2005 accident year from 36 to 48 months is a random draw from a lognormal distribution with a mean of 1.121 and a standard deviation of .0050 (the selected parameters in Chart 3.3). When applied to actual paid losses of $1,477, we obtain simulated cumulative claim payments at 48 months of $1,657. The same process simulates the development from 48 to ultimate, producing simulated ultimate losses for the 2005 accident year of $1,671.

To simulate losses for the pricing accident year, we take the average of the simulated past year loss ratios and perform a random draw from a lognormal distribution with that average as our mean. In the simulation trial shown, the average loss ratio is .673 (versus an original expectation of .669) and the simulated loss ratio is .744.

As a final step, the simulated 2008 losses are "spread back" to create simulated claim payments using the simulated RTRs for the current accident year.

With this step, each trial is an alternative realization of the ultimate claim liabilities for past and current accident years, reflecting only non-systematic risk, and with constant normative systematic levels. Note that the ultimate claims for the current accident year are dependent on what happens with the development of the past accident years. When simulated RTRs are high, ultimate claims for past years will be high, and the average historical loss ratio (used as the expected loss ratio for the current accident year) will be high. Thus, reserve and pricing risk are inter-linked in the model. In our simulated trial, the actual unpaid claim liabilities on past accident years are $1,944 (2.4% higher than expected), and the expected claim liabilities on the current accident year are $2,231 (11.1%

higher than expected). The latter difference reflects the combined effect of both a higher expected value, and a higher than expected random draw around the higher expected value.

Chart 3.4: Step 3 — Simulating Non-Systematic Risk

Simulated RTRs based on assumed lognormal distribution and selected parameters

Simulating future RTRs and ELRs

Accident Year	Earned Premium	Evaluation Date 12	24	36	48	Ultimate Loss Ratio
2004	2,243				1.006	0.631
2005	2,340			1.122	1.008	0.714
2006	3,288		1.208	1.119	1.005	0.613
2007	2,937	1.672	1.211	1.122	1.001	0.735
2008	3,000	1.668	1.223	1.120	1.003	
2008					Average	0.673
2008					Simulated	0.744

Expected ultimate loss ratio for latest accident year is average of prior years; simulated ultimate loss ratio is based on assumed lognormal distribution with selected CV

Squaring the triangle using simulated RTRs and ELRs

Accident Year	Earned Premium	Evaluation Date 12	24	36	48	Ultimate Losses
2004	2,243				1,406	1,415
2005	2,340			1,477	1,657	1,671
2006	3,288		1,483	1,791	2,003	2,014
2007	2,937	950	1,588	1,922	2,157	2,160
2008	3,000	973	1,623	1,985	2,224	2,231

Dark shaded: actual cumulative paid claims to-date
Light shaded: simulated values based on the paid chain ladder method and simulated RTRs

A key point of the model is that it is a mixture of a stochastic chain ladder and a stochastic loss ratio model. In the simple example, the former is applied to prior accident years and the latter is applied to the future accident year. As will be seen subsequently, the mixing of the two models can be generalized to include use of one or the other, or a mixture of the two, for any accident year.

To complete our example, all that is left to do is reintroduce future systematic variability into our model. This is step 4 in the process, displayed in Chart 3.5.

To reintroduce systematic risk, the simulated triangle from Step 3 is decumulated into incremental paid claims, the incremental paid claims are adjusted to reflect simulated inflation, and the triangle is re-accumulated. The adjustment factors are calculated in a manner analogous to Step 1. Future deviations from the long term expected inflation rate are calculated by comparing the constant inflation index based on our 2.5% inflation expectation to a simulated inflation index using an ESG. The net adjustment factors are applied to the simulated incremental payments; after re-accumulating these incremental amounts, we have the simulated ultimate claims reflecting both systematic and non-systematic risk.

To produce the final output, the simulated future claim payments are discounted for time value of money using a risk-free yield curve derived from swap rates. This gives us a distribution of the present value claim liabilities for past accident years, as well as the current accident year. Subtracting the expected values gives us the estimation error distribution.

The above example illustrates the structural approach to simulating ultimate losses, capturing both reserving and pricing risk in an integrated model. The example reflects a number of simplifications over the actual approach that we have implemented, including:

- The use of a simple calendar year inflation model for systematic risk;
- Exclusive reliance on the paid development projection method, without consideration of the information value of case-basis claim reserves or claim counts;
- Use of the historical average loss ratio for the current year, without consideration of changes in price levels or the underwriting cycle; and
- No consideration of parameter risk in the use of the lognormal model for the RTRs and the current year loss ratio.

In subsequent sections we will address these simplifications.

Chart 3.5: Step 4 — Overlaying Systematic Risk

Factors to adjust payments from expected to simulated inflation

Calendar Year	Expected Inflation Index	Simulated Inflation Index	Net Adjust. Factor
2007	1.000	1.000	1.0000
2008	1.025	1.030	1.0049
2009	1.051	1.063	1.0117
2010	1.077	1.095	1.0167
2011	1.104	1.133	1.0266
2012	1.131	1.171	1.0346

Simulated cumulative payments -- without systematic risk

Accident Year	Evaluation Date 12	24	36	48	60
2004	650	1,049	1,257	1,406	1,415
2005	755	1,218	1,477	1,657	1,671
2006	881	1,483	1,791	2,003	2,014
2007	950	1,588	1,922	2,157	2,160
2008	973	1,623	1,985	2,224	2,231

Simulated cumulative payments -- with systematic risk

Accident Year	Evaluation Date 12	24	36	48	60
2004	650	1,049	1,257	1,406	1,415
2005	755	1,218	1,477	1,658	1,672
2006	881	1,483	1,792	2,007	2,019
2007	950	1,591	1,929	2,168	2,171
2008	978	1,635	2,003	2,248	2,256

The simulated constant-inflation claim data is disturbed by re-introducing simulated inflation from the ESG; net inflation adjustment factors are applied to the incremental payments by calendar year diagonal

Finally, the example illustrates how the general structural framework can be applied to the run-off risk horizon to measure reserve and pricing risk. The implementation of the one-year risk horizon is significantly different from that of the run-off horizon, but the general structural concept of simulating systematic and non-systematic risk separately remains.

3.3 Removing Historical Systematic Risk

The starting point in building a structural simulation model is a body of historical claim experience, consisting of historical premiums and claim development triangles for a class of business. These could either be the available actual experience of a company, or they could be the experience of peer competitors, drawn from published Schedule P information. In our experience, the latter is often useful as a supplement to the company's own data. Our firm maintains a Schedule P database extending back in time more than twenty-five years for this purpose.

Typically, we make use of both a paid claim and a reported claim triangle. These triangles reflect the actual socio-economic conditions that existed in the historical period. Once the relevant social-economic variables are identified and their functional relationship to the claim development is specified, the historical triangles can be restated from their historical-conditions basis to a normative constant-conditions basis.

In the current implementation of the model, the only economic variable utilized is inflation. We use the inflation model based on that first proposed by Butsic [1]. The core of the Butsic approach is that the impact of inflation on claims paid in a particular period is a blend of the accident date and the payment date (i.e., calendar date). Mathematically, this can be expressed as:

$$Q_{a,d} = (I_a / I_0)^{\alpha} \times (I_{a+d} / I_d)^{(1-\alpha)} \tag{3.1}$$

Where: $Q_{a,d}$ is the blended inflation index for payments in accident year *a* and development period *d*;

I_t is the actual inflation index value at time *t*; and

α is the blending factor that determines the degree to which the accident year or payment year inflation predominates.

The extent of blending will vary with the characteristics of the line, depending on the degree to which claim costs are related to price levels at the time of the accident versus price levels at the time of payment. For example, since Workers Compensation wage loss benefits are keyed to the average weekly wage of the worker at the time of the injury, it is reasonable to assume that the accident date model predominates for indemnity claim payments. Conversely, since Workers Compensation medical benefits are paid at costs prevailing at the time of payment, it is reasonable to assume that the payment date model predominates for medical claim payments. If one is modeling Workers Compensation claim development on a combined basis, it would be reasonable to assume an intermediate value of the α blending factor.

The estimation of the α parameter from empirical data is difficult; our implementation therefore makes a judgmental selection of α. Since our goal is to measure volatility around the mean, and not predict the mean itself, the choice of α need not be based on "best fit" considerations. Sensitivity tests can demonstrate the sensitivity of the overall volatility to the selection of alternative values for α.

As noted above, the current implementation of the removal of systematic risk involves the substitution of a "steady-state" value of Z. Such an approach results in incremental inflation adjustments that take the form of:

$$A_{a,d} = [(I_a / I_0)/(R_a / R_0)]^{\alpha} \times [(I_{a+d} / I_d)/(R_{a+d} / R_d)]^{(1-\alpha)} \qquad (3.2)$$

Where: $A_{a,d}$ is the excess inflation over the long term expected inflation rate for accident year *i* and development period *d*;

R is the inflation index constructed from expected inflation; and

all other variables retain the same definition as above.

The adjustment factor above removes only the net deviation from the expected inflation rate.

The inflation index is chosen so that it relates specifically to the underlying drivers of the claim settlement costs for the class of insurance being modeled. For example, in Workers Compensation, the inflation index could be a mixture of medical cost inflation and wage inflation. In Auto Liability it could be a mixture of auto repair costs, legal costs, and general inflation. The inflation adjustment is applied to the paid development triangle to obtain an adjusted paid triangle. Typically, the actual case reserve triangle is then added to the adjusted paid triangle to produce an adjusted reported triangle; however, it would be possible to adjust the case reserves for inflation as well.

In addition to monetary inflation, if other systematic variables have been introduced, these would also be factored out of the claims experience in a similar manner.

3.4 Parameterizing Non-systematic Risk

The objective of removing systematic risk from the triangle is to reduce the remaining variability in the triangle to random noise of the type that would be generated from a stationary stochastic claim development process. Once the systematic risk has been removed, the adjusted triangle can be tested to see whether or not it reasonably meets the stationarity criterion, by inspection of the standardized residuals along the accident period, development period, and calendar period dimensions.

In some instances, the test results are not very good, indicating that there is still an unexplained factor influencing one or more of the three dimensions. At this stage one has three choices: (1) go back, and identify and add an additional systematic variable that explains the remaining pattern in the residuals; (2) go forward, but allow for correlation in the RTRs along the relevant dimensions; (3) go forward, accepting that the test results are "good enough". From our experience, the second option is sometimes a reasonable compromise, with explicit correlation in the factors along a calendar year diagonal. This introduces acceleration or deceleration in the calendar year payments or claim reporting.

In the simple example presented earlier, we illustrated how the non-systematic portion of the structural model combined a stochastic chain ladder model with a stochastic loss ratio model. The parameterization of each of those two components in the structural simulation model is described below.

3.4.1 The Stochastic Chain Ladder Method

Once suitably stationary adjusted paid and reported claim development triangles have been produced, the structural simulation model calculates RTRs at each maturity. These are assumed to be a random sample drawn from a stationary distribution.

At each maturity, one must choose a form of the distribution (e.g., lognormal) for the RTRs, and choose parameters for the selected distribution by treating the available observations as a sample. Typically, one would use the sample weighted mean across all observations as the mean RTR, as this is consistent with the assumption that the adjusted triangle reflects a stationary process. However, this is not required by the model; the user may select any suitable set of mean RTRs. To the extent that there are outliers or externalities that would suggest an alternative selection, the model allows the user to override any of the weighted average RTRs with a selected value.

The sample variances of the RTRs understate the overall risk, because they do not contemplate parameter risk. In order to properly consider total risk, we increase the sample variances of the RTRs to account for the parameter risk component of total risk. To do this, we used a by-product of the unpaid claims variability approach developed by Mack [8] and Murphy [10]. Murphy's equations, which were developed in a different but fundamentally equivalent form to those in the Mack paper, provide a quantification of process and parameter risk related to each RTR which are used in his final determination of overall unpaid claims risk. In our model, we calculate Murphy's parameter risk component only and combine it with the process risk determined above to estimate the total risk associated with each RTR.

The parameterization of the RTRs allows us to simulate future paid development or future reported development "across the row" for each historical accident year in the triangle. This gives us simulated estimates of ultimate claims using either a paid chain ladder or a reported chain ladder method. Based on experience, our preference is to use reported claim development data as the basis for the model.

3.4.2 The Stochastic Loss Ratio Method

The structural simulation model also uses a stochastic loss ratio method to project "down the column" to obtain ultimate claims. This is used for the current accident year, where there are no reported losses; and it can also be used for recent accident years in volatile classes such as excess liability, where the paid and reported chain ladder projections are highly volatile.

Rather than using the average loss ratio, the model uses a stochastic auto-regressive moving average (ARMA) model to project future loss ratios. The ARMA method is a simple time series method used to project future values based on past values.

The functional form of the ARMA model is:

$$LR_a = [LR_{a-1} + MRF \times (LR_\mu - LR_{a-1}) + MF \times (LR_{a-1} - LR_{a-2})] \times \varepsilon \qquad (3.3)$$

Where: LR_a is the simulated loss ratio for accident year *a*;

LR_μ is the selected normative long term average loss ratio;

MRF is the selected mean reversion factor;

MF is the momentum factor; and

ε is the percentage error factor, assumed to be an independent lognormal random variable with a mean of one.

In its traditional form, the ARMA model has an additive error term. In our model, we adjusted the error to be a multiplicative lognormal error factor to ensure that LR_a is never negative. While we have chosen to use a lognormal model for loss ratios, it would be relatively easy to use a different distributional form, for example to achieve a "fatter tail" in the simulated loss ratios.

Since the lognormal error term has a mean of one, the expected value of LR_a is given by equation 3.3 without the ε term.

The second term of equation 3.3 causes the loss ratios generated by the ARMA model to be mean reverting over time; the higher the value of the *MRF* parameter, the greater the mean reversion tendency. The third term causes the loss ratio to move in the same direction as it did between the prior two accident years; the higher the *MF* parameter, the greater the momentum tendency. In combination, the two terms can be used to introduce apparent "cyclicality" to the time series, a behavior that mimics that of actual loss ratio time series.

The *MRF* and *MF* parameters can be estimated in any number of ways. Our implementation makes judgmental selection for the *MRF* and *MF*. Since our goal is to measure volatility around the mean, and not predict the mean itself, the choices of *MRF* and *MF* need not be based on best-fit parameterization.

3.4.3 Mixing the Two Projection Methods

The chain ladder projections rely solely on the reported or paid claims within the given accident year to project its ultimate claims, ignoring all knowledge of the experience in adjacent accident years. Conversely, since the ARMA loss ratio projections rely solely on the prior two accident year loss ratios, the loss ratio method ignores all knowledge of the paid or reported claims in the given accident year to project its ultimate claims. This is obviously necessary for future years, where there are no paid or reported claims. It is equally useful for past accident years where the actual cumulative paid or reported claims have no credibility. Jing, Lebens and Lowe [5] outline a method for empirically measuring the skill of an actuarial projection method; when measured skill is equal to or less than zero, the projection is more volatile than the overall volatility around the expected loss ratio, and the expected loss ratio approach is a more accurate estimator. Skill can be measured by maturity, providing a basis for deciding when to use the chain ladder method and when to use the expected loss ratio method.

Between these two extremes, the structural simulation model allows for a blending of the two methods. For each accident year, the user selects the method that will determine the selected ultimate losses. For each year, the user may chose to set the ultimate claims equal to any of the following: (a) the results of the chain ladder method (either paid or reported); (b) the expected loss ratio method multiplied by the premium; or (c) the Bornhuetter-Ferguson result (either paid or reported), which is viewed as a weighted average of the two prior methods with weighting based on the expected reporting or payment pattern.

The simulation proceeds iteratively, starting with the oldest accident year, determining the selected ultimate losses for that year, and then using the result as an input to the next year. (The selection for the oldest two years must be determined by the chain ladder method.)

For example, for a model involving ten prior plus the current accident year, the selections might be:

{CL, CL, CL, CL, CL, CL, CL, BF, BF, ELR, ELR}

The intent is to imitate the actual reserving and pricing process, with the selected methods reflecting that which might actually be used on the specific class of business in the particular circumstances.

It is important to understand that the ARMA method uses the *selected* ultimate loss ratio from the two prior years to simulate the loss ratio for the next year. The error term in the ARMA model is estimated based on the variance of the errors between the chain ladder projected loss ratio and the ARMA projected loss ratio over all accident years that use the chain ladder method as the selected ultimate projection method. We do not use accident years where the selected ultimate is based on the Bornhuetter-Ferguson or loss ratio methods, as this would introduce a downward bias in the estimated variance since both of these latter methods rely on the ARMA projections. Thus, the error term can be interpreted as the error associated with the ARMA model's ability to accurately project our best estimate of ultimate losses.

In summary, non-systematic risk is introduced into the simulation in three ways:

1) In the chain ladder method through the simulated RTRs, where the selected distribution and parameters reflect non-systematic variability inherent in the claim reporting or payment process.
2) In the ARMA loss ratio method through the simulated loss ratios, where the selected variance reflects non-systematic variability inherent in the ultimate loss ratio over time.
3) In the interaction between the two methods, as the simulated ultimate claims from the chain ladder method enter into the ARMA method, affecting the expected loss ratio for the next year.

Once the ultimate claims are simulated for an accident year, the simulated paid RTRs are used to construct the claim payments that produce that ultimate value. The RTRs are scaled up or down to match the ultimate value.

3.5 Measuring and Simulating Systematic Risk

Systematic risk is incorporated via an economic scenario generator (ESG). The relevant socio-economic variables are simulated using the ESG. A detailed description of ESGs is beyond the scope of this paper. The reader is referred to Mulvey and Thorlacius [9] for a description of Towers Watson's Global CAP:Link ESG, which is used in the current structural model. Global CAP:Link uses a system of stochastic differential equations to generate plausible future paths of a variety of economic variables, including equity market returns, treasury yield curves, credit spreads above treasuries for several key debt instruments, GDP growth rates, unemployment rates, and several key inflation rates. It generates these indices for multiple economies, and also generates dynamic foreign exchange rates between the economies.

In addition to the standard economic variables, Global CAP:Link has the facility to generate customized indices that are related to the standard indices. This facility is used, for example, to generate medical inflation rates that are linked to the CPI and GDP growth rates.

Each scenario from the ESG represents a plausible future path of the economy. The parameters of the system of equations are selected to reproduce a stylized set of facts developed from the historical data. Examples would be the frequency of inverted yield curves, the volatility of equity returns over various holding periods, and the degree of correlation between interest rates and inflation rates over selected time horizons. Economic principles such as purchasing power parity also affect the form and parameterization of the equations.

Any of the generated indices can be used to induce systematic risk into the claim development simulation, to the extent that they are believed to be drivers of claim frequency or claim settlement costs. For example, the frequency of Workers Compensation claims might be related to levels of unemployment, or D&O claim experience might be linked to the behavior of the stock market. Many liability lines are subject to *social inflation* in addition to monetary inflation, which can also be modeled as a systematic risk variable.

The economic scenario indices can be generated on a monthly, quarterly, or annual basis.

The Global CAP:Link scenarios are *real-world* scenarios, suitable for modeling the risk of changes in future socio-economic conditions over time. They are not *risk-neutral* scenarios and therefore are not suitable for valuation of assets or liabilities with prices that are sensitive to the market price of risk.

In the current model, for each trial the simulated inflation index for each class of business is used to adjust the paid losses to account for the difference between expected cumulative inflation and actual cumulative inflation. As was indicated earlier, the process uses the Butsic formula to mix accident date and payment date inflation.

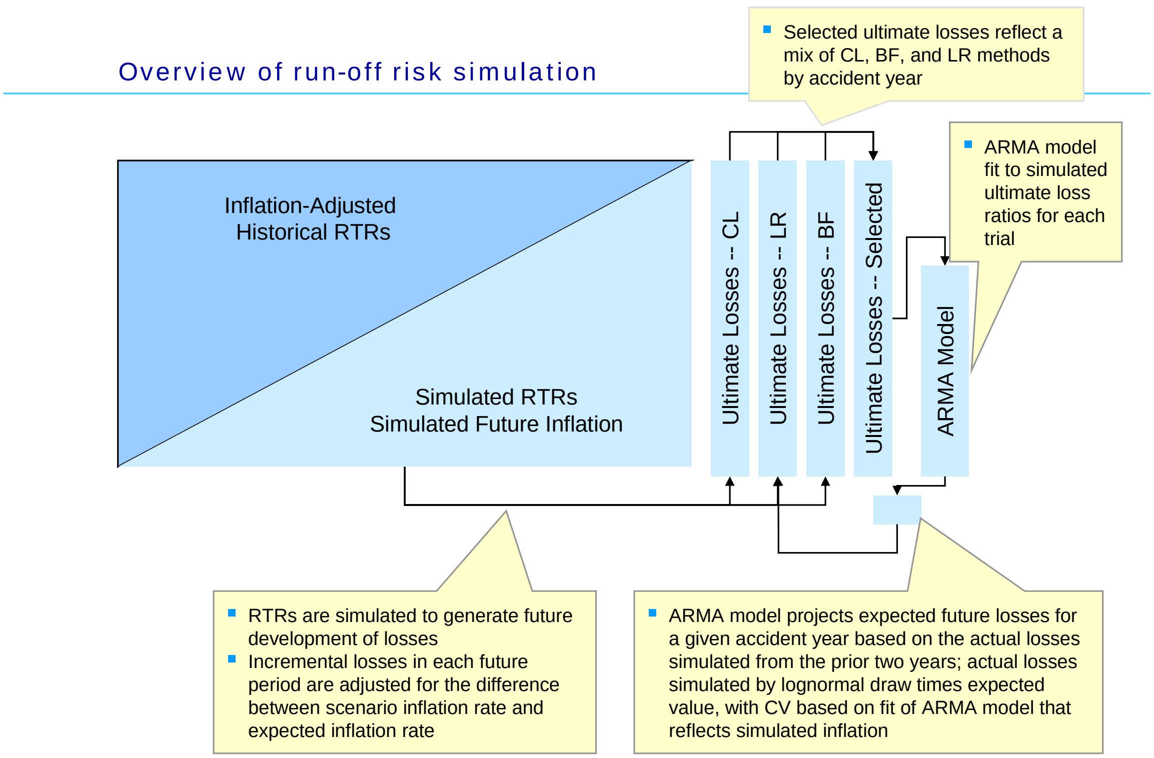
3.6 The Run-off Risk Horizon Model

Chart 3.6 portrays the general operation of the model for measuring risk using a run-off horizon.

Recall that the run-off model aims to answer the question "What is the potential variation in the ultimate cost of claim liabilities from the current actuarial central estimate?" The run-off model measures the variability of the actual ultimate claim liabilities around the current best estimate by simulating ultimate claims using the three stochastic methods: chain ladder (either paid or reported), ARMA loss ratio and Bornhuetter-Ferguson. The run-off model simulates the complete claim development process by generating all of the future missing values for the entire paid claim $C_{a,d}$ array. Systematic risk is overlaid on the simulated paid claim array to obtain an array that incorporates both systematic and non-systematic risk.

The simulated future claim payments from each trial are discounted for time value of money using the risk-free yield curve, and compared to the actuarial central estimate, to produce a distribution of the error term $e_r = \hat{L}^{(n)} - L^{(n)}$. This distribution is a representation of run-off insurance risk.

Chart 3.6: Schematic of the Run-Off Horizon Model

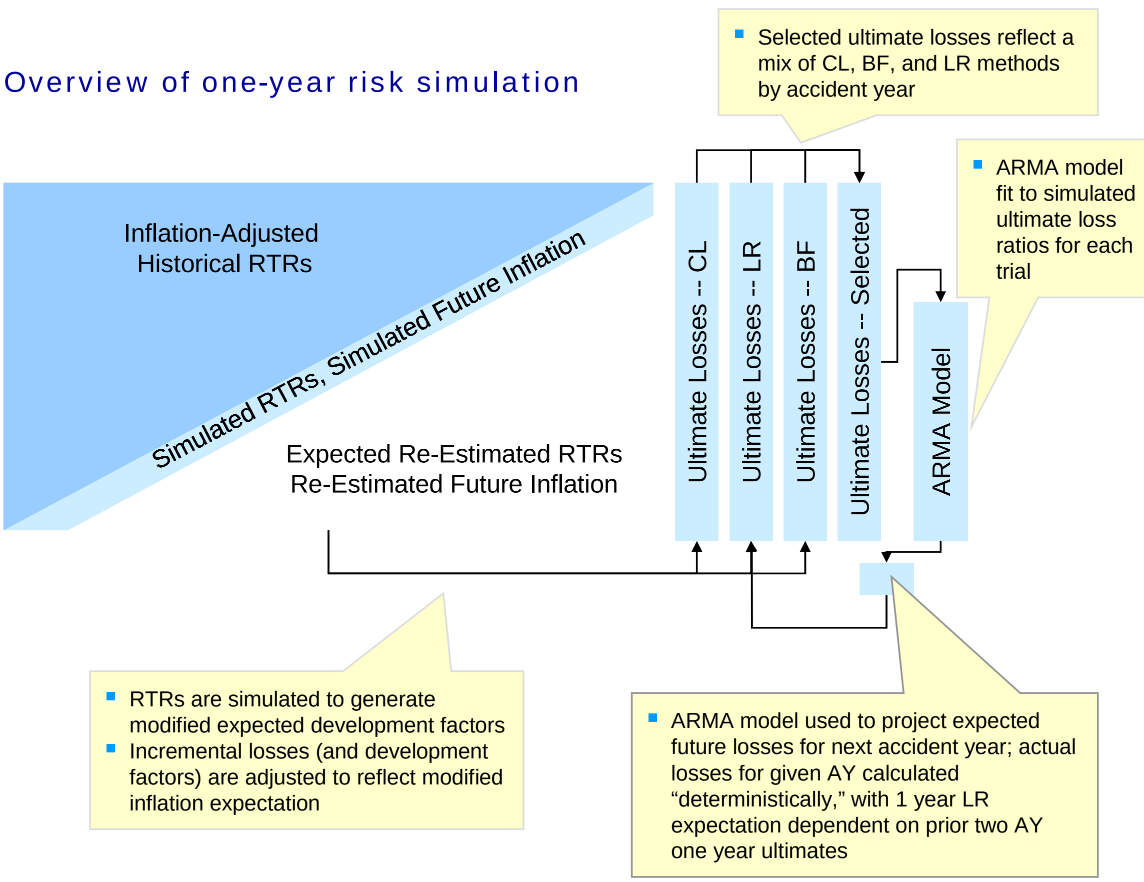


3.7 The One-year Risk Horizon Model

Chart 3.7 portrays the general operation of the model when measuring risk over a one-year horizon.

Recall that the one-year model aims to answer the question "What is the potential change in the actuarial central estimate of ultimate claim costs that could occur as a result of one additional year of actual claim emergence, other information, and changes in circumstances that affect the valuation?" The one-year model measures the variability of the one-year changes in actuarial central estimates by simulating how an additional "diagonal" of calendar year information would change the central estimate from its current value.

Chart 3.7: Schematic of the One-Year Horizon Model



The one-year chain ladder simulates RTRs and future inflation rates using the same model, parameters and inflation scenarios as in the run-off. The key difference is that only one calendar year (systematic and non-systematic risk) is simulated. That is, one diagonal of RTRs (non-systematic risk) and one calendar year of inflation (systematic risk) are simulated and combined to create a simulated realization of the next diagonal of the claim array. This simulated calendar year is then used to re-estimate the *expected* RTRs and the expected inflation rate. In general, this is done by examining two items:

1) ***The impact of the new simulated diagonal on the expected RTRs.*** The model selects new expected RTRs with the benefit of the new observed values along the simulated diagonal. The selection process for the new expected RTRs is a credibility weighted average of the observed simulated RTR, adjusted for the impact of the revised expected long term inflation, and the prior expected RTR. The credibility is a user input into the model.

2) ***The impact of the simulated inflation rate on the expected future inflation rate.*** The model uses an exponential smoothing technique to blend together the prior expectation with the new information, to obtain a revised expectation. The exponential smoothing equation is of the general form:

$$E' = E \times \beta + O \times (1 - \beta)$$

The β term determines the degree to which the new information alters the expectation.

E' is the revised expected long term inflation rate, E is the prior long term expected inflation rate and O is the simulated observed inflation rate

Finally, the development over the one-year simulation accretes into the re-estimated expectation of the remaining claim liabilities. That is, the simulated ultimate loss is a deterministic projection using the revised expected RTRs, applied to the simulated realization of the new diagonal. In addition, the entire triangle is adjusted to reflect the new level of expected inflation, so that the projection reflects the new inflation expectation.

This procedure is intended to exactly imitate the process that would be followed to derive the actuarial central estimate of ultimate losses one year hence, assuming we were using the inflation-adjusted chain ladder method. A comparison of the original estimate using this method to the simulated estimates using this method with one year of additional information is a measure of the one-year risk associated with the method.

The one-year ARMA loss ratio method follows a similar procedure. The new one-year chain ladder results are used to generate future expected loss ratios. As before, the ARMA model uses the selected loss ratio for the two prior years to project a given accident year's loss ratio. However, in this case the inputs for the ARMA model for a given accident year are the two prior years' *one-year* selected estimated ultimate losses. Since the one-year horizon is trying to measure the effect of one year's worth of information on the *expected* value, we are only concerned with the expected ARMA value and thus do not need to simulate an error factor (the expected error factor is 1.00). Again, this process can be considered a stepwise method since the one-year selected ultimate loss ratios for the two prior accident years are simulated and selected prior to projecting the given accident year loss ratio.

Given that all the inputs for the one-year ARMA model reflect the impact of one year's information, the one-year loss ratio method allows us to measure the impact of this extra year's worth of information on the given year's estimated ultimate losses.

The BF method is simply the credibility-weighted average of the chain ladder method and the loss ratio method, where the credibility and its complement are determined by the re-estimated expected percentage paid/reported pattern from the chain ladder method.

3.8 Correlation Between Classes of Business

This structural model's framework presents an intuitive way to measure the correlation between classes of business. Correlation between classes is introduced through the model in two ways:

1) *Through shared systematic risk.* Since each class of business being simulated will use the same scenario generated from the economic scenario generator, co-dependency in the results will be induced via the common systemic risk variables. For example, if two classes of business are both influenced by medical cost inflation, then in each trial of the simulation the two classes will be influenced by an identical medical inflation rate. Even if one class is affected by medical inflation and another is affected by wage inflation, a degree of co-dependency will be introduced to the extent that medical and wage inflation are correlated in the ESG.

2) *Through shared non-systematic risk.* Additional co-dependency can also be introduced by drawing all simulated values from a bivariate distribution. The dependency structure of the distribution is constructed by examining the empirical dependence structure of the historic RTRs, transformed to the cumulative probability space. Once this empirical dependence structure is determined, correlated bivariate values can be simulated by simulating a

correlated pair of values in the cumulative probability space and transforming that pair into their original bivariate space.

3.9 Final Calibration and Validation

In addition to the specification of the structure of systematic risk, the structural simulation requires the specification of many parameters, including:

- Means, variances (including process and parameter elements), and distributional form of the paid and reported claim RTRs;
- The Butsic mixing parameter for payment versus accident date inflation;
- ARMA mean reversion and momentum parameters;
- ARMA loss ratio model variance (including process and parameter elements), and distributional form;
- Choice of claim projection method (CL, BF, ELR) by accident year;
- The beta parameter that determines how inflation expectations adapt to new information; and
- All of the parameters within the ESG.

As a final calibration step, it is advisable to compare the results from the model to hindsight tests of historical claim estimation errors. For the run-off model, several comparisons can be made fairly readily. All of these comparisons should be on a present value basis, consistent with the risk measurements produced by the model.

- First, one can look back at historical reserve estimation errors, comparing actual emergence to original estimates. These can be done by accident year organized by maturity, or in the overall. The hindsight empirical reserve errors should look like a sample drawn from the modeled reserve risk distribution.
- Second, one can look at the volatility of historical ultimate loss ratios. These are cross-sectional rather than estimation errors, but they are nevertheless indicative of the volatility of loss ratios.

This volatility should look like a sample statistic drawn from the modeled pricing risk distribution.

- Alternatively, one can compare the initial estimate of loss ratios on the new accident year (for example, as is often used in reserving at the end of the first quarter) to the ultimate loss ratios. While this is technically a better comparison, few companies have the data available to make this comparison.

For the one-year model, several comparisons can be made, again on a present value basis, consistent with the model.

- One can look at historical one-year reserve development, either by accident year/maturity or in the overall. This data is generally available, and can be viewed as a sample from the model's one-year reserve risk distribution.
- One can look at the development of the estimated ultimate loss ratio on the current accident year from the beginning of the calendar year to the end of the year (i.e., at twelve months maturity). The latter values are typically available from year-end reserve analyses; the former values are sometimes available from planning or reserving work earlier in the year.

If the actual empirical reserve and pricing risk data do not compare favorably to the model, then the model's parameters should be reconsidered and adjusted where appropriate. Consistency tests across classes of business are also useful in this regard.

4. TESTING AND EMPIRICAL RESULTS

This section presents some illustrative empirical test results of the structural model. Unless noted to the contrary, the datasets used were the Personal Auto Liability and Commercial Auto Liability claim development data for a large U.S. insurer, as published in Schedule P of their statutory annual statements. This data is net of reinsurance (although little-to-no reinsurance is purchased by this insurer) and combines claim and claim defense costs together.

4.1 Comparison to Other Methods

Table 4.1 compares the results from the structural model to the results obtained from three other published stochastic reserving methods. Due to the limitations of those methods, the comparison relates only to reserve risk using a run-off risk horizon. In applying each method, the mean RTRs were selected as the volume weighted average of all years, and no tail factors or tail variability was used. Parameter risk was introduced into each method, except for the Practical Method, where only process risk is measured. To the extent possible, we strove to create an "apple-to-apples" comparison across the methods. In the case of the structural model, the results reflect a relatively modest level of expected inflation of 3.5%, with an alpha value of 0.5 (meaning that one-half of the inflation affects the development dimension). In each case, the methods were applied to both reported and paid claim development data.[10]

While the models generally produce a complete distribution of the possible outcomes, we have chosen to display the results in terms of the 99th percentile expressed as a ratio to the mean. Our choice of this statistic is merely one of convenience; it is an intuitive measure of the risk, relevant to economic capital issues.

Table 4.1: Indicated Reserve Risk from Structural Model versus Other Methods

	Ratio of 99th Percentile to Mean -- Run-Off Risk Horizon			
	Personal Auto Liability		Commercial Auto Liability	
	Reported	Paid	Reported	Paid
Structural Model	1.120	1.143	1.159	1.136
Bootstrap Method	1.085	1.087	1.166	1.151
Mack / Murphy Method	1.095	1.119	1.152	1.128
Practical Method	1.086	1.075	1.134	1.103

The results in Table 4.1 reflect the typical variations in indicated risk from the application of different methods to different datasets. The variations reflect differences in how the models measure the noise in the data. Note than in Commercial Auto Liability, across all methods, the indicated risk of adverse development is lower when the method is applied to paid claim development data. However, the reverse is the case in Personal Auto Liability, where the risk indications from the paid data are higher in three of the four methods. This shows that sometimes

[10] Note that the results are designed to facilitate comparisons, and that in actual practice the measured level of risk would be slightly higher than the values shown, due to the inclusion of tail factor variability.

the noise in the paid data is relatively higher, while at other times the noise in the reported data is higher; since all methods are keying off the same underlying historical development data, it is not surprising that all methods pick up the same relative relationship.

In general, from our work with these models we believe that the results obtained from application of the methods to the reported claim development data are a better indicator of reserve risk.

Focusing on the reported claim development results, we would observe that the structural model produces distinctly higher indicated reserve risk than the other three methods for the Personal Auto Liability class of business. Since other research has suggested that the Mack method may understate reserve risk and the Practical method excludes consideration of parameter risk, this is not an entirely surprising result. The indicated reserve risk for Commercial Auto Liability is more consistent across the different methods, with only the Practical Method lagging.

The structural model results are intuitively reasonable. They indicate that, at the 1-in-100 probability level, Personal Auto Liability reserves will develop adversely by as much as 12.0%; and that Commercial Auto Liability reserves will develop adversely by as much as 15.9%.

4.2 Integration of Pricing and Reserving Risk

Table 4.2 displays the indicated pricing and reserving risk when the structural model was applied to the reported claim development data for the two classes of business. Also shown are the indicated correlation between reserving and pricing risk generated by the model, and the indicated level of overall insurance risk (i.e., combined reserving and pricing risk).

Here, we show results on both a run-off and one-year risk horizon basis.

Table 4.2: Indicated Reserving, Pricing, and Insurance Risk from Structural Model

Ratio of 99th Percentile to Mean
Structural Model Applied to Reported Claim Development Data

	Personal Auto Liability		Commercial Auto Liability	
	Run-Off	**One-Year**	**Run-Off**	**One-Year**
Reserve Risk	1.120	1.093	1.159	1.129
Pricing Risk	1.195	1.118	1.289	1.125
Indicated Correlation	68%	87%	48%	74%
Insurance Risk	1.142	1.102	1.182	1.120

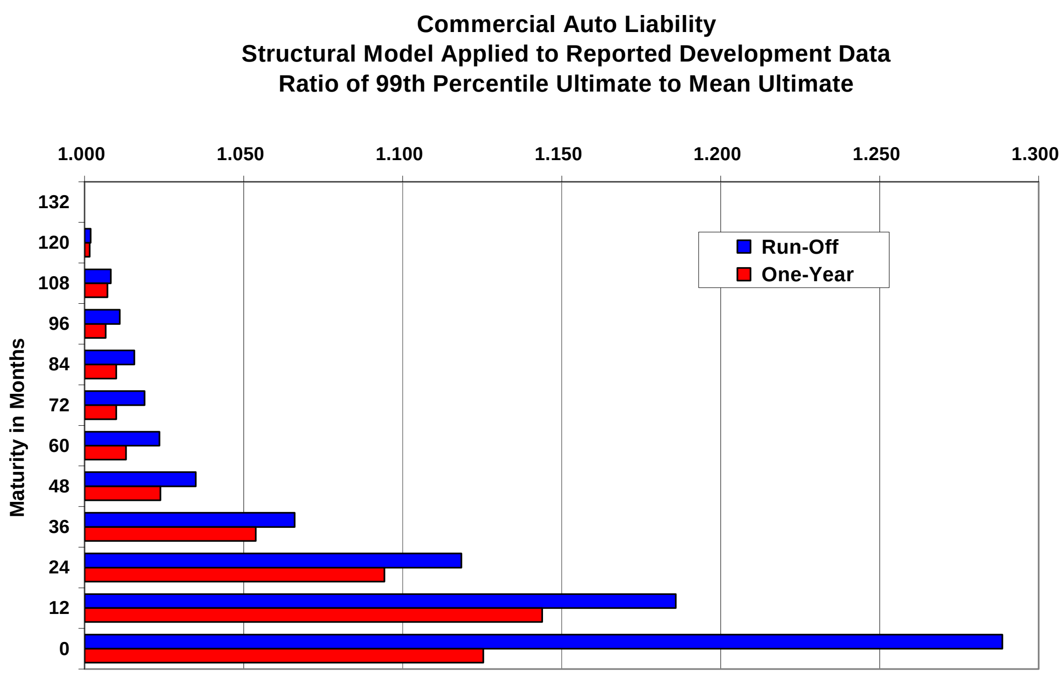
Not surprisingly, the indicated risk on a one-year horizon basis is distinctly lower than on a run-off horizon basis. This merely reflects the fact that the former focuses only on movements in estimates in a single year, rather than over the entire lifetime of the claim liabilities where annual movements can compound.

While indicated pricing risk is higher than reserve risk under the run-off risk horizon, the situation is less pronounced in the case of the one-year risk horizon. In fact, in the case of commercial auto liability, the one year pricing risk is less than the reserve risk. While somewhat counter-intuitive, this result also stems from the way that risk emerges. For longer-tailed liability lines, we sometimes learn more about the ultimate values of claims as they develop from 12 to 24 months than we do from the development from 0 to 12 months.

Finally, one can observe that the model produces relatively high correlation between reserve and pricing risk. This reflects the common pattern of inadequate reserves leading to inadequate prices.

The quoted results for reserve risk in Tables 4.1 and 4.2 are for all maturities combined. One can see the model results more clearly by examining results by individual maturity, as in Chart 4.3 below.

Chart 4.3: Standard Model Output – Risk by Maturity



Note that in Chart 4.3, risk is expressed as the ratio of the 1-in-100 ultimate claim amount (i.e., including paid-to-date) to the expected ultimate claim amount, rather than the ratio of the 1-in-100 liability to the expected liability. This is done to make the results more intuitive, showing how the uncertainty of the ultimate claims for an accident year resolves as the year matures. Chart 4.3 illustrates several important relationships, described below:

- The indicated risk ratios are much higher at early maturities, reflecting the "funnel of doubt" relationship referred to earlier in the paper.
- The indicated one-year and run-off risks are significantly different, with the one-year risk being smaller in every year. The run-off model is designed to capture the variability of all reasonably predictable future outcomes in all future calendar years. In contrast, the one-year model

contemplates the impact of outcomes for only one future calendar year. Consequently, one would expect that the one-year risks would be smaller than their run-off counterparts.

- In the chart, maturity zero is the prospective year. We refer to the indicated risk associated with this year as the pricing risk. Under the run-off horizon, the indicated pricing risk is much higher than all of the other years, as would be expected. However, under the one-year horizon the opposite is true; there is a slight dip in the pricing risk. Initially this may seem counter-intuitive, but recall that we are trying to answer the question "What is the potential change in my estimate of ultimate losses that could occur as a result of one additional year of actual claim emergence?" The dip in the one-year pricing risk can be explained by the fact that in many classes there may be very little expected emergence in the first year. With very little information, our estimate of ultimate losses will not change drastically one year hence. (This reversal is not necessarily the case, for example in short-tailed lines where much of the emergence occurs in the first year.)

4.3 Sensitivity to Inflation Parameters

We performed sensitivity tests on several of the key inflation parameters of the structural model, using the Commercial Auto Liability reported claim development data. For comparability with the results presented earlier, we focused on run-off reserve risk. Results are displayed in Table 4.4.

In the base case, the expected future inflation rate is at a constant level of 3.5%; and the Butsic alpha factor is 0.5, meaning that half of the calendar period inflation affects the development dimension. The base case incorporates systematic risk in the form of varying inflation rates about the expected value, as well as non-systematic risk due to variations in the future development factors. In the base case, the 1-in-100 adverse development is 15.9% above the best estimate.

In the first set of sensitivity tests, we turned one or the other of the two sources of risk off, so that we could observe the relative contribution of each risk. One can see that when the inflation volatility is turned off, the 1-in-100 adverse development drops slightly, to 15.6%. Conversely, when the development volatility is turned off, one can see that inflation alone contributes reserve risk of 4.7%. One can also see that the two risk sources are not additive, as the sum of the two stand-alone risks is substantially greater than the combined result. As expected, the introduction of inflation as a systematic risk adds to the risk that is otherwise modeled by the stochastic reserving method. In this case, the contribution is small, reflecting the relatively fast-paying nature of the claims and the low level of expected inflation.

Table 4.4: Selected Sensitivity Test Results

Structural Model Applied to Reported Claim Development Data

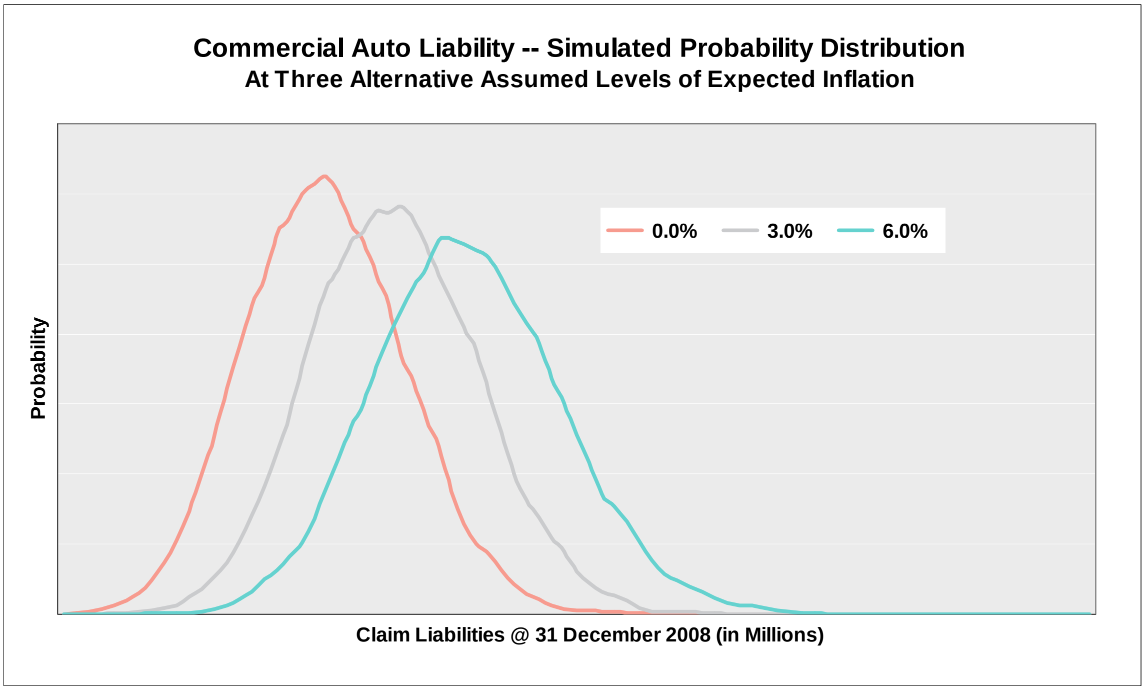
	Ratio of 99th Percentile to Mean Commercial Auto Liability Run-Off Reserve Risk
Base Case	1.159
No inflation volatility, just development volatility	1.156
No development volatility, just inflation volatility	1.047
Bustic alpha factor = 1, inflation does not affect development	1.160
Butsic alpha factor = 0, inflation fully affects development	1.173
Expected inflation lower, at constant 0.0%	1.137
Expected inflation higher, at constant 6.0%	1.174

We also tested the effect of varying the alpha factor from zero to one, around the base case value of 0.5. When the alpha value is decreased to a value of 0 (implying that inflation fully affects development), the risk increases to 17.3%. This is intuitive because the variability of inflation is allowed to fully affect development. In contrast, when the alpha factor is increased to a value of one (implying that inflation does not affect development), rather than decreasing as expected we obtain what initially appears to be an anomalous result; the 1-in-100 adverse development risk rises marginally from 15.9% to 16.0%. Because this is marginal increase, it could simply be attributed to noise. However, it turns out that these results are a consequence of the underlying data, and not just noise. When we set the alpha factor to one, we are implicitly assuming that inflation has no effect on the development dimension of the triangle. Hence the triangle is unadjusted for historical inflation. However, it turns out that some of the variation is actually explained by inflation. As a consequence, the unadjusted triangle exhibits more historical volatility than the adjusted triangle in the base case, and this higher level of volatility manifests itself in the marginal increase of the indicated level of risk.

Finally, we tested the sensitivity of the results to different expected inflation rates. Not unexpectedly, as the assumed inflation rate rises, the risk also rises. If expected inflation rises from 3.5% to 6%, then the 1-in-100 adverse reserve development rises from 15.9% to 17.4%. This result is qualitatively consistent with the experience of the late 1970s and early 1980s, where high inflation destabilized claim trends, resulting in significant adverse reserve development. This is an important result, as — consistent with history — the model implies that reserving risk is higher when inflation rates are higher.

The impact of varying the expected inflation rate can be seen graphically in Chart 4.5, where we show the claim reserve probability distributions generated by the model for Commercial Auto Liability at the three alternative assumptions regarding the expected levels of future inflation. Here, one sees the gradual flattening of the distribution as the expected inflation rate rises. In addition, the expected value of the claim liabilities also rises with the assumed inflation rate. This chart demonstrates a principal advantage of the structural simulation: the ability to delve into inflation as a potential driver of insurance risk.

Chart 4.5: Impact of the Assumed Inflation Rate on the Claim Liability Distribution



4.4 Validation Using Actual Historical Experience

As discussed in the previous section, validation against historical experience is a critical step in the calibration of all stochastic reserving methods. This is particularly true for the structural model, as it has many parameters that must be specified. While these parameters offer the user the advantage of a great deal of flexibility, they necessitate validation to assure that the model is producing results that are realistic.

Usually the available historical data on past reserving and pricing errors is limited, making it an insufficient resource for measuring reserve and pricing risk directly. (Simply stated, it is hard to estimate the 1-in-100 reserve error when the historical reserving database only has observations from about a dozen prior years.) The better approach is often to employ a stochastic model, and validate it by testing whether the historical data could reasonably be a sample drawn from the distributions produced by the stochastic model.

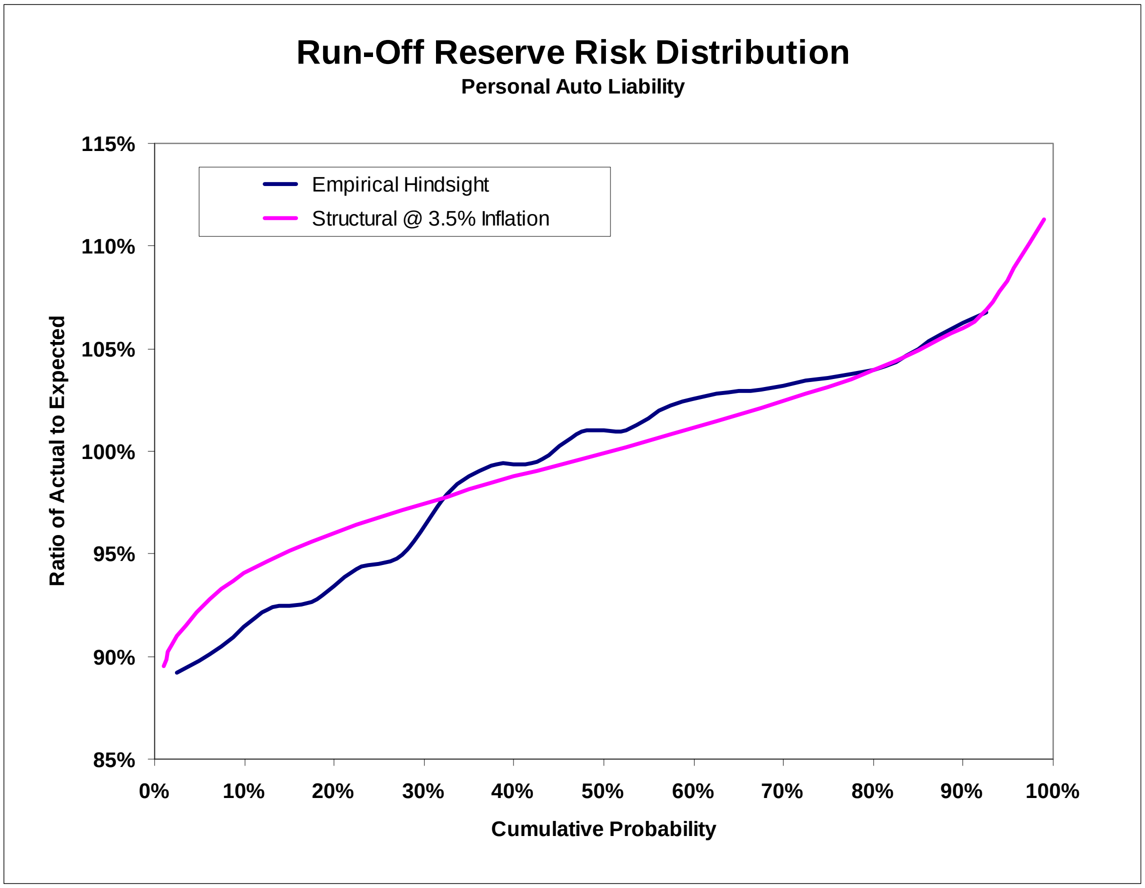
To illustrate some of the validation approaches, we performed hindsight tests on the Personal Auto Liability dataset. We developed best estimates of the claim liabilities for nineteen prior year-ends (i.e., from year-end 1989 to 2007) using only the information that would have been available at the time, and compared those estimates to the actual run-off experience through year-end 2008 to determine historical estimation errors. In addition to estimates of claim liabilities for past accident years, we also estimated claim liabilities for two prospective accident years based on premiums, price changes, and assumptions regarding claim cost trends for use in validating the pricing risk aspect of the structural model.

Over the nineteen year historical period, the overall reserve estimation errors for Personal Auto Liability ranged from a 10.8% redundancy to a 6.8% inadequacy. The highest redundancies occurred during the period from 1994 to 1998, and reflect the effects of disinflation that occurred around that time. While the estimation errors are slightly biased in magnitude towards redundancy, the number of years of redundancy is roughly equal to the number of years of inadequacy.

Chart 4.6 compares the cumulative probability distribution of reserve errors (expressed as a ratio of actual outcome to expected value) implied by the historical data to the cumulative probability distribution generated by the structural model for Personal Auto Liability. While the correspondence is not perfect, the model result compares favorably to the empirical evidence. This is not always the case; in unpublished tests of other models on other classes of business, the authors have observed a lack of correspondence between the model and the historical data that was sufficiently obvious to clearly reject the model. Even with the structural model, adjustments to one or more of the parameters are sometimes required to achieve good validation results.

The standard deviation of the historical errors is slightly higher than that produced by the structural model (5.4% versus 4.7%), driven largely by the higher historical redundancies. Note also that our comparisons are based on structural model output under an assumed expected inflation rate of 3.5%, which is consistent with the actual inflation rate in the latter part of the historical period. Use of a higher expected inflation rate would move the two standard deviations closer together.

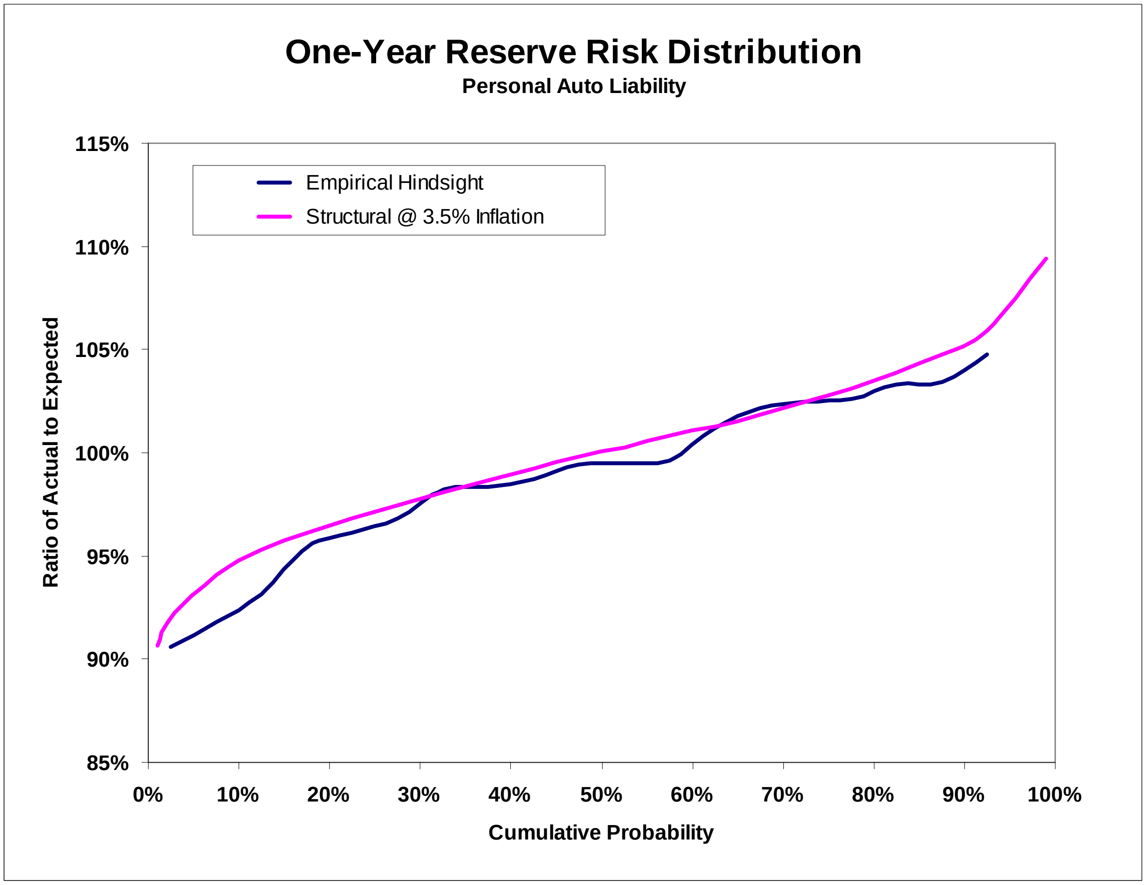
Chart 4.6: Validation of Overall Run-Off Reserve Risk in the Structural Model



The same historical datasets can be used to validate the one-year insurance risk model. Instead of comparing the ultimate run-off estimation errors to the model results, one compares the one-year historical development to the corresponding one-year structural model results. This comparison is displayed in Chart 4.7.

From the same analysis of Personal Auto Liability, we observed that, over the nineteen year historical period, the one-year movements in the best estimate of liabilities for past accident years ranged from a favorable development of 9.4% to an adverse development of 4.7%. As with the run-off results, high favorable development occurred in the calendar years from 1994 to 1998, reflecting disinflation. Once again, the model result compares favorably to this empirical evidence.

Chart 4.7: Validation of Overall One-Year Reserve Risk in the Structural Model



The standard deviation of the one-year historical movements in estimates is slightly higher than that produced by the structural model (4.2% versus 3.9%).

Rather than simply validating the overall result for reserve risk, one can validate the results by maturity. This is done in Chart 4.8, where we compare the ratio of the estimated ultimate claims to the actual ultimate claims, for each accident year at each maturity. The hindsight empirical data consists of the movement of the estimated ultimate claims for each of 25 past accident years, from 12 months prior to inception through to 120 months maturity.[11] The model output consists of the simulated 98% confidence interval for the same statistic at each maturity.

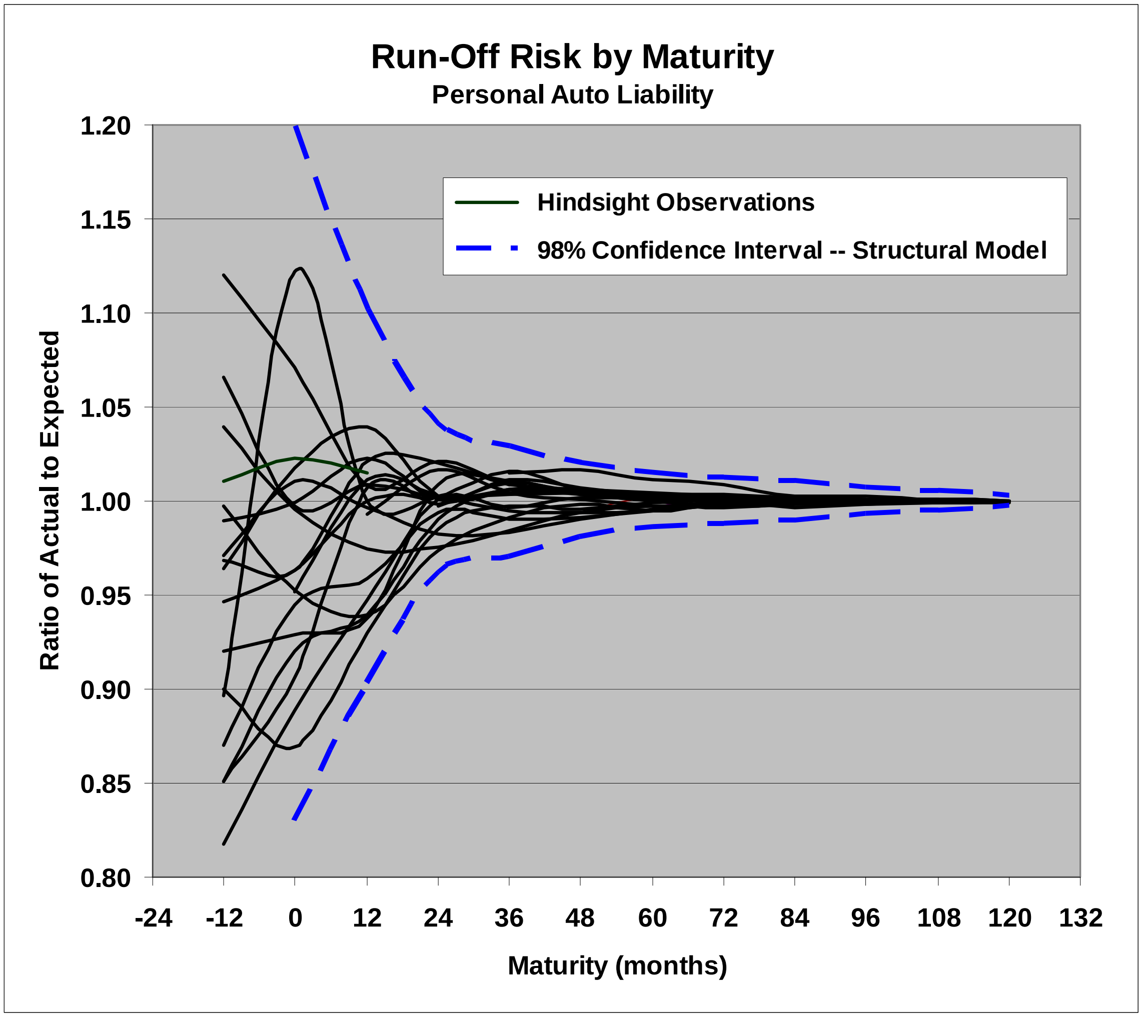
Collectively, the historical data paints an empirical "funnel of doubt", as described earlier in Section 2, except that the funnel here is an inverted mirror image of the funnel displayed in Chart 2.1.

The structural model results in Chart 4.8 compare reasonably well with the empirical hindsight data, with most empirical values falling within the simulated 98% confidence interval. The general shape of the structural model funnel is consistent with that of the hindsight data, with initial pricing risk dissipating relatively rapidly as actual claim experience becomes available. The results shown in Chart 4.8 are those of a "second iteration" of the model. Comparison of the initial results to the empirical data suggested that the model was understating reserve risk slightly at maturities in the 48 to 84 months range (the funnel was too narrow, and quite close to the empirical results. This was addressed by selecting slightly larger CVs for the development factors at these maturities. The final selected CVs are not inconsistent with the loss development data; the adjustments made can be attributed to sampling error.

Note that in Chart 4.8, the hindsight data includes estimates of two future accident years, whereas the structural model reflects only one prospective accident year. Some have suggested that, given the natural lag between available experience data and implementation of price changes, it may be more realistic to measure pricing risk by looking at estimation errors beyond the immediate prospective accident year. If this is the case, it would be easy to extend the structural model to include additional future years.

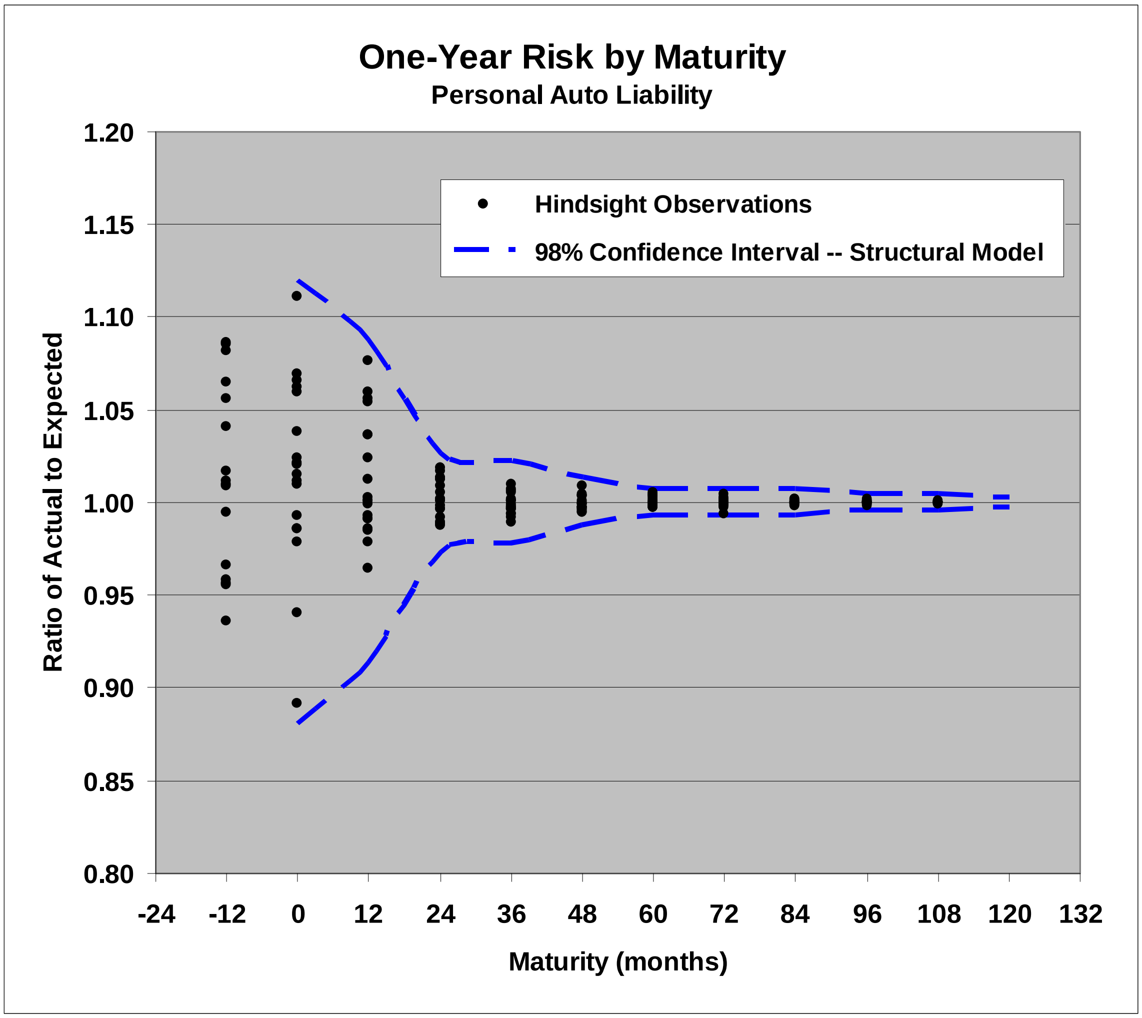
[11] Due to the limitations of the dataset, not all values are available for all accident years. For example the 1983 accident year is only available at valuations from 84 months onward. Similarly, the 2006 accident year is only available at valuations up to 24 months.

Chart 4.8: Validation of Run-Off Risk in the Structural Model by Maturity



A similar analysis to that depicted in Chart 4.8 can be produced to validate the one-year model, rather than the run-off model. Instead of comparing the estimated ultimate claims at one maturity to the true ultimate claims, one instead compares the estimated ultimate claims at one maturity to the estimate at the next maturity. One-year validation results are shown in Chart 4.9. These results are also those of the "second iteration" of the model, to be consistent.

***Chart 4.9:* Validation of One-Year Risk in the Structural Model by Maturity**



Other validation tests can also be devised to test various aspects of the model, for example focusing on run-off and one-year pricing risk. Tests can also be performed that examine the correlation between estimation errors between classes of business. In the interests of time and space these will not be presented here.

4.5 Sensitivity to Choice of Reserving Methods

A key set of parameters to the structural model is the choice of reserving method (i.e., chain ladder, Bornhuetter-Ferguson, or expected loss ratio) by accident year. The baseline approach is to use the chain ladder method for all historical accident years, introducing the other two methods only when the chain ladder method introduces more volatility than is actually present.

For both Personal and Commercial Auto Liability, we used the chain ladder method throughout the historical accident years, as this is what was indicated by the validation test results. To see the importance of this parameter to the final results, we performed some sensitivity tests on the Commercial Auto Liability dataset. These sensitivity test results are displayed in Table 4.10.

Table 4.10: Testing the Choice of Reserving Methods in the Structural Model

Structural Model Applied to Reported Claim Development Data

	Ratio of 99th Percentile to Mean Commercial Auto Liability Run-Off Reserve Risk
Actual claim development data	
Chain ladder for all years	1.159
B-F for latest accident year; CL for all prior	1.162
ELR for latest accident year; CL for all prior	1.215
Modified claim development -- highly volatile 12-to-24 months	
Chain ladder for all years	5.238
B-F for latest accident year; CL for all prior	2.977
ELR for latest accident year; CL for all prior	1.218

In the first set of sensitivity tests, we simply re-ran the structural model using either the BF or the ELR methods for the latest historical accident year (i.e., the year valued at 12 months maturity). Results are shown in the upper portion of Table 4.10. As one can see, the indicated reserve risk is slightly higher than the baseline case when the Bornhuetter-Ferguson method is employed and substantially higher than the baseline case when the expected loss ratio method is employed.

The reader may find these results somewhat counter-intuitive, as generally the purpose of introducing the Bornhuetter-Ferguson and/or expected loss ratio methods into any actuarial projection is to reduce the volatility of the chain ladder projection. Here, because the initial expected loss ratio is a random variable drawn from the ARMA model, the use of the latter methods actually raises the volatility. These results indicate that the chain ladder method does an adequate job at projecting ultimate losses, and the use of the latter two methods is injecting spurious volatility into the model.

Of course, there are situations where the chain ladder method is not the appropriate method to use. When the historic RTRs are highly volatile, the chain ladder method may have "negative skill" in that the projections it produces have greater volatility than that of the actual loss ratios around the expected. In such cases the chain ladder method is not an appropriate method for reserving, and its use in the structural model will overstate the reserve risk. This point is illustrated by considering the results in the lower half of Table 4.10. Here we have artificially modified the historical claim development data by dividing the historical reported claims at 12 months maturity by a lognormal error term with a mean of 1 and standard deviation of 1, making the development factor from 12 to 24 months highly volatile. (The RTRs in the remaining development periods are unchanged.) In this case, the use of the chain ladder method for the latest historical accident year in the structural model generates simulated losses for the latest year that are extremely volatile, causing reserve risk to be overstated. Even the introduction of the Bornhuetter-Ferguson method is insufficient to dampen the volatility of the projected ultimate claims for the latest historical accident year. Therefore, the correct choice is to use the expected loss ratio method for the latest accident year.

As discussed earlier, a simple validation test to determine whether the chain ladder method is appropriate is to compare the volatility of the historical chain ladder method projections at a given maturity to the volatility of historic ultimate loss ratios. Generally, the chain ladder method would be appropriate when the former is smaller than the latter.

4.6 Correlation Between Classes of Business

As noted earlier, the structural model can be used with multiple classes of business to generate an aggregate distribution across the classes. In order to compare the results of the aggregate distributions, we constructed several normal copula models using Bootstrap simulation results using differing correlation coefficient assumptions. Results are displayed in Table 4.11.[12]

Table 4.11: Structural Model Correlation Results versus a Normal Copula Structure

	Ratio of Indicated Percentile to Mean				
	Structural Method	Normal Copula - Bootstrap			
Confidence Intervals					
Mean	1.000	1.000	1.000	1.000	1.000
50.0th Percentile	1.000	0.999	0.999	0.999	1.000
75.0th Percentile	1.028	1.028	1.028	1.029	1.035
90.0th Percentile	1.055	1.054	1.055	1.057	1.070
95.0th Percentile	1.071	1.071	1.072	1.074	1.092
99.0th Percentile	1.102	1.103	1.104	1.108	1.131
99.5th Percentile	1.116	1.114	1.116	1.120	1.147
Correlation	17.6%	5.0%	10.0%	20.0%	100.0%

The structural model's risk margins are similar to those produced by the bootstrap normal copula model under the various correlation assumptions. Slight differences would be expected because the two methodologies are fundamentally different. However, the fact that the results are generally similar suggests that they are both producing reasonable conclusions.

A key strength of the structural model is that it does not require the user to specify the correlation between the classes nor the form of the bivariate distribution; the correlation is simply a by-product of the simulation. In contrast, an explicit correlation assumption and the assumption that the dependence structure follows a normal copula must be made when using the bootstrap models.

The structural model introduces correlation between classes through shared systematic risk and an empirically defined dependence in the non-systematic risk. The shared systematic risk component is of particular interest because it allows the user to reflect the impact of changing socio-economic factors on the aggregate book of business. Table 4.12 compares the structural correlation results under various inflation scenarios.

[12] To facilitate the comparison in this section, we adjusted the volume of the Commercial Auto data by a factor of 50 in order to prevent the Personal Automobile loss distribution from dominating the aggregate distribution. The 50-times adjustment resulted in the sizes of the two classes being approximately equal.

Table 4.12: Structural Model Correlation —Results Under Various Inflation Scenarios

Scenario	Correlation
0.0% mean future inflation	16.3%
3.5% mean future inflation	17.6%
6.0% mean future inflation	18.6%

As the expected future inflation increases, the correlation also increases. Consequently, the benefits of diversification decrease in highly inflationary environments. Unlike the normal copula bootstrap model, the structural model is able to capture this phenomenon.

5. REFERENCES

[1] Butsic, R., "The Effect of Inflation of Losses and Premiums for Property-Liability Insurers," Casualty Actuarial Society *Discussion Paper Program*, 1981, pp. 58-109.

[2] Christofides, S., "Regression Models Based on Log-Incremental Payments," Institute of Actuaries *Claims Reserving Manual*, Vol. 2, 1990, Section D5.

[3] England, P., and R. Verrall, "Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving," *Insurance: Mathematics and Economics* 25, 1999, pp. 281-293.

[4] Hodes, D., S. Feldblum, and G. Blumsohn, "Workers Compensation Reserve Uncertainty," Casualty Actuarial Society *Forum*, Summer 1996, pp. 61-150.

[5] Jing, Y., J. Lebens, and S. Lowe, "Claim Reserving: Performance Testing and the Control Cycle," *Variance* 3:2. 2009, pp. 161-193.

[6] Jing, Y., S. Lowe, and F. Morin, "Risk Horizon and the Measurement of General Insurance Risk," forthcoming, available in draft from the authors.

[7] Kelly, M., "Practical Loss Reserving Method with Stochastic Development Factors," Casualty Actuarial Society *Discussion Paper Program: Insurers Financial Solvency*, 1992, Volume 1, pp. 355-381.

[8] Mack, T., "Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates," *ASTIN Bulletin* 23, 1993, pp. 213-225.

[9] Mulvey, J. and E. Thorlacius, "The Towers Perrin Global Capital Market Scenario Generation System," Worldwide Asset and Liability Modeling, Cambridge University Press, 1998, pp. 286-312

[10] Murphy, D., "Unbiased Loss Development Factors," *Proceedings of the Casualty Actuarial Society* 81, 1994, pp. 154-222.

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

[12] ROC/GIRO Working Party, "Best Estimates and Reserving Uncertainty," 2007, available at www.actuaries.org.uk.

[13] Wacek, M., "The Path of the Ultimate Loss Ratio," *Variance* 1:2. 2007, pp. 173-192.

[14] Wuthrich, M.., M. Merz and N. Lysenko, "Uncertainty of the Claims Development Result in the Chain Ladder Method," Preprint, ETH Zurich, available at www.math.ethz.ch/~wueth/papers/cdr_wumely.pdf.

[15] Zehnwirth, B. and G. Barnett, "Best Estimates for Reserves", *Proceedings of the Casualty Actuarial Society* 87, 2000, pp. 245-321.

Abbreviations and Notations

α, Butsic mixing parameter, determines the degree to which inflation affects the development dimension
ARMA, Auto-Regressive Moving Average
CL, Chain Ladder projection method
CV, Coefficient of Variation
BF, Bornhuetter-Ferguson projection method
ELR, Expected Loss Ratio projection method
ESG, Economic Scenario Generator
RTR, Report-to-Report development factor

Biographies of the Authors

Timothy Gault is a consultant with the Risk and Financial Services segment of Towers Watson, located in the firm's Boston office. He holds a B.S. in Chemistry and a B.A. in Business Economics from the University of California, Santa Barbara. He is a Certified Public Accountant (inactive).

Mr. Gault specializes in deterministic and stochastic reserve modeling for use in risk management and control as well as for uses in internal and external financial reporting. He has performed numerous economic capital, fair value, and traditional reserving analyses for personal, commercial and non insurance lines of business.
Prior to joining Towers Watson, Mr. Gault was employed at a major audit firm, where he specialized in audit and financial reporting for financial services clients. In addition, he has experience in Sarbanes-Oxley internal control compliance.

Len Llaguno, FCAS, MAAA, is a consultant with the Risk and Financial Services segment of Towers Watson, located in the firm's Boston office. Since joining the firm, Mr. Llaguno has worked on a diverse range of insurance and risk management assignments for personal and commercial insurers as well as corporate clients. While his experience includes traditional reserve, ratemaking and funding studies, Mr. Llaguno specializes in developing models to help clients quantify, manage and leverage risk. He has assisted clients to develop economic capital models, stochastic reserving models, predictive models for personal and commercial insurance and models to value customer loyalty program liabilities.

Mr. Llaguno holds an Honors Bachelor of Science degree, magna cum laude, in computer science and actuarial science from the University of Toronto. He is a Fellow of the Casualty Actuarial Society and a Member of the American Academy of Actuaries.

Stephen Lowe, FCAS, MAAA, CERA, is a Director of the Global Property & Casualty Consulting Practice at Towers Watson. He lives and works in Hartford, CT.

With over 25 years of consulting experience, Mr. Lowe has participated in a wide range of assignments, advising both insurance company and corporate clients on a variety of financial, product, and strategic issues. His specialty is assisting clients in understanding the interplay between risk and capital, and how that interplay translates into the creation of value. He has assisted numerous clients in the development of financial models that measure risk and economic capital.

Mr. Lowe is a former member of the Casualty Actuarial Society's Board of Directors, and past Vice President of the American Academy of Actuaries. He has published several actuarial papers, including two that have won the Hachemeister prize.

Credibility for Experience Rating, A Minimum Variance Approach

Lawrence F. Marcus, F.C.A.S., M.A.A.A.

Abstract:

This paper attempts to provide a relatively simple, but still mathematically meaningful context for applying Bühlmann credibility to large account experience rating. It further extends this to rating excess layers. It also allows for the inclusion of an additional complement of credibility to the traditional weighting of excess experience and ILF derived indications. Finally this paper gives guidance as to when exposure rating indications should not be used.

Keywords. Credibility, Experience Rating

1. INTRODUCTION

Let's say you are using two, or potentially more, methods to estimate expected losses. It is well known that combining a number of different estimates will generally lead to a better estimate than any single estimate. The question becomes what is the best mix of these different estimators. One very simple answer is to use the weighting that minimizes the estimation error surrounding the composite estimate. The simplest and most common measure of error to minimize is variance, which of course is equivalent to minimizing standard deviation. As it turns out, the optimal weighting of two independent estimates is inversely proportional to their relative variances. In other words, if one estimator has half the variance of the other, than it should be weighted twice as heavily, and therefore has a credibility of two-thirds. This is relationship is true regardless of context, experience rating credibility, class based rating credibility, or even estimations of an utterly non-insurance nature.

This point is implicit in the original Bühlmann[1] formulation of least squares credibility. Boor[2] makes this point much more explicitly and extends it to the case where the errors between the two estimates are correlated. This paper takes this formulation to case of large account excess pricing, and extends it to include the introduction of a third estimate and to cases where some of the estimates may no longer be relevant.

In the context of large account pricing, primary layers are based on two basic estimates: manual rates and experience rates. These estimates are usually based on losses capped at a basic or working layer (here designated WL) limit in order to keep the variance of estimates to a manageable level. Manual rates, also known as class or exposure rates, are class averages which will be more [2]representative for some large risks than for others. Experience rating is based strictly on an account's own historical experience adjusted for development, trends and exposure changes. An estimate based on an account's own history has the advantage of being much more relevant to that account's future experience than the class averages. However, but for all but the largest risks the volatility of year to year experience will be too high for it to be the best estimate on its own.

Pricing excess layers is subject to even greater errors. Historical experience in excess layers is even more volatile, and excess experience rating requires larger development factors and larger trend factors bringing greater estimation errors. The second generally accepted method for estimating excess losses is by applying an increased limits factor to working layer loss estimates. However, by applying ILF's to an already volatile primary loss estimate greatly increases the error in the excess layer. In addition, the ILF Method estimate and the excess experience estimates are correlated, since a single large loss will increase both the working layer and excess loss experience. This correlation will only add to total variance of the estimates and will alter the optimal weighting. If a third estimate of excess losses is available, perhaps a manual excess rate like the Swiss Re or other reinsurance benchmarks or an internal company estimate, it can add accuracy to the overall estimate. This paper attempts to incorporate all these adjustments into an overall optimal weighting scheme.

2. APPLICATION

The theory for this has been in the literature for years. In reality, the tricky part of this calculation is assessing the errors of using each method. In the basic Bühlmann formulation the error associated with using experience rates was the Expected Process Variance, which is how much an account's historical experience is expected to vary year to year. It does not take into account how much a real life company changes year to year (a fact Bühlmann cited as future research), nor does it take into account how much error trend and development estimates bring into the process.

Ignoring these errors in the analysis of large companies can be problematic. Most calculations of the experience rating variance only increase proportionately with exposures, while the variance from manual ratings will increase with the square of exposure. Thus a comparison of relative errors will exaggerate the credibility of experience for large risk. Venter[3] suggests that errors in experience can be modeled with two components, one linear with exposures plus one quadratic. The quadratic factor in effect means process variance is not completely diversifiable. Errors from trending and developing losses, always very significant for excess layers, are certainly not diversifiable and could be approximately quadratic.

Errors from changing operations are an additional significant challenge. Here underwriting knowledge is essential for determining experience rates. The error behind estimates is probably impossible to determine, but are not likely to be significant. The error behind poor experience rating data likewise is probably unmeasurable as well, and can be more important than any other error.

The error associated with using manual rates in the original Bühlmann formula is only how different risks within the same rating class are. This is the Variance of Hypothetical Means (VHM). Techniques are available to estimate this. However, these estimates likely underestimate the errors associated with manual rates. The manual rates themselves may not have a sufficient loss volume to be fully credible. A perhaps even greater problem is that many large risks do not fit well into the class rating schemes. Many large risks are very unique entities, not represented well by any class and conglomerates which include many risk categories. Many of the rates coming from rating bureaus or company rating systems are based only on smaller, single class risks. It is well known that large risks behave differently from small risks and often demonstrate economies of scale when it comes to loss prevention.

Despite these difficulties, assessing these risks is well worthwhile. There will be some arbitrary judgments involved and inevitable there will be factors that cannot be taken into accounts. Breaking down the risks into these smaller components adds insight to the process compared with using more arbitrary measures (just selecting a k in n/n+k) that seem right for a subset of risks. At the very least one can check current credibility measures against this framework for reasonability. Given all these sources of risk, beyond those that can be reasonably estimate, it is always necessary for the actuary to use informed judgment for both designing rating tools and in helping underwriters to interpret and understand the results.

For most situations involving large account pricing a minimum variance approach is to be preferred. Limited fluctuation credibility (the experience has an x% chance of being within y% of the correct answer) implicitly assumes that that the manual rates are appropriate no matter what their relative reliability. So it probably makes little difference whether Minimum Variance or Limited Fluctuation Credibility is used if one cannot assess to any degree the reliability of manual rates. However, if one can come up with a reasonable estimate; say something of the form that manual rates are accurate for any given risk plus or minus z%, then it should be preferable to use a Minimum Variance approach.

True Bayesian approaches to credibility will yield better results when loss distributions are known with certainty, although for many common distributions they yield the same estimates. However, it is very unlikely that the true underlying loss distributions will that closely enough resemble those assumed in Bayesian analysis to yield better results. Thus this minimum variance approach, also known as Bayesian credibility, should yield just as good results without the added complexity and error added by assuming strict distribution forms. For most of this paper we do not need to assume any specific form of distribution, in many cases an empirical distribution will work just as well. We need only work with expected values, variances, covariances and correlations.

2.1 Working Layer Credibility

Let's define

$Exp_{WL} \equiv Experience\ based\ loss\ estimator\ limited\ to\ the\ working\ layer$

Man_{WL}
$\equiv Manual\ Rate, or\ Class$
$/Exposure\ based\ loss\ estimator\ limited\ to\ the\ working\ layer$

$Z_{WL} \equiv Credibility\ weighting\ of\ \ Exp_{WL}$

$Est_{WL} \equiv Z_{WL} \times\ Exp_{WL} + (1 - Z_{WL}) \times Man_{WL}$, this is the credibility weighted expected loss estimator limited to the working layer.

Let's determine the weighting w, that minimizes the variance of combination of the estimators Exp_{WL} and Man_{WL}. This quantity would be $Var(Est_{WL}) = Var(w \times Exp_{WL} + (1 - w) \times Man_{WL})$. To find the minimum variance estimator we set the derivative of $Var(Est_{WL})$ with respect to w equal to zero.

$$\frac{dVar(Est_{WL})}{dw} = \frac{dVar(w \times Exp_{WL} + (1 - w) \times Man_{WL})}{dw}$$
$$= 2w\sigma^2_{ExpXS} + 2(1 - 2w)\rho_{ExpWL,ManWL}\sigma_{ExpWL}\sigma_{ManWL} - 2(1 - w)\sigma^2_{ManWL} = 0$$

The correlation between manual rates and experience rates will generally be zero as they are separate calculations. While the experience of a given account is potentially in the manual rating database, in most cases this effect should be very small. In the next section we will analyze the situation where the correlation is unequal to zero. Assuming independence, the above equation simplifies to:

$$\frac{dVar(Est_{WL})}{dw} = 2w\sigma^2_{ExpWL} - 2(1 - w)\sigma^2_{ManWL} = 0$$

Rearranging terms and solving for w yields,

$w = \frac{\sigma^2_{ManWL}}{\left(\sigma^2_{ManWL} + \sigma^2_{ExpWL}\right)}$ and

$$1 - w = \frac{\sigma^2_{ExpWL}}{\left(\sigma^2_{ManWL} + \sigma^2_{ExpWL}\right)}$$

What this implies is that the minimum variance of the estimator Est_{WL} is achieved when the weight assigned to the experience rating is equal to the relative size of the variance of manual rate. For instance if the variance of the manual rate is twice variance of the experience rate, then the weighting assigned to the experience rate is twice that of the manual rate.

How do we estimate the variance of our manual rates and experience rating? In Bühlmann's original formulation, σ^2_{ManWL} was the Variance of Hypothetical Means. That is to say how different is the class mean or exposure rate from the true underlying mean of a given risk. We should amend that view to include error coming from less than perfectly credible manual rates and questions about the applicability of manual rates based generally on small risks to large risks. The variance in manual rates can come from the original class rating statistics, or an examination of the distance between manual rates and similar rating classes. This will generally require a decent amount of subjective judgment.

Estimation of σ^2_{ExpWL} also requires judgment. In the original Bühlmann formulation, this was the Expected Process Variance (EPV) of a single risk or period divided by the number of risks or periods observed. For large account rating n will generally be the number of years a risk is observed. Like in our ordinary calculation of sample variance, increasing the number of years observed decreases the variance of our estimate of the mean loss cost. Because the EPV is presumed to be known, we do not need to divide the EPV by (n-1), but can divide by n instead.

Calculation of σ^2_{ExpWL} could come from looking at the variance of annual loss costs in the history, adjusted for trends, development and exposure changes. In practice this estimate is too unstable. A better estimate is to use the manual rates for the risk. The implied severity distributions coming from ILF's can provide an estimate for both mean and variance of limited severities. Dividing the manual rates by the expected average limited severity provides an estimate of annual frequency. Making the usual Poisson assumptions about claim frequencies, we can set the frequency variance equal to the mean claim frequency. We could of course assume some contagion factor in our claims and use a negative binomial distribution. We can then estimate σ^2_{ExpWL} from the usual variance formula.

$$\sigma^2_{ExpWL} = \mu_{frequency}\,\sigma^2_{severity} + \mu^2_{severity}\sigma^2_{frequency}$$

Going back to the original Bühlmann equation we have, and substituting in VHM for σ^2_{ManWL} and EPV/n for σ^2_{ExpWL}

$$w = \frac{\sigma^2_{ManWL}}{\left(\sigma^2_{ManWL}+\sigma^2_{ExpWL}\right)} = \frac{VHM}{(VHM+EPV/n)} = \frac{VHM\times(\frac{n}{VHM})}{(VHM+EPV/n)\times(\frac{n}{VHM})} = \frac{n}{(n+EPV/VHM)} = \frac{n}{n+k} = Z_{WL}$$

So we can now call, w, our weight, the credibility factor

$$Z_{WL} = \frac{\sigma^2_{ManWL}}{\left(\sigma^2_{ManWL}+\sigma^2_{ExpWL}\right)} \; This\ is\ Equation\ 1.$$

This makes the total variance of our estimate:

$$\begin{aligned}\sigma^2_{EstWL} &\equiv Var(Est_{WL}) = Var(Z_{WL}\times\ Exp_{WL} + (1-Z_{WL})\times Man_{WL}) \\ &= \left[\frac{\sigma^2_{ManWL}}{\left(\sigma^2_{ManWL}+\sigma^2_{ExpWL}\right)}\right]^2 \times \sigma^2_{ExpWL} + \left[\frac{\sigma^2_{ExpWL}}{\left(\sigma^2_{ManWL}+\sigma^2_{ExpWL}\right)}\right]^2 \times \sigma^2_{ManWL} \\ &= \frac{\sigma^2_{ExpWL}\sigma^2_{ManWL}}{\left(\sigma^2_{ManWL}+\sigma^2_{ExpWL}\right)} \; This\ is\ Equation\ 2.\end{aligned}$$

Inspection of this formula shows that the variance of the combined estimate is lower than either of the individual variances. Thus we have shown that we have a working layer estimate with the minimum error.

We now move to a similar formula for excess rates. Instead of weighting primary experience rates and primary manual rates, we will now wish to weight excess experience rates and rates generated by multiplying an ILF to our working layer loss estimate. This becomes more complicated because use of ILF's adds error to our working layer estimate with its own error. It is also more complicated because the ILF Method estimate and the Experience Rating estimate are correlated, as both will be affected by random large losses. This correlation will alter our optimal proportions away from our inverse variance rule, towards a greater weighting to the lower variance estimate. This shift will be seen when we derive Equation 5 later in this essay, which is an expansion of Equation 1 to include correlation.

2.2 Variance of ILF Method for Estimating Excess Loss Rates

ILF (Increased Limits Factor) is generally defined as the relation between total limits pure premium and basic limits pure premium, and has also been shown to the be the ratio of the expected severity limited to total limits to the expected severity limited to basic limits. For the purposes of this paper, E(ILF) will be defined as the ratio between an expected Excess Layer severity (Sev_{XS}) and an expected Working Layer severity (Sev_{WL}). We are not defining it here, for these purposes, as the ratio of the random variables representing either total loss or average severities in the two layers.

$$ILF = E(Sev_{XS})/E(Sev_{WL})$$

Let $\widehat{ILF}$ represent our estimator of the ILF, perhaps as represented in company or ISO ILF tables. Therefore the ILF Method for determining excess losses will be defined as

$$ILF\ Method\ Expected\ Loss \equiv \widehat{ILF} \times Estimated\ Working\ Layer\ Expected\ Loss$$

As has already been shown the credibility weighted estimate of the working layer is

$$Est_{WL} = Z_{WL} \times Exp_{WL} + (1 - Z_{WL}) \times Man_{WL}$$

Therefore the ILF Method Expected Loss estimate is

$$\widehat{ILF} \times Est_{WL} = \widehat{ILF} \times \ [Z_{WL} \times Exp_{WL} + (1 - Z_{WL}) \times Man_{WL}]$$

The difference between the ILF and the $\widehat{ILF}$ will merely be the estimation error in measuring an ILF for a given account. For this paper we are not defining the ILF as the ratio of excess and working layer losses but instead the ratio of expected losses in the excess layer to expected losses in the working layer.

Estimating an excess layer pure premium with an ILF adds an additional source of variance to the overall estimate.

Therefore, $Var\left(\widehat{ILF} \times Est_{WL}\right) > \widehat{ILF}^2 \times Var(Est_{WL})$.

It is generally reasonable to assume that the working layer expected losses and $\widehat{ILF}$ are independent:

$$Var(\widehat{ILF} \times Est_{WL}) = E(\widehat{ILF}^2 \times {Est_{WL}}^2) - E^2(\widehat{ILF}) \times E^2(Est_{WL})$$
$$= E(\widehat{ILF}^2) \times E({Est_{WL}}^2) - E^2(\widehat{ILF}) \times E^2(Est_{WL})$$
$$= [Var(\widehat{ILF}) + E^2(\widehat{ILF})] \times [Var(Est_{WL}) + E^2(Est_{WL})] - E^2(\widehat{ILF}) \times E^2(Est_{WL})$$
$$= (Var(\widehat{ILF}) \times Var(Est_{WL})) + \left(\widehat{ILF}^2 \times Var(Est_{WL})\right)\left(Est_{WL}^2 \times Var(\widehat{ILF})\right)$$

Equivalently, we can write:

$$\sigma^2_{\text{ILF}\times\text{WL}} = [(\sigma^2_{\text{ILF}} \times \sigma^2_{\text{EstWL}}) + (\widehat{ILF}^2 \times \sigma^2_{\text{EstWL}}) + (Est_{WL}^2 \times \sigma^2_{\text{ILF}})], \qquad \text{(Equation 3)}$$

.

This is also our first major equation for determining excess layer credibilities.

Estimating the variability of $\widehat{ILF}$ is naturally a tricky proposition. One way to estimate these is too look at different ILF tables. One could estimate the standard deviation as half the distance between the current ILF and the next higher ILF table. Thus for a certain combination of limits the ISO Table A ILF is 1.4 and Table B is 1.6. A fair estimate of standard deviation might be 0.1. Thus for the majority of Table A risks the appropriate ILF to use would be between 1.3 and 1.5. Remember the error we are interested in is the appropriateness of using the expected ILF's found in our ILF tables. $Var(\widehat{ILF})$ is analogous to the Variance of Hypothetical means in our manual rates, where we (unlike Bühlmann's most basic formulation) do not have to assume they are perfectly accurate for a class as a whole. In fact, a very similar technique can be used for estimating the VHM of our manual rates.

2.3 Correlation between Excess Experience Rating and the ILF Method

Let us define:

Exp_{XS}; Experience based loss estimate of the excess layer, trended, developed and exposure adjusted.

The correlation between the excess experience rating and the ILF method is clearly greater than zero. Large losses will increase both the working layer experience rating and the excess layer experience rating. Assume we already have the correlation between aggregate working layer and excess losses ($\rho_{ExpXS,ExpWL}$). This can be estimated either empirically from a sample of loss projections or calculated explicitly from frequency and severity distributions. This is calculated in the Appendix. The rest will need to be broken down into its components.

$$\rho(Exp_{XS}, \widehat{ILF} \times Est_{WL})) = \rho(Exp_{XS}, \widehat{ILF} \times (Z_{WL} \times Exp_{WL} + (1 - Z_{WL}) \times Man_{WL}))$$

Examining the covariance between the excess experience and the ILF Method we have

$$Cov(Exp_{XS}, \widehat{ILF} \times (Z_{WL} \times Exp_{WL} + (1 - Z_{WL}) \times Man_{WL}))$$
$$= Cov(Z_{WL} \times \widehat{ILF} \times Exp_{WL}, Exp_{XS}) + Cov((1 - Z_{WL}) \times \widehat{ILF} \times Man_{WL}, Exp_{XS})$$

There is no apparent dependence between $\widehat{ILF}$, Z_{WL}, the manual rate or excess experience so we can set $Cov((1 - Z_{WL}) \times \widehat{ILF} \times Man, Exp_{XS})$ to zero. This yields:

$$Cov(Exp_{XS}, \widehat{ILF} \times (Z_{WL} \times Exp_{WL} + (1 - Z_{WL}) \times Man_{WL}))$$
$$= Cov(Z_{WL} \times \widehat{ILF} \times Exp_{WL}, Exp_{XS})$$

Since $\widehat{ILF}$ and Z_{WL} are estimators rather than the original random variables they can be treated as independent of our actual experience turning our equation into

$$Cov(Exp_{XS}, \widehat{ILF} \times (Z_{WL} \times Exp_{WL} + (1 - Z_{WL}) \times Man_{WL})) \\ = Z_{WL} \times \widehat{ILF} \times Cov(Exp_{WL}, Exp_{XS}) = Z_{WL} \times \widehat{ILF} \times \rho_{ExpXS,ExpWL} \times \sigma_{ExpXS} \times \sigma_{ExpWL}$$

This implies that

$$\rho(Exp_{XS}, \widehat{ILF} \times (Z_{WL} \times Exp_{WL} + (1 - Z_{WL}) \times Man_{WL})) \\ = \frac{Z_{WL} \times \widehat{ILF} \ \times \rho_{ExpXS,ExpWL}\sigma_{ExpXS}\sigma_{ExpWL}}{\sigma_{ExpXS}\sigma_{ILF((ZExpWL+(1-Z)Man)}} = \frac{Z_{WL} \times \widehat{ILF} \ \times \rho_{ExpXS,ExpWL}\sigma_{ExpWL}}{\sigma_{ILF \times \mathrm{EstWL}}}$$

Or

$$\rho(Exp_{XS}, \widehat{ILF} \times Est_{WL})) \equiv \rho_{ExpXS,ILF \times WL} = \frac{Z_{WL} \times \widehat{ILF} \ \times \rho_{ExpXS,ExpWL}\sigma_{ExpWL}}{\sigma_{ILF \times \mathrm{EstWL}}} \quad (Equation\ 4)$$

This is our second main equation to determine the credibility of our excess experience. We can see that it is a function of the credibility of the working layer pick. If the experience is little used in the working layer pick then there is little correlation between the excess experience and the working layer pick, in other words there is little correlation between the excess experience and the ILF times a manual rate.

2.4 Excess Credibility – ILF Method and Excess Experience

Let our Excess Rate, Est_{XS}, be defined as the weighted average of our Excess Experience and ILF Method Estimates. Let w be the weight assigned to excess experience rate.

$$Est_{XS} = \mathrm{w} \times Exp_{xs} + (1 - w) \times (\widehat{ILF} \ \times Est_{WL})$$

$$Var(Est_{XS}) = Var\big(\mathrm{w} \times Exp_{xs} + (1 - w) \times (\widehat{ILF} \ \times Est_{WL})\big) \\ = w^2\sigma^2_{ExpXS} + 2w(1 - w)\rho_{ExpXS,ILF \times WL}\sigma_{ExpXS}\sigma_{ILFxWL} + (1 - w)^2\sigma^2_{ILF \times WL}$$

Note that for brevity when we want to represent the ILF Method as a subscript we use ILF x WL, rather than the more precise $\widehat{ILF}$ x EstWL.

Minimize the variance of the above expression by taking its derivative with respect to w and setting it equal to zero:

$$\frac{dVar(Est_{XS})}{dw} = 2w\sigma^2_{ExpXS} + 2(1 - 2w)\rho_{ExpXS,ILF \times WL}\sigma_{ExpXS}\sigma_{ILF \times WL} - 2(1 - w)\sigma^2_{ILF \times WL} = 0$$

Regrouping terms yields:

$$w\big(\sigma^2_{ExpXS} - 2\rho_{ExpXS,ILF \times WL}\sigma_{ExpXS}\sigma_{ILF \times WL} + \sigma^2_{ILF \times WL}\big) + \rho_{ExpXS,ILF \times WL}\sigma_{ExpXS}\sigma_{ILF \times WL} \\ - \sigma^2_{ILF \times WL} = 0$$

Solving for w, we have

$$w = \frac{(\sigma^2_{ILFxWL} - \rho_{ExpXS,ILF \times WL}\sigma_{ExpXS}\sigma_{ILF \times WL}\big)}{\big(\sigma^2_{ExpXS} - 2\rho_{ExpXS,ILF \times WL}\sigma_{ExpXS}\sigma_{ILF \times EstWL} + \sigma^2_{ILF \times WL}\big)}$$

So w is the weight that minimizes the error when combining the excess experience and ILF method. We can redefine w as our excess credibility σ^2_{ExpXS} factor. This equation, in a more general context has previously been derived by Boor (op. cit.) and others.

$$Z_{xs} = \frac{\sigma^2_{ILF\times EstWL} - \rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}}{\left(\sigma^2_{ExpXS} - 2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL} + \sigma^2_{ILF\times WL}\right)}, \qquad \text{(Equation 5)}$$

From the above equation we need to be able to calculate the variance of the ILF method $\sigma^2_{ILF\times WL}$, the variance of the excess experience rating σ^2_{ExpXS}, and the correlation between the two $\rho_{ExpXS,ILF\times WL}$. Estimation of σ^2_{ExpXS} is very similar to the estimation of working layer experience variance. Here too we can break up the ILF Method expected excess into frequency and severity components. For the other quantities we have already derived Equations 3 and 4.

2.5 Three way credibility: Adding in excess manual rates

Additional reduction in the variance of excess can obtained if you have a third source of estimates, beyond excess experience and the ILF approach. In Europe there are a number of industry benchmarks, such as the Swiss Re curves in casualty. Alternatively companies may have their own excess rate exposure estimates.

Let's call this excess rate our excess manual rate Man_{XS}.

Our new excess estimate will be

$$Est_{XS} = \mathrm{w}_1 \times Exp_{xs} + w_2 \times (\mu_{ILF} \times Est_{WL}) + (1 - \mathrm{w}_1 - \mathrm{w}_2) \times Man_{XS}$$

Taking the variance of this estimate we get:

$$\begin{aligned} Var(Est_{XS}) &= Var\left(\mathrm{w}_1 \times Exp_{xs} + w_2 \times \left(\widehat{ILF} \times Est_{WL}\right) + (1 - \mathrm{w}_1 - \mathrm{w}_2) \times Man_{XS}\right) \\ &= w_1^2\sigma^2_{ExpXS} + w_2^2\sigma^2_{ILF\times WL} + (1 - \mathrm{w}_1 - \mathrm{w}_2)^2\sigma^2_{ManXS} \\ &\quad + 2\mathrm{w}_1\mathrm{w}_2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL} + 2\mathrm{w}_1(1 - \mathrm{w}_1 - \mathrm{w}_2)\rho_{ExpXS,ManXS}\sigma_{ExpXS}\sigma_{ManXS} \\ &\quad + 2\mathrm{w}_2(1 - \mathrm{w}_1 - \mathrm{w}_2)\rho_{ILFxWL,ManXS}\sigma_{ILFxWL}\sigma_{ManXS} \end{aligned}$$

Assuming that there is no correlation between the excess manual rates and either the excess experience and the ILF method, this simplifies to

$$Var(Est_{XS}) = w_1^2\sigma^2_{ExpXS} + w_2^2\sigma^2_{ILF\times WL} + (1 - \mathrm{w}_1 - \mathrm{w}_2)^2\sigma^2_{ManXS} + 2\mathrm{w}_1\mathrm{w}_2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}$$

To minimize this quantity we take its derivative with respect to both w_1 and w_2 and set the equations equal to zero.

$$\frac{dVar(Est_{XS})}{dw_1} = 2w_1\sigma^2_{ExpXS} - 2(1 - \mathrm{w}_1 - \mathrm{w}_2)\sigma^2_{ManXS} + 2\mathrm{w}_2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL} = 0$$

$$\frac{dVar(Est_{XS})}{dw_2} = 2w_2\sigma^2_{ILF\times WL} - 2(1 - \mathrm{w}_1 - \mathrm{w}_2)\sigma^2_{ManXS} + 2\mathrm{w}_1\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL} = 0$$

This yields two equations in two unknowns. Regrouping, and dividing by 2, we get:

$$w_1\left(\sigma^2_{ExpXS}+\sigma^2_{ManXS}\right)+\mathrm{w}_2\left(\sigma^2_{ManXS}+\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\right)=\sigma^2_{ManXS}$$

$$w_1\left(\sigma^2_{ManXS}+\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\right)+\mathrm{w}_2\left(\sigma^2_{ILF\times WL}+\sigma^2_{ManXS}\right)=\sigma^2_{ManXS}$$

Solving for w_1 and w_2, we get after some manipulation:

$$Z_{Exp}=w_1 = \frac{\sigma^2_{ManXS}\left(\sigma^2_{ILF\times WL}-\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\right)}{\sigma^2_{ExpXS}\sigma^2_{ManXS}+\sigma^2_{ILF\times WL}\sigma^2_{ManXS}-2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\sigma^2_{ManXS}+\sigma^2_{ExpXS}\sigma^2_{ILF\times WL}\left(1-\rho^2_{ExpXS,ILF\times WL}\right)}$$

$$Z_{ILF\times WL}=w_2 = \frac{\sigma^2_{ManXS}\left(\sigma^2_{ExpXS}-\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\right)}{\sigma^2_{ExpXS}\sigma^2_{ManXS}+\sigma^2_{ILF\times WL}\sigma^2_{ManXS}-2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\sigma^2_{ManXS}+\sigma^2_{ExpXS}\sigma^2_{ILF\times WL}\left(1-\rho^2_{ExpXS,ILF\times WL}\right)}$$

$$Z_{ManXS}=1-w_1-w_2 = \frac{\sigma^2_{ExpXS}\sigma^2_{ILF\times WL}\left(1-\rho^2_{ExpXS,ILF\times WL}\right)}{\sigma^2_{ExpXS}\sigma^2_{ManXS}+\sigma^2_{ILF\times WL}\sigma^2_{ManXS}-2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\sigma^2_{ManXS}+\sigma^2_{ExpXS}\sigma^2_{ILF\times WL}\left(1-\rho^2_{ExpXS,ILF\times WL}\right)}$$

If we assume correlation between the excess manual rates and the other two methods we could expand the above equations even further.

Note that if we could assume no correlation between the excess experience and the ILF Method, then the above equations become simpler. In essence, the credibility assigned to each method would have been proportional to the product of the variances of the other two methods. Thus we can extend our inverse variance rule for two estimates to an inverse product rule for three.

A couple of observations are to be noted. First, the uncorrelated risk gets a greater weighting than the other two risks. For instance if we set all the variances to be equal, the credibility of the manual rate is proportional to $1-\rho^2_{ExpXS,ILF\times WL}$ where the other estimates are proportional to $1-\rho_{ExpXS,ILF\times WL}$. Given a positive but imperfect correlation between excess experience and the ILF method (due to large losses) the excess manual rate gets a greater weighting.

3. A CLASSICAL STATISTICAL TWIST TO BAYESIAN CREDIBILITY

To this point this we have focused strictly on an expansion of the original Bühlmann approach to credibility, which is often called Bayesian Credibility. It is called this because even before Bühlmann the traditional credibility formula (Z x Average Experience + (1-Z) x Class Mean) could give the exact same answer as a formal Bayesian statistical approach, depending on the exact formula for Z. It was recognized that for certain combinations of prior probability distributions (here distribution where the manual rate is the mean) and likelihood functions (the distribution describing the experience) would yield a posterior distribution whose mean was equal to the results of our usual credibility formula. It was later shown that this works for a very broad class of distributions, known as the exponential family.[4]

Bühlmann's insight was that our usual credibility formula was the minimum least-squares estimator for any class, as long as Z= n/n+K, where K= EPV/VHM. Thus this result is pretty robust and does not depend on specifying any distribution. The extension of this method in this paper has the same advantage.

Implicit in this formulation is that the class rating is relevant to the risk we are rating. As was detailed above, for many large risks this will not necessarily be the case. Without adjustment, our credibility approach will always put some weight on our manual rates and our experience will never be considered fully credible. This is because neither the variance of manual rates will approach infinity nor will the variance of the experience rates approach zero. Even if the manual rate is a pretty good estimator for a risk, at some point an account should be sufficiently large that the experience rating alone will be the best estimator of the account's future experience. The NCCI experience rating plans recognize this, but make a somewhat arbitrary adjustment to arrive at a full credibility standard.

An approach which solves both these problems is to subject the manual rating to classical statistical testing. Specifically we will test the null hypothesis that the manual rate is valid, or more specifically that the manual rate is a valid mean for our risk given the actual experience. Our alternative hypothesis that the manual rate is invalid implies that the best estimate of rate for the risk is determined solely by the experience alone. In other words, if the null hypothesis is disproven, the proper credibility of the experience is 100%, and the account should be self-rated.

We chose as our null hypothesis that the manual rating is valid since class rating is the standard for most risks and because we want to temper the volatility of pure experience ratings. As with any hypothesis test a significance level needs to be assumed. The significance level we choose represents the probability that manual rate is either higher or lower than the experience due solely to chance, rather than being inappropriate for the risk being rated. Selection of the specific significance level is to some extent an arbitrary choice dependent on any number of considerations including:

1. Prior beliefs -a very high level of proof is required to disprove long held "facts" or to support a hypothesis that would be considered radical.
2. Consequences – a very high level of proof is required if the practical policy implied by the results is either expensive or risky. For example, there should be a high level of proof that a dangerous drug is effective before it is used to treat an illness.

In this case, we know that most manual rates have issues when applied to large risks, so we have good reason to not make our significance level too exacting. If we set our standards for rejecting the manual rates too low, we are at risk of seriously under-pricing business; however we will explore a way to limit this risk in the next section. Given the circumstances, a relatively commonly selected 5% significance level may not be a bad choice.

So we set up our hypothesis testing, for say the working layer:

$H_0: \mu_{WL} = Man_{WL};$

$H_0: \mu_{WL} < Man_{WL}, for\ the\ case\ where\ the\ manual\ rate\ is\ greater\ than\ the\ experience$

$\alpha = .05$

We could use the standard Student t-test:

$t = (\bar{x} - \mu_{WL})/(s/\sqrt{n})$, with n-1 degrees of freedom, where $\bar{x}$ is the average of your experience rating years and s is the sample standard deviation over n years

Perhaps a better formulation is to look at Man_{WL} and σ_{ManWL} from the prior section. As discussed before there are some better, and particularly more stable, estimates for the variance of the experience rating than just the sample standard deviation across what must be a limited sample of years. With these estimates one can go to either a standard normal test statistic, or better yet to choose an experience distribution that the actuary prefers. For instance with two moments in hand and knowledge that the experience should never be below zero we can do hypothesis testing assuming either a gamma or lognormal distribution for ease. Note that if we do this we will want to test aggregate losses over the experience period. Therefore we will look at $n \times Man_{WL}$ and $\sqrt{n} \times \sigma_{ManWL}$, which represent the first two moments for the aggregate loss distribution given that manual rate is valid. In practice, we will also want to do exposure adjustments to get to a true "As If" future basis. We compare this distribution to the actual, adjusted losses, $n \times Exp_{WL}$.

So again we look at if

$$Prob\left(\sum_{n} x = n \times Exp_{WL} \middle| n \times \mu_{WL} = n \times Man_{WL}\right) \leq .05$$

And then we reject the exposure rates.

No explicit distribution is required; we could also use a standard of two or three standard deviations to reject the manual rates. However, using a distribution gives some more flexibility, even if it adds some model specification error.

If we reject the manual rates, the experience rating data becomes the only valid source for expected losses for a given account. At this point the account would generally be self-rated. For smaller accounts, the exposure rating will only be rejected if the exposure rate is dramatically different from experience. For very large accounts the exposure rating may be rejected even if it's reasonably close to the experience. This occurs, of course, because standard deviation of losses grows much more slowly than expected losses as an account gets larger.

Just because we reject the exposure rating, does not mean we have to use the experience rate unmodified. To be specific, we do not have to use the mean of experience, even suitably adjusted. This is particularly the case when the experience rating is less than the exposure rating; by rejecting exposure rates we have now eliminated a measure of conservatism and stability in the loss rating process.

Implicitly this assumes that under-pricing business is more dangerous than over-pricing, as the underwriting loss due to under-pricing is worse than the revenue loss due to over-pricing. This will not in practice always be true. Later we will address the case of when the experience is worse than the exposure rating.

Having eliminated both the stability and conservatism of exposure rates, we are faced with the possibility that the experience rate is still understated. Even if the exposure rating is too high to be valid, it doesn't mean that the good experience isn't partially a product of luck. Let's say that we had determined that based on the exposure rating the experience was in the 1^{st} percentile or lower. We would reject the exposure rate. But to be conservative we could still say the experience was the 5^{th} or 10^{th} percentile, and explicitly make an assumption about how lucky we were.

If we did assume that the experience was the 10^{th} percentile, we would need want to know what the underlying mean was, since this will be our new expected loss pick for the working layer. For many typical distributions we can use numerical techniques to converge upon the proper answer.

Since we will often use two parameter aggregate distributions, we can extend the logic laid out in the Working Layer Credibility section. Let's first assume that the underlying mean of the "true distribution" is, say, twice our selected 10^{th} percentile. We can then imply a variance or standard deviation for that distribution by again assuming that selected ILF's have a reasonable claim severity distribution behind them. From this we can once again back into an implied frequency by dividing out the newly assumed mean and the ILF expected severity. Making the usual Poisson assumption, we assume that the frequency variance equals the mean. From this and our formula for the variance of an aggregate distribution, we can again parameterize the aggregate distribution. With the appropriate aggregate distribution we can then see what percentile our experience rate is. If our experience rating pick is below the 10th percentile, we can use a lower estimate for mean of the aggregate distribution; otherwise we use a higher estimate of the mean. We can through multiple iterations to get to a mean that is close enough.

Unfortunately adding this margin of error over the experience rating mean does mean that we will add some margin of error over the mean for even the largest risks. Surely that margin will be smaller as a percentage of the experience mean for larger risks, but it will always exist. Another solution is to use the newly determined mean as a new complement of credibility to the actual experience mean. The minimum variance credibility formulas of this paper can not be used, since we don't really have two separate variances to minimize, plus as discussed under minimum variance credibility (and all related approaches) you never have full credibility. Here, for this limited case, we can use a full credibility standard derived from a limited fluctuation approach.

Now let's examine the case where the experience rate is so much worse than the exposure rate, and so we reject the exposure rate as being too low. This removes a more liberal element from our calculation as well as a source of stability. We could use the procedure laid out above to come up with an estimate lower than experience, but still above what a credibility calculation including the exposure rate would have yielded. This becomes questionable, because we justified building in the margin of error above based on a principle of conservatism. We may do this if we believe there is some information left in the *direction* of exposure rating, even if we reject the magnitude of exposure rating. If we believe that the exposure rating tells us nothing about a specific account, we can use the experience rating unmodified or even *add* a margin of error to our estimate. This decision becomes a matter of actuarial or underwriting judgment.

4.0 RESULTS AND CONCULUSIONS

In this essay we are looking at the optimal way to combine a number of different estimates in calculating both primary and excess large account losses. At its base we need to compare the relative size of estimation errors in a structured way to come up with the best possible composite estimate. For the simple case of comparing two uncorrelated estimates, the optimal weighting is in inverse proportion to each method's estimation error. The bulk of the paper is expanding that frame work to take into account a third estimate, correlations, and connections between the errors coming from primary and excess ratings.

We have the following main sources of variance:

1. Working Layer Manual Rates – These need to be estimated, either from in informed judgment or from a close examination primary ratemaking techniques. This is more than just the Variance of Hypothetical Means, but need to incorporate additional sources of error including errors in the manual rate estimation as it applies to large accounts.

2. Working Layer Experience Rates – This can be estimated from actual year over year historical experience. However a more stable estimate, which does not bias against risks with better expected experience, will be to look at manual rates and come up with an expected frequency and severity based on ILF's. The Expected Process Variance can be calculated this way. However, as was the case with the Variance of Hypothetical Means, additional sources of error should be contemplated, including trend and development.

3. ILF's – ILF's should not be looked at solely as point estimates, but should be viewed as random variable requiring estimation. Assessing the error in ILF's avoids overweighting the ILF's approach to excess rates.

4. Excess Manual Rates – Like working layer manual rates these will have an uncertainty associated with them, however we would expect these errors to be greater. Excess manual rates may not be based on sufficient data, often have their own sets of implied ILF's, and require trending and developing of data.

5. Excess Experience Rates – Excess experience rates suffer all the same errors as working layer rates, however these too are much greater. A similar approach to determining variance can be estimated for excess rates; however we are even more dependent on imperfect measures to base these upon.

With these estimates, we can use the algebra presented here to come up with appropriate credibility weightings.

1. The Working Layer credibility for experience when combined with a manual rate:

$$Z_{WL} = \frac{\sigma^2_{ManWL}}{\left(\sigma^2_{ManWL} + \sigma^2_{ExpWL}\right)} \quad (Equation\ 1)$$

2. The total variance of the working layer, when using the above weighting to determine the optimal weighting for the experience and manual rates:

$$\sigma^2_{EstWL} = \frac{\sigma^2_{ExpWL}\sigma^2_{ManWL}}{\left(\sigma^2_{ManWL} + \sigma^2_{ExpWL}\right)} \quad (Equation\ 2)$$

3. The credibility assigned to an excess layers experience, when compared to the ILF method for determining excess losses, including recognition of the correlation between the ILF Method and the excess experience due to large losses being part of both estimates.

$$Z_{xs} = \frac{\sigma^2_{ILF\times WL} - \rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}}{\left(\sigma^2_{ExpXS} - 2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL} + \sigma^2_{ILF\times EstWL}\right)} \quad (Equation\ 5)$$

4. Equation 5 above requires an estimated variance for the ILF method, derived from the errors of both the working layer estimate and the error derived from the use of ILF's themselves:

$$\sigma^2_{\mathrm{ILF\times WL}} = (\sigma^2_{\mathrm{ILF}} \times \sigma^2_{\mathrm{EstWL}} + \widehat{ILF}^2 \times \sigma^2_{\mathrm{EstWL}} + Est^2_{WL} \times \sigma^2_{\mathrm{ILF}}) \quad (Equation\ 3)$$

5. Equation 5 above also requires an estimate for the correlation between excess experience and the ILF Method.

$$\rho_{ExpXS,ILF\times WL} = \frac{Z_{WL} \times \widehat{ILF} \times \rho_{ExpXS,ExpWL}\sigma_{ExpWL}}{\sigma_{ILF\times \text{EstWL}}} \quad (Equation\ 4)$$

6. Equation 4 requires a further estimation of $\rho_{ExpXS,ExpWL}$. The best way to estimate this is from severity distributions. From the Appendix we have:

$$\rho_{ExpXS,ExpWL} \cong \rho_{WL,XS} = \frac{(Limit)\mu_{SevXS}}{\sqrt{\sigma^2_{SevWL} + \mu^2_{SevWL}} \times \sqrt{\sigma^2_{SevXS} + \mu^2_{SevXS}}} \quad (Equation\ 6)$$

where the Limit refers to the Working Layer Limit of Liability.

7. If we have a third estimator for excess loss we can then expand the credibility formulas to:

$$Z_{ExpXS} = \frac{\sigma^2_{ManXS}(\sigma^2_{ILF\times WL} - \rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL})}{\sigma^2_{ExpXS}\sigma^2_{ManXS} + \sigma^2_{ILF\times WL}\sigma^2_{ManXS} - 2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\sigma^2_{ManXS} + \sigma^2_{ExpXS}\sigma^2_{ILF\times WL}\left(1 - \rho^2_{ExpXS,ILF\times WL}\right)}$$

$$Z_{ILF\times WL} \frac{\sigma^2_{ManXS}(\sigma^2_{ExpXS} - \rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL})}{\sigma^2_{ExpXS}\sigma^2_{ManXS} + \sigma^2_{ILF\times WL}\sigma^2_{ManXS} - 2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\sigma^2_{ManXS} + \sigma^2_{ExpXS}\sigma^2_{ILF\times WL}\left(1 - \rho^2_{ExpXS,ILF\times WL}\right)}$$

$$Z_{ManXS} = \frac{\sigma^2_{ExpXS}\sigma^2_{ILF\times WL}(1 - \rho^2_{ExpXS,ILF\times WL})}{\sigma^2_{ExpXS}\sigma^2_{ManXS} + \sigma^2_{ILF\times WL}\sigma^2_{ManXS} - 2\rho_{ExpXS,ILF\times WL}\sigma_{ExpXS}\sigma_{ILF\times WL}\sigma^2_{ManXS} + \sigma^2_{ExpXS}\sigma^2_{ILF\times WL}\left(1 - \rho^2_{ExpXS,ILF\times WL}\right)}$$

It is well known that rating casualty layers in difficult, especially for excess layers, because of the errors laid out above and more. However, with the framework presented by this paper, plus some more clever solutions to estimating the variances above, a methodical approach to determining credibility can be established. While the mechanics look complicated, most of these factors can be easily be put into a spreadsheet or a computer program. If need be, correlations could be ignored to simplify the calculations.

This method, a generalization of Bühlmann's Bayesian Credibility, does come up with an optimized weighting unlike limited fluctuation credibility, and does not require any kind of distributional assumptions like pure Bayesian analysis. To get answers we need to only estimate variances and correlations, which although tricky, is far easier than estimating appropriate distributions.

We have also examined a procedure for examining the case of when exposure rating should no longer be used at all. We do so by hypothesis testing the exposure rates as the underlying mean of a distribution that yielded our experience. If we reject the exposure rates, then we can use the experience rating mean. Alternatively we can use the experience mean plus a statistically determined margin of error. If desired this new pick can become the complement of credibility in a limited fluctuation calculation.

Lawrence Marcus is Vice President at Zurich Financial Services. He is a Fellow of the CAS and a Member of the American Academy of Actuaries. He is also a CFA Charterholder.

Appendix A

Estimating the Correlation between Working Layers and Excess Layer Experience

In order to estimate the correlation between the ILF Method and Excess Experience, $\rho_{ExpXS,ExpWL}$, we need to estimate the correlation between aggregate Working Layer experience and Excess Layer experience. We will approximate this with $\rho_{XS,WL}$ which is ultimately calculated by looking at Working Layer and Excess Severities.

To do this we need to devise a formula for aggregate losses then calculate the correlation between Working Layer severities and Excess severities. Let's look at the covariance between the aggregate losses of in the Working Layer (WL) and Excess Layer (XS). The number of claims is unknown but represented by the random variable N. Lets designate each individual claim $SevWL_i$ for the Severity in the Working Layer of Claim i. We will make the usual assumption that individual claims are independently drawn from the same ground up severity distribution.

$$Cov(WL, XS) = Cov\left(\sum_{i=1}^{N} SevWL_i\,, \sum_{i=1}^{N} SevXS_i\right)$$
$$= E\left(\sum_{i=1}^{N} SevWL_i \sum_{i=1}^{N} XS_i\right) - E\left(\sum_{i=1}^{N} SevWL_i\right) \times E\left(\sum_{i=1}^{N} SevXS_i\right)$$

They key here is to examine the $E(\sum_{i=1}^{N} SevWL_i \sum_{i=1}^{N} SevXS_i)$, by looking at the conditional expectation of that quantity holding N constant, and then take expected value of that quantity with respect to N.

$$E\left(\sum_{i=1}^{N} SevWL_i \sum_{i=1}^{N} SevXS_i\right) = E_N\left(E\left(\sum_{i=1}^{n} SevWL_i \sum_{i=1}^{n} SevXS_i \middle| N = n\right)\right)$$
$$= E_N\left(E\left(\sum_{i=1}^{n} SevWL_i \times SevXS_i + \sum_{i=1}^{n}\sum_{i \neq j}^{n} SevWL_i \times SevXS_j \middle| N = n\right)\right)$$

In the expectation above $E(SevWL_i \times SevXS_i)$ can be broken into $E(SevWL_i) \times E(SevXS_j)$ only when $i \neq j$ since we assume separate claims are independent. When $i = j$, $SevWL_i$ *and* $SevXS_i$ will be highly correlated because they are the limited and excess portion of the same claim.

$$E\left(\sum_{i=1}^{N} SevWL_i \sum_{i=1}^{N} SevXS_i\right)$$
$$= E_N\left(N \times E(SevWL \times SevXS) + \left(N \times (N-1)\right) \times E(SevWL)E(SevXS)\right)$$
$$= E(N)E(SevWL \times SevXS) + E(N^2 - N)E(SevWL)E(SevXS)$$
$$= E(N)E(SevWL \times SevXS) + Var(N)E(SevWL)E(SevXS)$$
$$+ E^2(N)E(SevWL)E(SevXS) - E(N)E(SevWL)E(SevXS),$$

$$as\ E(N^2) = Var(N) + E^2(N)$$

We also have by similar logic

$$E\left(\sum_{i=1}^{N} SevWL_i\right) = E(N)E(SevWL) \text{ and } E\left(\sum_{i=1}^{N} SevXS_i\right) = E(N)E(SevXS)$$

We can then substitute these last three equations into our equation for covariance above.

$$Cov(WL, XS) = E(N)E(SevWL \times SevXS) + Var(N)E(SevWL)E(SevXS) - E(N)E(SevWL)E(SevXS)$$

Now we know that $Cov(WL, XS) = E(WL \times XS) - E(WL)E(XS)$, so rearranging this and inserting this in to the above equation yields:

$$Cov(WL, XS) = E(N)Cov(SevWL \times SevXS) + Var(N)E(SevWL)E(SevXS)$$

$$Corr(WL, XS) = \frac{Cov(\sum_{i=1}^{N} SevWL_i, \sum_{i=1}^{N} SevXS_i)}{\sqrt{Var(\sum_{i=1}^{N} SevWL_i)Var(\sum_{i=1}^{N} SevXS_i)}}$$

$$= \frac{E(N)Cov(SevWL \times SevXS) + Var(N)E(SevWL)E(SevXS)}{\sqrt{Var(SevWL)E(N) + E^2(SevWL)Var(N)} \times \sqrt{Var(SevXS)E(N) + E^2(SevXS)Var(N)}}$$

The equation within the square root is our usual equation for the variance of a compound process combining frequency and severity. For ease we write the above as

$$\rho_{WL,XS} = \frac{\mu_N \sigma_{SevWL,SevXS} + \sigma_N^2 \mu_{SevWL} \mu_{SevXS}}{\sqrt{\sigma_{SevWL}^2 \mu_N + \mu_{SevWL}^2 \sigma_N^2} \times \sqrt{\sigma_{SevXS}^2 \mu_N + \mu_{SevXS}^2 \sigma_N^2}}$$

Substituting in $\sigma_{SevWL,SevXS} = \rho_{SevWL,SevXS} \sigma_{SevWL} \sigma_{SevWL}$ yields

$$\rho_{WL,XS} = \frac{\mu_N \rho_{SevWL,SevXS} \sigma_{SevWL} \sigma_{SevWL} + \sigma_N^2 \mu_{SevWL} \mu_{SevXS}}{\sqrt{\sigma_{SevWL}^2 \mu_N + \mu_{SevWL}^2 \sigma_N^2} \times \sqrt{\sigma_{SevXS}^2 \mu_N + \mu_{SevXS}^2 \sigma_N^2}}$$

If we assume a Poisson process for claims, with the variance of N equal to mean of N, then all the N's cancel out and we have the following:

$$\rho_{WL,XS} = \frac{\rho_{SevWL,SevXS} \sigma_{SevWL} \sigma_{SevXS} + \mu_{SevWL} \mu_{SevXS}}{\sqrt{\sigma_{SevWL}^2 + \mu_{SevWL}^2} \times \sqrt{\sigma_{SevXS}^2 + \mu_{SevXS}^2}}$$

So now we have an equation for the correlation between aggregate losses in the working and excess layers. Presumably we have estimates for both expected severity and the variance of severity for both our working layer. Thus all we are missing is an estimate for the correlation between the working layer severity and the excess severity.

$$\rho_{SevWL,SevXS} = \frac{\sigma_{SevWL,SevXS}}{\sigma_{SevWL} \sigma_{SevXS}} = \frac{E(SevWL \times SevXS) - \mu_{SevWL} \mu_{SevXS}}{\sigma_{SevWL} \sigma_{SevXS}}$$

$$E(SevWL \times SevXS) = \iint_{0}^{Policy\ Limit} SevWL \times SevXS \times f(SevWL \times SevXS)\, d(SevWL)\, d(SevXS)$$

Let our ground up claims be designated SevGU and let the Working layer limit be designated Limit.

If SevGU < Limit then SevWL = SevGU and SevXS =0

If SevGU >= Limit then SevWL = Limit and SevXS =GU-Limit

We can then rewrite the above equation in terms of SevGU

$$E(SevWL, \mathrm{Sev}XS) = \int_0^{\mathrm{Limit}} \mathrm{SevGU} \times 0 \ \mathrm{f(SevGU)\,d(SevGU)} + \int_{\mathrm{Limit}}^{\mathrm{Policy\ Limit}} \mathrm{Limit} \times (\mathrm{SevGU} - \mathrm{Limit})\ \mathrm{f(SevGU)d(SevGU)}$$

The first integral is easily valued as zero, the second integral can be identified as Limit times the expected XS Claim severity.

$$E(SevWL \times SevXS) = \mathrm{Limit} \times \mathrm{E(SevXS)}$$

Thus we can rewrite $\rho_{SevWL,SevXS}$ as

$$\rho_{SevWL,SevXS} = \frac{\sigma_{SevWL,SevXS}}{\sigma_{SevWL}\sigma_{SevXS}} = \frac{E(SevWL, SevXS) - \mu_{SevWL}\mu_{SevXS}}{\sigma_{SevWL}\sigma_{SevXS}} = \frac{(Limit)\mu_{SevXS} - \mu_{SevWL}\mu_{SevXS}}{\sigma_{SevWL}\sigma_{SevXS}} = \frac{(Limit - \mu_{SevWL})\mu_{SevXS}}{\sigma_{SevWL}\sigma_{SevXS}}$$

Then we can substitute the above equation into our aggregate equation:

$$\rho_{WL,XS} = \frac{\rho_{SevWL,SevXS}\sigma_{SevWL}\sigma_{SevXS} + \mu_{SevWL}\mu_{SevXS}}{\sqrt{\sigma^2_{SevWL} + \mu^2_{SevWL}} \times \sqrt{\sigma^2_{SevXS} + \mu^2_{SevXS}}} = \frac{\rho_{SevWL,SevXS}\sigma_{SevWL}\sigma_{SevXS} + \mu_{SevWL}\mu_{SevXS}}{\sqrt{\sigma^2_{SevWL} + \mu^2_{SevWL}} \times \sqrt{\sigma^2_{SevXS} + \mu^2_{SevXS}}} = \frac{(Limit - \mu_{SevWL})\mu_{SevXS} + \mu_{SevWL}\mu_{SevXS}}{\sqrt{\sigma^2_{SevWL} + \mu^2_{SevWL}} \times \sqrt{\sigma^2_{SevXS} + \mu^2_{SevXS}}} = \frac{(Limit)\mu_{SevXS}}{\sqrt{\sigma^2_{SevWL} + \mu^2_{SevWL}} \times \sqrt{\sigma^2_{SevXS} + \mu^2_{SevXS}}}$$

This is Equation 6.

5. REFERENCES

[1] Hans Bühlmann, ASTIN Bulletin International Actuarial Association - Brussels, Belgium 1967: 4:3 p.199-207.

[2] Joseph Boor, Proceedings of the Casualty Actuarial Society Casualty Actuarial Society - Arlington, Virginia 1992: LXXIX p.166-185

[3] Venter, Gary G., Casualty Actuarial Society Forum Casualty Actuarial Society - Arlington, Virginia 2003: Winter 621-267

[4] Jewell, William S., ASTIN Bulletin International Actuarial Association - Brussels, Belgium 1974: 8:1, 77-90