SYLLABUS OF BASIC EDUCATION

Complete Online Text References

Spring 2023

Exam 7

V02 2023_01_31

Loss Development Using Credibility

Eric Brosius

March 30, 1993

Abstract

Actuaries use development techniques to estimate future losses. Unfortunately, real data is subject to both random fluctuations and systematic distortions; only in textbooks can we expect smooth, stable development patterns. To correct for this, developed losses are often weighted with a prior estimate to stabilize the results.

This paper describes a method that applies credibility directly to the loss development process. The approach appeals to our intuition, but it also has a sound theoretical base. While it requires little more data than the familiar link ratio method and is almost as easy to use, it responds more gracefully to situations in which the data is thin and random fluctuations are severe.

Introduction

The method of least squares development is worth considering whenever random year to year fluctuations in loss experience are significant. This paper provides both a practical guide to its use and a discussion of its theoretical underpinnings. The goal is to provide actuaries with the familiarity and confidence they need to use the method in their work. Along the way we will uncover some related methods which may be used to evaluate losses for new or rapidly changing lines of business, and we will establish a conceptual framework that broadens our understanding of loss development.

Least squares development was proposed by Simon, in his 1957 discussion[1] of a paper by Tapley,[2] as a way to establish loss reserves for automobile bodily injury claims. More recently Clarke has used it to develop reinsurance losses.[3] Both Simon and Clarke justify the method on practical grounds—it works. DeVylder[4] and Robbin[5] apply credibility techniques to loss development, and though these authors approach the subject from a slightly different direction, this paper owes much to their ideas.

We will begin the paper with a simple example that shows how least squares development works. This will help the reader to get a feel for the method, and to compare it with more traditional approaches. We will then apply the method to several loss models; it often proves to be the right tool for the job, although a non-linear Bayesian development function is (in theory) preferable in some cases. The next part of the paper develops credibility formulas, similar to those of Bühlmann, which describe the best linear approximation to the Bayesian estimate in terms of the means and variances of the loss and loss reporting distributions. In the final part we examine the implications of the method for practical work, warn of its limitations, and work out a complete example.

[1] Simon, L.J., *PCAS* 44 (1957), pp. 100–110.

[2] Tapley, D.A., "Month of Loss Deficiency Reserves for Automobile Bodily Injury Losses Including Reserves for Incurred But Not Reported Claims," *PCAS* 43 (1956), pp. 166–198.

[3] Clarke, H.E., "Recent Developments in Reserving for Losses in the London Reinsurance Market," *PCAS* 75 (1988), pp. 9–12, 15–18.

[4] DeVylder, F., "Estimation of IBNR Claims by Credibility Theory," *Insurance Mathematics and Economics* (January 1982), pp. 35–40.

[5] Robbin, I., "A Bayesian Credibility Formula for IBNR Counts," *PCAS* 73 (1986), pp. 129–164.

How the method works—an example

The data in Table 1, while hypothetical, is typical of what one might face in developing losses for a small state. We will assume that the book of business is reasonably stable from year to year, and we will ignore inflation for the time being. Even so, the data is so thin that there are serious fluctuations—fluctuations that make it hard to apply the link ratio method. We are reluctant to give

	Incurred Loss		Link Ratios
AY	15 mo.	27 mo.	15–27
1985	19,039	23,279	1.223
1986	33,040	41,560	1.258
1987	14,637	18,937	1.294
1988	2,785	5,185	1.862
1989	51,606	54,206	1.050
1990	5,726	15,726	2.746
1991	$x =$ 40,490	$y =$?	

Table 1: State AA, Line BB: Losses limited to $10,000 per occurrence.

full credibility to the observed loss for 1991 (which is high already) by applying a large factor to it. On the other hand, we do not wish to ignore it altogether.

Let's take a step back. Focus for a moment upon the 15- and 27-month columns of the table. We wish to predict the 27-month value for the 1991 accident year. We may base our prediction (if we deem it appropriate) upon the 15-month value, which is already known.

Call the value in the 15-month column x and the value in the 27-month column y. We wish to predict y based on x. In this task we are guided by the (x, y) pairs from previous years. For any value of x—even if it had not been $x = 40{,}490$ as we see here—we would have determined in some way a corresponding y-value. Let $L(x)$ be our estimate of y, given that we have already observed x.

The link ratio method The traditional link ratio method estimates y as $L(x) = cx$, where c is a "selected link ratio". The value of c is chosen after a review of the observed link ratios from previous years—as an average of several years, perhaps, or as a weighted average. The choice is not easy in situations like this one, where the observed link ratios vary greatly from year to year.

The budgeted loss method If the fluctuation is extreme, or if past data is not available, the value of x is sometimes ignored. That is, a value k is chosen, and y is estimated as $L(x) = k$ no matter what x may happen to be. This method is known as the "budgeted loss" (or "pegged") method because it fixes the forecast loss y without reference to the observed value x. The estimate k may be chosen either as an average of y values from past years, or by multiplying earned premium for the year by an expected loss ratio, or by a number of other methods.[6]

The problem is depicted graphically in Figure 1.[7] The observed (x, y) values form a collection of points in the (x, y)-plane (Figure 1a). The link ratio method fits a line through the origin to these points; as the observed value x increases, the estimate $L(x)$ increases in direct proportion (Figure 1b). The budgeted loss method, on the other hand, fits a horizontal line; as x increases, $L(x)$ remains unchanged (Figure 1c).

[6] For instance, one can multiply earned exposures by an estimated pure premium. Or, if the data is for a minor coverage which is sold in conjunction with a major coverage, one can multiply developed losses for the major coverage by a ratio determined from the experience of previous years. Different techniques may be appropriate in different situations.

[7] See J.C. Narvell's review of Clarke's paper (*PCAS* 76 (1989), pp. 197–200.) Our approach here parallels Narvell's.

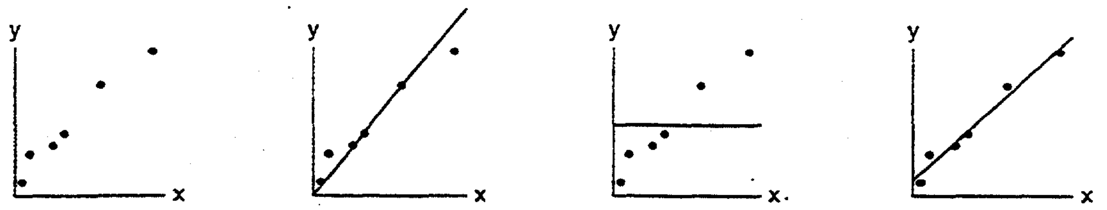


Figure 1: Fitting a line to the loss data from Table 1—a comparison of methods.

The least squares method This method estimates $L(x)$ by fitting a line to the points (x, y) using the method of least squares. The resulting line is not (except by coincidence) either a horizontal line or a line through the origin. Instead it is of the form $L(x) = a + bx$, where the constants a and b are determined by the least squares fit (Figure 1d).

Recall how the least squares coefficients a and b are determined. One first computes the four averages $\overline{x}$, $\overline{y}$, $\overline{x^2}$, and $\overline{xy}$. One then sets

$$b = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - \overline{x}^2} \qquad \text{and} \qquad a = \overline{y} - b\overline{x}.$$

For the 15–27 month development under consideration, and for accident years 1985–1990, we have $\overline{x} = 21{,}139$, $\overline{y} = 26{,}482$, $\overline{x^2} = 7.287 \times 10^8$, and $\overline{xy} = 8.326 \times 10^8$. This gives us $b = 0.968$ and $a = 6{,}023$, which implies that $L(x) = 0.968\,x + 6{,}023$. For the 1991 accident year we estimate $y = 0.968\,(40{,}490) + 6{,}023 = 45{,}217$.

The least squares fit is flexible enough to include the link ratio and budgeted loss methods as special cases, as follows:

- When x and y are totally uncorrelated, b will be zero. In this case the estimate is identical to a budgeted loss estimate. This makes sense; we should not make y dependent on x if we observe no relationship between the two.

- It is also possible for a to be zero—most obviously, when the observed link ratios y/x are all equal. In this case the estimate is identical to a link ratio estimate.

This flexibility is an important advantage of the method. As we shall see below, the least squares method is at heart a credibility weighting system in which the weights are determined by the properties of the loss and loss reporting distributions. It can thus adapt to the data at hand, giving more or less weight to the observed value of x as appropriate.[8]

The Bornhuetter-Ferguson method A third special case is the Bornhuetter-Ferguson method,[9] which estimates ultimate loss as "expected unobserved loss plus actual observed loss"; that is, it sets $L(x) = a + x$ for some a. This method, like ours, seeks a compromise between the link ratio and budgeted loss methods. However, our approach allows b, the coefficient of x, to vary as needed.

[8] Narvell observes that the ~~least-squares~~ estimate is essentially a weighted average and points out the need to understand the nature of the weights. This paper provides such an understanding.

[9] Bornhuetter, R.L. and Ferguson, R.E., "The Actuary and IBNR," *PCAS* 59 (1972), p. 181.

Bornhuetter and Ferguson always have $b = 1$, which can be a real limitation; in particular, Salzmann warns against using the Bornhuetter-Ferguson method when losses develop downward.[10]

Potential problems in parameter estimation Least squares development, like any method that uses observed values to estimate underlying parameters, is subject to parameter estimation errors. If there is a significant change in the nature of the loss experience, the use of unadjusted data can lead to serious errors. Furthermore, even when the book of business is stable, sampling error can lead to values for a and b which do not reflect its true character.[11]

In two cases the mismatch is obvious: if either $a < 0$ or $b < 0$. In the former case, our estimate of y will be negative for small values of x. In the latter case, our estimate of y gets smaller as x increases. The actuary should intervene when either of these situations arises: one might substitute the link ratio method if $a < 0$ and the budgeted loss method if $b < 0$.

Hugh White's question

It is not hard to come up with a variety of loss development methods. The challenge is in deciding which method to use in a given situation. In his review of the Bornhuetter-Ferguson paper, Hugh White asks:[12]

> I offer the following problem. You are trying to establish the reserve for commercial automobile bodily injury and the reported proportion of expected losses as of statement date for the current accident year period is 8% higher than it should be. Do you:
>
> 1. Reduce the bulk reserve a corresponding amount (because you sense an acceleration in the rate of report);
> 2. Leave the bulk reserve at the same percentage level of expected losses (because you sense a random fluctuation such as a large loss); or
> 3. Increase the bulk reserve in proportion to the increase of actual reported over expected reported (because you don't have 100% confidence in your "expected losses")?
>
> Obviously, none of the three suggested "answers" is satisfactory without further extensive investigation, and yet, all are reasonable. While it is a gross over-simplification of the question the reserve actuary will face, it still illustrates the limitations of the effectiveness of expected losses.

We can identify the three "answers" described above as the budgeted loss method, the Bornhuetter-Ferguson method, and the link ratio method, respectively. These three options lie on a continuum—a continuum which also includes the many other options implied by the expression $L(x) = a + bx$.

Let us try to answer Mr. White's question—in which direction, and by how much, should we change our estimate of outstanding losses when reported losses are not what we expected? Each of the above options can be correct in the right circumstances. But how do we know which one to choose? The least squares fit makes sense intuitively, but is there any theoretical justification for its use?

The credibility formulas which we shall develop in this paper are analytical tools that guide us in making these decisions. They lend credence to the least squares method, and they provide the understanding we need to make adjustments when problems arise. Of course, no actuarial formula can serve as a substitute for the actuary him- or herself, or for a thorough knowledge of the book of business; these techniques should supplement, rather than replace, informed judgment.

[10] Salzmann, R.E., *Estimated Liabilities for Losses and Loss Adjustment Expenses* (1984), p. 41.
[11] This problem is not unique to least squares development; the link ratio method is subject to similar errors.
[12] White, H.G., *PCAS* 60 (1973), p. 166.

Loss and loss reporting distributions—using models to test the method

Although the above example is instructive, we need more than experimental evidence if we wish to evaluate the method's theoretical soundness. The fit in Figure 1d looks good, but we may have been lucky. We must know the form of the underlying distributions if we wish to prove that the method works.

For this reason we will test the method using various theoretical models. Our first example is designed for simplicity and not realism. Later examples use the Poisson and negative binomial distributions to model claim counts. If the method handles these latter distributions successfully, we can apply it with some confidence to real-life problems.

A simple model Our first model is designed to clarify the techniques we plan to use. Suppose

- The number of claims incurred each year is a random variable Y which is either 0 or 1 with equal probability.
- If there is a claim, there is a 50% chance that it will be reported by year end.

(Many of our examples involve claim counts. The techniques also apply to incurred losses or claim severity, but the exposition is simplest for claim counts. Note that x and y are integers in this case.)

Question: If x claims have been reported by year end, what is the expected number outstanding?

Let the random variable X represent the number of claims (either 0 or 1) reported by year end. If $Q(x)$ represents the expected total number of claims, and $R(x)$ the expected number of claims outstanding, both given that $X = x$, we have

$$\begin{aligned} Q(x) &= E(Y|X=x), \\ R(x) &= E(Y-X|X=x) \\ &= Q(x)-x. \end{aligned}$$

We begin with the case $x = 0$. Bayes' Theorem tells us[13] that

$$\begin{aligned} P(Y=0|X=0) &= \frac{P(Y=0)P(X=0|Y=0)}{P(Y=0)P(X=0|Y=0)+P(Y=1)P(X=0|Y=1)} \\ &= \frac{(1/2)(1)}{(1/2)(1)+(1/2)(1/2)} \\ &= 2/3, \qquad \text{and similarly} \\ P(Y=1|X=0) &= 1/3. \end{aligned}$$

This means

$$Q(0) = E(Y|X=0) = (0)(2/3)+(1)(1/3) = 1/3;$$

that is, if no claims have been reported by year end, the expected total number of claims is $1/3$. When $x = 1$, our job is even easier. Since in this case y must also have been 1, we must have $Q(1) = 1$. Putting the two together, we have $Q(x) = (2/3)x + 1/3$ where $x = 0$ or 1, and $R(x) = -x/3 + 1/3$.

Return now to the graphical viewpoint (Figure 2.) There are but three possibilities for the point (x, y): it will be $(0,0)$ half the time, $(0,1)$ one quarter of the time, and $(1,1)$ one quarter of the time. The best (Bayesian) estimate of y, given x, is a line with slope $b = 2/3$ and y-intercept $a = 1/3$.

[13] The student may wish to refer to Herzog, T.N., *An introduction to Bayesian credibility and related topics* (CAS, 1985) for an excellent introduction to Bayesian probability.

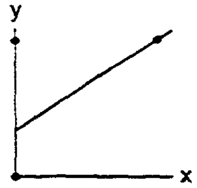


Figure 2: The simple model

Since we have neither $a = 0$, $b = 0$, nor $b = 1$, this relationship is compatible with neither the link ratio method, the budgeted loss method, nor the Bornhuetter-Ferguson method. It is, however, compatible with the least squares method; with enough observations, the least squares estimator will approach $Q(x)$.[14]

A Poisson-Binomial example We now consider a more realistic example. Suppose claim counts for a small book of business have the following properties:

- The number of claims incurred each year is a random variable Y which is Poisson distributed with mean and variance 4.
- Any given claim has a 50% chance of being reported by year end.
- The chance of any claim being reported by year end is independent of the reporting of any other claim, and is also independent of the number of claims incurred.

A sample data set, generated at random, is shown in Table 2. Even though each year's experience is taken from the same distribution, the observed values differ greatly.

	Claims Reported		
	At year end	At ultimate	Link ratio
1984	1	1	1.00
1985	2	9	4.50
1986	1	2	2.00
1987	0	2	—
1988	6	7	1.17
1989	2	5	2.50
1990	1	3	3.00
1991	x	y	?

Table 2: Poisson-Binomial example with $\mu = 4$ and $d = 1/2$.

Here X is a binomial random variable with parameters $(y, 1/2)$. This means X is produced by a Poisson-Binomial mixed process—a Poisson process which produces y followed by a binomial process with y as the first parameter.

Again we ask for the expected number of outstanding claims, given that x claims have been reported by year end. We will solve this problem in two ways: the long way and the short way. We

[14] This example also demonstrates an often overlooked fact: although the least squares line $x = y/2$ expressing x as a function of y passes through the origin, the line expressing y as a function of x does not.

will also consider the link ratio method, but as we shall see, it does not offer an entirely satisfactory solution.

The long way (Bayesian analysis) Bayes' Theorem tells us that, for $y > x$,

$$\begin{aligned} P(Y=y|X=x) &= \frac{P(Y=y)P(X=x|Y=y)}{\sum_i P(Y=i)P(X=x|Y=i)} \\ &= \frac{(4^y e^{-4}/y!)(2^{-y}\binom{y}{x})}{\sum_{i=x}^{\infty}(4^i e^{-4}/i!)(2^{-i}\binom{i}{x})} \\ &= \frac{2^{y-x}e^{-2}}{(y-x)!}. \end{aligned}$$

It follows that

$$\begin{aligned} Q(x) &= \sum_{y=x}^{\infty} y\left(\frac{2^{y-x}e^{-2}}{(y-x)!}\right) \\ &= x\left[\sum_{y=x}^{\infty}\frac{2^{y-x}e^{-2}}{(y-x)!}\right] + \left[\sum_{y=x}^{\infty}(y-x)\frac{2^{y-x}e^{-2}}{(y-x)!}\right] \\ &= x+2 \end{aligned}$$

(where we use our knowledge of the Poisson distribution with mean 2 to evaluate the expressions in square brackets.) The expected number of outstanding claims is thus $R(x) = Q(x) - x = 2$. This may seem surprising, but it is true in general: when the claim distribution is Poisson and the claim reporting distribution is binomial, the expected number of outstanding claims *does not depend* on the number already reported.

The short way Once we know that $R(x) = 2$, the special properties of the Poisson distribution lead us to a quicker derivation. Consider the Poisson process that generates Y to be composed of the sum of two independent Poisson processes with mean 2: one process generating claims that will be reported by year end, and the other generating claims that will not be reported by year end. Regardless of the result of the first process, the expected value of the result of the second process is 2; this is $R(x)$.

Unfortunately, this shortcut will not work for other distributions; in most cases we will have to return to the method that we used above.

The link ratio method Let us now apply the familiar link ratio method to the above problem. To use the link ratio method, one selects a ratio c and uses it to obtain estimates

$$\begin{aligned} E(Y|X=x) &\approx cx, \\ E(Y-X|X=x) &\approx (c-1)x. \end{aligned}$$

Since there is no c for which $cx \equiv x+2$, this method cannot possibly produce the correct Bayesian estimate $Q(x)$ for every value of x. However, there are several options for c.

Option 1. If we wish to obtain an unbiased estimate, we must ask that $E((c-1)X) = 2$. This implies that $c = 1 + 2/E(X) = 2$.

Option 2. Instead we can minimize the mean squared error (MSE) of our estimate. This is equivalent to the problem of minimizing $E(((c-1)X-2)^2) = (c-1)^2\,Var(X) + ((c-1)E(X)-2)^2 = 6c^2 - 20c + 18$. The minimum is found at $c = 5/3$. Unfortunately, as we can see by comparison with Option 1, this estimate is biased low. The biased estimate can have a lower MSE than the unbiased estimate because its variance is lower.

Option 3. One commonly used method uses $E(Y/X)$ (or an estimate thereof) for the link ratio.[15] This presents problems when the data is thin, as in Table 2, since Y/X is not defined where $X = 0$. If we throw these cases out and compute instead $c =$

$$\begin{aligned} E(Y/X\,|X \neq 0) &= (1-P(X=0))^{-1}\sum_{x=1}^{\infty} P(X=x)\frac{E(Y|X=x)}{x} \\ &= (1-e^{-2})^{-1}\sum_{x=1}^{\infty}\frac{2^x e^{-2}}{x!}\frac{x+2}{x} \\ &\approx 2.153, \end{aligned}$$

we obtain an estimate which is biased high, despite the exclusion of cases in which $x = 0$.

Option 4. A better approach (described by Salzmann[16] as the "iceberg technique") selects

$$d = E(X/Y\,|Y \neq 0) = 1/2, \qquad c = d^{-1} = 2.$$

This is the same value of c that produced the unbiased estimate of Option 1; in this example, it is clearly superior to Option 3.

While some values of c are better than others, no link ratio estimate is as good as the Bayesian estimate $Q(x)$. For $c = 5/3$ the MSE is $10/3$, for the unbiased estimate $c = 2$ it is 4, and for $c = 2.153$ it is approximately 4.752. In comparison, for $Q(x)$ (which is also unbiased) the MSE is 2.

The general Poisson-Binomial case If we generalize our example to the situation where Y is Poisson distributed with mean μ, and where any given claim has probability d of being reported by year end, the methods described above yield

$$\begin{aligned} Q(x) &= x + \mu(1-d), \\ R(x) &= \mu(1-d). \end{aligned}$$

The expected number of outstanding claims is simply the total number of claims originally expected times the expected percentage outstanding; as noted above, it does not depend upon the number of claims already reported. We conclude that the Bornhuetter-Ferguson estimate—and hence Mr. White's second answer—is optimal in the Poisson-Binomial case.

The Negative Binomial-Binomial case Although the Poisson distribution is often used to model claim counts, the negative binomial distribution is a better choice in some situations.[17] Let us therefore consider the situation where the distribution of Y is negative binomial with parameters (r, p), and where any given claim has probability d of being reported by year end. Using the techniques of Bayesian analysis described above, we compute

$$\begin{aligned} P(Y=y|X=x) &= \frac{\left[\binom{r+y-1}{y}p^r(1-p)^y\right]\left[\binom{y}{x}d^x(1-d)^{y-x}\right]}{\sum_{i=x}^{\infty}\left[\binom{r+i-1}{i}p^r(1-p)^i\right]\left[\binom{i}{x}d^x(1-d)^{i-x}\right]} \\ &= \binom{(x+r)+(y-x)-1}{y-x}[(1-d)(1-p)]^{y-x}[1-(1-d)(1-p)]^{x+r}, \end{aligned}$$

[15] This method seems to be based on the heuristic assumption that $E(Y)$ can be approximated by $E(X)E(Y/X)$. The problem is that the random variables X and Y/X are often negatively correlated in practice, so that $E(Y) < E(X)E(Y/X)$. This issue is discussed by J.N. Stanard in "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques," *PCAS* 72 (1985), p. 124.

[16] *Op. cit.*, p. 31.

[17] See, for example, Dropkin, L., "Some Considerations on Automobile Rating Systems Utilizing Individual Driving Records", *PCAS* 46 (1959), pp. 165–176.

which is a negative binomial distribution in y with parameters $(x+r, 1-(1-d)(1-p))$, shifted by x. This implies that

$$R(x) = \frac{(1-d)(1-p)}{1-(1-d)(1-p)}(x+r).$$

Except in the trivial case where $d = 1$, this is an increasing linear function in x. Take for example $r = 4$ and $d = p = 1/2$, so that $E(Y) = 4$ and $Var(Y) = 8$. Here $R(x) = x/3 + 4/3$ and $Q(x) = (4/3)x + 4/3$. This does not correspond exactly to any of Mr. White's answers—while an increase in reported claims does lead to an increase in our estimate of outstanding claims, the relationship is not proportional. Since $a = b = 4/3$, neither the link ratio method, the budgeted loss method, nor the Bornhuetter-Ferguson method gives the correct estimate.

How can we make intuitive sense of this result? The negative binomial distribution has more variance than the Poisson distribution with the same mean; as a result, we have less confidence in our prior estimate of expected losses. Given a value of x that is larger than predicted, we are thus relatively more willing to increase our estimated ultimate claim count than we were when Y was Poisson; this implies a larger b.

The fixed prior case Suppose the random variable Y is not random at all; that is, there is some value k such that Y is sure to equal k (perhaps we are selling single-premium whole life policies.) In this case, $Q(x) = k$ for any value of x (regardless of the distribution of X.) The expected number of outstanding claims is then $R(x) = k - x$.

This situation corresponds perfectly to White's first answer—we decrease our estimate of outstanding claims by an amount equal to the increase in reported claims, leaving the total incurred count for the year unchanged.

The fixed reporting case For the other extreme, suppose there is a number $d \neq 0$ such that the percentage of claims reported by year end is always d; that is, $P(X = dy|Y = y)$ is 1 for all y. In this case $Q(x) = d^{-1}x$ and the expected number of outstanding claims is $R(x) = (d^{-1} - 1)x$.

This is our old friend the link ratio method, which corresponds perfectly to White's third answer.[18]

A non-linear example In each of the examples considered above, the Bayesian estimate $Q(x)$ is linear in x, and is thus of the form $a + bx$. This is not always true. The following example, which illustrates a pragmatic approach, leads to a non-linear $Q(x)$.

Company management believes the number of claims Y for the year is uniformly distributed on $\{2, 3, 4, 5, 6\}$—that is, $P(Y = y) = 1/5$ for $y = 2, 3, 4, 5, 6$. (Here $E(Y) = 4$ and $Var(Y) = 2$.) Any given claim has a 50% chance of being reported by year end. Armed with these assumptions, we proceed to compute $Q(x)$. The calculations (Table 3) correspond exactly to those in our first model.

In this example $R(x) = Q(x) - x$ is not linear. It is also not monotonic; it is generally decreasing, but it increases slightly between $x = 1$ and $x = 2$. It makes sense that $R(x)$ should decrease; since Y has less variance than a Poisson distribution with the same mean, we have more confidence in our prior estimate of expected losses, and we are relatively less willing to revise our estimated ultimate claim count based on what has been reported so far.

This example corresponds somewhat to White's third answer, although not as much as the fixed prior example discussed above. It also models real-life pressures in a convincing, if simplistic, way—as long as the losses remain within a "comfort range", the analysis is permitted to take its course, but when the indication strays outside the bounds, there is a tendency to ignore it. The variance of Y here seems unreasonably low; it probably reflects management psychology better than it reflects reality.

The method of Bayesian development Despite the difficulties involved, the technique used in this section has considerable practical applicability. If we are willing to estimate the distributions of

[18] Note, however, that this model is extremely unrealistic: the behavior described could hardly occur in real life unless the claims department were making the claims up!

y	2	3	4	5	6		
$x =$			$P(X = x \text{ and } Y = y)$			*Total*	
0	16/320	8/320	4/320	2/320	1/320	31/320	
1	32/320	24/320	16/320	10/320	6/320	88/320	
2	16/320	24/320	24/320	20/320	15/320	99/320	
3		8/320	16/320	20/320	20/320	64/320	
4			4/320	10/320	15/320	29/320	
5				2/320	6/320	8/320	
6					1/320	1/320	
$x =$			$P(Y = y \mid X = x)$			$Q(x)$	$R(x)$
0	16/31	8/31	4/31	2/31	1/31	$83/31 = 2.677$	2.677
1	32/88	24/88	16/88	10/88	6/88	$256/88 = 2.909$	1.909
2	16/99	24/99	24/99	20/99	15/99	$390/99 = 3.939$	1.939
3		8/64	16/64	20/64	20/64	$308/64 = 4.812$	1.812
4			4/29	10/29	15/29	$156/29 = 5.379$	1.379
5				2/8	6/8	$46/8 = 5.750$	0.770
6					1/1	$6/1 = 6.000$	0.000

Table 3: Y uniform on $\{2, 3, 4, 5, 6\}$ and $d = 1/2$.

Y and $X|Y$, we can produce Bayesian estimates of ultimate claim costs. Even if the equations cannot be solved exactly, it is not hard to approximate the answer to any desired degree of accuracy. We can also test the sensitivity of the answer to changes in the distributions chosen.

The linear approximation (Bayesian credibility)

The final example in the previous section brings us to a fork in the road. While it is certainly possible for the actuary to compute a pure Bayesian estimate Q based on assumed distributions for Y and for $X|Y$, such a procedure requires a good deal of knowledge about the loss and loss reporting processes—knowledge we may not be willing to assume. For this reason we shall now consider a linear estimate that is based on the concept of Bayesian credibility.

Bayesian credibility as described by Bühlmann[19] uses not the Bayesian estimate itself, but the best linear approximation to it. The approximation, though less accurate than the pure Bayesian estimate, is simpler to compute, easier to understand and explain, and less dependent upon the underlying distributions. As we study the application of Bayesian credibility to loss development, our approach will follow the path laid down by Bühlmann.

Let $Q(x)$ be the Bayesian estimate discussed in the previous section, and let L be the best linear approximation to Q; that is, L is the linear function that minimizes $E_X([Q(X) - L(X)]^2)$. If $L(x) = a + bx$, we must minimize

$$E_X([Q(X) - a - bX]^2).$$

The following is a standard statistical result:[20]

Development Formula 1 *Given random variables Y describing ultimate losses and X describing reported losses, let $Q(x) = E(Y|X = x)$. Then the best linear approximation to Q (in the sense*

[19] Bühlmann, H., "Experience Rating and Credibility", *The ASTIN Bulletin* 4 (1967), pp. 199–207.

[20] See for instance Meyers, G., *Report of the Credibility Subcommittee: Development and Testing of Empirical Bayes Credibility Procedures for Classification Ratemaking*, ISO (1980), p. 61.

described above) is the function

$$L(x) = (x - E(X))\frac{Cov(X,Y)}{Var(X)} + E(Y).$$

This equation agrees with our expectations; if $x = E(X)$, we have $L(x) = E(Y)$, but if x differs from $E(X)$, our estimate differs by a proportional amount. This formula provides us with an answer to Mr. White's question, at least if we are willing to make do with the linear approximation:

1. If $Cov(X,Y) < Var(X)$, a large reported amount should lead to a *decrease* in the reserve.
2. If $Cov(X,Y) = Var(X)$, a change in the reported amount should *not effect* the reserve.
3. If $Cov(X,Y) > Var(X)$, a large reported amount should lead to an *increase* in the reserve.

We conclude that each of the three answers is correct in the right circumstances.

Practical application of the first formula—least-squares development

If we had hoped by using Bayesian credibility to avoid making assumptions about the distributions of Y and X, we may be disillusioned to see terms involving these random variables in our formula. This concern is not entirely justified; if we have a series of past years for which we are willing to assume a common Y and X, we can estimate the means, variance, and covariance from the data. Taking the simple-minded approach, we estimate $Cov(X,Y)$ by $\overline{XY} - \overline{X}\,\overline{Y}$, $Var(X)$ by $\overline{X^2} - \overline{X}^2$, $E(X)$ by $\overline{X}$, and $E(Y)$ by $\overline{Y}$. This gives us

$$L(x) = (x - \overline{X})\frac{\overline{XY} - \overline{X}\,\overline{Y}}{\overline{X^2} - \overline{X}^2} + \overline{Y}.$$

Turning back to the data in Table 2, we have $\overline{X} = 13/7$, $\overline{Y} = 29/7$, $\overline{XY} = 76/7$, and $\overline{X^2} = 47/7$. Thus $b \approx 0.969$, $a \approx 2.344$, and $L(x) \approx 0.969\,x + 2.344$. Of course, this is only an approximation to the true Bayesian estimate $Q(x) = x + 2$; sampling error makes it unlikely that we will reproduce Q exactly. Even so, the MSE of our estimate is approximately 2.081—better than the best link ratio estimate and not much worse than the true Bayesian estimate.

As the reader has no doubt recognized, this is the least squares procedure that was introduced at the start of the paper. If it were not for sampling error, the least squares method would give us the best linear approximation to the Bayesian estimate. This is true regardless of the distributions of X and Y.

Note, however, that even if the method is working perfectly, the least squares fit may not yield a high correlation. The points (x, y) can be expected to lie above and below the fitted line $y = L(x)$ because $Var_X(Y|X)$ is not zero.

A simulation test of least-squares development The fit that we obtained in the previous section using data from Table 2 is remarkably good; we will not always do so well. To test the effectiveness of this method, and to compare it to the traditional link ratio method, we will use a simulation test.

For each trial, seven y-values and corresponding x-values were generated at random using the distributions used for Table 2. Two estimates were then produced: one exactly as outlined above, and one using the link ratio method with $c = \overline{Y}/\overline{X}$. The MSE was computed for each.

The results are shown in Table 4. This comparison is "fair": neither method uses prior assumptions about the underlying distributions, since both work solely with the observed data. As we see, when the data fluctuates as much as it does here, either method can go astray. Even so, the least squares method produces a superior estimate in the great majority of cases. In addition, some of its poorer

Trial	$\hat{b}$	$\hat{a}$	*MSE*	*Link Ratio*	*MSE*
1	0.167	4.095	3.573	2.214	5.133
2	2.605	1.079	12.395	3.444	22.296
3	0.308	3.462	2.964	3.000	14.000
4	1.362	1.447	2.291	1.895	3.645
5	1.500	1.429	2.684	2.214	5.133
6	−0.175	4.450	4.771	1.556	3.407
7	0.750	1.643	2.860	1.571	3.388
8	1.356	1.422	2.271	1.941	3.785
9	0.750	2.750	2.188	1.882	3.612
10	1.500	1.500	2.750	3.000	14.000
11	0.130	3.815	3.521	2.800	11.040
12	1.574	−0.704	5.079	1.385	3.811
13	0.939	0.970	3.333	1.462	3.586
14	0.464	4.773	5.465	1.800	3.440
15	0.957	1.787	2.092	2.000	4.000
16	1.138	1.319	2.202	1.600	3.360
17	0.667	1.476	3.639	2.143	4.694
18	1.542	0.708	2.630	1.923	3.728
19	1.958	0.500	4.010	2.250	5.375
20	0.537	2.870	2.432	2.364	6.248
Average	1.001	2.040	3.658	2.122	6.384

Table 4: Comparison of the least squares method with the link ratio method.

performances (trials 6 and 12) can be identified by the appearance of a negative coefficient and judgmentally weeded out as suggested previously. This correction would further increase the accuracy of this method.

Note too that the link ratio method is biased. The average link ratio of 2.122 in Table 4 is higher than the unbiased value of 2.000. This is no accident; we can prove using a power series approximation that the expected link ratio produced by this method is about 2.085. The least squares method may have some sampling bias as well in the determination of $\hat{a}$ and $\hat{b}$, but the bias appears to be significantly less than for the link ratio method.

When is least-squares development appropriate? The careful reader will have noticed the caveat put forth above: the least squares fit makes sense "*if* we have a series of years for which we are willing to assume a common Y and X." For what real-life book of business can it truly be said that a single pair of distributions is appropriate for all years? And what good is a method that relies on such an unlikely assumption?

From a practical point of view the issue is one of relativity: if year to year changes are due largely to systematic shifts in the book of business, other methods may be more appropriate.[21] On the other hand, if random chance is the primary cause of fluctuations, then the present method commends itself to our attention. And it is in this very case that the actuary is in most need of an objective approach; one can correct for systematic distortion, but the temptation when facing variability like that in Table 2 is to throw up one's hands in despair and ignore the data entirely.

Furthermore, one can adjust for known or suspected distortions before using least squares development. If we are studying incurred loss data, a correction for inflation is almost certainly advisable; we should fit our line only after putting the years on a constant-dollar basis. Similarly, if the book of

[21] An excellent discussion of the types of approaches one might take in these situations appears in Berquist, J.R., and Sherman, R.E., "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," *PCAS* 64 (1978), p. 10.

business expands, but does not change in character, we can divide each year's losses by an exposure measure to eliminate the resulting distortion.[22] Other adjustments may be made using techniques such as those discussed in the Berquist-Sherman paper cited above.

A credibility form of the development formula

In this section we consider an alternative form of Development Formula 1 that provides us with additional insight. Following Bühlmann, we seek to express L in terms of

$$\begin{aligned} E_Y(Var(X|Y)) &= \text{"Expected value of the process variance" } (EVPV) \quad \text{and} \\ Var_Y(E(X|Y)) &= \text{"Variance of the hypothetical mean" } (VHM) \end{aligned}$$

(basically, $EVPV$ represents variability resulting from the loss reporting process while VHM represents variability resulting from the loss occurrence process.) Bayesian credibility as it is customarily presented uses one or more observations of a random variable to predict future values of that same variable.[23] Here our task is slightly different: we wish to estimate the value of the random variable Y by observing X, a differently distributed, though related, random variable. This leads to a formula that differs slightly in form from the usual formula for Bayesian credibility, and that requires an additional hypothesis. The proof is given in the Appendix.

Development Formula 2 *Suppose there is a real number $d \neq 0$ such that $E(X|Y = y) = dy$ for all y. Then the best linear approximation to Q (in the sense described previously) is the function*

$$\begin{aligned} L(x) &= \frac{x - E(X)}{d} \frac{VHM}{VHM + EVPV} + E(Y) \\ &= Z\frac{x}{d} + (1 - Z)E(Y), \end{aligned}$$

where

$$Z = \frac{VHM}{VHM + EVPV}.$$

This formula views L as a credibility weighting of the link ratio estimate x/d with the budgeted loss estimate $E(Y)$. If $EVPV = 0$ we give full weight to the link ratio estimate, as in the fixed reporting example discussed above. If $VHM = 0$, as in the fixed prior example, we set $L(x) = E(Y)$. But when there is uncertainty about both the reporting pattern and the prior estimate, we use a weighted average, with weights $EVPV$ and VHM.[24]

Let us apply Formula 2 to some of the other examples discussed above.

- For our simple model with at most one claim per year, the process variance is 0 when $Y = 0$ and $1/4$ when $Y = 1$. (Recall that a binomial process with parameters (n, d) has mean nd and variance $nd(1 - d)$.) Thus $EVPV = (1/2)\,0 + (1/2)(1/4) = 1/8$. The hypothetical mean is 0 when $Y = 0$ and $1/2$ when $Y = 1$, so $VHM = 1/16$. Thus $Z = VHM/(VHM + EVPV) = 1/3$ and $L(x) = (1/3)(x/d) + (2/3)E(Y) = (2/3)x + 1/3$. Of course, this agrees with our previous estimate since $L(x)$ must equal $Q(x)$ whenever Q is linear.
- In the Poisson-Binomial case with parameters μ and d, we have $EVPV = E(yd(1 - d)) = \mu d(1 - d)$ and $VHM = Var(yd) = \mu d^2$. This gives us $Z = \mu d^2/(\mu d^2 + \mu d(1 - d)) = d$ and $L(x) = x + \mu(1 - d)$.

[22] If we assume that the new business is homogenous with the old, both $E(X)$ and $E(Y)$ will increase in proportion to exposure, while $Var(X)$ and $Cov(X, Y)$ will increase in proportion to the square of the exposure. This implies we can divide by exposures to adjust data for use in Development Formula 1.

[23] To be precise, we should speak of a *sequence* of independent, identically distributed, random variables.

[24] A cynic might claim that VHM measures our distrust of the underwriter while $EVPV$ measures our distrust of the claims department!

- More generally, we have $Z \equiv d$ whenever the least squares estimate coincides with the Bornhuetter-Ferguson estimate. This makes sense in that Z should increase from 0 to 1 over time, but there is no reason to expect that it will always do so in exact proportion to d.[25]

- In the Negative Binomial-Binomial case with parameters (r, p) and d, we have $\mu = E(Y) = r(1-p)/p$. Thus $EVPV = \mu d(1-d)$ while $VHM = Var(Yd) = \mu d^2/p$. In this case, $Z = d/(d+p(1-d))$ and $L(x) = x/(d+p(1-d)) + \mu p(1-d)/(d+p(1-d))$. Since VHM is larger here than in the Poisson-Binomial case, while $EVPV$ is the same, Z is larger, and the link ratio estimate receives more weight.

$x =$	$Q(x)$	$L(x)$
0	2.677	2.667
1	2.909	3.333
2	3.939	4.000
3	4.812	4.667
4	5.379	5.333
5	5.750	6.000
6	6.000	6.667

Table 5: Linear approximation: Y uniform on $\{2, 3, 4, 5, 6\}$ and $d = 1/2$.

- Next consider the non-linear example worked out in Table 3. We have $d = 1/2$ and $EVPV = E(Y)d(1-d) = 1$. With $VHM = Var(Yd) = 1/2$, we obtain $Z = (1/2)/(3/2) = 1/3$ and $L(x) = (2/3)x + 8/3$. Since VHM is smaller than in the Poisson-Binomial case, while $EVPV$ is the same, Z is smaller, and the link ratio estimate receives less weight. Here L does not equal Q, but it is the best linear approximation to it. As Table 5 demonstrates, the fit is reasonably good considering the rather artificial distribution of Y.

- Finally, let us return to the example of Table 1, with $b = 0.968$, $a = 6{,}023$, $\overline{x} = 21{,}139$, and $\overline{y} = 26{,}482$. If we set $d = \overline{x}/\overline{y} = 0.798$, then $Z = bd = 0.773$. The least squares estimate which we obtained for this problem can thus be seen to assign a weight of 0.773 to the link ratio estimate (with link ratio $d^{-1} = 1.253$) and a weight of 0.227 to the budgeted loss estimate.

A different application of Bayesian credibility The underlying assumption of the least squares method—that year to year changes in loss and loss reporting distributions are small, or can be corrected for—will sometimes fail. When this happens we can apply Bayesian credibility methods by estimating the terms $EVPV$ and VHM in Development Formula 2.

Consider an example. We wish to develop personal automobile losses for a state which has just instituted a strict verbal tort threshold. Suppose

- Expected losses under the old system would have been \$20 million, but industry studies estimate that the reform should save 40% in the first year.

- In the past about 62% of incurred losses have been reported by year end, but under no fault this figure is expected to rise to 75%.

We are thus expecting an ultimate loss of \$12 million, with \$9 million reported by year end.

[25] I would like to thank Dr. Robbin for pointing out to me that the Bornhuetter-Ferguson estimate is a weighted average of the link ratio and budgeted loss estimates.

When the year-end data is available, however, the reported loss is only \$6 million. This presents us with a dilemma. The savings resulting from the reform may be greater than expected; if so, we should reduce our estimate of ultimate loss. On the other hand, there may be temporary reporting delays as claim adjusters become familiar with the new coverages. In this case, it would be a mistake to reduce our estimate. What do we do while we await better information?

Neither the least squares method nor the link ratio method makes sense here. Both methods assume that past experience is a reliable guide to the future. This assumption is not justified when there has been a major change in coverage. On the other hand, our doubts about the estimated savings make the budgeted loss estimate uncertain.

The Bayesian credibility method provides us with a reasonable solution to this problem. To use this method we must estimate the means and standard deviations of two random variables: the loss Y and the reporting ratio X/Y.[26]

We already have estimates of the means: $E(Y)$ is \$12 million and $E(X/Y)$ is 75%. Suppose we estimate $\sigma(Y)$ to be \$3 million and $\sigma(X/Y)$ to be 14%.[27]

We can then compute

$$\begin{aligned} VHM &= Var(0.75Y) &&= (0.75 \times \$3 \text{ million})^2 &&= 5.06, \\ EVPV &= E((0.14)^2Y^2) &&= (0.14)^2[Var(Y) + E(Y)^2] &&= 3.00. \end{aligned}$$

Thus $Z = 5.06/(5.06 + 3.00) = 0.628$ and $L(x) = 0.628\,(x/0.75) + (1 - 0.628)(\$12 \text{ million}) =$ \$9.5 million.

The estimate is larger than the link ratio estimate \$6 million/(0.75) = \$8 million and smaller than the budgeted loss estimate \$12 million. This reflects our relative uncertainty concerning these two estimates. It is also slightly larger than the Bornhuetter-Ferguson estimate, which would be \$9 million, because $b = 0.628/0.75$ is less than 1. This implies that we have placed slightly less confidence in the low reported loss (or, equivalently, more confidence in the high prior estimate) than if we had used the Bornhuetter-Ferguson method.

To use this method we must be willing to select the means and standard deviations. Fortunately, the answer is not extremely sensitive to changes in these selections. For instance, if we change $\sigma(X|Y)$ to 10% in the example above, $L(x)$ becomes \$8.9 million. If instead we change $\sigma(Y)$ to \$2 million, $L(x)$ becomes \$10.3 million.

The caseload effect

In Development Formula 2, we assumed that the expected number of claims reported is proportional to the number of claims incurred. This might be seen as a flaw in our analysis; since a claim is more likely to be reported in a timely fashion when the caseload is low, we expect the development ratio $E(X|Y = y)/y$ to be not a constant but a decreasing function of y.

Fortunately, a constant development ratio is not essential for a credibility-based development formula. In this section we make the more general assumption that $E(X|Y = y) = dy + x_0$, where $d \neq 0$ (one can presume that both d and x_0 are positive.) This gives a development ratio of $d + x_0/y$, which does indeed decrease as y gets larger. On the other hand, it gives us $E(X|Y = 0) = x_0 > 0$. This may perhaps be undesirable, but no one who has had dealings with a real-life claims department is likely to be shocked by this assumption. When $x_0 = 0$ we obtain Development Formula 2 as a special case. The proof is given in the Appendix.

[26] We assume for the purposes of this example that the mean and standard deviation of X/Y do not depend on Y. This may not be strictly true, but it is likely to work well enough in practice.

[27] It is wise to validate such assumptions by discussing the situation with underwriters, claims officers, and company management.

Development Formula 3 *Suppose there are real numbers $d \neq 0$ and x_0 such that $E(X|Y = y) = dy + x_0$ for all y. Then the function L defined above can be written as*

$$L(x) \quad = \quad Z\frac{x - x_0}{d} + (1 - Z)E(Y),$$

where

$$Z = \frac{VHM}{VHM + EVPV}.$$

We conclude that the least squares method can make sense even in cases where the development ratio varies with the caseload. It may be impossible in practice to determine the values of x_0 and of d, but we do not need these values to apply the least squares method.

A final example

In this section we will look at a fully worked out example based on real data that has been disguised slightly. Suppose we are given earned premium and incurred losses for a small book of business.

		Reported Loss ($000)				
AY	EP ($000)	12 mo.	24 mo.	36 mo.	48 mo.	60 mo.
1985	4260	102	104	209	650	847
1986	5563	0	543	1309	2443	3003
1987	7777	412	2310	3083	3358	4099
1988	8871	219	763	1637	1423	
1989	10465	969	4090	3801		
1990	11986	0	3467			
1991	12873	932				

Table 6: State CC, Line DD: Total limits losses.

One could use link ratios to develop these losses, but the least squares method is the better choice if we believe that the changes in the book of business are accurately reflected in the earned premiums. Because of the significant growth in volume, we will divide the losses by the premium to put the accident years on a more nearly equal basis. This gives us a triangle of reported loss ratios:

	Reported Loss Ratio				
AY	12 mo.	24 mo.	36 mo.	48 mo.	60 mo.
1985	0.024	0.024	0.049	0.153	0.199
1986	0.000	0.098	0.235	0.439	0.540
1987	0.053	0.297	0.396	0.432	0.527
1988	0.025	0.086	0.185	0.160	
1989	0.093	0.391	0.363		
1990	0.000	0.289			
1991	0.072				

Table 7: Reported loss ratios.

Unlike the data in Table 1, this data includes accident years at many different maturities. Following Clarke, we begin by developing the most mature years to ultimate. We then use the information obtained from those years to develop successively less mature years, ending with the 1991 year.

Losses may continue to develop after sixty months; to assume development stops at the end of the triangle is to assume the world ends at the horizon. For this line of business, we believe that losses will increase by an additional 10% from sixty months to ultimate. Based on this assumption, we estimate the ultimate loss ratios for accident years 1985, 1986, and 1987 to be 0.219, 0.594, and 0.580 respectively.

We next turn our attention to the 1988 year. We shall estimate the ultimate loss ratio for this year by looking at the relationship between the reported loss ratio at 48 months (our x value) and the ultimate loss ratio (our y value.) We base this relationship upon the observed 48-month and projected ultimate values for accident years 1985–1987. For these three years we have $\overline{x} = 0.341$, $\overline{y} = 0.464$, $\overline{x^2} = 0.134$, and $\overline{xy} = 0.181$ (it will be convenient to display these values directly beneath the 48-month column of the triangle.) This gives us $b = 1.301$, $a = 0.020$, and $y = 0.020 + (1.301)(0.160) = 0.229$ as the ultimate loss ratio for 1988.

AY	Reported Loss Ratio 12 mo.	24 mo.	36 mo.	48 mo.	60 mo.	Ultimate
1985	0.024	0.024	0.049	0.153	0.199	0.219
1986	0.000	0.098	0.235	0.439	0.540	0.594
1987	0.053	0.297	0.396	0.432	0.527	0.580
1988	0.025	0.086	0.185	*0.160*		*0.229*
1989	0.093	0.391	0.363			
1990	0.000	0.289				
1991	0.072					
$\overline{x}$				0.341		
$\overline{y}$				0.464		
$\overline{x^2}$				0.134		
$\overline{xy}$				0.181		
b				1.301		
a				0.020		
c				1.360		
Z				0.957		

Table 8: Estimation of the ultimate loss ratio for 1988.

We can also compute some supplemental values that, while not essential to our analysis, help us to understand the results. Our estimated ultimate loss ratio for 1988 is the weighted average of a link ratio estimate and a budgeted loss estimate. We have $c = \overline{y}/\overline{x} = 1.360$, giving a link ratio estimate of $y = cx = (1.360)(0.160) = 0.218$. For the budgeted loss estimate we have $y = \overline{y} = 0.464$. The credibility assigned to the link ratio estimate is $Z = b/c = 0.957$, giving a least squares estimate of $y = (0.957)(0.218) + (0.043)(0.464) = 0.229$. We expect a high credibility for the link ratio estimate here; at this stage of maturity, only a small portion of the variance in x arises from the reporting process. In fact, it is not uncommon for a to be negative in this part of the triangle; when this happens we set $Z = 1$ and use a simple link ratio estimate, ignoring the budgeted loss estimate.

We move next to the 1989 accident year, this time using the relationship between the reported loss ratio at 36 months and that at ultimate. We can now base the computation of a and b upon the values for 1985–1988, building on the work done in the previous step. When the ultimate loss ratio for 1989 has been determined, we continue working backwards to determine those for 1990 and 1991.

AY	Reported Loss Ratio 12 mo.	24 mo.	36 mo.	48 mo.	60 mo.	Ultimate
1985	0.024	0.024	0.049	0.153	0.199	0.219
1986	0.000	0.098	0.235	0.439	0.540	0.594
1987	0.053	0.297	0.396	0.432	0.527	0.580
1988	0.025	0.086	0.185	0.160		0.229
1989	0.093	0.391	*0.363*			*0.576*
1990	0.000	*0.289*				*0.537*
1991	*0.072*					*0.497*
$\overline{x}$	0.032	0.179	0.216	0.341		
$\overline{y}$	0.456	0.439	0.405	0.464		
$\overline{x^2}$	0.002	0.052	0.062	0.134		
$\overline{xy}$	0.016	0.096	0.106	0.181		
b	1.027	0.884	1.162	1.301		
a	0.422	0.281	0.154	0.020		
c	14.078	2.452	1.873	1.360		
Z	0.073	0.361	0.620	0.957		

Table 9: Estimation of ultimate loss ratios.

In this example Z increases steadily as the accident years mature and reported losses become more credible. The value of c decreases, as one would expect. Similarly, the value of a (which is what our estimate of ultimate losses would have been if no losses had been reported) decreases over time. These patterns provide a way to cross-check the work; data fluctuations can lead to unusual results, and one should not believe the analysis if it makes no sense.

In the final step we apply the ultimate loss ratios to earned premium to obtain ultimate losses.

AY	EP	Ultimate Loss Ratio	Ultimate Loss ($000)
1985	4260	0.219	932
1986	5563	0.594	3303
1987	7777	0.580	4509
1988	8871	0.229	2030
1989	10465	0.576	6028
1990	11986	0.537	6434
1991	12873	0.497	6396

Table 10: Computation of ultimate losses.

The procedure used in this section is easy to use and requires only commonly available data. It is less fragile than the link ratio method, as this example demonstrates—a link ratio analysis of this data would require a great deal of judgment in selecting the factors. In addition, we can present the analysis in a convenient tabular form which allows us to examine the assumptions that lie beneath it.

Conclusion

Least squares development as presented by Simon and Clarke is not only practically useful, but also justifiable on theoretical grounds. When random year to year fluctuations in loss experience are severe, it tends to produce more reasonable estimates of ultimate loss than the more familiar link ratio method, and it does so without requiring a great deal of additional data.

Least squares development is by no means a panacea. Like any method, it works best when it is used with a clear understanding of its limitations, and in conjunction with other appropriate methods. When there are significant exposure changes or other shifts in the loss history, one can go astray unless one makes the necessary corrections. Even under favorable circumstances the method is subject to the type of sampling errors that are always present when one estimates parameters from observed data.

Nevertheless, least squares development is a method that deserves a place in every actuary's toolbox. At my own company we now use this method in certain analysis situations; it can be most helpful in developing losses for small states, or for lines that are subject to serious fluctuations. This is especially true if one can use earned premium to adjust losses from past years to a level consistent with the current year.

Finally, the ideas presented here provide us with a conceptual framework that also helps us to understand more traditional development methods, and to see the relationships between them. Such an understanding must be our goal as we seek to deal intelligently with reserving and ratemaking issues.

Appendix—Proof of Development Formulas 2 and 3

Proof of Development Formula 2: As usual, $Var(X) = VHM + EVPV$. Since $E(X|Y = y) = dy$ by hypothesis, it follows that $VHM = Var_Y(E(X|Y = y)) = Var(dY) = d^2\,Var(Y)$. This means that $Cov(X,Y) = Cov(E_Y(X|Y),Y) = Cov(dY,Y) = d\,Var(Y) = VHM/d$.

The result now follows from Development Formula 1. We have

$$\begin{aligned} L(x) &= (x - E(X))\frac{Cov(X,Y)}{Var(X)} + E(Y) \\ &= (x - dE(Y))\frac{VHM/d}{VHM + EVPV} + E(Y) \\ &= Z\frac{x}{d} + (1 - Z)E(Y), \end{aligned}$$

where

$$Z = \frac{VHM}{VHM + EVPV}.$$

Proof of Development Formula 3: If we let $W = X - x_0$, then W and X share a common $EVPV$ and VHM. We can thus apply Development Formula 2 to W and Y to prove the formula.

LDF Curve-Fitting and Stochastic Reserving: A Maximum Likelihood Approach

David R. Clark, FCAS, MAAA

LDF Curve-Fitting and Stochastic Reserving: A Maximum Likelihood Approach

or

How to Increase Reserve Variability with Less Data

David R. Clark
American Re-Insurance

2003 Reserves Call Paper Program

Abstract

An application of Maximum Likelihood Estimation (MLE) theory is demonstrated for modeling the distribution of loss development based on data available in the common triangle format. This model is used to estimate future loss emergence, and the variability around that estimate. The value of using an exposure base to supplement the data in a development triangle is demonstrated as a means of reducing variability. Practical issues concerning estimation error and extrapolation are also discussed.

The author gratefully acknowledges the help and encouragement of the following people: Dick Currie, Jeff Davis, Leigh Halliwell, Don Mango, Dave Spiegler, and Chuck Thayer.

Introduction

Many papers have been written on the topic of statistical modeling of the loss reserving process. The present paper will focus on one such model, making use of the theory of maximum likelihood estimation (MLE) along with the common Loss Development Factor and Cape Cod techniques. After a review of the underlying theory, the bulk of this paper is devoted to a practical example showing how to make use of the techniques and how to interpret the output.

Before beginning a discussion of a formal model of loss reserving, it is worth re-stating the objectives in creating such a model.

The primary objective is to provide a tool that describes the loss emergence (either reporting or payment) phenomenon in simple mathematical terms as a guide to selecting amounts for carried reserves. Given the complexity of the insurance business, it should never be expected that a model will replace a knowledgeable analyst, but the model can become one key indication to assist them in selecting the reserve.

A secondary objective is to provide a means of estimating the range of possible outcomes around the "expected" reserve. The range of reserves is due to both random "process" variance, and the uncertainty in the estimate of the expected value.

From these objectives, we see that a statistical loss reserving model has two key elements:

- The expected amount of loss to emerge in some time period
- The distribution of actual emergence around the expected value

These two elements of our model will be described in detail in the first two sections of this paper. The full paper is outlined as follows:

Section 1:	Expected Loss Emergence
Section 2:	The Distribution of Actual Loss Emergence and Maximum Likelihood
Section 3:	Key Assumptions of the Model
Section 4:	A Practical Example
Section 5:	Comments and Conclusion

The practical example includes a demonstration of the reduction in variability possible from the use of an exposure base in the Cape Cod reserving method. Extensions of the model for estimating variability of the prospective loss projection or of discounted reserves are discussed more briefly.

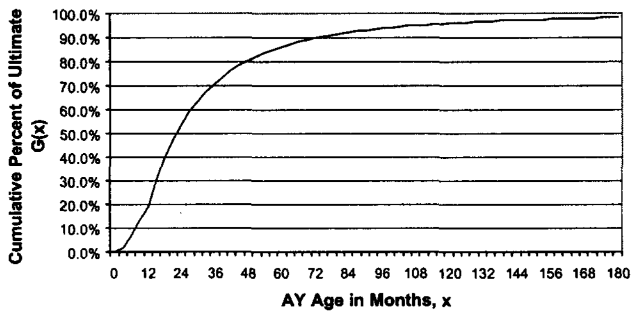
Most of the material presented in this paper makes use of maximum likelihood theory that has already been described more rigorously elsewhere. The mathematics presented here is sufficient for the reader to reproduce the calculations in the examples given, but the focus will be on practical issues rather than on the statistical theory itself.

Section 1: Expected Loss Emergence

Our model will estimate the expected amount of loss to emerge based on a) an estimate of the ultimate lo ss by year, and b) an estimate of the pattern of loss emergence.

For the expected emergence pattern, we need a pattern that moves from 0 to 100% as time moves from 0 to 8. For our model, we will assume that this pattern is described using the form of a cumulative distribution function[1] (CDF), since a library of such curves is readily available.

$$G(x) \;=\; 1/LDF_x \quad = \text{cumulative \% reported (or paid) as of time x}$$



We will assume that the time index "x" represents the time from the "average" accident date to the evaluation date. The details for approximating different exposure periods (e.g., accident year versus policy year) are given in Appendix B.

For convenience, the model will include two familiar curve forms: Weibull and Loglogistic. Each of these curve forms can be parameterized with a scale θ and a shape ω ("warp"). The Loglogistic curve is familiar to many actuaries under the name "inverse

[1] We are using the form of the distribution function, but do not mean to imply any probabilistic model. The paper by Weissner [9] makes the report lag itself the random variable. By contrast, the loss dollars will be the random variable in our application.

power" (see Sherman[2] [8]), and will be considered the benchmark result. The Weibull will generally provide a smaller "tail" factor than the Loglogistic.

The Loglogistic curve has the form:

$$G(x \mid \omega, \theta) = \frac{x^{\omega}}{x^{\omega} + \theta^{\omega}} \qquad LDF_x = 1 + \theta^{\omega} \cdot x^{-\omega}$$

The Weibull curve has the form:

$$G(x \mid \omega, \theta) = 1 - \exp\left(-(x/\theta)^{\omega}\right)$$

In using these curve forms, we are assuming that the expected loss emergence will move from 0% to 100% in a strictly increasing pattern. The model will still work if some actual points show decreasing losses, but if there is real expected negative development (e.g., lines of business with significant salvage recoveries) then a different model should be used.

There are several advantages to using parameterized curves to describe the expected emergence pattern. First, the estimation problem is simplified because we only need to estimate the two parameters. Second, we can use data that is not strictly from a triangle with evenly spaced evaluation dates – such as the frequent case in which the latest diagonal is only nine months from the second latest diagonal. Third, the final indicated pattern is a smooth curve and does not follow every random movement in the historical age-to-age factors.

The next step in estimating the amount of loss emergence by period is to apply the emergence pattern G(x), to an estimate of the ultimate loss by accident year.

Our model will base the estimate of the ultimate loss by year on one of two methods: either the LDF or the Cape Cod method. The LDF method assumes that the ultimate loss

[2] Sherman actually applies the inverse power curve to the link ratios between ages. Our model will apply this curve to the age-to-ultimate pattern.

amount in each accident year is independent of the losses in other years. The Cape Cod method assumes that there is a known relationship between the amount of ultimate loss expected in each of the years in the historical period, and that this relationship is identified by an exposure base. The exposure base is usually onlevel premium, but can be any other index (such as sales or payroll), which is reasonably assumed to be proportional to expected loss.

The expected loss for a given period will be denoted:

$\mu_{AY;x,y}$ = expected incremental loss dollars in accident year AY between ages x and y

Then the two methods for the expected loss emergence are:

Method #1: "Cape Cod"

$$\mu_{AY;x,y} = \text{Premium}_{AY} \cdot ELR \cdot [G(y \mid \omega,\theta) - G(x \mid \omega,\theta)]$$

Three parameters: $ELR,\ \omega,\ \theta$

Method #2: "LDF"

$$\mu_{AY;x,y} = ULT_{AY} \cdot [G(y \mid \omega,\theta) - G(x \mid \omega,\theta)]$$

n+2 Parameters: n Accident Years (one ULT for each AY) + $\omega,\ \theta$

While both of these methods are available for use in estimating reserves, Method #1 will generally be preferred. Because we are working with data summarized into annual blocks as a development triangle, there will be relatively few data points included in the model (one data point for each "cell" in the triangle). There is a real problem with overparameterization when the LDF method is used.

For example, if we have a triangle for ten accident years then we have provided the model with 55 data points. The Cape Cod method requires estimation of 3 parameters, but the LDF method requires estimation of 12 parameters.

The Cape Cod method may have somewhat higher process variance estimated, but will usually produce a significantly smaller estimation error. This is the value of the information in the exposure base provided by the user[3]. In short: *the more information that we can give to the model, the smaller the reserve variability due to estimation error.*

The fact that variance can be reduced by incorporating more information into a reserve analysis is, of course, the point of our ironic subtitle: <u>How to Increase Reserve Variability with Less Data</u>. The point is obvious, but also easy to overlook. The reduction in variability is important even to those who do not explicitly calculate reserve ranges because it still guides us towards better estimation methods: lower variance implies a better reserve estimate.

[3] Halliwell [2] provides additional arguments for the use of an exposure index. See especially pages 441 - 443.

Section 2: The Distribution of Actual Loss Emergence and Maximum Likelihood

Having defined the model for the expected loss emergence, we need to estimate the "best" parameters for that model and, as a secondary goal, estimate the variance around the expected value. Both of these steps will be accomplished making use of maximum likelihood theory.

The variance will be estimated in two pieces: process variance (the "random" amount) and parameter variance (the uncertainty in our estimator).

2.1 Process Variance

The curve $G(x \mid \omega,\theta)$ represents the <u>expected</u> loss emergence pattern. The <u>actual</u> loss emergence will have a distribution around this expectation.

We assume that the loss in any period has a constant ratio of variance/mean[4]:

$$\frac{Variance}{Mean} = \sigma^2 \approx \frac{1}{n-p} \cdot \sum_{AY,t}^{n} \frac{\left(c_{AY,t} - \mu_{AY,t}\right)^2}{\mu_{AY,t}}$$

where p = # of parameters

$c_{AY,t}$ = actual incremental loss emergence

$\mu_{AY,t}$ = expected incremental loss emergence

(this is recognized as being equivalent to a chi-square error term)

For estimating the parameters of our model, we will further assume that the actual incremental loss emergence "c" follows an over-dispersed Poisson distribution. That is, the loss dollars will be a Poisson random variable times a scaling factor equal to σ^2.

---

[4] This assumption will be tested by analysis of residuals in our example.

Standard Poisson: $\Pr(x) = \dfrac{\lambda^x \cdot e^{-\lambda}}{x!}$ $E[x] = Var(x) = \lambda$

Actual Loss: $c = x \cdot \sigma^2$ $\Pr(c) = \dfrac{\lambda^{c/\sigma^2} \cdot e^{-\lambda}}{(c/\sigma^2)!}$ $E[c] = \lambda \cdot \sigma^2 = \mu$

$Var(c) = \lambda \cdot \sigma^4 = \mu \cdot \sigma^2$

The "over-dispersed Poisson" sounds strange when it is first encountered, but it quickly proves to have some key advantages. First, inclusion of the scaling factor allows us to match the first and second moments of any distribution, which gives the model a high degree of flexibility. Second, maximum likelihood estimation exactly produces the LDF and Cape Cod estimates of ultimate, so the results can be presented in a format familiar to reserving actuaries.

The fact that the distribution of ultimate reserves is approximated by a discretized curve should not be cause for concern. The scale factor σ^2 is generally small compared to the mean, so little precision is lost. Also, the use of a discrete distribution allows for a mass point at zero, representing the cases in which no change in loss is seen in a given development increment.

Finally, we should remember that this maximum likelihood method is intended to produce the mean and variance of the distribution of reserves. Having estimated those two numbers, we are still free to switch to a different distribution form when the results are used in other applications.

2.2 The Likelihood Function – Finding the "Best" Parameters

The likelihood function is:

$$\text{Likelihood} = \prod_i \Pr(c_i) = \prod_i \frac{\lambda_i^{c_i/\sigma^2} \cdot e^{-\lambda_i}}{(c_i/\sigma^2)!} = \prod_i \frac{\left(\mu_i/\sigma^2\right)^{c_i/\sigma^2} \cdot e^{-\mu_i/\sigma^2}}{(c_i/\sigma^2)!}$$

This can be maximized using the logarithm of the likelihood function:

$$\text{LogLikelihood} = \sum_i \left(c_i/\sigma^2\right)\cdot\ln\left(\mu_i/\sigma^2\right) - \mu_i/\sigma^2 - \ln\left(\left(c_i/\sigma^2\right)!\right)$$

Which is equivalent to maximizing:

$$\boxed{\ell \;=\; \sum_i c_i \cdot \ln(\mu_i) - \mu_i} \qquad \text{if } \sigma^2 \text{ is assumed to be known}$$

Maximum likelihood estimators of the parameters are found by setting the first derivatives of the loglikelihood function ℓ equal to zero:

$$\frac{\partial \ell}{\partial ELR} \;=\; \frac{\partial \ell}{\partial \theta} \;=\; \frac{\partial \ell}{\partial \omega} \;=\; 0$$

For "Model #1: Cape Cod", the loglikelihood function becomes:

$$\ell \;=\; \sum_{i,t} \left(c_{i,t} \cdot \ln\left(ELR \cdot P_i \cdot \left[G(x_t) - G(x_{t-1})\right]\right) - ELR \cdot P_i \cdot \left[G(x_t) - G(x_{t-1})\right]\right)$$

where $c_{i,t}$ = actual loss in accident year i, development period t

P_i = Premium for accident year i

x_{t-1} = beginning age for development period t

x_t = ending age for development period t

$$\frac{\partial \ell}{\partial ELR} \;=\; \sum_{i,t}\left(\frac{c_{i,t}}{ELR} - P_i \cdot \left[G(x_t) - G(x_{t-1})\right]\right)$$

For $\dfrac{\partial \ell}{\partial ELR} = 0, \qquad ELR = \dfrac{\sum_{i,t} c_{i,t}}{\sum_{i,t} P_i \cdot [G(x_t) - G(x_{t-1})]}$

The MLE estimate for ELR is therefore equivalent to the "Cape Cod" Ultimate. It can be set based on θ and ω, and so reduce the problem to be solved to two parameters instead of three.

For "Model #2: LDF", the loglikelihood function becomes:

$$\ell = \sum_{i,t} \left(c_{i,t} \cdot \ln\left(ULT_i \cdot [G(x_t) - G(x_{t-1})]\right) - ULT_i \cdot [G(x_t) - G(x_{t-1})] \right)$$

$$\frac{\partial \ell}{\partial ULT_i} = \sum_t \left(\frac{c_{i,t}}{ULT_i} - [G(x_t) - G(x_{t-1})] \right)$$

For $\dfrac{\partial \ell}{\partial ULT_i} = 0, \qquad ULT_i = \dfrac{\sum_t c_{i,t}}{\sum_t [G(x_t) - G(x_{t-1})]}$

The MLE estimate for each ULT_i is therefore equivalent to the "LDF Ultimate"[5]. It can also be set based on θ and ω, and to again reduce the problem to be solved to two parameters instead of $n+2$.

A final comment worth noting is that the maximum loglikelihood function never takes the logarithm of the actual incremental development $c_{i,t}$. The model will work even if some of these amounts are zero or negative.

[5] See Mack [5], Appendix A, for a further discussion of this relationship.

2.3 *Parameter Variance*[6]

The second step is to find the variance in the estimate of the parameters. This is done based on the Rao-Cramer approximation, using the second derivative information matrix I, and is commonly called the "Delta Method" (c.f. Klugman, et al [3], page 67).

The second derivative information matrix for the "Cape Cod Method" is 3x3 and assumes the same ELR for all accident years:

$$I = \begin{bmatrix} \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ELR^2} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ELR \, \partial \omega} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ELR \, \partial ?} \\ \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial \omega \, \partial ELR} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial \omega^2} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial \omega \, \partial ?} \\ \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ? \, \partial ELR} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ? \, \partial \omega} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ?^2} \end{bmatrix}$$

The covariance matrix is calculated using the inverse of the Information matrix:

$$\Sigma = \begin{bmatrix} Var(ELR) & Cov(ELR,\omega) & Cov(ELR,\theta) \\ Cov(\omega, ELR) & Var(\omega) & Cov(\omega,\theta) \\ Cov(\theta, ELR) & Cov(\theta,\omega) & Var(\theta) \end{bmatrix} \geq -\sigma^2 \cdot I^{-1}$$

The scale factor σ^2 is again estimated as above:

$$\sigma^2 \approx \frac{1}{n-p} \cdot \sum_{AY,t}^{n} \frac{\left(c_{AY,t} - \hat{\mu}_{AY,t}\right)^2}{\hat{\mu}_{AY,t}}$$

The second derivative matrix for "LDF Method" is (n+2)x(n+2) and assumes that there is a different ULT for each accident year. The information matrix, I, is given as:

[6] To be precise, we are calculating the variance in the estimator of the parameter; the parameter itself does not have any variance. Nonetheless, we will retain the term "parameter variance" as shorthand.

$$
\left[\begin{array}{cccc|cc}
\sum_t \frac{\partial^2 \ell_{1,t}}{\partial ULT_1^2} & 0 & \cdots & 0 & \sum_t \frac{\partial^2 \ell_{1,t}}{\partial ULT_1\, \partial\omega} & \sum_t \frac{\partial^2 \ell_{1,t}}{\partial ULT_1\, \partial ?} \\
0 & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial ULT_2^2} & \cdots & 0 & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial ULT_2\, \partial\omega} & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial ULT_2\, \partial ?} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial ULT_n^2} & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial ULT_n\, \partial\omega} & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial ULT_n\, \partial ?} \\
\hline
\sum_t \frac{\partial^2 \ell_{1,t}}{\partial\omega\, \partial ULT_1} & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial\omega\, \partial ULT_2} & \cdots & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial\omega\, \partial ULT_n} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial\omega^2} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial\omega\, \partial ?} \\
\sum_t \frac{\partial^2 \ell_{1,t}}{\partial ?\, \partial ULT_1} & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial ?\, \partial ULT_2} & \cdots & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial ?\, \partial ULT_n} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ?\, \partial\omega} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ?^2}
\end{array}\right]
$$

The covariance matrix Σ is again calculated using the inverse of the Information matrix, but for the LDF Method this matrix is larger.

2.4 The Variance of the Reserves

The final step is to estimate the variance in the reserves. The variance is broken into two pieces: the process variances and the estimation error (loosely "parameter variance"). For an estimate of loss reserves R for a given period $\mu_{AY;x,y}$, or group of periods $\sum \mu_{AY;x,y}$, the process variance is given by:

Process Variance of R: $\quad \sigma^2 \cdot \sum \mu_{AY;x,y}$

The estimation error makes use of the covariance matrix Σ calculated above:

Parameter Variance of R: $\quad Var(E[R]) = (\partial R)' \cdot \Sigma \cdot (\partial R)$

where

$$\partial R = \left\langle \frac{\partial R}{\partial ELR}, \frac{\partial R}{\partial \theta}, \frac{\partial R}{\partial \omega} \right\rangle \qquad \text{or} \qquad \partial R = \left\langle \left\{ \frac{\partial R}{\partial ULT_i} \right\}_{i=1}^{n}, \frac{\partial R}{\partial \theta}, \frac{\partial R}{\partial \omega} \right\rangle$$

The future reserve R, under the Cape Cod method is given by:

$$\text{Reserve:} \qquad R = \sum \text{Premium}_i \cdot ELR \cdot \left(G(y_i) - G(x_i)\right)$$

The derivatives needed are then easily calculated:

$$\frac{\partial R}{\partial ELR} = \sum \text{Premium}_i \cdot \left(G(y_i) - G(x_i)\right)$$

$$\frac{\partial R}{\partial \theta} = \sum \text{Premium}_i \cdot ELR \cdot \left(\frac{\partial G(y_i)}{\partial \theta} - \frac{\partial G(x_i)}{\partial \theta}\right)$$

$$\frac{\partial R}{\partial \omega} = \sum \text{Premium}_i \cdot ELR \cdot \left(\frac{\partial G(y_i)}{\partial \omega} - \frac{\partial G(x_i)}{\partial \omega}\right)$$

For the LDF Method, let $\text{Premium}_i = 1$ and $ELR = ULT_i$.

All of the mathematics needed for the estimate of the process and parameter variance is provided in Appendix A. For the two curve forms used, all of the derivatives are calculated analytically, without the need for numerical approximations.

Section 3: Key Assumptions of this Model

- Incremental losses are independent and identically distributed (iid)

The assumption that all observed points are independent and identically distributed is the famous "iid" of classical statistics. In introductory textbooks this is often illustrated by the problem of estimating the proportion of red and black balls in an urn based on having "randomly" selected a sample from the urn. The "independence" assumption is that the balls are shaken up after each draw, so that we do not always pull out the same ball each time. The "identically distributed" assumption is that we are always taking the sample from the same urn.

The "independence" assumption in the reserving context is that one period does not affect the surrounding periods. This is a tenuous assumption but will be tested using residual analysis. There may in fact be positive correlation if all periods are equally impacted by a change in loss inflation. There may also be negative correlation if a large settlement in one period replaces a stream of payments in later periods.

The "identically distributed" assumption is also difficult to justify on first principles. We are assuming that the emergence pattern is the same for all accident years; which is clearly a gross simplification from even a rudimentary understanding of insurance phenomenon. Different risks and mix of business would have been written in each historical period, and subject to different claims handling and settlement strategies. Nonetheless, a parsimonious model requires this simplification.

- The Variance/Mean Scale Parameter σ^2 is fixed and known

In rigorous maximum likelihood theory, the variance/mean scale parameter σ^2 should be estimated simultaneously with the other model parameters, and the variance around its estimate included in our covariance matrix.

Unfortunately, including the scale parameter in the curve-fitting procedure leads to mathematics that quickly becomes intractable. Treating the scale parameter as fixed and known is an approximation made for convenience in the calculation, and the results are sometimes called "quasi-likelihood estimators". McCullough & Nelder [7] give support for the approximation that we are using.

In effect, we are ignoring the variance on the variance.

In classical statistics, we usually relax this assumption (e.g., in hypothesis testing) by using the Student-T distribution instead of the Normal distribution. Rodney Kreps' paper [4] provides additional discussion on how reserve ranges could increase when this additional source of variability is considered.

- Variance estimates are based on an approximation to the Rao-Cramer lower bound.

  The estimate of variance based on the information matrix is only exact when we are using linear functions. In the case of non-linear functions, including our model, the variance estimate is a Rao-Cramer lower bound.

  Technically, the Rao-Cramer lower bound is based on the true expected values of the second derivative matrix. Since we are using approximations that plug in the estimated values of the parameters, the result is sometimes called the "observed" information matrix rather than the "expected" information matrix. Again, this is a limitation common to many statistical models and is due to the fact that we do not know the true parameters.

All of the key assumptions listed above need to be kept in mind by the user of a stochastic reserving model. In general, they imply that there is potential for more variability in future loss emergence than the model itself produces.

Such limitations should not lead the user, or any of the recipients of the output, to disregard the results. We simply want to be clear about what sources of variability we are able to measure and what sources cannot be measured. That is a distinction that should not be lost.

Section 4: A Practical Example

4.1 The LDF Method

For the first part of this example, we will use the "LDF Method" (referred to above as "Method 2"). The improvements in the model by moving to the Cape Cod method will be apparent as the numbers are calculated.

The triangle used in this example is taken from the 1993 Thomas Mack paper [6]. The accident years have been added to make the display appear more familiar.

	12	24	36	48	60	72	84	96	108	120
1991	357,848	1,124,788	1,735,330	2,182,708	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
1992	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	
1993	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315		
1994	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268			
1995	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311				
1996	396,132	1,333,217	2,180,715	2,985,752	3,691,712					
1997	440,832	1,288,463	2,419,861	3,483,130						
1998	359,480	1,421,128	2,864,498							
1999	376,686	1,363,294								
2000	344,014									

The incremental triangle, calculated by taking differences between cells in each accident year, is given by:

	12	24	36	48	60	72	84	96	108	120
1991	357,848	766,940	610,542	447,378	562,888	574,398	146,342	139,950	227,229	67,948
1992	352,118	884,021	933,894	1,183,289	445,745	320,996	527,804	266,172	425,046	
1993	290,507	1,001,799	926,219	1,016,654	750,816	146,923	495,992	280,405		
1994	310,608	1,108,250	776,189	1,562,400	272,482	352,053	206,286			
1995	443,160	693,190	991,983	769,488	504,851	470,639				
1996	396,132	937,085	847,498	805,037	705,960					
1997	440,832	847,631	1,131,398	1,063,269						
1998	359,480	1,061,648	1,443,370							
1999	376,686	986,608								
2000	344,014									

This incremental triangle is actually better arranged as a table of values, rather than in the familiar triangular format (see Table 1.1). In the tabular format, the column labeled "Increment" is the value that we will be approximating with the expression

$$\mu_{AY;x,y} = ULT_{AY} \cdot [G(y \mid \omega,\theta) - G(x \mid \omega,\theta)].$$

The x and y values are the "From" and "To" dates.

Before calculating the fitted values, it is worth showing the flexibility in this format. First, if we have only the latest three evaluations of the triangle, we can still use this method directly.

The original triangle becomes:

	12	24	36	48	60	72	84	96	108	120
1991								3,606,286	3,833,515	3,901,463
1992							4,647,867	4,914,039	5,339,085	
1993						4,132,918	4,628,910	4,909,315		
1994					4,029,929	4,381,982	4,588,268			
1995				2,897,821	3,402,672	3,873,311				
1996			2,180,715	2,985,752	3,691,712					
1997		1,288,463	2,419,861	3,483,130						
1998	359,480	1,421,128	2,864,498							
1999	376,686	1,363,294								
2000	344,014									

and the incremental triangle is:

	12	24	36	48	60	72	84	96	108	120
1991								3,606,286	227,229	67,948
1992							4,647,867	266,172	425,046	
1993						4,132,918	495,992	280,405		
1994					4,029,929	352,053	206,286			
1995				2,897,821	504,851	470,639				
1996			2,180,715	805,037	705,960					
1997		1,288,463	1,131,398	1,063,269						
1998	359,480	1,061,648	1,443,370							
1999	376,686	986,608								
2000	344,014									

The tabular format then collapses from 55 rows down to 27 rows, as shown in Table 1.2.

Another common difficulty in working with development triangles is the use of irregular evaluation periods. For example, we may have accident years evaluated at each year-end

- producing ages 12, 24, 36, etc – but the most recent diagonal is only available as of the end of the third quarter (ages 9, 21, 33, etc). This is put into the tabular format by simply changing the evaluation age fields ("Diag Age") as shown in Table 1.3.

Returning to the original triangle, we calculate the fitted values for a set of parameters ULT_{AY}, ω, θ and the MLE term to be maximized.

Fitted Value: $$\mu_{AY;x,y} = ULT_{AY} \cdot [G(y \mid \omega,\theta) - G(x \mid \omega,\theta)]$$

MLE Term: $$c_{AY;x,y} \cdot \ln(\mu_{AY;x,y}) - \mu_{AY;x,y}$$

In Table 1.4, these numbers are shown as additional columns. These values also have the desired unbiased property that the sum of the actual incremental dollars $c_{AY;x,y}$ equals the sum of the fitted values $\hat{\mu}_{AY;x,y}$.

The fitted parameters for the Loglogistic growth curve are:

ω	1.434294
θ	48.6249

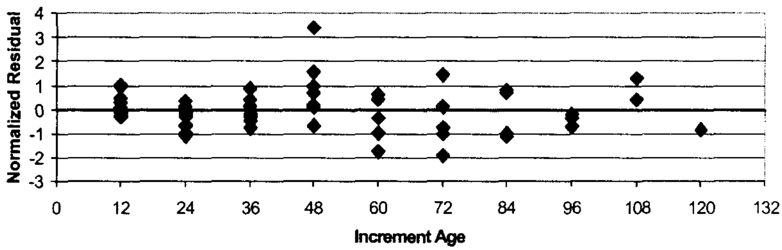
The fitted parameters are found by iteration, which can easily be accomplished in the statistics capabilities of most software packages. Once the data has been arranged in the tabular format, the curve-fitting can even be done in a spreadsheet.

The scale parameter σ^2 is also easily calculated. We recall that the form of this calculation is the same as a Chi-Square statistic, with 43 degrees of freedom (55 data points minus 12 parameters). The resulting σ^2 is 65,029. This scale factor may be thought of as the size of the discrete intervals for the over-dispersed Poisson, but is better thought of simply as the process variance-to-mean ratio. As such, we can calculate the process variance of the total reserve, or any sub-segment of the reserve, by just multiplying by 65,029.

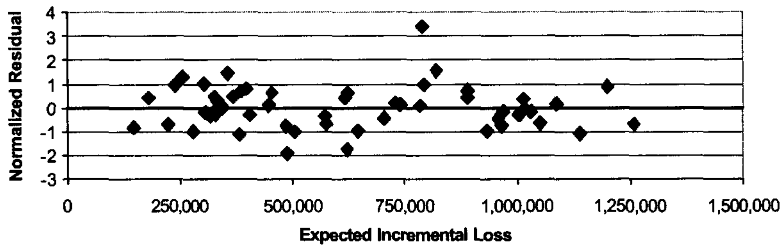
The scale factor σ^2 is also useful for a review of the model residuals (error terms).

$$\text{Normalized Residual:} \qquad r_{AY;x,y} = \frac{\left(c_{AY;x,y} - \hat{\mu}_{AY;x,y}\right)}{\sqrt{\sigma^2 \cdot \hat{\mu}_{AY;x,y}}}$$

The residuals can be plotted in various ways in order to test the assumptions in the model. The graph below shows the residuals plotted against the increment of loss emergence. We would hope that the residuals would be randomly scattered around the zero line for all of the ages, and that the amount of variability would be roughly constant. The graph below tells us that the curve form is perhaps not perfect for the early 12 and 24 points, but the pattern is not enough to reject the model outright.



A second residual plot of the residuals against the expected loss in each increment (the fitted values) is shown below. This graph is useful as a check on the assumption that the variance/mean ratio is constant. If the variance/mean ratio were not constant, then we would expect to see the residuals much closer to the zero line at one end of the graph.



The residuals can also be plotted against the accident year, the calendar year of emergence (to test diagonal effects), or any other variable of interest. The desired outcome is always that the residuals appear to be randomly scattered around the zero line. Any noticeable pattern or autocorrelation is an indication that the some of the model assumptions are incorrect.

Having solved for the parameters ω and θ, and the derived ultimates by year, we can estimate the needed reserves.

Accident Year	Reported Losses	Age at 12/31/2000	Average Age (x)	Growth Function	Fitted LDF	Ultimate Losses	Estimated Reserves
1991	3,901,463	120	114	77.24%	1.2946	5,050,867	1,149,404
1992	5,339,085	108	102	74.32%	1.3456	7,184,079	1,844,994
1993	4,909,315	96	90	70.75%	1.4135	6,939,399	2,030,084
1994	4,588,268	84	78	66.32%	1.5077	6,917,862	2,329,594
1995	3,873,311	72	66	60.78%	1.6452	6,372,348	2,499,037
1996	3,691,712	60	54	53.75%	1.8604	6,867,980	3,176,268
1997	3,483,130	48	42	44.77%	2.2338	7,780,515	4,297,385
1998	2,864,498	36	30	33.34%	2.9991	8,590,793	5,726,295
1999	1,363,294	24	18	19.38%	5.1593	7,033,659	5,670,365
2000	344,014	12	6	4.74%	21.1073	7,261,205	6,917,191
Total	34,358,090					69,998,708	35,640,618

From this initial calculation, we can quickly see the impact of the extrapolated "tail" factor. Our loss development data only includes ten years of development (out to age 120 months), but the growth curve extrapolates the losses to full ultimate. From this data, the Loglogistic curve estimates that only 77.24% of ultimate loss has emerged as of ten years.

Extrapolation should always be used cautiously. For practical purposes, we may want to rely on the extrapolation only out to some finite point – an additional ten years say.

Accident Year	Reported Losses	Age at 12/31/2000	Average Age (x)	Growth Function	Fitted LDF	Truncated LDF	Losses at 240 mo	Estimated Reserves
		240	234	90.50%	1.1050	1.0000		
1991	3,901,463	120	114	77.24%	1.2946	1.1716	4,570,810	669,347
1992	5,339,085	108	102	74.32%	1.3456	1.2177	6,501,273	1,162,188
1993	4,909,315	96	90	70.75%	1.4135	1.2792	6,279,848	1,370,533
1994	4,588,268	84	78	66.32%	1.5077	1.3644	6,260,358	1,672,090
1995	3,873,311	72	66	60.78%	1.6452	1.4888	5,766,692	1,893,381
1996	3,691,712	60	54	53.75%	1.8604	1.6836	6,215,217	2,523,505
1997	3,483,130	48	42	44.77%	2.2338	2.0215	7,041,021	3,557,891
1998	2,864,498	36	30	33.34%	2.9991	2.7140	7,774,286	4,909,788
1999	1,363,294	24	18	19.38%	5.1593	4.6689	6,365,149	5,001,855
2000	344,014	12	6	4.74%	21.1073	19.1012	6,571,068	6,227,054
Total	34,358,090						63,345,723	28,987,633

As noted above, the process variance for the estimated reserve of 28,987,633 is found by multiplying by the variance-to-mean ratio of 65,029. The process standard deviation around our reserve is therefore 1,372,966 for a coefficient of variation (CV = SD/mean) of about 4.7%.

As an alternative to truncating the tail factor at a selected point, such as age 240, we could make use of a growth curve that typically has a lighter "tail". The mathematics for the Weibull curve is provided for this purpose. An example including a fit of the Weibull curve is shown below.

Accident Year	Reported Losses	Age at 12/31/2000	Average Age (x)	Growth Function	Weibull LDF	Ultimate Losses	Estimated Reserves
1991	3,901,463	120	114	95.01%	1.0525	4,106,189	204,726
1992	5,339,085	108	102	92.54%	1.0806	5,769,409	430,324
1993	4,909,315	96	90	89.00%	1.1237	5,516,376	607,061
1994	4,588,268	84	78	84.01%	1.1904	5,461,745	873,477
1995	3,873,311	72	66	77.14%	1.2963	5,020,847	1,147,536
1996	3,691,712	60	54	67.95%	1.4717	5,433,242	1,741,530
1997	3,483,130	48	42	56.01%	1.7853	6,218,284	2,735,154
1998	2,864,498	36	30	41.19%	2.4277	6,954,204	4,089,706
1999	1,363,294	24	18	23.94%	4.1764	5,693,693	4,330,399
2000	344,014	12	6	6.37%	15.6937	5,398,863	5,054,849
Total	34,358,090					55,572,851	21,214,761

The fitted Weibull parameters θ and ω are 48.88453 and 1.296906, respectively. The lower "tail" factor of 1.0525 (instead of 1.2946 for the Loglogistic) may be more in line with the actuary's expectation for casualty business. The difference between the two curve forms also highlights the danger in relying on a purely mechanical extrapolation formula. The selection of a truncation point is an effective way of reducing the reliance on the extrapolation when the thicker-tailed Loglogistic is used.

The next step is our estimate of the parameter variance.

The parameter variance calculation is more involved than what was needed for process variance. As discussed in Section 2.3, we need to first evaluate the Information Matrix, which contains the second derivatives with respect to all of the model parameters, and so is a 12x12 matrix. The mathematics for all of these calculations is given in Appendix A, and is not difficult to program in most software. For purposes of this example, we will simply show the resulting variances:

Accident Year	Reported Losses	Estimated Reserves	Process Std Dev	CV	Parameter Std Dev	CV	Total Std Dev	CV
1991	3,901,463	669,347	208,631	31.2%	158,088	23.6%	261,761	39.1%
1992	5,339,085	1,162,188	274,911	23.7%	257,205	22.1%	376,471	32.4%
1993	4,909,315	1,370,533	298,537	21.8%	298,628	21.8%	422,260	30.8%
1994	4,588,268	1,672,090	329,749	19.7%	356,827	21.3%	485,860	29.1%
1995	3,873,311	1,893,381	350,891	18.5%	401,416	21.2%	533,160	28.2%
1996	3,691,712	2,523,505	405,094	16.1%	518,226	20.5%	657,768	26.1%
1997	3,483,130	3,557,891	481,005	13.5%	704,523	19.8%	853,064	24.0%
1998	2,864,498	4,909,788	565,047	11.5%	968,806	19.7%	1,121,545	22.8%
1999	1,363,294	5,001,855	570,321	11.4%	1,227,880	24.5%	1,353,867	27.1%
2000	344,014	6,227,054	636,348	10.2%	2,838,890	45.6%	2,909,336	46.7%
Total	34,358,090	28,987,633	1,372,966	4.7%	4,688,826	16.2%	4,885,707	16.9%

From this table, one conclusion should be readily apparent: the parameter variance component is much more significant than the process variance. The chief reason for this is that we have overparameterization of our model; that is, the available 55 data points are really not sufficient to estimate the 12 parameters of the model. The 1994 Zehnwirth paper ([10], p. 512f) gives a helpful discussion of the dangers of overparameterization.

The main problem is that we are estimating the ultimate loss for each accident year independently from the ultimate losses in the other accident years. In effect, we are saying that knowing the ultimate loss for accident year 1999 provides no information about the ultimate loss for accident year 2000. As such, our model is fitting to what may just be "noise" in the differences from one year to the next.

This conclusion is unsettling, because it indicates a high level of uncertainty not just in our maximum likelihood model, but in the chain-ladder LDF method in general.

4.2 The Cape Cod Method

A natural alternative to the LDF Method is the Cape Cod method. In order to move on to this method, we need to supplement the loss development triangle with an exposure base that is believed to be proportional to ultimate expected losses by accident year. A natural candidate for the exposure base is onlevel premium – premium that has been adjusted to a common level of rate per exposure.

Unadjusted historical premium could be used for this exposure base, but the impact of the market cycle on premium is likely to distort the results. We prefer onlevel premium so that the assumption of a constant expected loss ratio (ELR) across all accident years is reasonable.

A further refinement would include an adjustment for loss trend net of exposure trend, so that all years are at the same cost level as well as rate level.

There may be other candidates for the exposure index: sometimes the original loss projections by year are available; the use of estimated claim counts has also been suggested. In practice, even a judgmentally selected index may be used.

For the example in the Mack paper, no exposure base was supplied. For this exercise, we will use a simplifying assumption that premium was \$10,000,000 in 1991 and increased by \$400,000 each subsequent year.

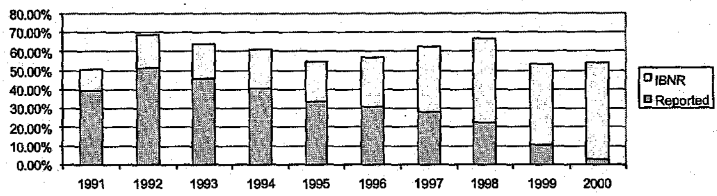
The tabular format of our loss data is shown in Table 2.1. This is very similar to the format used for the LDF Method but instead of the "AY Total" column (latest diagonal), we display the onlevel premium for each accident year. The expected ultimate loss by year is calculated as the ELR multiplied by the onlevel premium.

Accident Year	Onlevel Premium	Age at 12/31/2000	Average Age (x)	Growth Function	Premium x Growth Func	Reported Losses	Ultimate Loss Ratio
1991	10,000,000	120	114	77.76%	7,775,733	3,901,463	50.17%
1992	10,400,000	108	102	74.85%	7,784,279	5,339,085	68.59%
1993	10,800,000	96	90	71.29%	7,699,022	4,909,315	63.77%
1994	11,200,000	84	78	66.87%	7,489,209	4,588,268	61.27%
1995	11,600,000	72	66	61.31%	7,112,024	3,873,311	54.46%
1996	12,000,000	60	54	54.24%	6,508,439	3,691,712	56.72%
1997	12,400,000	48	42	45.17%	5,600,712	3,483,130	62.19%
1998	12,800,000	36	30	33.60%	4,301,252	2,864,498	66.60%
1999	13,200,000	24	18	19.46%	2,568,496	1,363,294	53.08%
2000	13,600,000	12	6	4.69%	638,334	344,014	53.89%
Total	118,000,000				57,477,500	34,358,090	59.78%

The Loglogistic parameters are again solved for iteratively in order to maximize the value of the log-likelihood function in Table 2.1. The resulting parameters are similar to those produced by the LDF method.

ω	1.447634
θ	48.0205

One check that should be made on the data before we proceed with the reserve estimate is a quick test on the assumption that the ELR is constant over all accident years. This is best done with a graph of the estimated ultimate loss ratios:



From this graph, the ultimate loss ratios by year do not appear to be following a strong autocorrelation pattern, or other unexplained trends. If we had observed an increasing or decreasing pattern, then there could be a concern of bias introduced in our reserve estimate.

The following calculation shows the method of estimating reserves out to the 240 month evaluation point. As in the LDF method, this truncation point is used in order avoid undue reliance on a mechanical extrapolation formula.

The Cape Cod method works much like the more familiar Bornhuetter-Ferguson formula. Estimated reserves are calculated as a percent of the premium and the calculated expected loss ratio (ELR).

Accident Year	Onlevel Premium	Age at 12/31/2000	Average Age (x)	Growth Function	90.83% minus Growth Func	Premium x ELR	Estimated Reserves
		240	234	90.83%			
1991	10,000,000	120	114	77.76%	13.07%	5,977,659	781,218
1992	10,400,000	108	102	74.85%	15.98%	6,216,765	993,281
1993	10,800,000	96	90	71.29%	19.54%	6,455,872	1,261,416
1994	11,200,000	84	78	66.87%	23.96%	6,694,978	1,604,006
1995	11,600,000	72	66	61.31%	29.52%	6,934,085	2,046,646
1996	12,000,000	60	54	54.24%	36.59%	7,173,191	2,624,620
1997	12,400,000	48	42	45.17%	45.66%	7,412,297	3,384,400
1998	12,800,000	36	30	33.60%	57.22%	7,651,404	4,378,344
1999	13,200,000	24	18	19.46%	71.37%	7,890,510	5,631,298
2000	13,600,000	12	6	4.69%	86.13%	8,129,616	7,002,255
Total	118,000,000					70,536,377	29,707,484

For the variance calculation, we again begin with the process variance/mean ratio, which follows the chi-square formula. The sum of chi-square values is divided by 52 (55 data points minus 3 parameters), resulting in a σ^2 of 61,577. This turns out to be less than the 65,029 calculated for the LDF method because there we divided by 43 (55 data points minus 12 parameters).

The covariance matrix is estimated from the second derivative Information Matrix, and results in the following:

$$\begin{array}{cc} & \begin{array}{ccc} \text{ELR} & \omega & \theta \end{array} \\ \begin{array}{c} \text{ELR} \\ \omega \\ \theta \end{array} & \begin{pmatrix} 0.002421 & -0.002997 & 0.242396 \\ -0.002997 & 0.007853 & -0.401000 \\ 0.242396 & -0.401000 & 33.021994 \end{pmatrix} \end{array}$$

The standard deviation of our reserve estimate is calculated in the following table.

Accident Year	Reported Losses	Estimated Reserves	Process Std Dev	CV	Parameter Std Dev	CV	Total Std Dev	CV
1991	3,901,463	781,218	219,329	28.1%	158,913	20.3%	270,848	34.7%
1992	5,339,085	993,281	247,312	24.9%	192,103	19.3%	313,156	31.5%
1993	4,909,315	1,261,416	278,701	22.1%	229,523	18.2%	361,047	28.6%
1994	4,588,268	1,604,006	314,277	19.6%	270,790	16.9%	414,846	25.9%
1995	3,873,311	2,046,646	355,002	17.3%	314,629	15.4%	474,360	23.2%
1996	3,691,712	2,624,620	402,015	15.3%	358,200	13.6%	538,445	20.5%
1997	3,483,130	3,384,400	456,510	13.5%	396,353	11.7%	604,563	17.9%
1998	2,864,498	4,378,344	519,235	11.9%	421,934	9.6%	669,054	15.3%
1999	1,363,294	5,631,298	588,862	10.5%	430,873	7.7%	729,664	13.0%
2000	344,014	7,002,255	656,641	9.4%	439,441	6.3%	790,118	11.3%
Total	34,358,090	29,707,484	1,352,515	4.6%	3,143,967	10.6%	3,422,547	11.5%

In the earlier LDF example, the standard deviation on the overall reserve was 4,885,707 and this reduces to 3,422,547 when we switch to the Cape Cod method. The reduction is primarily seen in the more recent years 1999 and 2000, but is generally true for the full loss history. The reduction in the variance (the standard deviations squared) is even more extreme – <u>the overall variance in reserves is cut in half</u>.

This conclusion is at first surprising, since the two methods are very familiar to most actuaries. The difference is that we are making use of more information in the Cape Cod method, namely the onlevel premium by year, and this information allows us to make a significantly better estimate of the reserve.

4.3 Other Calculations Possible with this Model

Once the maximum likelihood calculations have been done, there are some other uses for the statistics besides the variance of the overall reserve. We will briefly look at three of these uses.

4.3.1 Variance of the Prospective Losses

Reserve reviews always focus on losses that have already occurred, but there is an intimate connection to the forecast of losses for the prospective period. The variability estimates from the Cape Cod method help us make this connection.

If the prospective period is estimated to include 14,000,000 in premium, we have a ready estimate of expected loss as 8,369,200 based on our 59.78% ELR. The process variance is calculated using the variance/mean multiplier 61,577, producing a CV of 8.6%.

The parameter variance is also readily calculated using the covariance matrix from the earlier calculation.

$$\begin{array}{c c} & \begin{array}{ccc} \text{ELR} & \omega & \theta \end{array} \\ \begin{array}{c} \text{ELR} \\ \omega \\ \theta \end{array} & \begin{pmatrix} 0.002421 & -0.002997 & 0.242396 \\ -0.002997 & 0.007853 & -0.401000 \\ 0.242396 & -0.401000 & 33.021994 \end{pmatrix} \end{array}$$

The .002421 variance on the ELR translates to a standard deviation of 4.92% (by taking the square root) around our estimated ELR of 59.78%. Combined with the process variance, we have a total CV of 11.9%.

The CV from this estimate can then be compared to numbers produced by other prospective pricing tools.

4.3.2 Calendar Year Development

The stochastic reserving model can also be used to estimate development or payment for the next calendar year period beyond the latest diagonal. An example, using the LDF method is shown below.

Accident Year	Reported Losses	Age at 12/31/2000	Growth Function	Age at 12/31/2001	Growth Function	Estimated Ultimate	Est. 12 month Development
1991	3,901,463	120	77.24%	132	79.67%	5,050,867	122,450
1992	5,339,085	108	74.32%	120	77.24%	7,184,079	210,145
1993	4,909,315	96	70.75%	108	74.32%	6,939,399	247,928
1994	4,588,268	84	66.32%	96	70.75%	6,917,862	305,811
1995	3,873,311	72	60.78%	84	66.32%	6,372,348	353,146
1996	3,691,712	60	53.75%	72	60.78%	6,867,980	482,859
1997	3,483,130	48	44.77%	60	53.75%	7,780,515	699,093
1998	2,864,498	36	33.34%	48	44.77%	8,590,793	981,372
1999	1,363,294	24	19.38%	36	33.34%	7,033,659	981,996
2000	344,014	12	4.74%	24	19.38%	7,261,205	1,063,384
Total	34,358,090					69,998,708	5,448,182

The estimated development for the next 12-month calendar period is calculated by the difference in the growth functions at the two evaluation ages times the estimated ultimate losses. The standard deviation around this estimated development is:

Accident Year	Reported Losses	Est. 12 month Development	Process Std Dev	CV	Parameter Std Dev	CV	Total Std Dev	CV
1991	3,901,463	122,450	89,234	72.9%	24,632	20.1%	92,572	75.6%
1992	5,339,085	210,145	116,900	55.6%	37,767	18.0%	122,849	58.5%
1993	4,909,315	247,928	126,974	51.2%	42,716	17.2%	133,967	54.0%
1994	4,588,268	305,811	141,020	46.1%	50,260	16.4%	149,708	49.0%
1995	3,873,311	353,146	151,541	42.9%	57,208	16.2%	161,980	45.9%
1996	3,691,712	482,859	177,200	36.7%	74,987	15.5%	192,413	39.8%
1997	3,483,130	699,093	213,217	30.5%	106,043	15.2%	238,131	34.1%
1998	2,864,498	981,372	252,621	25.7%	158,978	16.2%	298,482	30.4%
1999	1,363,294	981,996	252,702	25.7%	225,920	23.0%	338,966	34.5%
2000	344,014	1,063,384	262,965	24.7%	480,861	45.2%	548,068	51.5%
Total	34,358,090	5,448,182	595,223	10.9%	635,609	11.7%	870,798	16.0%

A major reason for calculating the 12-month development is that the estimate is testable within a relatively short timeframe. If we project 5,448,182 of development, along with a standard deviation of 870,798, then one year later we can compare the actual development and see if it was within the forecast range.

4.3.3 Variability in Discounted Reserves

The mathematics for calculating the variability around discounted reserves follows directly from the payout pattern, model parameters and covariance matrix already calculated. The details are provided in Appendix C. This calculation is, of course, only appropriate if the analysis is being performed on paid data.

For the Cape Cod calculation of reserves, along with the 240 month truncation point, the discounted reserve using a 6.0% rate is provided below.

Accident Year	Estimated Reserves	Discounted Reserves	Process Std Dev	C.V.	Parameter Std Dev	C.V.	Total Std Dev	C.V.
1991	781,218	632,995	179,807	28.4%	125,961	19.9%	219,538	34.7%
1992	993,281	796,674	201,069	25.2%	149,689	18.8%	250,670	31.5%
1993	1,261,416	1,003,816	225,216	22.4%	175,899	17.5%	285,767	28.5%
1994	1,604,006	1,269,446	252,987	19.9%	204,084	16.1%	325,043	25.6%
1995	2,046,646	1,614,650	285,275	17.7%	232,952	14.4%	368,305	22.8%
1996	2,624,620	2,068,611	323,114	15.6%	259,904	12.6%	414,672	20.0%
1997	3,384,400	2,669,559	367,518	13.8%	280,605	10.5%	462,394	17.3%
1998	4,378,344	3,459,057	418,912	12.1%	289,876	8.4%	509,427	14.7%
1999	5,631,298	4,449,320	475,291	10.7%	286,857	6.4%	555,147	12.5%
2000	7,002,255	5,490,513	526,186	9.6%	284,582	5.2%	598,213	10.9%
Total	29,707,484	23,454,641	1,089,311	4.6%	2,198,224	9.4%	2,453,322	10.5%

From Section 4.2 above, we saw that the full-value reserve of 29,707,486 had a CV of 11.5%. The discounted reserve of 23,454,641 has a CV of 10.5%. The smaller CV for the discounted reserve is because the "tail" of the payout curve has the greatest parameter variance and also receives the deepest discount.

Section 5: Comments and Conclusion

5.1 Comments

Having worked through an example of stochastic reserving, a few practical comments are in order.

1) Abandon your triangles!

The maximum likelihood model works most logically from the tabular format of data as shown in tables 1.1 and 2.1. It is possible to first create the more familiar triangular format and then build the table, but there is no need for that intermediate step. All that is really needed is a consistent aggregation of losses evaluated at more than one date; we can skip the step of creating the triangle altogether.

2) The CV Goes with the Mean

The question of the use of the standard deviation or CV from the MLE is common. If we select a carried reserve other than the maximum likelihood estimate, then can we still use the CV from the model?

The short answer is "no". The estimate of the standard deviation in this model is very explicitly the standard deviation around the maximum likelihood estimate. If you do not trust the expected reserve from the MLE model, then there is even less reason to trust the standard deviation.

The more practical answer is an equivocal "yes". The final carried reserve is a <u>selection</u>, based on many factors including the use of a statistical model. No purely mechanical model should be the basis for setting the reserve, because it cannot take into account all of the characteristics of the underlying loss phenomenon. The standard deviation or CV around the selected reserve must therefore also be a selection, and a reasonable basis for that selection is the output of the MLE model.

The selection of a reserve range also needs to include consideration about changes in mix of business and the process of settling claims. These types of considerations might better be labeled "model variance", since by definition they are factors outside of the assumptions of the model.

3) Other Curve Forms

This paper has applied the method of maximum likelihood using growth curves that follow the Loglogistic and Weibull curve forms. These curves are useful in that they smoothly move from 0% to 100%, they often closely match the empirical data, and the first and second derivatives are calculable without the need for numerical approximations. However, the method in general is not limited to these forms and a larger library of curves can be investigated.

In this paper the Loglogistic and Weibull curves were applied to the <u>average</u> evaluation age, rather than the age from inception of the historical policy period. This was done for practical purposes, and is one way of improving the fit at immature ages. When evaluation ages fall within the period being developed (that is the period is not yet fully earned), then a further annualizing adjustment is needed. The formulas for this adjustment are given in Appendix B.

5.2 Conclusion

The method of maximum likelihood is a very useful technique for estimating both the expected development pattern and the variance around the estimated reserve. The use of the over-dispersed Poisson distribution is a convenient link to the LDF and Cape Cod estimates already common among reserving actuaries.

The chief result that we observe in working on practical examples is that the "parameter variance" component is generally larger than the "process variance" – most of the uncertainty in the estimated reserve is related to our inability to reliably estimate the expected reserve, not to random events. As such, our most pressing need is not for more sophisticated models, but for more complete data. Supplementing the standard loss development triangle with accident year exposure information is a good step in that direction.

Table 1.1
Original Triangle in Tabular Format

AY	From	To	Increment	Diag Age	AY Total
1991	0	12	357,848	120	3,901,463
1991	12	24	766,940	120	3,901,463
1991	24	36	610,542	120	3,901,463
1991	36	48	447,378	120	3,901,463
1991	48	60	562,888	120	3,901,463
1991	60	72	574,398	120	3,901,463
1991	72	84	146,342	120	3,901,463
1991	84	96	139,950	120	3,901,463
1991	96	108	227,229	120	3,901,463
1991	108	120	67,948	120	3,901,463
1992	0	12	352,118	108	5,339,085
1992	12	24	884,021	108	5,339,085
1992	24	36	933,894	108	5,339,085
1992	36	48	1,183,289	108	5,339,085
1992	48	60	445,745	108	5,339,085
1992	60	72	320,996	108	5,339,085
1992	72	84	527,804	108	5,339,085
1992	84	96	266,172	108	5,339,085
1992	96	108	425,046	108	5,339,085
1993	0	12	290,507	96	4,909,315
1993	12	24	1,001,799	96	4,909,315
1993	24	36	926,219	96	4,909,315
1993	36	48	1,016,654	96	4,909,315
1993	48	60	750,816	96	4,909,315
1993	60	72	146,923	96	4,909,315
1993	72	84	495,992	96	4,909,315
1993	84	96	280,405	96	4,909,315
1994	0	12	310,608	84	4,588,268
1994	12	24	1,108,250	84	4,588,268
1994	24	36	776,189	84	4,588,268
1994	36	48	1,562,400	84	4,588,268
1994	48	60	272,482	84	4,588,268
1994	60	72	352,053	84	4,588,268
1994	72	84	206,286	84	4,588,268
1995	0	12	443,160	72	3,873,311
1995	12	24	693,190	72	3,873,311
1995	24	36	991,983	72	3,873,311
1995	36	48	769,488	72	3,873,311
1995	48	60	504,851	72	3,873,311
1995	60	72	470,639	72	3,873,311
1996	0	12	396,132	60	3,691,712
1996	12	24	937,085	60	3,691,712
1996	24	36	847,498	60	3,691,712
1996	36	48	805,037	60	3,691,712
1996	48	60	705,960	60	3,691,712
1997	0	12	440,832	48	3,483,130
1997	12	24	847,631	48	3,483,130
1997	24	36	1,131,398	48	3,483,130
1997	36	48	1,063,269	48	3,483,130
1998	0	12	359,480	36	2,864,498
1998	12	24	1,061,648	36	2,864,498
1998	24	36	1,443,370	36	2,864,498
1999	0	12	376,686	24	1,363,294
1999	12	24	986,608	24	1,363,294
2000	0	12	344,014	12	344,014

Table 1.2
Triangle Collapsed for Latest Three Diagonals

AY	From	To	Increment	Diag Age	AY Total
1991	0	96	3,606,286	120	3,901,463
1991	96	108	227,229	120	3,901,463
1991	108	120	67,948	120	3,901,463
1992	0	84	4,647,867	108	5,339,085
1992	84	96	266,172	108	5,339,085
1992	96	108	425,046	108	5,339,085
1993	0	72	4,132,918	96	4,909,315
1993	72	84	495,992	96	4,909,315
1993	84	96	280,405	96	4,909,315
1994	0	60	4,029,929	84	4,588,268
1994	60	72	352,053	84	4,588,268
1994	72	84	206,286	84	4,588,268
1995	0	48	2,897,821	72	3,873,311
1995	48	60	504,851	72	3,873,311
1995	60	72	470,639	72	3,873,311
1996	0	36	2,180,715	60	3,691,712
1996	36	48	805,037	60	3,691,712
1996	48	60	705,960	60	3,691,712
1997	0	24	1,288,463	48	3,483,130
1997	24	36	1,131,398	48	3,483,130
1997	36	48	1,063,269	48	3,483,130
1998	0	12	359,480	36	2,864,498
1998	12	24	1,061,648	36	2,864,498
1998	24	36	1,443,370	36	2,864,498
1999	0	12	376,686	24	1,363,294
1999	12	24	986,608	24	1,363,294
2000	0	12	344,014	12	344,014

Table 1.3
Latest Diagonal Representing only 9 Months of Development

AY	From	To	Increment	Diag Age	AY Total
1991	0	12	357,848	117	3,901,463
1991	12	24	766,940	117	3,901,463
1991	24	36	610,542	117	3,901,463
1991	36	48	447,378	117	3,901,463
1991	48	60	562,888	117	3,901,463
1991	60	72	574,398	117	3,901,463
1991	72	84	146,342	117	3,901,463
1991	84	96	139,950	117	3,901,463
1991	96	108	227,229	117	3,901,463
1991	108	117	67,948	117	3,901,463
1992	0	12	352,118	105	5,339,085
1992	12	24	884,021	105	5,339,085
1992	24	36	933,894	105	5,339,085
1992	36	48	1,183,289	105	5,339,085
1992	48	60	445,745	105	5,339,085
1992	60	72	320,996	105	5,339,085
1992	72	84	527,804	105	5,339,085
1992	84	96	266,172	105	5,339,085
1992	96	105	425,046	105	5,339,085
1993	0	12	290,507	93	4,909,315
1993	12	24	1,001,799	93	4,909,315
1993	24	36	926,219	93	4,909,315
1993	36	48	1,016,654	93	4,909,315
1993	48	60	750,816	93	4,909,315
1993	60	72	146,923	93	4,909,315
1993	72	84	495,992	93	4,909,315
1993	84	93	280,405	93	4,909,315
1994	0	12	310,608	81	4,588,268
1994	12	24	1,108,250	81	4,588,268
1994	24	36	776,189	81	4,588,268
1994	36	48	1,562,400	81	4,588,268
1994	48	60	272,482	81	4,588,268
1994	60	72	352,053	81	4,588,268
1994	72	81	206,286	81	4,588,268
1995	0	12	443,160	69	3,873,311
1995	12	24	693,190	69	3,873,311
1995	24	36	991,983	69	3,873,311
1995	36	48	769,488	69	3,873,311
1995	48	60	504,851	69	3,873,311
1995	60	69	470,639	69	3,873,311
1996	0	12	396,132	57	3,691,712
1996	12	24	937,085	57	3,691,712
1996	24	36	847,498	57	3,691,712
1996	36	48	805,037	57	3,691,712
1996	48	57	705,960	57	3,691,712
1997	0	12	440,832	45	3,483,130
1997	12	24	847,631	45	3,483,130
1997	24	36	1,131,398	45	3,483,130
1997	36	45	1,063,269	45	3,483,130
1998	0	12	359,480	33	2,864,498
1998	12	24	1,061,648	33	2,864,498
1998	24	33	1,443,370	33	2,864,498
1999	0	12	376,686	21	1,363,294
1999	12	21	986,608	21	1,363,294
2000	0	9	344,014	9	344,014

Table 1.4
Original Triangle along with Fitted Values – LDF Method

AY	From	To	Increment	Diag Age	AY Total	Est. ULT	Fitted	MLE Term	Chi-Square
1991	0	12	357,848	120	3,901,463	5,050,868	239,295	4,192,814	58,734
1991	12	24	766,940	120	3,901,463	5,050,868	739,686	9,624,727	1,004
1991	24	36	610,542	120	3,901,463	5,050,868	705,171	7,516,507	12,698
1991	36	48	447,378	120	3,901,463	5,050,868	576,987	5,357,739	29,114
1991	48	60	562,888	120	3,901,463	5,050,868	453,829	6,878,055	26,208
1991	60	72	574,398	120	3,901,463	5,050,868	355,106	6,985,799	135,422
1991	72	84	146,342	120	3,901,463	5,050,868	279,911	1,555,543	63,737
1991	84	96	139,950	120	3,901,463	5,050,868	223,278	1,500,370	31,098
1991	96	108	227,229	120	3,901,463	5,050,868	180,455	2,569,751	12,124
1991	108	120	67,948	120	3,901,463	5,050,868	147,745	661,056	43,099
1992	0	12	352,118	108	5,339,085	7,184,081	340,360	4,144,834	406
1992	12	24	884,021	108	5,339,085	7,184,081	1,052,089	11,206,001	26,848
1992	24	36	933,894	108	5,339,085	7,184,081	1,002,997	11,902,020	4,761
1992	36	48	1,183,289	108	5,339,085	7,184,081	820,675	15,293,216	160,220
1992	48	60	445,745	108	5,339,085	7,184,081	645,502	5,317,578	61,817
1992	60	72	320,996	108	5,339,085	7,184,081	505,083	3,710,390	67,094
1992	72	84	527,804	108	5,339,085	7,184,081	398,131	6,407,657	42,235
1992	84	96	266,172	108	5,339,085	7,184,081	317,579	3,054,416	8,321
1992	96	108	425,046	108	5,339,085	7,184,081	256,670	5,037,510	110,456
1993	0	12	290,507	96	4,909,315	6,939,401	328,768	3,361,574	4,453
1993	12	24	1,001,799	96	4,909,315	6,939,401	1,016,256	12,840,263	206
1993	24	36	926,219	96	4,909,315	6,939,401	968,836	11,798,028	1,875
1993	36	48	1,016,654	96	4,909,315	6,939,401	792,724	13,016,722	63,256
1993	48	60	750,816	96	4,909,315	6,939,401	623,517	9,394,719	25,990
1993	60	72	146,923	96	4,909,315	6,939,401	487,881	1,436,491	238,280
1993	72	84	495,992	96	4,909,315	6,939,401	384,571	5,993,828	32,282
1993	84	96	280,405	96	4,909,315	6,939,401	306,763	3,235,826	2,265
1994	0	12	310,608	84	4,588,268	6,917,864	327,748	3,616,974	896
1994	12	24	1,108,250	84	4,588,268	6,917,864	1,013,102	14,312,364	8,936
1994	24	36	776,189	84	4,588,268	6,917,864	965,829	9,730,631	37,236
1994	36	48	1,562,400	84	4,588,268	6,917,864	790,264	20,427,319	754,424
1994	48	60	272,482	84	4,588,268	6,917,864	621,582	3,013,334	196,065
1994	60	72	352,053	84	4,588,268	6,917,864	486,366	4,123,668	37,092
1994	72	84	206,286	84	4,588,268	6,917,864	383,377	2,268,795	81,803
1995	0	12	443,160	72	3,873,311	6,372,350	301,903	5,289,828	66,093
1995	12	24	693,190	72	3,873,311	6,372,350	933,213	8,595,646	61,734
1995	24	36	991,983	72	3,873,311	6,372,350	889,668	12,699,114	11,767
1995	36	48	769,488	72	3,873,311	6,372,350	727,947	9,658,589	2,371
1995	48	60	504,851	72	3,873,311	6,372,350	572,566	6,120,690	8,008
1995	60	72	470,639	72	3,873,311	6,372,350	448,014	5,676,214	1,143
1996	0	12	396,132	60	3,691,712	6,867,982	325,384	4,702,625	15,382
1996	12	24	937,085	60	3,691,712	6,867,982	1,005,797	11,945,927	4,694
1996	24	36	847,498	60	3,691,712	6,867,982	958,865	10,714,153	12,935
1996	36	48	805,037	60	3,691,712	6,867,982	784,566	10,142,109	534
1996	48	60	705,960	60	3,691,712	6,867,982	617,100	8,795,314	12,796
1997	0	12	440,832	48	3,483,130	7,780,518	368,618	5,281,753	14,147
1997	12	24	847,631	48	3,483,130	7,780,518	1,139,436	10,681,663	74,730
1997	24	36	1,131,398	48	3,483,130	7,780,518	1,086,268	14,638,194	1,875
1997	36	48	1,063,269	48	3,483,130	7,780,518	888,809	13,675,465	34,244
1998	0	12	359,480	36	2,864,498	8,590,795	407,006	4,236,247	5,550
1998	12	24	1,061,648	36	2,864,498	8,590,795	1,258,098	13,652,867	30,675
1998	24	36	1,443,370	36	2,864,498	8,590,795	1,199,393	19,003,928	49,629
1999	0	12	376,686	24	1,363,294	7,033,660	333,234	4,456,931	5,666
1999	12	24	986,608	24	1,363,294	7,033,660	1,030,060	12,629,654	1,833
2000	0	12	344,014	12	344,014	7,261,202	344,014	4,041,627	0
			34,358,090				34,358,090		2,796,260

Table 2.1
Original Triangle along with Fitted Values – Cape Cod Method

AY	From	To	Increment	Diag Age	Premium	Est. ULT	Fitted	MLE Term	Chi-Square
1991	0	12	357,848	120	10,000,000	5,977,659	280,569	4,208,482	21,285
1991	12	24	766,940	120	10,000,000	5,977,659	882,582	9,617,292	15,152
1991	24	36	610,542	120	10,000,000	5,977,659	845,554	7,486,969	65,319
1991	36	48	447,378	120	10,000,000	5,977,659	691,227	5,324,318	86,024
1991	48	60	562,888	120	10,000,000	5,977,659	542,171	6,889,829	792
1991	60	72	574,398	120	10,000,000	5,977,659	422,833	7,018,339	54,329
1991	72	84	146,342	120	10,000,000	5,977,659	332,202	1,528,317	103,985
1991	84	96	139,950	120	10,000,000	5,977,659	264,171	1,483,014	58,412
1991	96	108	227,229	120	10,000,000	5,977,659	212,900	2,574,877	964
1991	108	120	67,948	120	10,000,000	5,977,659	173,860	646,001	64,519
1992	0	12	352,118	108	10,400,000	6,216,765	291,792	4,139,189	12,472
1992	12	24	884,021	108	10,400,000	6,216,765	917,885	11,219,571	1,249
1992	24	36	933,894	108	10,400,000	6,216,765	879,376	11,902,801	3,380
1992	36	48	1,183,289	108	10,400,000	6,216,765	718,876	15,238,302	300,023
1992	48	60	445,745	108	10,400,000	6,216,765	563,858	5,338,946	24,742
1992	60	72	320,996	108	10,400,000	6,216,765	439,746	3,731,261	32,068
1992	72	84	527,804	108	10,400,000	6,216,765	345,490	6,385,446	96,207
1992	84	96	266,172	108	10,400,000	6,216,765	274,738	3,058,687	267
1992	96	108	425,046	108	10,400,000	6,216,765	221,416	5,009,964	187,273
1993	0	12	290,507	96	10,800,000	6,455,872	303,015	3,363,630	516
1993	12	24	1,001,799	96	10,800,000	6,455,872	953,188	12,839,147	2,479
1993	24	36	926,219	96	10,800,000	6,455,872	913,198	11,798,887	186
1993	36	48	1,016,654	96	10,800,000	6,455,872	746,525	13,001,875	97,746
1993	48	60	750,816	96	10,800,000	6,455,872	585,545	9,385,515	46,648
1993	60	72	146,923	96	10,800,000	6,455,872	456,660	1,457,996	210,084
1993	72	84	495,992	96	10,800,000	6,455,872	358,778	5,985,187	52,477
1993	84	96	280,405	96	10,800,000	6,455,872	285,305	3,236,950	84
1994	0	12	310,608	84	11,200,000	6,694,978	314,238	3,617,409	42
1994	12	24	1,108,250	84	11,200,000	6,694,978	988,491	14,309,720	14,509
1994	24	36	776,189	84	11,200,000	6,694,978	947,020	9,734,175	30,816
1994	36	48	1,562,400	84	11,200,000	6,694,978	774,174	20,411,270	802,533
1994	48	60	272,482	84	11,200,000	6,694,978	607,232	3,021,320	184,538
1994	60	72	352,053	84	11,200,000	6,694,978	473,573	4,127,077	31,182
1994	72	84	206,286	84	11,200,000	6,694,978	372,066	2,273,929	73,866
1995	0	12	443,160	72	11,600,000	6,934,085	325,460	5,299,568	42,565
1995	12	24	693,190	72	11,600,000	6,934,085	1,023,795	8,569,280	106,759
1995	24	36	991,983	72	11,600,000	6,934,085	980,842	12,704,721	127
1995	36	48	769,488	72	11,600,000	6,934,085	801,823	9,659,092	1,304
1995	48	60	504,851	72	11,600,000	6,934,085	628,919	6,111,729	24,475
1995	60	72	470,639	72	11,600,000	6,934,085	490,486	5,676,368	803
1996	0	12	396,132	60	12,000,000	7,173,191	336,683	4,704,848	10,497
1996	12	24	937,085	60	12,000,000	7,173,191	1,059,098	11,941,015	14,056
1996	24	36	847,498	60	12,000,000	7,173,191	1,014,664	10,706,291	27,541
1996	36	48	805,037	60	12,000,000	7,173,191	829,472	10,142,011	720
1996	48	60	705,960	60	12,000,000	7,173,191	650,606	8,799,134	4,710
1997	0	12	440,832	48	12,400,000	7,412,297	347,906	5,276,973	24,821
1997	12	24	847,631	48	12,400,000	7,412,297	1,094,401	10,692,516	55,643
1997	24	36	1,131,398	48	12,400,000	7,412,297	1,048,487	14,635,924	6,556
1997	36	48	1,063,269	48	12,400,000	7,412,297	857,121	13,668,552	49,581
1998	0	12	359,480	36	12,800,000	7,651,404	359,129	4,239,137	0
1998	12	24	1,061,648	36	12,800,000	7,651,404	1,129,704	13,666,979	4,100
1998	24	36	1,443,370	36	12,800,000	7,651,404	1,082,309	18,972,750	120,451
1999	0	12	376,686	24	13,200,000	7,890,510	370,351	4,459,595	108
1999	12	24	986,608	24	13,200,000	7,890,510	1,165,008	12,616,168	27,319
2000	0	12	344,014	12	13,600,000	8,129,616	381,574	4,039,715	3,697
			34,358,090				34,358,090		3,202,001

References:

[1] England & Verrall; *A Flexible Framework for Stochastic Claims Reserving*; Proceedings of the Casualty Actuarial Society, 2001.

[2] Halliwell, Leigh; *Loss Prediction by Generalized Least Squares*; Proceedings of the Casualty Actuarial Society, 1996.

[3] Klugman, Panjer & Wilmot; Loss Models: From Data to Decisions; John Wiley & Sons, 1998.
See mainly pages 60-68, for discussion of maximum likelihood.

[4] Kreps, Rodney; *Parameter Uncertainty in (Log)Normal Distributions*; Proceedings of the Casualty Actuarial Society, 1997.

[5] Mack, Thomas; *A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserve;* 1991 ASTIN Bulletin 21 (1), 93-109.

[6] Mack, Thomas; *Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates;* 1993 ASTIN Bulletin 23 (2), 213-225.

[7] McCullagh, P. & Nelder, J.A.; Generalized Linear Models; Chapman & Hall/CRC, second edition 1989.
See Chapter 6 on Log-linear Models, and the over-dispersed Poisson.

[8] Sherman, Richard E.; *Extrapolating, Smoothing, and Interpolating Development Factors*; Proceedings of the Casualty Actuarial Society, 1984.

[9] Weissner, Edward W.; *Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood*; Proceedings of the Casualty Actuarial Society, 1978.

[10] Zehnwirth, Ben; *Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals and Risk Based Capital.* CAS Forum Spring 1994, Volume Two.

Appendix A: Derivatives of the Loglikelihood Function

The loglikelihood function for the over-dispersed Poisson is proportional to

$$\ell = \sum_i c_i \cdot \ln(\mu_i) - \mu_i$$

$$\text{where } \mu_{i,t} = ELR \cdot P_i \cdot [G(x_t \mid \omega,\theta) - G(x_{t-1} \mid \omega,\theta)]$$

as described in section 2.2 of this paper. The derivatives below are then used to complete the Information Matrix needed in the parameter variance calculation.

The derivatives of the exact loglikelihood function would require dividing all of these numbers by the constant scale factor σ^2, but it is easier to omit that here and apply it to the final covariance matrix at the end.

$$\frac{\partial^2 \ell}{\partial ELR^2} = \sum_{i,t}\left(\frac{-c_{i,t}}{ELR^2}\right)$$

$$\frac{\partial^2 \ell}{\partial ELR\,\partial\omega} = -\sum_{i,t} P_i \cdot \left[\frac{\partial G(x_t)}{\partial\omega} - \frac{\partial G(x_{t-1})}{\partial\omega}\right]$$

$$\frac{\partial^2 \ell}{\partial ELR\,\partial\theta} = -\sum_{i,t} P_i \cdot \left[\frac{\partial G(x_t)}{\partial\theta} - \frac{\partial G(x_{t-1})}{\partial\theta}\right]$$

$$\frac{\partial \ell}{\partial\omega} = \sum_{i,t}\left\{\left[\frac{c_{i,t}}{G(x_t) - G(x_{t-1})} - ELR \cdot P_i\right] \cdot \left[\frac{\partial G(x_t)}{\partial\omega} - \frac{\partial G(x_{t-1})}{\partial\omega}\right]\right\}$$

$$\frac{\partial^2 \ell}{\partial\omega^2} = \sum_{i,t}\left\{\left[\frac{-c_{i,t}}{(G(x_t) - G(x_{t-1}))^2}\right] \cdot \left[\frac{\partial G(x_t)}{\partial\omega} - \frac{\partial G(x_{t-1})}{\partial\omega}\right]^2 + \right.$$

$$\left.\left[\frac{c_{i,t}}{G(x_t) - G(x_{t-1})} - ELR \cdot P_i\right] \cdot \left[\frac{\partial^2 G(x_t)}{\partial\omega^2} - \frac{\partial^2 G(x_{t-1})}{\partial\omega^2}\right]\right\}$$

$$\frac{\partial^2 \ell}{\partial\omega\,\partial\theta} = \sum_{i,t}\left\{\left[\frac{-c_{i,t}}{(G(x_t)-G(x_{t-1}))^2}\right]\cdot\left[\frac{\partial G(x_t)}{\partial\omega}-\frac{\partial G(x_{t-1})}{\partial\omega}\right]\left[\frac{\partial G(x_t)}{\partial\theta}-\frac{\partial G(x_{t-1})}{\partial\theta}\right]+\left[\frac{c_{i,t}}{G(x_t)-G(x_{t-1})}-ELR\cdot P_i\right]\cdot\left[\frac{\partial^2 G(x_t)}{\partial\omega\,\partial\theta}-\frac{\partial^2 G(x_{t-1})}{\partial\omega\,\partial\theta}\right]\right\}$$

$$\frac{\partial \ell}{\partial\theta} = \sum_{i,t}\left\{\left[\frac{c_{i,t}}{G(x_t)-G(x_{t-1})}-ELR\cdot P_i\right]\cdot\left[\frac{\partial G(x_t)}{\partial\theta}-\frac{\partial G(x_{t-1})}{\partial\theta}\right]\right\}$$

$$\frac{\partial^2 \ell}{\partial\theta^2} = \sum_{i,t}\left\{\left[\frac{-c_{i,t}}{(G(x_t)-G(x_{t-1}))^2}\right]\cdot\left[\frac{\partial G(x_t)}{\partial\theta}-\frac{\partial G(x_{t-1})}{\partial\theta}\right]^2+\left[\frac{c_{i,t}}{G(x_t)-G(x_{t-1})}-ELR\cdot P_i\right]\cdot\left[\frac{\partial^2 G(x_t)}{\partial\theta^2}-\frac{\partial^2 G(x_{t-1})}{\partial\theta^2}\right]\right\}$$

For the LDF Method, these same formulas apply but replacing:

$$ELR \to ULT_i \quad \text{and} \quad P_i \to 1.$$

Weibull Distribution

$$G(x) \;=\; F(x) \;=\; 1 - \exp\left[-(x/\theta)^{\omega}\right]$$

$$f(x) \;=\; \frac{\omega}{x}\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\exp\left[-(x/\theta)^{\omega}\right]$$

$$E[x^k] \;=\; \theta^k\cdot\Gamma(1+k/\omega)$$

θ is approximately the 63.2%-tile $=1-\exp[-1]$, $LDF_{\theta}\approx 1.582$

$$\frac{\partial G(x)}{\partial\omega} \;=\; \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\ln\left(\tfrac{x}{\theta}\right)$$

$$\frac{\partial G(x)}{\partial\theta} \;=\; \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\left(\frac{-\omega}{\theta}\right)$$

$$\frac{\partial^2 G(x)}{\partial\omega^2} \;=\; \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\ln\left(\tfrac{x}{\theta}\right)^2\cdot\left[1-\left(\frac{x}{\theta}\right)^{\omega}\right]$$

$$\frac{\partial^2 G(x)}{\partial\omega\partial\theta} \;=\; \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\left(\frac{-1}{\theta}\right)\cdot\left\{1+\omega\cdot\ln\left(\tfrac{x}{\theta}\right)\cdot\left[1-\left(\frac{x}{\theta}\right)^{\omega}\right]\right\}$$

$$\frac{\partial^2 G(x)}{\partial\theta^2} \;=\; \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\left(\frac{\omega}{\theta^2}\right)\cdot\left\{1+\omega\cdot\left[1-\left(\frac{x}{\theta}\right)^{\omega}\right]\right\}$$

Loglogistic Distribution (for "inverse power" LDFs)

$$G(x) \;=\; F(x) \;=\; \frac{x^{\omega}}{x^{\omega}+\theta^{\omega}} \;=\; 1-\left(\frac{1}{1+(x/\theta)^{\omega}}\right)$$

$$f(x) \;=\; \frac{\omega}{x}\cdot\left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{\theta^{\omega}}{x^{\omega}+\theta^{\omega}}\right)$$

$$E[x^{k}] \;=\; \theta^{k}\cdot\Gamma(1+k/\omega)\cdot\Gamma(1-k/\omega)$$

θ is the median of the distribution $LDF_{\theta}=2.000$

$$\frac{\partial G(x)}{\partial\omega} \;=\; \left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{\theta^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\ln\left(\tfrac{x}{\theta}\right)$$

$$\frac{\partial G(x)}{\partial\theta} \;=\; \left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{\theta^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{-\omega}{\theta}\right)$$

$$\frac{\partial^{2} G(x)}{\partial\omega^{2}} \;=\; \left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{\theta^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\ln\left(\tfrac{x}{\theta}\right)^{2}\cdot\left[1-2\cdot\left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\right]$$

$$\frac{\partial^{2} G(x)}{\partial\omega\,\partial\theta} \;=\; \left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{\theta^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{-1}{\theta}\right)\cdot\left\{1+\omega\cdot\ln\left(\tfrac{x}{\theta}\right)\cdot\left[1-2\cdot\left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\right]\right\}$$

$$\frac{\partial^{2} G(x)}{\partial\theta^{2}} \;=\; \left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{\theta^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\cdot\left(\frac{\omega}{\theta^{2}}\right)\cdot\left\{1+\omega\cdot\left[1-2\cdot\left(\frac{x^{\omega}}{x^{\omega}+\theta^{\omega}}\right)\right]\right\}$$

Appendix B: Adjustments for Different Exposure Periods

The percent of ultimate curve is assumed to be a function of the average accident date of the period being developed to ultimate.

$$G^{*}(x \mid \omega, \theta) \quad = \text{cumulative percent of ultimate as of average date } x$$

Further, we will assume that this is the percent of ultimate for the portion of the period that has already been earned. For example, if we are 9 months into an accident year, then the quantity $G^{*}(4.5 \mid \omega, \theta)$ represents the cumulative percent of ultimate of the 9-month period only. The loss development factor $LDF_9^{*} = 1 / G^{*}(4.5 \mid \omega, \theta)$ is the adjustment needed to calculate the ultimate loss dollars for the 9-month period (before annualizing).

In order to estimate the cumulative percent of ultimate for the full accident year, we also need to multiply by a scaling factor representing the portion of the accident year that has been earned.

The AY cumulative percent of ultimate as of 9 months is

$$G_{AY}(9 \mid \omega, \theta) \quad = \quad \left(\frac{9}{12}\right) \cdot G^{*}(4.5 \mid \omega, \theta)$$

We find therefore that we need to make two calculations:

1) Calculate the percent of the period that is exposed; *Expos(t)*
2) Calculate the average accident date given the age from inception t; *AvgAge(t)*

These functions can be easily calculated for accident year or policy year periods.

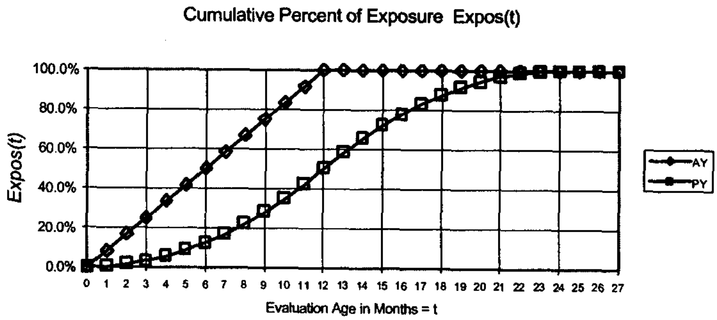
1) Calculate the percent of the period that is exposed: $Expos(t)$

For accident years (AY):

$$Expos(t) = \begin{cases} t/12 & t \le 12 \\ 1 & t > 12 \end{cases} \quad \text{or}$$

For policy years (PY):

$$Expos(t) = \begin{cases} \tfrac{1}{2}\cdot(t/12)^2 & t \le 12 \\ 1-\tfrac{1}{2}\cdot \max(2-t/12,0)^2 & t > 12 \end{cases}$$



2) Calculate the average accident date of the period that is earned: $AvgAge(t)$

For accident years (AY):

$$AvgAge(t) = \begin{cases} t/2 & t \le 12 \\ t-6 & t > 12 \end{cases} \qquad \text{or} \qquad AvgAge(t) = \max(t-6, t/2)$$

For policy years (PY):

$$AvgAge(t) = \begin{cases} t/3 & t \le 12 \\ \dfrac{(t-12)+\frac{1}{3}\cdot(24-t)\cdot(1-Expos(t))}{Expos(t)} & t > 12 \end{cases}$$

The final cumulative percent of ultimate curve, including annualization, is given by:

$$\boxed{G_{AY\,or\,PY}(t \mid \omega,\theta) = Expos(t)\cdot G^{*}\left(AvgAge(t) \mid \omega,\theta\right)}$$

Appendix C: Variance in Discounted Reserves

The maximum likelihood estimation model allows for the estimation of variance of discounted reserves as well as the variance of the full-value reserves. These calculations are a bit more tedious, and so are given just in this appendix.

Calculation of Discounted Reserve

We begin by recalling that the reserve is estimated as a sum of portions of all the historical accident years, and is calculated as:

$$\text{Reserve:} \qquad R = \sum_{AY} \mu_{AY;x,y} = \sum_{AY} ULT_{AY} \cdot \left(G(y) - G(x)\right)$$

This expression can be expanded as the sum of individual increments.

$$R = \sum_{AY} \sum_{k=1}^{y-x} ULT_{AY} \cdot \left(G(x+k) - G(x+k-1)\right)$$

To be even more precise, we could write this as a continuous function.

$$R = \sum_{AY} ULT_{AY} \cdot \int_x^y g(t)\,dt \qquad \text{where} \quad g(t) = \frac{\partial G(t)}{\partial t}$$

The value of the discounted reserve R_d would then be written as follows.

$$R_d = \sum_{AY} ULT_{AY} \cdot \int_x^y v^{t-x} \cdot g(t)\,dt \qquad \text{where} \quad v = \frac{1}{1+i}$$

For purposes of this paper, we will assume that the discount rate i is constant. There is also some debate as to what this rate should be (cost of capital?, market yield?), but we will avoid that discussion here.

An interesting note on this expression is seen in the case of $x=0$ and $y=\infty$, in which the form of the discounted loss at time zero is directly related to the moment generating function of the growth curve.

$$\int_0^\infty v^t \cdot g(y)\,dt = \int_0^\infty e^{-t \cdot \ln(1+i)} \cdot g(t)\,dt = MGF\left(-\ln(1+i)\right)$$

Unfortunately, for the Loglogistic and Weibull growth curves, the moment generating function is intractable and so does not simplify our calculation. For practical purposes we will use the incremental approximation instead.

$$R_d \approx \sum_{AY}\sum_{k=1}^{y-x} ULT_{AY} \cdot v^{k-1/2} \cdot (G(x+k) - G(x+k-1))$$

The variance can then be calculated for the discounted reserve in two pieces: the process variance and the parameter variance.

Process Variance

The process variance component is actually trivial to calculate. We already know that the variance of the full value reserve is estimated by multiplying by the scale factor σ^2. We then need to recall that the variance for some random variable times a constant is given by $\quad Var(v^k \cdot R) = v^{2k} \cdot Var(R)$.

The process variance of the discounted reserve is therefore:

$$Var(R_d) \approx \sigma^2 \cdot \sum_{AY}\sum_{k=1}^{y-x} ULT_{AY} \cdot v^{2k-1} \cdot (G(x+k) - G(x+k-1))$$

Parameter Variance

The parameter variance again makes use of the covariance matrix of the model parameters Σ. The formula is then given below.

$$Var(E[R_d]) = (\partial R_d)' \cdot \Sigma \cdot (\partial R_d)$$

where

$$\partial R_d = \left\langle \frac{\partial R_d}{\partial ELR}, \frac{\partial R_d}{\partial \omega}, \frac{\partial R_d}{\partial \theta} \right\rangle \quad \text{for the Cape Cod method}$$

or

$$\partial R_d = \left\langle \left\{ \frac{\partial R_d}{\partial ULT_{AY}} \right\}_{AY=1}^{n}, \frac{\partial R_d}{\partial \omega}, \frac{\partial R_d}{\partial \theta} \right\rangle \quad \text{for the LDF method}$$

In order to calculate the derivatives of the discounted reserves, we make use of the same mathematical expressions as for the full value reserves. That is,

$$\frac{\partial R}{\partial \omega} = \sum_{AY,x} \frac{\partial \mu_{AY,x}}{\partial \omega} \qquad \text{becomes} \qquad \frac{\partial R_d}{\partial \omega} = \sum_{AY,x} v_{AY,x} \cdot \frac{\partial \mu_{AY,x}}{\partial \omega}$$

The calculation is similar to the variance calculation for the full value reserve, but now it is expanded for each increment so that the time dimension is included. The complexity of the calculations does not change, but the number of times they are performed greatly increases.

The combination of the process and parameter variances is simple addition, the same as for the full value reserves, since we make the assumption that the two sources of variance are independent.

CAS Study Note – Exam 7
Reserving for Reinsurance

Jacqueline Friedland, FCIA, FCAS, FSA

© 2023 Casualty Actuarial Society. All Rights Reserved.

Table of Contents

Preface

As the author of the Casualty Actuarial Society's (CAS's) text on reserving, I am honored to prepare this new text on reserving for reinsurance. In many ways, I view this text as a supplement to my earlier work, *Estimating Unpaid Claims Using Basic Techniques*, and I strongly encourage readers to be familiar with that text prior to this one.

With the goal of having this text used by actuaries and actuarial candidates around the world, I strive to present concepts in a simple and straightforward manner, particularly for those for whom English may not be their first spoken language. With this global mindset, I also chose not to use any currency in the examples.

I wish to express sincere thanks to all the members of the CAS educational committee who helped guide this text in its initial development and through countless reviews: Arthur Zaremba, Eric Blancke, Jill Labbadia, Jonathan Schreck, and Fran Sarrel. Additional thanks to these CAS members for reviewing this document: Jackie Ruan, Zora Law, Eric Lam, Meg Glenn, Kenneth Hsu and Joseph Lindner.

I also express sincere thanks to Wesley Griffiths, who worked with me as a Staff Actuary at the CAS.

Readers should be aware that figures in the supporting tables and exhibits are often carried to a greater number of decimals than shown. Thus, totals and calculations may not agree exactly due to rounding differences.

Please notify the CAS of any errors so that this text can be corrected in subsequent editions.

—Jacqueline Frank Friedland, FCAS, FCIA, FSA

Chapter 1 – Introduction

The objective of this text is to address the estimation of unpaid losses[1] from the perspective of reinsurance. Reinsurance, which is insurance for insurers, is critical for the operation of the insurance industry as a whole. Through reinsurance, the cost of risk is spread across the marketplace, often globally, and the financial effect of an insured event is lessened for a single insurer[2] or economy. This text is intended for actuaries working with reinsurers as well as for actuaries working with primary insurers who estimate losses that are ceded to reinsurers. The text is also intended for actuaries working with self-insurers and captive insurers[3] who utilize reinsurance.

It is assumed that the reader of this text is knowledgeable about basic reserving, including typical data requirements, key assumptions, and traditional methodologies (such as the chain ladder, expected loss, and Bornhuetter-Ferguson techniques). Thus, this text focuses on the differences in reserving for reinsurance versus primary insurance and not on the detailed mechanics of the traditional unpaid losses projection techniques.[4]

Like insurers, reinsurers do not know the true cost of goods sold during a financial reporting period until years, possibly decades, later – after all claims are settled. Thus, it is critical that insurers and reinsurers maintain robust processes for the estimation of unpaid losses. Most frequently, the actuary estimates unpaid losses by subtracting paid losses from projections of ultimate losses. This text explores numerous considerations for such projections and issues related to data, understanding the environment (internal and external to the reinsurer), and the selection of methodology and assumptions. In this text, the term **reserves** refers to an amount booked in a financial statement, which may differ from the actuary's estimate of unpaid losses.[5]

Appropriate estimates of unpaid losses and reserves are essential for the internal management of a reinsurer as well as for its key stakeholders.

---

[1] The estimation of unpaid losses is also referred to as reserving.

[2] In actuarial and accounting literature and standards, the term insurer is often used to refer to primary insurers, reinsurers, captive insurers, and self-insurers. Given that this text focuses specifically on reinsurance, the term reinsurer is generally used to differentiate reinsurers from other insurers.

[3] The International Risk Management Institute (IRMI) Glossary defines a captive insurer as "an insurance company that has as its primary purpose the financing of the risks of its owners or participants. Typically licensed under special purpose insurer laws and operated under a different regulatory system than commercial insurers. The intention of such special purpose licensing laws and regulations is that the captive provides insurance to sophisticated insureds that require less policyholder protection than the general public" (See https://www.irmi.com/term/insurance-definitions/captive.)

[4] For further information, see Jacqueline Friedland, *Estimating Unpaid Claims Using Basic Techniques* (Arlington, VA: Casualty Actuarial Society, 2010).

[5] This use of the term reserves is consistent with the U.S. Actuarial Standards Board's *Actuarial Standard of Practice (ASOP) 43–Property/Casualty Unpaid Claim Estimates.*

- Internal management requires sound reserves because they affect virtually every area of a reinsurer's operations, including but not limited to pricing, underwriting, strategic planning, and financial decision making.
- Investors require appropriate reserves because they are essential to the evaluation of a company's financial health. Reserves that are either inadequate or excessive can lead to misstated balance sheets and income statements for the reinsurer, and key financial metrics used by investors could be misleading. A reinsurer with insufficient reserves could present itself in a stronger position than it truly is. Conversely, a reinsurer with excessive reserves may appear to be in a weaker position than its true state. Both situations could affect investors' decisions related to the reinsurer.
- Insurance regulators rely on the financial statements of reinsurers to carry out their supervisory role. Inappropriate reserves could result in a misstatement of the true financial position of a reinsurer. If a financially struggling reinsurer is masking its true state with inadequate reserves, a regulator may not become involved until it is too late to help the reinsurer regain its strength and protect the public's interests.
- Rating agencies evaluate movement over time in reinsurers' reserves. A reinsurer who reports significant adverse reserve development that results in reduced capital and a weakened financial position could face a downgrade from rating agencies. A rating downgrade, or even the threat of a downgrade, threatens a reinsurer's ability to attract and retain business because primary insurers typically have requirements for minimum ratings of their reinsurers.

Further requirements for appropriate reserves emanate from jurisdictional law (i.e., state, provincial, or national), the National Association of Insurance Commissioners for U.S. reinsurers, accounting standards such as the U.S. Generally Accepted Accounting Principles (GAAP) and International Financial Reporting Standards (IFRS), and actuarial standards of practice.

This chapter is organized in the following sections:

- Basic reinsurance terminology
- Functions of reinsurance
- Major types of reinsurance
- Reinsurance concepts and contract provisions influencing the estimation of unpaid losses

Basic Reinsurance Terminology

Reinsurance has its own vocabulary, so it is important to start with basic reinsurance terms before a discussion of the functions and types of reinsurance. New terms are shown in bold when defined, which may not be at the term's first use.

Reinsurance is a form of insurance in which the reinsurer, in consideration of a premium, agrees to indemnify the reinsured for part or all of the loss that the reinsured may sustain under the policy or policies that it has issued. The **reinsured**, which is the insurer that cedes its business (i.e., reinsures its liability) with another, is also referred to as the **ceding company**, or the **cedent**. Reinsurance is used by primary insurers, captive insurers, self-insurers, and even by reinsurers. Given the range of organizations that purchase reinsurance, the term ceding company is typically used in this text to refer to those who purchase reinsurance. The **reinsurer** is the insurer that accepts all or part of the insurance liabilities of the ceding company for a stated premium.

In the context of reinsurance, insurers and reinsurers refer to business that is ceded and assumed. For business **ceded**, the risk is transferred from the ceding company to the reinsurer. Ceded insurance policies are referred to as the **subject policies** or the **underlying policies**. In the context of IFRS 17–Insurance Contracts, ceded reinsurance contracts are referred to as **reinsurance contracts held**. A reinsurer **assumes** the business transferred through reinsurance from the insurer.

A reinsurer can transfer risks it has reinsured to another reinsurer through a **retrocession**, which is the reinsuring of reinsurance. In a retrocession, the ceding reinsurer is known as the **retrocedent,** and a **retrocessionaire** is the assuming reinsurer.

When working with data and reporting on financial results, the terms **gross**, **net**, and **ceded** (losses and premiums) have slightly different meanings when used with primary insurers and reinsurers. When used for a primary insurer,

- Gross experience refers to the sum of direct and assumed business,
- Ceded experience refers to business transferred through reinsurance, and
- Net experience is equal to gross less ceded experience.

In a reinsurance context,

- Gross experience refers to assumed business,
- Ceded experience refers to business transferred through retrocessions, and
- Net experience is equal to gross less ceded experience.

In a reinsurance context, the **retention** is the amount of insurance liability or loss that the ceding company retains for its own account after consideration for reinsurance. Depending on the type of reinsurance, the retention can be expressed as a percentage or a dollar amount. The ceding company's retention may also be referred to as the **attachment point**, which is the point at which reinsurance begins to apply.

The **working layer** is a dollar range in which the insurer (or reinsurer) expects relatively predictable loss experience with a fairly high level of loss frequency. The determination of the boundary of a working layer is subjective and depends on an organization's unique risk appetite.[6] A layer that the ceding company determines to be a working layer would typically be different from a layer that a reinsurer determines to be a working layer. Frequently, a ceding company retains losses within its working layer and cedes losses (or a portion of losses) in excess of such a working layer.

Reinsurers often receive data by **bordereau** (plural **bordereaux**) from ceding companies or the brokers of their ceding companies. Bordereau is defined by the International Risk Management Institute (IRMI) as follows:

> Furnished periodically by the reinsured, a detailed report of insurance premiums or losses affected by reinsurance. A premium bordereau contains a detailed list of policies (or bonds) reinsured under a reinsurance treaty during the reporting period, reflecting such information as the name and address of the primary insured, the amount and location of the risk, the effective and termination dates of the primary insurance, the amount reinsured and the reinsurance premium applicable thereto. A loss bordereau contains a detailed list of claims and claims expenses outstanding and paid by the reinsured during the reporting period, reflecting the amount of reinsurance indemnity applicable thereto. Bordereau reporting is primarily applicable to pro rata reinsurance arrangements and to a large extent has been supplanted by summary reporting.[7]

Chapter 2 expands on issues related to reinsurance bordereaux.

The final term to be defined in this section is counterparty default risk, or simply default risk. In a reinsurance context, **counterparty default risk** is the risk that the reinsurer is unable to meet its contractual obligations. In all situations, to the extent that a reinsurer is unable to meet its obligations, the assumed liability falls back to the ceding company who has the contractual relationships with the underlying insured or policyholder.

Functions of Reinsurance

Reinsurance is used to spread risk by transferring some of the risk from the ceding company to the reinsurer or reinsurers. In *Foundations of Casualty Actuarial Science,* Gary Patrik states:

> The nature and purpose of reinsurance is to reduce the financial cost to insurance companies arising from the potential occurrence of specified insurance claims, thus further enhancing innovation, competition, and efficiency in the marketplace. The cession of shares of liability

[6] The IRMI Glossary defines risk appetite as "the degree to which an organization's management is willing to accept the uncertainty of loss for a given risk when it has the option to pay a fixed sum to transfer that risk to an insurer" (see https://www.irmi.com/term/insurance-definitions/risk-appetite.)

[7] Robert Strain, "Reinsurance Terminology Explained: Bordereau and Other Terms of Art," IRMI Expert Commentary, https://www.irmi.com/articles/expert-commentary/reinsurance-terminology-explained-bordereau.

> spreads risk further throughout the insurance system. Just as an individual or company purchases an insurance policy from an insurer, an insurance company may purchase fairly comprehensive reinsurance from one or more reinsurers.[8]

Ceding companies purchase reinsurance for five primary reasons:

- Promote stability.
- Increase capacity.
- Protect against catastrophe.
- Manage capital and solvency margin.
- Access technical expertise.

Promote Stability

Reinsurance is used to help ceding companies stabilize their loss experience within a year and from year to year. Ceding companies typically retain smaller, more predictable claims and cede those claims that are more unusual and infrequent. In this manner, reinsurance can protect the ceding company from shocks associated with large unforeseeable losses. Some ceding companies use reinsurance with relatively low attachment points to provide stability even for losses that are not considered large or unforeseeable. With reinsurance, results can be stabilized by limiting a ceding company's losses following a single event or the accumulation of losses arising from multiple events. By promoting stability, reinsurance can decrease the probability of ruin for a ceding company.

Increase Capacity

Reinsurance expands a ceding company's ability to assume risk by ceding a portion of all its policies or simply its larger policies. Ceding companies often purchase reinsurance to increase their capacity for accepting more business, particularly higher limit policies. For example, assume a large primary insurer was approached to write commercial property insurance for a sports stadium with policy limits of 500 million. Further assume that the primary insurer's risk appetite framework established a net retention of 5 million. Thus, to be able to offer an insurance solution for the stadium, the primary insurer could seek reinsurance from one or more reinsurers to provide the additional 495 million limits of coverage.

The ability for a cedent to offer more capacity on an individual account can be very important, especially for quality accounts that the ceding company might otherwise not be able to write. Furthermore, by providing capacity, reinsurers help facilitate the competition of small insurers with large insurers who, by their nature, can and do generally accept more risk.

[8] Patrik, "Reinsurance," in *Foundations of Casualty Actuarial Science*, 4th ed. (Arlington, VA: CAS, 2001), 344.

Protect Against Catastrophes

Protection from catastrophes, both natural and man-made, is a major purpose of reinsurance. Reinsurance is used to protect ceding companies from a single catastrophic loss event (such as a hurricane or typhoon, earthquake, or wildfire) as well as multiple large loss events (such as multiple hurricanes or typhoons within a single year or a season of multiple wildfires in a single state, province, or country). Reinsurance is also used to protect against casualty losses in which multiple insureds are involved in one occurrence (such as terrorism attacks or vehicle accidents in which many people are injured).

Manage Capital and Solvency Margin

A ceding company can avoid large losses by passing risk to a reinsurer and thus freeing up additional capital. Insurers (including reinsurers) are required by law and regulation to have sufficient capital for potential future claims on all policies written. According to the Insurance Information Institute, "If the insurer can reduce its responsibility, or liability, for these claims by transferring a part of the liability to another insurer, it can lower the amount of capital it must maintain to satisfy regulators that it is in good financial health and will be able to pay the claims of its policyholders."[9]

Through the purchase of some types of reinsurance, a ceding company can accept new risks and avoid the need to raise additional capital. Patrik describes the reinsurance function of managing financial results as follows:

> Reinsurance can alter the timing of income, enhance statutory and/or GAAP surplus, and improve various financial ratios by which insurers are judged. An insurance company with a growing book of business whose growth is stressing their surplus can cede part of their liability to a reinsurer to make use of the reinsurer's surplus. This is essentially a loan of surplus from the reinsurer to the cedant until the cedant's surplus is large enough to support the new business.[10]

Financial results of the ceding company are managed because the ceded commission on the unearned premium reserve transfers statutory surplus from the reinsurer to the cedent. The premium ceded also reduces the ceding company's net premium-to-surplus ratio, referred to as the solvency margin. With a lower premium-to-surplus ratio, the ceding company can write more business.

Access Technical Expertise

An important function of reinsurance is access to the technical expertise of reinsurers, particularly in areas of underwriting, marketing, claims, loss prevention, and pricing. In an IRMI Expert Commentary article on reinsurance, Larry Schiffer states, "Quality reinsurers provide special expertise to their direct

[9] Quoted in Bethan Moorcraft, "Facultative and Treaty Reinsurance: The Differences Explained," *Insurance Business Canada*, June 3, 2019, https://www.insurancebusinessmag.com/ca/guides/facultative-and-treaty-reinsurance-the-differences-explained-168931.aspx.

[10] Patrik, "Reinsurance," 345–46.

insurer clients and assist the direct insurer in providing the best possible protection and risk management for the direct insurer's own clients."[11]

This can be particularly important for small insurers, for whom reinsurers often provide engineering, actuarial, and claims expertise and training. Insurers seeking to enter new lines of business or regions where they do not have experience often turn to reinsurers with market leadership for insight and knowledge. The expertise of reinsurers can be used to help ceding companies explore their underwriting opportunities and ultimately their book of business.

Other Functions of Reinsurance

Reinsurance can be used to facilitate a ceding company's withdrawal from a line of business, geographic area, or a production source. Finally, there are certain market conditions where reinsurance is used for arbitrage when a ceding company believes that additional profits can be garnered by purchasing reinsurance for a value less than the premium the cedent collects from its policyholders.

Different types of reinsurance serve these varied purposes to different degrees.

[11] Schiffer, "Reinsurance Matters," IRMI Expert Commentary, March 2000, https://www.irmi.com/articles/expert-commentary/reinsurance-matters.

Types of Reinsurance

Insurers frequently purchase a variety of reinsurance contracts to serve the functions of stability, capacity, catastrophe protection, financing, and expertise. It is critical for the actuary to understand details of the types of reinsurance used to cede and assume business as there are likely implications on actuarial work, particularly on the data required, the selection of methodology, and the development of assumptions.

An important characteristic of reinsurance contracts is their manuscript nature, whereby reinsurance contracts are developed to meet the specific needs of the ceding company and the reinsurer(s). This is quite different from many primary insurance contracts, particularly personal auto[12] and personal property[13] policies, where the contract is the same for all insureds, with the exception minors such as deductible and policy limits and the use of standard endorsements. Given the tailored nature of reinsurance contracts, it can be challenging to generalize about the types of reinsurance. Thus, it should be understood that exceptions to the material presented in this section are common.

Reinsurance is typically categorized as treaty or facultative and as proportional or non-proportional.

Treaty and Facultative Reinsurance

Treaty Reinsurance

Treaty reinsurance is a type of reinsurance in which the ceding company enters into a contractual agreement with one or more reinsurers to cede all business arising from certain lines of business as specified in the contract. The treaty may span one year or multiple years. In treaty reinsurance, the ceding company agrees to cede and the reinsurers agree to assume all the business written by the ceding company that falls within the terms of the treaty, subject to the limits specified in the treaty. With treaty reinsurance, the reinsurer agrees to accept policies that the ceding company has not yet written to the extent that the risks fall within the treaty's terms.

The most important characteristic of treaty reinsurance is the absence of individual underwriting by the reinsurer. In essence, treaty reinsurance transfers underwriting risk from the ceding company to the reinsurer, leaving the reinsurer exposed to the possibility that the initial underwriting process did not adequately evaluate the risks insured.

Facultative Reinsurance

Facultative reinsurance differs from treaty reinsurance in that a facultative cession is not automatic. The word facultative connotes that both the ceding company and the reinsurer usually have the faculty (i.e., option) of accepting or rejecting the individual submission. Facultative reinsurance is distinguished from

[12] Auto insurance is also referred to as motor and car insurance.

[13] Personal property insurance is also referred to as homeowners, home, and household insurance.

treaty reinsurance where there is an obligation for the cedant to cede a risk or for the reinsurer to accept the ceding risk. In facultative reinsurance, a submission, acceptance, and resulting agreement are required for each individual risk or a defined group of risks that the ceding company wants to reinsure, and the ceding company negotiates an individual reinsurance agreement for each policy it reinsures.

For facultative coverage, a certificate of reinsurance is frequently used. The **certificate** is a record of reinsurance coverage pending replacement by a formal reinsurance contract. With facultative reinsurance, the ceding company can acknowledge acceptance of terms, with the reinsurer's obligation contingent on validity of key information that is stated in the certificate.[14]

The primary purpose of facultative reinsurance is capacity. Facultative contracts can be tailored to the specific circumstances, and thus are typically used for high-value and hazardous commercial risks. Facultative reinsurance has the potential for adverse selection. However, unlike treaty reinsurance, a reinsurer may conduct its own underwriting with facultative reinsurance and thus mitigate the risk of adverse selection.

Examples of Treaty and Facultative Reinsurance

Generalizing about reinsurance is challenging given the tailored nature of most reinsurance contracts. Nevertheless, the following examples help demonstrate common uses of facultative and treaty reinsurance:

- A ceding company maintains property treaty reinsurance for all policyholders with total insured values (TIV) less than 25 million. Reinsurance coverage for all policyholders with TIV of 25 million or more is placed through the facultative market.
- A ceding company maintains casualty treaty reinsurance for automobile risks and uses facultative reinsurance for environmental liability risks.
- A ceding company maintains workers' compensation treaty reinsurance for employers with less than 1,000 employees. Workers' compensation policies for employers with more than 1,000 employees are protected with facultative reinsurance.

For the treaty reinsurance mentioned above, all ceded risks would be subject to the terms and limits of each treaty (i.e., property, casualty, and workers' compensation). For the facultative reinsurance, terms and conditions would be tailored to meet the unique situations of the ceded risks.

Hybrid of Treaty and Facultative Reinsurance

Hybrid contracts, which blend characteristics of treaty and facultative reinsurance, can be used to provide capacity and some degree of stabilization as they can cover many underlying policies. Patrik

[14] "Certificate of Reinsurance," IRMI Glossary, https://www.irmi.com/term/insurance-definitions/certificate-of-reinsurance.

notes that "because of the many special cases and exceptions, it is difficult to make correct generalizations about reinsurance."[15] This is particularly true of hybrid agreements.

The IRMI Glossary contains the following two definitions of hybrid reinsurance arrangements:

> **Facultative Automatic** – a form of property and casualty (P&C) reinsurance that is a hybrid between facultative and treaty. A bordereau of risks ceded is submitted to the reinsurer, which has limited rights to decline individual risks.
>
> **Facultative Obligatory Treaty** – the hybrid between the facultative versus treaty approach. It is a treaty under which the primary insurer has the option to cede or not cede individual risks. However, the reinsurer must accept any risks that are ceded.[16]

Guy Carpenter defines **facultative semi-obligatory treaty** as "a reinsurance contract under which the ceding company may or may not cede exposures or risks of a defined class to the reinsurer, which is obligated to accept if ceded."[17] Finally, Patrik describes **non-obligatory agreements** where "either the cedant may not be required to cede or the reinsurer may not be required to assume every single policy of the specified type."[18]

Given the manuscript nature of most reinsurance contracts, it is incumbent on the actuary working with reinsurance to understand the details of these specialized agreements.

Proportional and Non-Proportional Reinsurance

Both treaty reinsurance and facultative reinsurance can be written on either a proportional or non-proportional basis. Proportional reinsurance is intended to provide capacity and surplus relief to ceding companies, while non-proportional reinsurance is intended to provide stability by protecting the risks insured by the ceding company's losses above a limit.

Proportional reinsurance, which is also known as **pro rata reinsurance** and **participating reinsurance**, is given its name because both premiums and losses (payments and liabilities) are shared between the ceding company and the reinsurers based on the cession percentage. With proportional reinsurance, the reinsurer typically pays a **ceding commission** to the ceding company to reimburse for expenses associated with issuing the underlying policy (e.g., acquisition and underwriting expenses). This commission can be reduced if there is uncertainty about the expected profitability of the business.

---

[15] Patrik, "Reinsurance," 344.

[16] See IRMI Glossary, https://www.irmi.com/term/insurance-definitions/facultative-automatic and https://www.irmi.com/term/insurance-definitions/facultative-obligatory-treaty.

[17] "Facultative Semi-Obligatory Treaty," Guy Carpenter Glossary, https://www.guycarp.com/content/guycarp/en/home/the-company/media-resources/glossary/f.html.

[18] Patrik, "Reinsurance," 347.

Proportional reinsurance is generally quite easy to administer and offers protection to the ceding company against both the frequency and severity of losses. The two types of proportional reinsurance are quota share and surplus share.

Quota Share Reinsurance

With **quota share reinsurance**, the ceding company cedes to the reinsurer an agreed percentage of each risk it insures (i.e., each subject or underlying policy) that falls within the line(s) of business subject to the reinsurance contract. In return, the reinsurer receives a fixed percentage of premium and losses for all risks ceded to the quota share arrangement.

A simplistic example of quota share reinsurance follows. Assume a quota share reinsurance treaty applicable to a single line of business with a cession percentage of 60% (i.e., the ceding company retains 40% and the reinsurer assumes 60%). Table 1. 1 presents the retained and ceded premium and losses for two underlying policies that are subject to the quota share reinsurance.

Table 1. 1. Quota Share Reinsurance Example

Insured	Gross of Reinsurance		Retained (Net of Reinsurance)		Ceded	
	Earned Premium	Ultimate Loss	Earned Premium	Ultimate Loss	Earned Premium	Ultimate Loss
#1	1,000	600	400	240	600	360
#2	1,000	3,000	400	1,200	600	1,800
Total	2,000	3,600	800	1,440	1,200	2,160

The gross, net of reinsurance, and ceded loss ratios are summarized in Table 1. 2.

Table 1. 2. Quota Share Reinsurance Example (Continued)

Insured	Ultimate Loss Ratio		
	Gross	Net of Reinsurance	Ceded
#1	60%	60%	60%
#2	300%	300%	300%
Total	180%	180%	180%

Observe that with quota share reinsurance, the loss ratios (i.e., the losses divided by the premium) are the same for both the ceding company and the reinsurer.

Variable quota share reinsurance is a special form of quota share reinsurance in which the cession percentage varies based on explicit risk characteristics, such as limit, geography, or type of risk.

Typically, but not always, quota share reinsurance is on a treaty basis. Quota share reinsurance usually applies to the ceding company's net retained account (i.e., after deducting all other reinsurance except perhaps excess of loss catastrophe reinsurance), but practices vary.

Surplus Share Reinsurance

With **surplus share reinsurance**, the ceding reinsurer only reinsures losses that exceed the "surplus" amount after the cedant's retention. The ceding company cedes the surplus amount of risk above its retained line subject to a maximum ceded percentage and limit. In surplus share reinsurance, the **line** describes the amount of the ceding company's retained risk; the reinsurer's share is typically expressed as a multiple of the ceding company's retained line. For example, a three-line surplus share treaty provides reinsurance for three times the ceding company's retained liability, enabling the ceding company to write four times as much insurance as was possible before reinsurance. Continuing with a three-line surplus share reinsurance example, assume the following:

- A ceding company wants to write commercial automobile insurance policies to a maximum limit of 10 million per policy, but its risk appetite framework sets a net retention of 2.5 million per policy.
- A three-line surplus share treaty meets the ceding company's objective by providing 7.5 million surplus share reinsurance.
- Losses arising from policy limits of 2.5 million and lower are retained fully by the ceding company.
- For losses arising from policies with limits greater than 2.5 million, the proportion of each loss covered by the surplus share reinsurance is determined by the formula

Proportion Ceded = [Policy Limit – Retained Line] / [Policy Limit].

Table 1. 3 demonstrates the different proportions ceded based on three different insureds with different policy limits assuming each insured incurs a 2.5 million loss.

Table 1. 3. Surplus Share Reinsurance Example

Insured	Policy Limits(M)	Ultimate Loss (M)	Proportion Ceded	Ultimate Loss (M)	
				Retained	Ceded
#1	2.5	2.5	0%	2.5	0
#2	5	2.5	50% = (5 M – 2.5 M) /5 M	1.25	1.25
#3	10	2.5	75% = (10 M – 2.5 M) /10 M	0.625	1.875

Given the different proportions ceded, surplus share reinsurance can be described as variable quota share reinsurance. In her definition of surplus treaty, Ana J. Mata explains the difference between quota share and surplus share reinsurance:

> The main difference between a surplus treaty and quota share reinsurance (or standard proportional reinsurance) is that in a quota share the insurer and the reinsurer share in a fixed proportion each and every risk of the portfolio (losses and premiums), for example, 80% of every risk may be ceded to the reinsurer. In a surplus treaty, the ceding company retains a fixed maximum amount for each risk and this amount defines the retained proportion depending on the total size of the underlying policy. For example, if the retained line is $100 000 per risk, for a $500 000 policy limit the ceding company retains 20%, while for a $200 000 policy limit it retains 50%.[19]

With surplus share reinsurance, the ceding company limits its net exposure to one line regardless of the amount of insurance written. In practice, there are many variations in how surplus share reinsurance operates, with different numbers of lines that may be in separate reinsurance contracts with different reinsurers.

Functions of Proportional Reinsurance

Of the five primary functions of reinsurance described previously, proportional reinsurance is frequently used to manage capital and solvency margins and to increase capacity. In their 2012 CAS Study Note on reinsurance accounting, Ralph Blanchard and Jim Klann present a detailed example of how a quota share reinsurance contract provides surplus relief, and they comment, “Net leverage ratios [written premium-to-surplus] are significantly improved, although ceded reinsurance leverage ratios are significantly increased. Hence, the insurer’s solvency becomes more reliant on its reinsurers’ solvency.”[20]

Ceding companies often use proportional reinsurance to support their need to write larger risks than they are comfortable with (i.e., increase capacity), and surplus share reinsurance does this most effectively. Depending on the cession percentage and the exposure to event or catastrophic risk, proportional reinsurance can also protect against catastrophes.

Non-Proportional Reinsurance

In non-proportional reinsurance, which is also referred to as excess of loss reinsurance, the reinsurer’s response to a loss is determined by the size of the loss. This type of reinsurance is called non-proportional because the premium is not proportional to the limits of coverage. Like proportional reinsurance, non-proportional reinsurance may be written on a treaty or facultative basis.

---

[19] Ana J. Mata, “Surplus Treaty,” in *Encyclopedia of Actuarial Science* (Wiley Online Library, 2006), https://doi.org/10.1002/9780470012505.tas047.

[20] Ralph S. Blanchard III and Jim Klann, “Basic Reinsurance Accounting – Selected Topics” (CAS Study Note, Arlington, VA, 2012), https://www.casact.org/library/studynotes/Blanchard-Klann-Basic-Rein-Accounting.pdf.

Excess of loss reinsurance describes a form of reinsurance that, subject to a specified limit, indemnifies the ceding company against all or a portion of the amount of loss in excess of the ceding company's retention. The main types of excess of loss reinsurance include the following:

- excess per risk
- excess per occurrence and catastrophe
- annual aggregate excess of loss
- clash.

To understand the differences between these types of reinsurance, it is helpful to focus on the **subject loss**, which are the losses that are relevant to the reinsurance cover.

Excess Per Risk Reinsurance

Excess per risk reinsurance, which is also referred to as **excess per policy reinsurance**, is a form of excess of loss reinsurance that, subject to a specified limit, indemnifies the ceding company against the amount of loss in excess of a specified retention with respect to each risk involved in each loss. A "risk" in this form of reinsurance could be the coverage on one building or a group of buildings for fire or flood or the insurance coverage under a single policy that the ceding company treats as a single risk. Excess per risk insurance is typically less exposed than excess per occurrence and catastrophe reinsurance to accumulations of exposures and losses but can still be impacted by natural catastrophes including earthquakes, wildfires, floods, etc.

An example of excess per risk reinsurance is a ceding company that sells property policies with a 10 million limit and maintains excess per risk reinsurance with a 3 million attachment point and reinsurance limit of 7 million. For a loss of 3 million, the ceding company retains the full loss (i.e., there is no coverage from the excess per risk reinsurance). For a 6.5 million loss, the ceding company retains losses of 3 million, and the reinsurer assumes losses of 3.5 million.

Excess per risk reinsurance is primarily used to protect property exposures, although it can be used for casualty lines of business. Like proportional reinsurance, excess per risk reinsurance enables ceding companies to write larger risks (i.e., increase capacity). While some excess per risk treaties have ceding commissions, the expense and surplus relief tend to be less than proportional reinsurance because the premiums tend to be significantly less.

Excess Per Occurrence Reinsurance and Catastrophe Reinsurance

Excess per risk and excess per occurrence are similar in that the ceding company retains the first portion of loss and the reinsurer assumes the excess of the retention, subject to the reinsurance limit.

Excess per occurrence reinsurance differs from excess per risk as it protects a ceding company from an accumulation of losses due to a single occurrence or event. The subject loss in excess per occurrence reinsurance is the sum of all losses arising from an insured event for all subject policies.

Catastrophe reinsurance, which is also referred to as **catastrophe excess of loss** and **catastrophe cover**, is a form of excess of loss reinsurance that, subject to a specified limit, indemnifies the ceding company for the accumulation of losses in excess of a specified retention arising from a single catastrophic event or a series of events. Catastrophe reinsurance protects against property as well as casualty losses that arise due to natural events (e.g., hurricanes and earthquakes) and man-made events (e.g., terrorist attacks and airplane accidents). Catastrophe reinsurance is offered on a worldwide basis as well as in limited regions.

In the event of a loss, which may be a full limit loss or other amount (e.g., 50% of limit) that is specified in the reinsurance contract, most catastrophe reinsurance contracts provide for a reinstatement of the policy limit. A **reinstatement** is the restoration of the policy limit following payment of a full limit loss. One or more reinstatements may be automatic as part of the reinsurance terms or may be available on request. Depending on the terms, the reinstatement may be included with or without additional premium. Premium paid for a reinstatement is referred to as **reinstatement premium**.

It is important for the actuary to track reinstatement premiums separately, as the accounting treatment of reinstatement premiums may differ from other reinsurance premium in that reinstatement premium may be considered earned immediately. Furthermore, reinstatement premium can distort historical relationships between premium and losses and should be recognized in the determination of expected loss ratios, which are critical assumptions for some loss projection techniques.

An example of catastrophe reinsurance is a ceding company that maintains catastrophe reinsurance of 35 million. Assume a flood results in total personal property and commercial property losses of 42 million. The ceding company would retain losses of 35 million, and the reinsurer would assume losses of 7 million.

Example of Excess Per Risk and Catastrophe Reinsurance

It is critically important to understand how multiple reinsurance contracts, both treaty and facultative, interact. In reinsurance, one refers to how a contract inures to the benefit of another. Guy Carpenter’s Glossary of Reinsurance Terms defines **inure to the benefit of** as follows:

> To take effect for the benefit of either the reinsurer or the reinsured. With respect to a given reinsurance contract (usually treaty), other reinsurances which are first applied to reduce the loss subject to the given contract are said to inure to the benefit of the reinsurer of that given contract. If the other reinsurances are to be disregarded as respects loss to the given contract, they are said to inure to the benefit of the reinsured.[21]

[21] “Inure to the Benefit of,” Guy Carpenter Glossary, https://www.guycarp.com/content/guycarp/en/home/the-company/media-resources/glossary/i.html.

An example helps clarify the application of excess per risk reinsurance and catastrophe reinsurance as well as how one contract inures to the benefit of another contract. Assume a ceding company writes 200 personal property policies each with a 2 million limit. Further, assume that the ceding company purchases excess per risk reinsurance with a retention of 1 million and reinsurance policy limit of 1 million. The ceding company also purchases catastrophe reinsurance with a retention of 20 million and reinsurance policy limit of 150 million. The per risk excess reinsurance inures to the benefit of the catastrophe reinsurance. After a major wildfire, the ceding company's total insured losses (prior to any reinsurance) and the losses ceded to the per risk reinsurance are summarized in Table 1. 4

Table 1. 4. Excess Per Risk Reinsurance Example

Individual Losses Expressed as Proportion of 2 Million Policy Limits	Individual Losses Per Policy	# Insureds Suffering Losses	Total Insured Losses	Losses Ceded Excess Per Risk Reinsurance
10%	200,000	35	7 million	0
50%	1 million	10	10 million	0
100%	2 million	5	10 million	5 million

The ceding company's retained losses after the excess per risk reinsurance are 22 million, and the catastrophe reinsurance then applies with a cession of 2 million (22 million minus retention of 20 million). Recall that the ceding company's net retention is 20 million.

The situation would be quite different if all 200 homes were totally destroyed by the wildfire, which is a highly unlikely situation. Nevertheless, the losses for such an event would be as follows:

- Total insured losses of 400 million (200 insureds x 2 million policy limits).
- Total losses ceded to excess per risk of 200 million (200 insureds x 1 million excess per risk policy limits).
- Total losses ceded to catastrophe reinsurance of 150 million.
- Total losses retained by ceding company of 50 million, which are equal to
    - 20 million retention of catastrophe reinsurance, and
    - 30 million of losses above the 150 million policy limit of the catastrophe reinsurance.

If the ceding company were to incur a full limit loss under the catastrophe reinsurance, reinstatement of the policy limit could be very important, especially if the losses were to occur when there is significant time remaining in the contract period.

Annual Aggregate Excess of Loss Reinsurance

Aggregate excess of loss reinsurance, which is also referred to as **aggregate stop-loss reinsurance**, is a form of excess of loss reinsurance that provides the ceding company with a guarantee that their losses will not exceed a predetermined threshold, which can be specified as a percentage of premiums (i.e., loss ratio) or a fixed dollar amount. The reinsurer indemnifies the ceding company for the amount of losses that are greater than a specified aggregate value.

For example, assume a captive insurer writing medical malpractice coverage seeks aggregate excess of loss reinsurance. Alternatives for the aggregate excess of loss reinsurance coverage could include:

- 20% loss ratio excess of the captive's retention of a 90% loss ratio, and
- 10 million limits excess of the captive's retention of 50 million.

Continuing this example, assume the aggregate excess of loss reinsurance is stated in terms of loss ratio and that the captive has subject premium of 10 million. Thus, the aggregate excess of loss reinsurance would provide coverage of 2 million (10 million premium x 20%) excess of 9 million losses (10 million premium x 90%).

Aggregate excess of loss reinsurance generally applies to all or part of the ceding company's net retention and protects net results (i.e., other reinsurance inures to the benefit of the aggregate excess of loss reinsurance), although claims occurring from natural catastrophes may be excluded or have per occurrence limits. For a ceding company seeking to protect its capital, aggregate excess of loss reinsurance best achieves this objective. However, this type of reinsurance is often unavailable and, when available, can be very expensive.

Clash

Clash reinsurance is a casualty reinsurance contract that attaches above all other policy limits. IRMI describes clash coverage as a type of reinsurance that protects a ceding company "from the loss of its normal reinsurance recoveries when it is faced with multiple claims from multiple insureds arising out of the same catastrophe and where its reinsurance does not fully reimburse the insurer for these related losses."[22] The objective of clash coverage is to protect the ceding company burdened by multiple claims arising from exceptional events that are beyond the types of claims contemplated by traditional primary insurance and excess of loss reinsurance policies.

The definition of clash event is a critical aspect of a clash reinsurance contract and varies according to the intentions of the insurer and reinsurer. IRMI notes that the core definition of clash event generally has three components:

[22] Larry Schiffer, "Clash Cover Reinsurance and Economic Catastrophe Losses," IRMI Expert Commentary, March 2009, https://www.irmi.com/articles/expert-commentary/clash-cover-reinsurance-and-economic-cat-losses.

- The loss must arise out of multiple policies held by one insured or similar policies held by multiple insureds.
- All damages are traceable to and the direct consequence of a specific event.
- The event must take place in its entirety within a specific timeframe.[23]

Finite Risk Reinsurance

The Insurance Information Institute describes **finite risk reinsurance** as “a form of reinsurance that specifically incorporates the time value of money. Unlike most reinsurance contracts, finite risk contracts are usually multi-year. In other words, they spread risk over time and generally take into account the investment income generated over the period.”[24]

Finite reinsurance products typically have the following features:

- Risk transfer and risk financing combined in a multi-year contract.
- Emphasis on the time value of money, with investment income explicitly included in the contract.
- Limited assumption of risk by the reinsurer.
- Sharing of the results with the ceding company.[25]

The Insurance Information Institute uses the term **run-off** to refer to a special segment of solutions and products focused on the full-scale transfer of reserve development risks. They state:

> Run-off solutions are tools that address a firm’s earnings volatility arising from past activities. There are a number of special situations that motivate a company to choose a run-off option, like corporate restructuring, mergers & acquisitions, discontinuation of lines of business, erratic changes in the valuation or cost of a liability, or regulatory, accounting or tax changes. The biggest run-off transactions to date in the United States have involved either asbestos & environmental (A&E) or workers’ compensation liabilities. Most transactions have involved insurers, but the economics also work for corporations and captives.[26]

Loss Portfolio Transfers

While most primary P&C insurance contracts are written for a one-year policy term, losses frequently pay out over many years. As a result, insurers hold large loss reserves that are associated with payments in future years for policies written in prior years. At times, insurers want to be relieved of the uncertainty associated with such loss reserves and relief in the capital that must be held for these

---

[23] Schiffer, “Clash Cover Reinsurance.”

[24] “Finite Risk Reinsurance,” Insurance Information Institute, https://www.iii.org/article/finite-risk-reinsurance.

[25] “Finite Risk Reinsurance,” https://www.iii.org/article/finite-risk-reinsurance.

[26] “Finite Risk Reinsurance,” https://www.iii.org/article/finite-risk-reinsurance.

reserves. A **loss portfolio transfer** (LPT) is a form of reinsurance that transfers, at a specified accounting date, from the ceding company to the reinsurer all or a portion of the liability for future loss payments. The IRMI Glossary provides the following definition of an LPT:

> A financial reinsurance transaction in which loss obligations that are already incurred and will ultimately be paid are ceded to a reinsurer. In determining the premium paid to the reinsurer, the time value of money is considered, and the premium is therefore less than the ultimate amount expected to be paid. The cedent's statutory surplus increases by the difference between the premium and the amount that had been reserved. An insurer seeking to withdraw from writing workers' compensation coverage in a given state could, for example, use a loss portfolio transfer to meet its obligations under policies it has written, without the need to continue the day-to-day management of the claims resolution function.[27]

Typically, LPTs are used with long-tail lines of business (such as medical malpractice, asbestos, and pollution liability) where there are significant delays in the reporting of claims and the losses may not be settled for years. Timing is the main element of risk. If claims are settled earlier than expected, then investment income could be lower than anticipated, and the reinsurer could lose money on the contract. In an LPT, the ultimate total nominal losses are usually limited by the finite reinsurance contract.

Adverse Development Cover

An alternative to an LPT is **adverse loss development cover** (or simply **adverse development cover**), where the ceding company receives reimbursement from the reinsurer for losses in excess of a pre-agreed retention level. Unlike an LPT, there is no transfer of loss reserves from the ceding company to the reinsurer providing the adverse loss development cover. Instead, reinsurance is set at the level of the reserves held or at some higher level (often expressed as a multiple) of the held reserves. A key use of adverse development cover is mergers and acquisitions where the ceding company can transfer risks associated with both timing and adverse reserve development.

Reinsurance Concepts and Contract Provisions Influencing the Estimation of Unpaid Losses

Losses-Occurring-During and Risks-Attaching

Given the tailor-made nature of reinsurance contracts, it is critically important that the contract wording appropriately reflects the intent of the parties and that the ceding company and reinsurer fully understand what risks are being reinsured. The **business-covered clause**[28] describes "whether the reinsurance contract is covering risks or policies written by the reinsured that attach to the reinsurance

[27] "Loss Portfolio Transfer (LPT)," IRMI Glossary, https://www.irmi.com/term/insurance-definitions/loss-portfolio-transfer.

[28] This clause is also known as the reinsuring clause, cover clause, business reinsured clause, or the application of agreement clause.

contract or whether losses on policies issued by the reinsured occurring during the life of the reinsurance contract are being reinsured."[29]

There are two primary approaches of reinsurance coverage: losses-occurring-during and risks-attaching (also known as policies-attaching). Losses-occurring-during contracts provide reinsurance coverage for all losses that occur between the contract inception and expiration dates regardless of when the ceding company issued the underlying policy that resulted in the loss. Risks-attaching contracts provide reinsurance coverage only for those policies that incepted during the reinsurance contract effective period; the underlying policies that are covered by risks-attaching reinsurance can have a policy expiration that is later than the expiration date of the reinsurance contract.

For example, assume a ceding company has a property per risk excess of loss reinsurance contract with an attachment point of 2 million and policy limits of 10 million. Further assume that the reinsurance contract is losses-occurring-during with an inception date of January 1, 2020 and expiration date of December 31, 2020.

- A 3 million fire loss that occurred on February 15, 2020 arising from an underlying policy with effective dates of July 1, 2019 to June 30, 2020 would have reinsurance coverage of 1 million (i.e., 3 million total loss less 2 million retention of the ceding company) because the occurrence date of the loss is within the effective period of the reinsurance contract.
- Similarly, a 3 million fire loss that occurred on February 15, 2020 arising from an underlying policy with effective dates of February 1, 2020 to January 31, 2021 would have reinsurance coverage of 1 million.
- A 3 million fire loss that occurred on February 15, 2021 arising from an underlying policy with effective dates of July 1, 2020 to June 30, 2021 would not have reinsurance coverage, because the date of loss (i.e., February 15, 2021) is after the reinsurance contract expiry date of December 31, 2020. This assumes that the reinsurance contract was not renewed or replaced with other applicable coverage.

Next, assume a ceding company has a liability quota share risks-attaching contract with a 60% ceding percentage (i.e., the reinsurer assumes 60% of premium and losses). Further assume that the inception date of the contract is July 1, 2020 and the expiration date is June 30, 2021.

- A 2 million liability loss that occurred on February 15, 2021 arising from an underlying policy with effective dates of June 1, 2020 to May 31, 2021 would not have reinsurance coverage because the underlying policy began before the inception date of the reinsurance contract (i.e., July 1, 2020).

[29] Larry Schiffer, "Understanding the Business-Covered Clause in a Reinsurance Contract," IRMI Expert Commentary, November 2003, https://www.irmi.com/articles/expert-commentary/understanding-the-business-covered-clause.

- A 2 million liability loss that occurred on February 15, 2021 arising from an underlying policy with effective dates of July 15, 2020 to July 14, 2021 would have reinsurance coverage because the inception date of the underlying policy is within the reinsurance contract effective dates.
- A 2 million liability loss that occurred on August 15, 2021 arising from an underlying policy with effective dates of September 1, 2020 to August 31, 2021 would have reinsurance coverage because the underlying policy incepted during the reinsurance contract period even though the loss occurred after the expiry of the reinsurance contract period.

While losses-occurring-during and risks-attaching are the two most common types of reinsurance contracts, coverage can be tailored to meet unique circumstances of the parties to the contract. Thus, it is incumbent on the actuary to understand details of the contract provisions.

Subscription Percentage

Some reinsurance placements are shared by multiple reinsurers through subscription policies. In the context of reinsurance, a **subscription policy** is a reinsurance policy in which multiple reinsurers share the risk associated with providing the reinsurance coverage. Subscriptions can be used when the amount of coverage is more than any one reinsurer is willing to assume and when the primary insurer is seeking to diversify its risk, particularly credit risk. For losses subject to reinsurance placed with multiple reinsurers, it is important that the actuary be aware of the percentage subscribed, as there can be situations in which the full coverage is not placed, and thus the primary insurer would bear responsibility for losses that had been intended for reinsurance.

Commutation Clause

Commutation refers to the cancellation or dissolution of a reinsurance contract. With a commutation, the reinsurer pays funds (at present value) that are not yet due to the ceding company in exchange for full termination of all future obligations related to the reinsurance contract.

Some reinsurance contracts contain a **commutation clause**, also known as a **commutation agreement**, that sets out the terms and conditions for the estimation, payment, and complete discharge of all obligations of the parties to a reinsurance contract. This clause is common in reinsurance contracts covering U.S. workers' compensation and can be optional or mandatory.

Ceding companies use commutations for many reasons. For example, a ceding company may commute a reinsurance contract because it wants to:

- Exit a line of business or geographic region.
- Manage reserves for transfer or sale.
- Avoid the credit risk associated with its reinsurer, particularly if the reinsurer has suffered a ratings downgrade.
- Better manage claims and claims-related expenses and believes that its own staff has the expertise required.

Similarly, reinsurers use commutations for a variety of reasons. For example, a reinsurer may commute a reinsurance contract because it wants to

- Terminate a relationship with a ceding company that is in run-off or one with which it no longer conducts business.
- Protect itself from the potential insolvency of a ceding company.
- Avoid disputes when there are significant differences of opinion with respect to future loss development of subject losses.

Understanding commutations is important for the actuary estimating unpaid losses for several reasons. First, actuaries are frequently involved in the analysis of reinsurance contracts that are subject to commutation. Second, an actuary at a ceding company must be aware of contracts that are commuted, as such affects the estimation of unpaid ceded losses. Similarly, an actuary at a reinsurer must be aware of contracts that are commuted as there is no longer liability associated with such contracts. Finally, actuaries working for both primary insurers and reinsurers should track commuted reinsurance contracts, as the loss development patterns for such contracts could be different from other contracts that remain in force. Thus, actuaries frequently choose to exclude commuted contracts from historical data.

Conclusion

This text is meant to serve as an introduction to reinsurance with a focus on basic reserving methodologies. Reinsurance, which is foundational to a sound global insurance market, can be exceptionally complex. This text is not intended to address these complexities – neither those seen in the commercial market between insurers and reinsurers nor those used within an insurance group through the use of internal reinsurance agreements. Similarly, it is not intended to describe the sophisticated reinsurance arrangements that are frequently created by combining different types of reinsurance with manuscript terms and conditions. Examples and descriptions of complex reinsurance towers can be found readily through internet searches. Instead, the objective is to provide a foundation for the actuary that aids in further study as well as experience working with reinsurance.

Chapter 2 – Data Requirements

This chapter is organized as follows:

- Introduction
- Sufficient and Reliable Data
- Homogeneity and Credibility of Data
- Organization of Data by Experience Period
- Knowledge of Reinsurance Terms and Conditions
- Types of Data
- Sources of Data

Introduction

In *Actuarial Standards of Practice* (ASOP) *23–Data Quality*, the U.S. Actuarial Standards Board (ASB-US) defines data as: "numerical, census, or classification information, or information derived mathematically from such items, but not general or qualitative information. Assumptions are not data, but data are commonly used in the development of actuarial assumptions." [30] The *International Actuarial Standard of Practice (ISAP) Glossary* has a slightly different definition of data and states that data "are usually quantitative but may be qualitative."[31]

Many considerations related to data (quantitative and qualitative) are similar for actuaries working with insurers and those working with reinsurers. Actuaries seek data that are sufficient and reliable. They strive to aggregate data in segments that are homogeneous and credible. They organize data by experience periods that best meet their needs from operational as well as user perspectives. There are important differences, however, in each of these areas as well as in the types and sources of data used by actuaries working in primary insurance versus reinsurance. Many of these issues are explored in this chapter.

Sufficient and Reliable Data

The requirements for sufficient and reliable data for actuarial work are typically set out in actuarial standards of practice. The Canadian actuarial standards of practice describe sufficient and reliable data

[30] ASB-US, *ASOP 23* (revised edition, December 2016), section 2.3, http://www.actuarialstandardsboard.org/wp-content/uploads/2017/01/asop023_185.pdf.

[31] International Actuarial Association, *ISAP Glossary* (November 2019), 2, https://www.actuaries.org/iaa/IAA/Publications/ISAPs/IAA/Publications/05ISAPs.aspx.

as follows: "Data are sufficient if they include the needed information for the work ... Data are reliable if they are sufficiently complete, consistent, and accurate for the purposes of the work."[32]

The International Actuarial Association's *ISAP 1 – General Actuarial Practice* has similar descriptions. *ASOP 23* uses the term appropriate data and defines the term as: "Data suitable for the intended purpose of an analysis and relevant to the system or process being analyzed."[33]

Sufficiency

To determine if data are sufficient for the estimation of unpaid losses, it is helpful to review the key assumptions of the development method, which is one of the most common methods used to project ultimate values. Key assumptions of the development method include the following:

- Losses recorded to date (reported or paid) will continue to develop in a similar manner in the future.
- The relative change in a given year's losses from one evaluation point to the next is similar to the relative change in prior years' losses at similar evaluation points.
- For an immature year, the losses observed to date are valuable for projecting the losses yet to be observed.
- Throughout the experience period, there has been consistency in the mix of business, attachment points and policy limits, and claim processing (which includes the reporting, establishment of case estimates, and settlement of claims).

Ensuring the sufficiency of data can be particularly challenging for actuaries working with reinsurers due in large part to the manuscript nature of many reinsurance contracts, where terms can differ from one ceding company to the next and can change from year to year. Furthermore, operational and strategic changes that were implemented at the ceding companies, the reinsurer, or both can lead to violation of the assumption of consistency in the mix of business, attachment points and limits, and claims processing.

Reliability

With respect to the accuracy of data, the actuary has an obligation to validate the data. *ISAP 1* sets out the following requirements for data validation:

> Data Validation – The actuary should take reasonable steps to review the consistency, completeness, and accuracy of the data used. These might include:

[32] Canadian Institute of Actuaries, *Standards of Practice* (January 2020), Section 1440.04 and .05, https://www.cia-ica.ca/publications/standards-of-practice.

[33] *ASOP 23*, section 2.1.

a. Undertaking reconciliations against audited financial statements, trial balances, or other relevant records, if these are available;
b. Testing the data for reasonableness against external or independent data;
c. Testing the data for internal consistency and consistency with other relevant information; and
d. Comparing the data to those for a prior period or periods.

The actuary should describe this review in any report.[34]

ASOP 23 sets out the following requirements for the review of data:

> A review of data may not always reveal defects. Nevertheless, the actuary should perform a review, unless, in the actuary's professional judgment, such review is not necessary or not practical. In exercising such professional judgment, the actuary should take into account the purpose and nature of the assignment, any relevant constraints, and the extent of any known checking, verification, or audit of the data that has already been performed.[35]

ASOP 23 describes the requirements for the actuary to make a reasonable effort to determine the definition of each data element used in the analysis, to identify questionable data values, and to review prior data.

Actuaries working for reinsurers can face more challenges than those working with primary insurers in the validation of data due to the following:

- For each ceding company and broker reporting on behalf of a ceding company, different it systems that capture different types of data and use different terminology for similar types of data.
- Use of bordereau reporting that can differ (by ceding company and broker) in the types of data reported, the labeling of such data, and the frequency of submission to the reinsurer.
- Lags in reporting related to:
    - The inherent delay in claims that must first be reported to the ceding company before they are reported to the reinsurer;
    - The long-tailed nature of certain types of reinsurance such as excess per risk (where it takes time to know that a specific claim has breached the ceding company's retention) and catastrophe reinsurance (where it can take time before aggregated losses exceed the ceding company's retention); and

[34] International Actuarial Association, *ISAP 1* (December 2018), section 2.5.2, https://www.actuaries.org/iaa/IAA/Publications/ISAPs/IAA/Publications/05ISAPs.aspx.

[35] *ASOP 23*, section 3.3.

  - Bordereau reporting, where losses are only reported on a quarterly or more infrequent basis.
- Gaps in reporting critical information from the ceding companies about claims (including loss payments and case reserves) and claims-management expenses (e.g., investigation, legal, and expert witness expenses).
- Manuscript nature of reinsurance policies that can lead to different coverage for similar loss events with different ceding companies.
- Issues related to data coding for the reinsurer itself.

Nevertheless, the obligations related to using reliable data and validating data that stem from professionalism requirements as well as insurance law and regulation are equally applicable to actuaries working with reinsurers as primary insurers.

Homogeneity and Credibility of Data

Considerations related to the homogeneity and credibility of data are important for all actuaries estimating unpaid losses.

Homogeneity

The term **homogeneous risk group** (**HRG**) used in the European Union's Solvency II Directive is helpful in explaining the key characteristics that underlie the actuary's segmentation of data. HRG is described as:

> Set of (re)insurance obligations which are managed together and which have similar risk characteristics in terms of, for example, underwriting policy, claims settlement patterns, risk profile of policyholders, likely policyholder behaviour, product features (including guarantees), future management actions and expense structure. The risks in each group should be sufficiently similar to allow for a reliable valuation of technical provisions[36] (including a meaningful statistical analysis). The classification is undertaking-specific.[37]

The goal in segmenting data is to improve the robustness of the estimates of unpaid losses by subdividing experience into groups that exhibit similar characteristics. As a result, when separating data into groups for an analysis of unpaid losses, actuaries working for primary insurers and reinsurers focus on similar considerations, such as

---

[36] The term **technical provisions** is used widely outside of the U.S. and Canada. Technical provisions is defined in the International Association of Insurance Supervisors' Glossary as: "The amount that an insurer sets aside to fulfil its insurance obligations and settle all commitments to policyholders and other beneficiaries arising over the lifetime of the portfolio, including the expenses of administering the policies, reinsurance and of the capital required to cover the remaining risks." (see https://www.iaisweb.org/page/supervisory-material/glossary).

[37] Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS), *CEIOPS' Advice for Level 2 Implementing Measures on Solvency II: Technical Provisions – Lines of Business on the Basis of which (Re)Insurance Obligations Are to Be Segmented* (October 2009), section 3.6, https://register.eiopa.europa.eu/CEIOPS-Archive/Documents/Advices/CEIOPS-L2-Final-Advice-Technical-Provisions-Segmentation.pdf.

- Consistency of the coverage triggered by the losses in the group.
- Length of time to report the claim once an insured event has occurred (i.e., reporting patterns).
- Ability to develop an appropriate case outstanding estimate from earliest report through the life of the claim.
- Length of time to settle the claim once it is reported (i.e., settlement, or payment, patterns).
- Likelihood of claim to reopen once it is settled.
- Average settlement value (i.e., severity).
- Volume of losses in the group.

Actuaries strive to determine HRGs in which the claims display similar traits with respect to these characteristics.

Credibility

The goal for the actuary is to divide the data into sufficiently homogeneous risk groups without compromising credibility. The ASB-US's *ASOP 25–Credibility Procedures* defines credibility as: "A measure of the predictive value in a given application that the actuary attaches to a particular set of data (predictive is used here in the statistical sense and not in the sense of predicting the future)."[38] Increasing the homogeneity of the group of data and increasing the volume of data in the group tend to increase credibility. If, however, the actuary divides the data into too many homogeneous groupings, there is a risk that the volume of data in the individual groups becomes insufficient to perform a reliable analysis.

Differences in Considerations Related to Homogeneity and Credibility for Reinsurance versus Insurance

While many of the considerations are similar for actuaries working with primary insurance and reinsurance, there are some important differences. In particular, there are notable differences in how actuaries working with primary insurance and reinsurance segment data. For example, actuaries working with primary insurance frequently aggregate data by line or sub-line of business, as claims within such lines are typically subject to the same or similar laws, policy terms, claims-management expense, etc. For reinsurance, however, there can be important differences within a line of business based on the type of reinsurance contract (e.g., treaty versus facultative and proportional versus non-proportional) that require further segmentation.

Using auto insurance as an example to differentiate reinsurance from primary insurance, an actuary working with a large insurer may have a sufficient volume of credible experience to segment data by the following:

[38] ASB-US, *ASOP 25* (revised edition, December 2013), section 2.1, http://www.actuarialstandardsboard.org/wp-content/uploads/2014/02/asop025_174.pdf.

- Personal lines auto separate from commercial lines auto;
- Jurisdiction (e.g., state, province, or region); and
- Sub-coverage, including:
    - Third-party liability, which may be further separated for bodily injury (BI) and property damage (PD);
    - No-fault benefits (known as personal injury protection, or PIP, in the United States and accident benefits, or AB, in Canada), which may be further separated for various types of benefits including medical and rehabilitation, disability income, funeral, etc.; and
    - Physical damage, which may be further separated for type of coverage, such as collision and comprehensive.

In contrast, an actuary working with a large reinsurer may segment auto reinsurance data by:

- Personal lines auto separate from commercial lines auto.
- Treaty separate from facultative.
- Pro rata separate from excess.
- Aggregate stop-loss and finite risk covers separate from all other segments.

One notable difference with the segmentation for reinsurers when compared to primary insurers is that losses are generally not segmented at a sub-coverage level or jurisdiction level, although a global reinsurer would likely segment data by country or region. Furthermore, a reinsurer may segment excess of loss per risk and excess of loss per occurrence at various attachment points, where a primary insurer may segment losses at alternative limits (e.g., losses limited to 1 million, losses limited to 2.5 million, etc.).

In his chapter on reinsurance, Patrik discusses partitioning the reinsurance portfolio into reasonably homogeneous exposure groups that are relatively consistent over time with respect to the mix of business. For partitioning a reinsurance portfolio, he provides a list of the important variables that affect the pattern of claim report lags to the reinsurer and the development of individual case amounts. Patrik's priority-ordered list includes:

- Line of business (property, casualty, bonding, ocean marine, etc.);
- Type of contract (facultative, treaty, finite or financial);
- Type of reinsurance cover (quota share, surplus share, excess per risk, excess per occurrence, aggregate excess, catastrophe, loss portfolio transfer, etc.);
- Primary line of business for casualty;
- Attachment point for casualty;
- Contract terms (flat-rated, retro-rated, sunset clause, share of loss adjustment expense, claims-made or occurrence coverage, etc.);
- Type of ceding company (small, large, or excess and surplus; and

- Intermediary (i.e., broker).[39]

Patrik notes that it is likely not possible to separate data by all of the above criteria, as the resulting segments would lack sufficient volume to produce credible results. A critical factor in determining how to segment data is related to the credibility of the data. Noting that there is no "typical reinsurer," he nevertheless provides the following example of segmentation for a reinsurer:

- Treaty casualty excess
- Treaty casualty proportional
- Treaty property excess
- Treaty property proportional
- Treaty property catastrophe
- Facultative casualty
- Facultative property
- Surety
- Fidelity
- Ocean marine
- Inland marine
- Construction risks
- Aviation
- Finite or nontraditional reinsurance
- Miscellaneous special contracts, pools, and associations
- Asbestos, pollution, and other health hazard or mass tort claims[40]

A large global reinsurer may further segregate some of the above groups by major region such as Americas, Europe, Asia, and rest of world.

Another consideration regarding the homogeneity and the grouping of data relates to changes in the portfolio. In some circumstances, it may be appropriate to combine data from treaty and facultative reinsurance even if these types of reinsurance typically exhibit different underlying loss patterns. However, if the relative volume of business is changing between these two types of reinsurance and underlying development patterns differ, then the grouping may not be appropriate. *Estimating Unpaid Claims Using Basic Techniques* contains a detailed example of the effect on various projection techniques of analyzing a portfolio where the growth of personal automobile and commercial automobile differ, and the consequence of the changing proportions on the various estimation techniques is significant.

[39] Patrik, "Reinsurance," 443.

[40] Patrik, "Reinsurance," 444.

Organization of Data by Experience Period

For estimating unpaid losses, reinsurers typically rely on aggregation by accident year or underwriting year. Underwriting year is also referred to as treaty year and contract year. In this text, the terms underwriting year and treaty year are used interchangeably.

The requirements for financial reporting as well as internal management reporting and planning are important considerations for selecting an approach to aggregating data. For example, reinsurers operating in the United States and Canada require accident year results for statutory financial reporting. That said, reinsurers may analyze data by treaty year and then use allocation approaches to derive accident year results for statutory financial reporting purposes.

Accident Year Aggregation

Accident year data refer to losses grouped according to the date of occurrence (i.e., the accident date or the coverage triggering event). For example, accident year 2020 consists of all losses with an occurrence date in 2020. Aggregation by accident year is the most common grouping of loss data for the actuarial analysis of unpaid losses for primary insurers. Accident year aggregation is also used extensively by many reinsurers in the United States and Canada because of financial and statistical reporting requirements.

Calendar year earned premiums are used to provide an approximate matching of the losses that occur during the year with the insurance premiums earned by an insurer during the year in which the insurance coverage is effective.

Accident year aggregation has become the accepted norm for P&C insurers (including reinsurers) in the United States and Canada. Accident year grouping is easy to achieve and easy to understand. It represents losses occurring over a shorter time frame than for underwriting year aggregation, implying that ultimate accident year losses should become reliably estimable sooner than those for an underwriting year. Industry benchmarks, including data from the Reinsurance Association of America (RAA) and AM Best, are based on accident year experience. Finally, tracking losses by accident year is valuable when there are changes due to economic or regulatory forces (such as inflation or law amendments) or major loss events (such as atypical weather or a major catastrophe) that can influence loss experience.

A significant disadvantage of accident year aggregation is the potential mismatch between losses and premiums. Accident year aggregation includes losses from policies underwritten and priced at more varied times than underwriting year aggregation.

Underwriting (Treaty) Year Aggregation

Underwriting year data, which is frequently used by European reinsurers and Lloyds of London, refer to losses grouped by the year in which the reinsurance policy became effective (i.e., the year in which the contract was incepted). Underwriting year for reinsurance is similar conceptually to policy year for primary insurance.

Losses arising from an underwriting year can extend over many calendar years. For example, if the reinsurance contract is for a 12-month duration and on a risks-attaching basis, the losses arising from such an underwriting year can extend over three calendar years. Continuing this example, underwriting treaty year 2020 for a reinsurer writing proportional risks-attaching reinsurance contracts refers to all reinsurance policies with beginning effective dates between January 1, 2020 and December 31, 2020. For annual reinsurance policies with a January 1, 2020 effective date, covered policies will have effective dates between January 1, 2020 and December 31, 2020 and thus accident dates between January 1, 2020 and December 30, 2021. For annual reinsurance policies with a December 31, 2020 effective date, covered policies will have effective dates between December 31, 2020 and December 30, 2021 and thus accident dates between December 31, 2020 and December 29, 2022. Thus, for this example, treaty year 2020 includes losses arising from three calendar years.

The primary advantage of underwriting year aggregation is a true match between losses and premiums. Underwriting year experience can be important when underwriting or pricing changes occur, such as

- A shift in attachment points or limits.
- A new emphasis on certain classes of business or regions.
- A change in the types of ceding company.
- An increase or decrease in the price.

All of the above can lead to a significant change in expected loss ratios, and many of the above can lead to changes in loss development patterns.

The primary disadvantage of underwriting year aggregation is the extended time frame. As seen in our previous example, an underwriting year can extend over a 36-month period, generally resulting in a longer time until all the losses are reported and a longer time until the ultimate losses can be reliably estimated. This disadvantage can present challenges in the projection of ultimate losses for the most immature underwriting years where cumulative development factors are highly leveraged and the written premium is not fully earned. (Chapter 3 includes examples of possible solutions to these challenges.) Underwriting year data can also make it difficult to understand and isolate the effect of a single large event such as a major court ruling that changes how insurance contracts are interpreted.

Allocation to Accident Year from Underwriting Year

Reinsurers often use underwriting year aggregation for the development of best estimates of ultimate losses and unpaid losses and rely on accident year aggregation for financial reporting and to track how ultimate losses (i.e., reported losses plus incurred but not reported, IBNR, losses) develop over time.

Actuaries who conduct their analysis of unpaid losses using data aggregated by treaty year may need to allocate results to accident year for financial reporting or other purposes. Allocation processes are typically based on how premium is earned over the contract period.

When the reinsurer receives from the ceding company (or broker) detailed loss data including key dates (such as date of loss and policy effective date), then accurate assignment to accident year or underwriting year can occur. However, there are times, particularly for treaty proportional reinsurance, when such details are not available to the reinsurer. In such situations, the reinsurer would typically use earnings profiles to allocate estimates of unpaid losses to accident year. (See Chapter 3 for a detailed example of earning premium.)

Knowledge of Reinsurance Terms and Conditions

It is critically important that actuaries understand the key terms and conditions of reinsurance programs. This is true for actuaries working with reinsurers and those working with primary insurers with responsibility for estimating the ultimate losses and unpaid losses ceded to reinsurers. For example, actuaries need to know the following:

- Business covered, exclusions, and limitations.
- Ceding percentage for quota share reinsurance.
- Retention (i.e., first line) and number of lines for surplus share reinsurance.
- Retention and limits for excess of loss reinsurance and whether excess insurance is per risk or per occurrence.
- Attachment point and limits for stop-loss reinsurance.
- Treatment of loss adjustment expenses and recoveries (such as salvage and subrogation).

It is common for reinsurance terms and conditions, including ceding percentages and retentions, to change from time to time. Thus, it is the actuary’s responsibility to maintain documentation of historical terms as well as be familiar with current terms. Actuaries work closely with underwriters and claims professionals to ensure knowledge of qualitative information that can influence the estimation of unpaid losses.

Types of Data

Actuaries working with reinsurers typically rely on paid losses, case reserves, and reported losses (i.e., the sum of paid losses and case reserves) as well as written and earned premiums. Case reserves often include the case reserves set by the primary insurer as well as **additional case reserves** (ACR) that are set by the reinsurer. Unlike actuaries working with primary insurers, actuaries working with reinsurers usually do not have access to detailed claim count data nor earned exposure information, such as the number of insured vehicles for auto insurance or number of insured properties for homeowners insurance.

The absence of claim count and exposure data leads to far fewer options for triangle-based diagnostics, as the actuary is not able to calculate triangles of average claim values (e.g., average paid, average case outstanding, and average reported) nor count-based ratio triangles (e.g., ratios of closed-to-reported counts and closed with pay-to-closed counts). Thus, the actuary should turn to other types of investigation, particularly interviews with management of the reinsurer and ceding companies to understand the environment and any changes therein. Chapter 4 of *Estimating Unpaid Claims Using Basic Techniques* includes significant detail about meetings with management to understand the environment and includes sample questions for interviews with senior leaders and the underwriting, claims, data processing, and pricing departments.

Bordereau Reporting

Reinsurers often receive data from ceding companies by **bordereaux**, which Robert W. Strain defined as:

> Furnished periodically by the reinsured, a detailed report of insurance premiums or losses affected by reinsurance. A premium bordereau contains a detailed list of policies (or bonds) reinsured under a reinsurance treaty during the reporting period, reflecting such information as the name and address of the primary insured, the amount and location of the risk, the effective and termination dates of the primary insurance, the amount reinsured and the reinsurance premium applicable thereto. A loss bordereau contains a detailed list of claims and outstanding expenses and paid by the reinsured during the reporting period, reflecting the amount of reinsurance indemnity applicable thereto. Bordereau reporting is primarily applicable to pro rata reinsurance arrangements and to a large extent has been supplanted by summary reporting.[41]

There are numerous challenges associated with bordereau reporting, including how data are cumulated by the ceding company or the broker and absorbed by the reinsurer. There are also issues related to the frequency with which reinsurers receive bordereaux. Bordereaux can be submitted by ceding companies or brokers on a monthly, quarterly, semi-annual, or annual basis. The more infrequent the reporting, the greater the lag in reporting and settlement loss development patterns of the reinsurer.

Ceding companies typically have relationships with multiple reinsurers; similarly, reinsurers work with multiple ceding companies as well as multiple brokers. Each of these companies and brokers will have different IT systems that generate different types of reports. Ceding companies and brokers often struggle to access data from existing systems and extract data in the formats suitable for reinsurers. Similarly, reinsurers have difficulty efficiently and accurately absorbing the data to transform into the format required for actuarial purposes. The creation, distribution, and absorption of data via bordereaux files remains a manually intensive process. Another challenge with bordereau reporting is that the loss detail on a bordereau does not contain near as complete details as are available on the claim files of the ceding company.

[41] Quoted in Larry Schiffer, "Reinsurance Terminology Explained: Bordereau and Other Terms of Art," IRMI Expert Commentary, March 2021, https://www.irmi.com/articles/expert-commentary/reinsurance-terminology-explained-bordereau.

While the insurance industry has made great strides in defining standardized data sets to be used by ceding companies and their reinsurers, the adoption of these data sets has been slow. Even when standardized formats for reporting are used, the issue of data disparity still exists. Many stakeholders have not fully implemented standardized data standards in their IT systems due to the high cost and effort required to update existing systems and the higher priority of other IT transformation initiatives.

Loss Adjustment Expenses

One area that requires the actuary's close attention is the treatment of **loss adjustment expenses** (**LAE**), which are expenses associated with the investigation, management, and settlement of claims. This text uses similar terminology to *Estimating Unpaid Claims Using Basic Techniques*. **Allocated loss adjustment expenses** (**ALAE**) correspond to those costs the insurer (or reinsurer) can assign to a particular claim, such as legal and expert witness expenses. **Unallocated loss adjustment expenses** (**ULAE**) are expenses that cannot be easily allocated to a specific claim. Examples of ULAE include the payroll, rent, and computer expenses for the claims department of an insurer (or reinsurer).

It is important that the actuary working with reinsurance (ceded and assumed) understand the treatment of LAE in reinsurance contracts. Frequently, although not always, ULAE are excluded from reinsurance coverage. For ALAE, there are generally three possible treatments:

1. Included with the claim amount in determining excess of loss coverage, which is a common treatment;
2. Included on a pro rata basis (i.e., the ratio of the excess portion of the loss to the total loss amount determines coverage for ALAE); and
3. Not included in the coverage.

For example, assume a ceding company issues liability policies with limits of 5 million and maintains liability excess per occurrence reinsurance with a retention of 2 million and limits of 3 million. Table 2. 1 presents the primary insurer's loss and ALAE on a gross of reinsurance and ceded basis for three occurrences assuming the three different options for the treatment of ALAE.

Table 2. 1. Examples of ALAE Treatment Under Reinsurance

Occurrence	**Gross of Reinsurance**		**Ceded Loss and ALAE based on Reinsurance Treatment of ALAE**		
	Loss	**ALAE**	**ALAE Included with Loss**	**ALAE Included Pro Rata Basis**	**ALAE Not Included**
#1	2	2	2	0	0
#2	3	2	3	1.67	1
#3	0	3	1	0	0

In this example, the loss and ALAE are each 2 million for occurrence #1. If ALAE are included with the loss amount covered by the reinsurance contract, then the total subject loss is 4 million, of which 2 million is retained by the ceding company and 2 million is assumed by the reinsurer. If ALAE are included on a pro rata basis for occurrence #1, there is no assumption of losses by the reinsurer, as the subject loss (i.e., 2 million) does not exceed the ceding company's retention and there are no losses to enter into a pro rata calculation. Finally, for occurrence #1, if ALAE are not included in the reinsurance contract, then there is no assumption by the reinsurer as the subject loss (i.e., 2 million) does not exceed the ceding company's retention.

For occurrence #2, the loss of 3 million exceeds the ceding company's retention even before consideration of ALAE. If ALAE are included with the loss amount covered by the reinsurance contract, then the total subject loss is 5 million, of which 2 million is retained by the ceding company and 3 million is assumed by the reinsurer. If ALAE are included on a pro rata basis for occurrence #2, there is an assumption of ALAE by the reinsurer as well as losses. The calculation for assumed ALAE (i.e., ALAE ceded to the reinsurer) is equal to:

(1 million loss assumed / 3 million total loss) x 2 million ALAE = 0.67 million ALAE assumed.

If, for occurrence #2, ALAE are not included in the reinsurance contract, then assumed losses by the reinsurer are 1 million, and the ceding company retains 2 million losses and 2 million ALAE.

Finally, for occurrence #3, the sum of the loss of 0 and ALAE of 3 million exceeds the ceding company's retention when ALAE are included. Thus, there is a recovery from the reinsurance of 1 million if ALAE are included with the loss amount covered by the reinsurance contract. Given that there are no losses that exceed the retention, there is no recovery from the reinsurer for ALAE for occurrence #3 if ALAE are covered on a pro rata basis. Finally, if for occurrence #3, ALAE are not included in the reinsurance contract, then assumed losses by the reinsurer are nil, and the ceding company retains the full ALAE of 3 million.

Given the large amounts that can be paid for ALAE, particularly for legal and expert witness fees on liability classes of business such as medical malpractice, asbestos and environmental, and directors and officers, the treatment of ALAE and changes in such treatment over time can influence development patterns and relationships in the data and thus have implications for projections of future losses.

Multiple Currencies

Loss data for some ceding companies may exist in the IT systems in different currencies. For example, global reinsurers aggregate data across U.S. dollars, Canadian dollars, Euros, Japanese yen, Chinese Yuan, etc. Depending on the volume of losses in differing currencies, the actuary may need to adjust the data prior to the analysis. One approach is to separate the data by currency and then combine the data after translating data to a common currency using the appropriate exchange rates at a single point in time; such an approach avoids the influence of fluctuations in exchange rates over time. Another approach can be used when writing catastrophe reinsurance in a region with numerous countries and currencies (e.g., South and Central America) where losses are aggregated based on the ceding company's currency of origin.

Large Losses

It is important for the actuary to be aware of how large losses influence the different projection techniques. The presence of unusually large losses, such as those arising from a natural catastrophe event or a class action suit, can distort some of the methods used for estimating unpaid losses. In these situations, the actuary may choose to exclude the large losses from the initial projection and, at the end of the unpaid loss analysis, add a case-specific projection for the reported portion of large losses and a smoothed provision for the IBNR portion of large losses. Given the nature of reinsurance and in particular coverage on an excess of loss basis, both for individual occurrences and catastrophe events, adjusting data, methodology, and assumptions for large losses can be particularly important for the actuary working with reinsurance. When faced with unusually large losses, reinsurers frequently rely on the expertise of claims adjusters as well as input from catastrophe models to supplement traditional loss development and other basic projection methodologies.

Recoveries

Given that reinsurance is insurance for insurers, recoveries (such as deductibles, salvage, and subrogation) that are applicable to the subject loss generally apply before the cession for both proportional and excess of loss reinsurance. It is important for the actuary working with reinsurance to understand the processes related to the recording of payment and case outstanding for recoverables. Some primary insurers establish a case outstanding net of the deductible, while others do not consider the deductible in the establishment of the case outstanding. Even within the same insurer, practices may vary between lines of business. Similar differences in procedures can exist with respect to the establishment of case outstanding for salvage and subrogation recoveries.

Actuaries working with primary insurers and reinsurers should take care to understand how recoveries are applied, particularly for large property losses that can take time to settle all aspects of the claim, especially business interruption losses than can extend over multiple years. For example, assume the following:

- For calendar year 2019, a primary insurer wrote 10 million limit commercial property policies and maintained commercial property excess per risk reinsurance with a retention of 2 million and limits of 8 million.
- An insured incurred a major fire due to an explosion of the boiler on January 2, 2019, which resulted in property losses as well as substantial business interruption losses for a total loss of 7 million gross of salvage and subrogation recoveries.[42]
- The primary insurer paid losses of 2 million in 2019, 3 million in 2020, and the final 2 million in 2021.
- During 2019, expected salvage recoveries of 0.25 million were received.
- During 2022, the ceding company received an unexpected subrogation recovery from the boiler manufacturer of 1.5 million. At year-end 2019, carried reserves reflected the losses net of salvage but without the subrogation that was received in 2022.

For year-end 2019, the ceding company would report losses net of reinsurance and salvage of 2 million and ceded losses of 4.75 million to the reinsurer (total gross loss of 7 million minus salvage of 0.25 million minus the retention of 2 million). In 2022, the primary company receives the subrogation payment of 1.5 million and would transfer this entirely to the reinsurer. Thus, there is no benefit to the ceding company (or change in financial results on a net of reinsurance basis) of the unexpected subrogation, and the benefit is solely for the reinsurer.

If the total losses net of salvage were only 2.75 million instead of 6.75 million, then a subrogation recovery of 1.5 million would reduce the total value of the claim below the reinsurance retention. Any payments by the reinsurer would be returned, and then the remaining subrogation recovery would accrue to the benefit of the ceding company. In this revised example, the ceding company would report losses net of reinsurance and salvage of 2 million for year-end 2019 and cede losses of 0.75 million to the reinsurer. In 2022, the reinsurer would receive reimbursement of 0.75 million from the unexpected subrogation, and the ceding company would also report favorable development of 0.75 million, the balance of the 1.5 million subrogation recovery.

Challenges with Data for Reinsurer

Influence of Change in Operations and the Environment

The actuary working for a reinsurer can face greater challenges than the actuary working for a primary insurer in understanding the effects of operational changes on the estimation of unpaid losses. This is in part because operational changes can take place at the reinsurer as well as at the ceding companies,

[42] For purpose of this example, assume the loss values are accurate and there is no further development on the claim.

and both can influence the projection of ultimate losses and resulting estimates of unpaid losses. Over the past 20 years, many insurers have instituted significant transformational projects to modernize systems including the implementation of new policy administration and claims administration systems. Many insurers have increased the use of analytics and big data to influence pricing, marketing, and underwriting. These transformational initiatives can affect the operations of the ceding companies, their target markets, how risks are underwritten and how claims are reported and settled, as well as the types of data available. All of these changes can influence the reporting and payment patterns of ceded losses. Similarly, reinsurers have undertaken major transformational initiatives that influence loss reporting and settlement practices.

Further changes arise when ceding companies acquire and divest business (companies and large portfolios), and the actuary needs to understand how such activities affect losses historically and in the future. Finally, actuaries need to understand the legal and economic environments of the ceding companies. For example, major reforms in a large jurisdiction (such as tort reform or product reform in coverages such as automobile or workers' compensation insurance) can have major implications on the loss experience of ceding companies that is passed on to reinsurers.

Other Experience Typically Excluded from Development Analyses

Changes in the operations and environment may lead the actuary to choose to exclude discontinued business (i.e., business in run-off) from the analysis because such data could distort historical patterns and relationships, particularly for more recent years. Discontinued business may not be representative of the portfolio of ongoing business, and thus development patterns and loss ratios, which are key assumptions of basic actuarial techniques, should be selected that reflect the characteristics of the ongoing business. This is true when selecting assumptions for reporting and settlement of losses as well as with frequency and severity of losses (albeit reinsurers often do not have sufficient data to project frequencies and severities). Furthermore, some types of discontinued business (such as asbestos, environmental impairment liability, and abuse) may not be suited to development triangle analyses.

Reporting Lags

As described in Chapter 1, reinsurance is insurance for insurers. Thus, claims must first be reported and investigated by the ceding company before loss data can be reported to the reinsurer. As a result, loss data for reinsurers lag those of the ceding companies, and, at times, the lag can be significant. Delayed reporting is particularly true for excess of loss reinsurance, where there is not only a lag because of the need to report to the primary insurer first but also because these claims often take time for the insurer to realize that the claim may exceed its retention, especially for liability claims.

Reinsurers recognize the challenges associated with lags in reporting and often incorporate reporting requirements in the reinsurance contract. For example, the ceding company may be required to report a claim once it reaches a certain threshold, which may be expressed as a dollar value or a percentage of the ceding company's retention (i.e., the reinsurer's attachment point). Alternatively, a ceding company may be required to report certain types of claims that are known to have a higher likelihood of resulting in large losses (such as an abuse claim or a class action suit) regardless of amount.

Heterogeneity of Contract Wordings

The manuscript nature of reinsurance contracts is mentioned repeatedly in this chapter. Patrik states that the "heterogeneity of contract wordings also means that whenever you are accumulating, analyzing, and comparing various reinsurance data, you must be careful that the reinsurance coverages producing the data are reasonably similar."[43] This concern is true when using internal and external data.

Sources of Data

With respect to sources of data for actuarial work, *ISAP 1* states:

> To the extent possible and appropriate when setting assumptions, the actuary should consider using data specific to the organization or the subject of the actuarial services. Where such data are not available, relevant, or sufficiently credible, the actuary should consider industry data, data from other comparable sources, population data, or other published data, adjusted as appropriate. The data used, and the adjustments made, should be described in any report.[44]

Actuaries working for large reinsurers are typically able to rely on detailed loss and premium data from their own IT systems. Internal data may be based on the experience of an individual reinsurer or aggregated experience from affiliated reinsurers within a group.

Smaller reinsurers, however, can face more challenges with data due to IT limitations as well as limitations in the volume and homogeneity of losses. Thus, actuaries working with small reinsurers often need to seek external data sources. External data can be valuable when analyzing development factors (particularly tail factors), trend rates, and expected loss ratios, as well as when the actuary evaluates and attempts to reconcile the results of various projection methods.

There are not nearly as many external data sources for reinsurance as there are for primary insurance. For reinsurance, actuaries can turn to the following:

- Reinsurance Association of America (RAA)
- *Best's Aggregates & Averages*
- Reports from global brokers, such as Guy Carpenter, Aon, and Willis Towers Watson
- Reports from global reinsurers, such as Swiss Re, Munich Re, and SCOR S.E.
- Other internet searches

---

[43] Patrick, "Reinsurance," 344.

[44] *ISAP 1*, section 2.5.3.

Reinsurance Association of America (RAA)

The RAA is the leading trade association of P&C reinsurers doing business in the United States. Members of the RAA include reinsurance underwriters and intermediaries licensed in the United States and those that conduct business on a cross-border basis. Since 1969, the RAA has published a biannual study of loss development triangles. The RAA study includes historical loss development patterns by accident year for reinsurers writing casualty excess reinsurance for automobile liability, general liability, and medical malpractice. In addition, the RAA study does the following:

- Organizes patterns separately by treaty and facultative business and five ranges of attachment points.
- Presents data of broad historical loss development composites by a cross-section of reinsurers.
- Discusses how loss development patterns have changed over the last few years and suggests possible reasons for those changes.
- Discusses how loss development has varied depending on the circumstances and the nature of the business being considered.[45]

Best's Aggregates & Averages

The data available in *Best's Aggregates & Averages*[46] exemplify the differences in segmentation of insurance and reinsurance data. Schedule P, which contains data for U.S. insurers, separately presents the loss and premium data for major lines of business including three non-proportional reinsurance segments:

- Reinsurance – non-proportional assumed property;
- Reinsurance – non-proportional assumed liability; and
- Reinsurance – non-proportional assumed financial lines.

Schedule P–Part 1 contains 10 years of data sorted by the year in which premiums were earned and losses incurred. The types of data include earned premiums, loss and expense payments and reserves, and salvage and subrogation received and anticipated. Unlike primary insurance, Schedule P–Part 1 for the three reinsurance segments does not include data for the number of reported claims and the number of claims outstanding.

Schedule P–Part 2 contains incurred (which includes sum of paid, case outstanding, and IBNR) net losses and defense and cost containment expenses, and Schedule P–Part 3 contains cumulative paid losses and defense and cost containment expenses. Bulk and IBNR reserves on net losses and defense and cost

---

[45] "Historical Loss Development Study," RAA, https://www.reinsurance.org/ProductDetail.aspx?id=147.

[46] *Best's Aggregates & Averages* is an annual publication that benchmarks the performance of individual insurance companies and insurance groups against industry totals, segments, and composites. The publication includes balance sheet, summary of operations, and annual statement. For further information, see http://www.ambest.com/sales/AggAvg/default.asp.

containment expenses are included in Schedule P–Part 4. The reinsurance triangles include data for 10 accident years and evaluations from 12 to 120 months.

While actuaries working with reinsurers may find some value in the aggregated industry data contained in Schedule P, there are important limitations including but not limited to:

- An experience period of only 10 years, which is typically not long enough for excess of loss reinsurance.
- Segmentation that is not sufficiently refined by major line of business and type of reinsurance.
- The combination of experience that may not reflect targets markets, terms and conditions, and operations of any individual reinsurer.

Reinsurance data that are aggregated by accident year for Schedule P tend to look and behave more like primary insurance data, which is generally not an accurate portrayal of the volatility and long-tail nature of many reinsurance losses. Reinsurance actuaries who rely on data aggregated by treaty year will view data much differently than the lines of business included in Schedule P of the U.S. annual statement.

Internet Searches

Another potential source for external data can be found through online searches of publicly available reinsurer data. Generally, these triangles are presented on a worldwide basis and are highly aggregated by major line of business.

It is important to note that many of the reinsurers who publish triangles based on worldwide consolidated experience state that, in practice, their actuaries review between 50 to 500 separate segments for reserving purposes. One global reinsurer describes the governance process around segmentation and the objective to form segments that are "based on a variety of criteria (proportional basis or not, underlying risks typology, geography, pricing environments, legislative environments)."[47] It is important to recognize that data aggregated across many countries, lines of business, and types of reinsurance would likely not be deemed sufficient without some modification (that should be documented in accordance with professionalism requirements) for actuarial work related to a single reinsurer in a particular jurisdiction.

Shortcomings of External Data

Actuaries need to be aware of the potential shortcomings in the use of external data. While similar considerations apply to actuaries working with primary insurance, the issues are heightened for actuaries working with reinsurance. There is a risk that external data may be misleading or irrelevant due to differences in the following:

[47] *SCOR's Loss Development Triangles and Reserves* (SCOR, December 2010), 9, https://www.scor.com/sites/default/files/2011_trianglesdisclosure.pdf.

- Manuscript wording and terms and conditions, where contracts can vary significantly.
- Mix of assumed business, particularly differences by major industry, region, attachment points, and policy limits.
- Types of reinsurance (e.g., treaty, facultative, proportional, and non-proportional).
- Underwriting processes, including engineering and risk control services.
- Claims management philosophies and processes.
- Coding and IT systems.

Thus, the actuary must carefully evaluate the relevance and value of external data.

Conclusion – Importance of Understanding the Data

In conclusion, it is critically important for actuaries to fully appreciate their obligations with respect to data. Actuaries should understand the different types of data that are inputs to and outputs from the insurer's and reinsurer's information systems. Ceding companies and brokers who report on behalf of ceding companies may use the same term to mean different things. The actuary is responsible for knowing the true meaning of the types of loss data contained in the loss reports and information systems that are used as inputs for the estimation of unpaid losses. The importance of understanding the data is equally applicable to actuaries working with primary insurance and reinsurance.

Chapter 3 – Methods Frequently Used to Estimate Unpaid Losses for Reinsurance

This chapter addresses three of the most frequently used methods for estimating unpaid losses: development, expected, and Bornhuetter-Ferguson methods. The chapter is organized in the following major sections:

- Introductory Comments
- Review of the Development, Expected, and Bornhuetter-Ferguson Methods
- Background About Examples
- Comparison of Age-to-Age Factors and Development Patterns
- Implications of the Volatility in Loss Development Experience
- Quota Share and Stop-Loss Reinsurance Examples

As noted in Chapter 1, it is assumed that readers of this text are knowledgeable about basic reserving including typical data requirements, key assumptions, and the traditional methodologies (such as the development, expected loss, and Bornhuetter-Ferguson techniques). Thus, the focus of this chapter is on differences in reserving for reinsurance versus primary insurance and not on detailed mechanics of the traditional projection techniques.[48]

Introductory Comments

For financial reporting, planning, and risk management purposes, actuaries estimate unpaid losses on a gross, ceded, and net of reinsurance basis. For primary insurers, ceded losses reflect business transferred to reinsurers. For reinsurers, gross losses represent the business they assume, and ceded losses reflect the business that they retrocede. The two basic approaches for determining these three estimates of unpaid losses include the following:

- Projecting ultimate losses and the resulting unpaid losses (i.e., ultimate losses minus paid losses) on a gross of reinsurance basis and net of reinsurance basis, then estimating ceded unpaid losses as the difference; and
- Projecting ultimate losses and the resulting unpaid losses on a gross of reinsurance basis and ceded basis, then estimating net unpaid losses as the difference.

Ceded data often have limited credibility due to a lower volume of losses, higher volatility associated with large claims and catastrophe events, and frequent changes in terms and conditions (such as attachment points, limits, participation percentages, and treatment of ALAE) that result in data that are

[48] For further information, see Friedland, Estimating Unpaid Claims Using Basic Techniques.

not homogeneous. Thus, actuaries typically use the first approach and select development patterns and expected loss ratios, which are key assumptions of the projection methods, gross and net of reinsurance rather than gross and ceded.

To project ultimate values and estimate unpaid losses, actuaries frequently use the development, expected, and Bornhuetter-Ferguson methods.

Review of the Development, Expected, and Bornhuetter-Ferguson Methods

The following descriptions of key assumptions and the major steps of the three projection methods are based on those in *Estimating Unpaid Claims Using Basic Techniques*.

Development Method

Key Assumptions

The distinguishing characteristic of the development method is that ultimate values for each year[49] in the experience period are produced from recorded values assuming that future development is similar to prior years' development. For reinsurers, the development method is used most frequently with reported and paid losses as well as with premiums. The underlying assumption in the development method is that values recorded to date will continue to develop in a similar manner in the future (i.e., the past is indicative of the future).

An implicit assumption in the development technique is that, for an immature year, the losses (or premiums) observed thus far tell the actuary something about the losses (or premiums) yet to be observed. This contrasts with the primary assumption underlying the expected method and the Bornhuetter-Ferguson method, where the unrecorded (unreported or unpaid) losses are based on an *a priori* (or initial) estimate of losses.

Other important assumptions of the development method include consistency throughout the experience period in claim processing, the mix of business (and resulting losses), policy limits, and reinsurance coverage (e.g., retention, participation percentage, and policy limits).

Mechanics

The development method consists of seven basic steps:

1. Compile development data in a development triangle.
2. Calculate age-to-age factors.

[49] For insurers, the "years" are typically accident years. For reinsurers, the years are often treaty (or underwriting) years, although accident years are used by reinsurance actuaries in the United States and Canada due to regulatory financial reporting requirements.

3. Calculate average age-to-age factors.
4. Select development factors for each age-to-age interval.
5. Select tail factor.
6. Calculate cumulative development factors.
7. Project ultimate values.

One of the major differences in projecting ultimate losses for primary insurance and reinsurance is the credibility of the reinsurance data that, as noted previously, tends to be lower for reinsurance due to volume, volatility, and heterogeneity of the data. By their nature, losses associated with excess of loss reinsurance can be substantially more volatile than ground-up losses. This is true for catastrophe coverage as well as reinsurance at high attachment points, where significant frequency of claims is not expected.

Considerations in Selecting Age-to-Age Factors

In *Estimating Unpaid Claims Using Basic Techniques*, there is an important discussion about the characteristics the actuary looks for in the selection of age-to-age factors:

- *Smooth progression of individual age-to-age factors and average factors across development periods.* Ideally, the pattern should demonstrate steadily decreasing incremental development from valuation to valuation, especially in the later valuations. Such decreases are seen in many, although not all, of the examples presented later in this chapter.
- *Stability of age-to-age factors for the same development period.* Ideally, there should be a relatively small range of factors (small variance) within each development interval (i.e., down the columns). The actuary looks for stability within the age-to-age factors themselves as well as within the various averages for the same development period. For both reported and paid losses, the greatest variability in age-to-age factors is typically seen at early age-to-age intervals, where losses are at their most immature state (i.e., when the claims professionals have the least amount of information about the circumstances of the insured event and the potential damages and injuries of claimants). There tends to be much greater volatility in the age-to-age factors for reinsurance when compared with primary insurance and for non-proportional reinsurance when compared with proportional reinsurance, and such differences are seen repeatedly in the examples included in this chapter.
- *Credibility of the experience.* Actuaries generally determine credibility based on the volume and the homogeneity of the experience for a given year and maturity age. If the loss development experience has low credibility because of the limited volume of losses, organizational changes, or other factors, it may be necessary to use benchmark development factors. (See the discussion in Chapter 2 about the use of external data.)
- *Changes in patterns and applicability of the historical experience.* Actuaries determine the appropriateness of historical age-to-age factors for projecting future development based on quantitative and qualitative information regarding changes in the book of business and operations over time. There are numerous reasons why historical development experience may not be appropriate, such as

- Dramatic changes in volume of premiums and claims.
- Presence of large claims that distort the development experience.
- Significant changes in the portfolio that are not captured by trend rates.
- Changes in claims processing that affect the manner in which claims are reserved and paid.

Actuaries also consider the effect of changes in external factors that have not yet manifested themselves in the recorded experience (i.e., reported losses, paid losses, or premiums).

All of these considerations are equally applicable to actuaries working with primary insurance and reinsurance.

Expected Method

The expected method is frequently used when:

- Entering a new line of business or new region.
- Changes in strategy, operations, or the environment that make recent historical loss data irrelevant for projecting future loss activity for a particular cohort of losses.
- The development method is not appropriate for less mature periods because the development factors to ultimate are too highly leveraged.
- Data are unavailable for other methods.

Each of these situations is equally applicable to actuaries working with primary insurance and reinsurance.

Key Assumptions

The key assumption of the expected method is that the actuary can better estimate total unpaid losses based on an a priori estimate than from loss experience observed to date. In certain circumstances, the losses reported to date may provide little information about ultimate losses, especially when compared with the a priori estimate.

Mechanics

The most common approach for estimating expected losses associated with reinsurance is an expected loss ratio multiplied by earned premium. The expected loss ratio is often based on pricing information, industry data, and historical experience adjusted to the conditions of the year under review. In selecting the expected loss ratio, the actuary seeks input from management and considers changes in market conditions, pricing, terms and conditions, underwriting, claims emergence, and other factors that could influence expected ultimate losses.

In addition to the expected loss ratio, actuaries working with primary insurance also use frequency-severity and exposure-loss cost approaches to estimate expected losses. In contrast, actuaries working with reinsurers typically do not have access to detailed claim count and exposure information. For a reinsured, estimating ceded losses can be complicated by reinsurance coverage that spans across multiple lines of business or years, which can complicate the assignment of claim counts and exposure units with losses. Actuaries can also use complex stochastic models to estimate expected losses; such models are outside the scope of this text.

Bornhuetter-Ferguson Method

Actuaries rely on the Bornhuetter-Ferguson method almost as often as they rely on the development method. The Bornhuetter-Ferguson method is essentially a blend of the development and expected methods. In the development method, the actuary multiplies actual losses by a cumulative development factor. This method can lead to erratic, unreliable projections when the cumulative development factor is large because a relatively small swing in reported losses or the reporting of an unusually large loss could result in a very large swing in projected ultimate losses. In the expected method, the unpaid loss estimate is equal to the difference between a predetermined estimate of expected losses and the actual payments. This has the advantage of stability but completely ignores actual results as reported. The Bornhuetter-Ferguson method combines the two methods by splitting ultimate losses into two components: actual reported (or paid) losses and expected unreported (or unpaid) losses. As experience matures, more weight is given to the actual losses and the expected losses become gradually less important.

Key Assumptions

The key assumption of the Bornhuetter-Ferguson method is that unreported (or unpaid) losses will develop based on expected losses. In other words, the losses reported to date contain no information about the amount of losses yet to be reported. This is different from the development method where the primary assumption is that unreported (or unpaid) losses will develop based on reported (or paid) losses to date.

Mechanics

As noted, the Bornhuetter-Ferguson method is a blend of the development and expected methods. The following two formulae represent the reported and paid Bornhuetter-Ferguson methods, respectively:

$$\text{Ultimate Losses} = \text{Actual Reported Losses} + \text{Expected Unreported Losses}$$
$$= \text{Actual Reported Losses} + (\text{Expected Losses}) \times (\%\ \text{Unreported})$$
$$\text{Ultimate Losses} = \text{Actual Paid Losses} + \text{Expected Unpaid Losses}$$
$$= \text{Actual Paid Losses} + (\text{Expected Losses}) \times (\%\ \text{Unpaid})$$

Given that the actual reported and paid losses are both known quantities, the challenge of the Bornhuetter-Ferguson method is to calculate the expected unreported and expected unpaid losses. To complete the Bornhuetter-Ferguson method, the actuary must select loss development patterns and develop an expected loss estimate. The development factors are typically based on the selection of age-to-age factors from the development method applied to the insurer's historical data, but they can also be based on external data.

Further Comments about the Development, Expected, and Bornhuetter-Ferguson Methods

Detailed Calculations

Detailed step-by-step explanations and calculations for the development, expected, and Bornhuetter-Ferguson methods are included in *Estimating Unpaid Claims Using Basic Techniques* and are not repeated in this text. The three methods can be used with reported losses, paid losses, and claim counts, although claim counts are used far less with reinsurance than with primary insurance. In carrying out each of these methods, issues related to the types of data required, considerations regarding the selection of assumptions, and the mathematical steps to project ultimate values are similar for primary insurance and reinsurance.

Differences in Assumptions for Reinsurance and Primary Insurance

While the mechanics for each of the methods are the same for actuaries working with primary insurance and reinsurance, there are important differences in assumptions. For example, for reinsurance:

- For a similar line of business, loss development factors in the earlier maturity age intervals are often higher for reinsurance than for primary insurance due to reporting lags. (See Chapter 2 for further discussion about the drivers of reporting lags in reinsurance). Tail factors can also be higher, particularly for non-proportional reinsurance when compared with primary insurance and for non-proportional when compared with proportional reinsurance for a similar line of business.
- Loss trend factors tend to be higher for excess of loss reinsurance than primary insurance.
- There is often less precision in premium on-level factors that adjust for rate changes. Actuaries working with primary insurance regularly maintain detailed information about historical rate changes by major jurisdiction and line of business, especially where rates are highly regulated. These actuaries use premium on-level factors to adjust historical premiums to current rate levels. The rate change information available for reinsurers can be far more challenging to quantify given the manuscript nature of reinsurance arrangements and the changes in coverage that can occur from year to year. Nevertheless, reflecting rate changes is important when determining expected loss ratios for the expected and Bornhuetter-Ferguson methods for reinsurance.[50]
- In reinsurance, there is more limited use of adjustment factors for changes such as tort and product reform than that seen with primary insurance.

---

[50] For examples of the calculation of premium on-level factors, see chapter 5 of Geoff Werner and Claudine Modlin, *Basic Ratemaking* (CAS, 2016), 64–89, https://www.casact.org/library/studynotes/Werner_Modlin_Ratemaking.pdf.

The use of professional judgment is particularly important for actuaries working in reinsurance. In selecting assumptions, actuaries should consider professionalism requirements as set forth in applicable actuarial standards of practice, which should be reviewed on a regular basis.

Effect of Changes in Currency Exchange Rates

Changes in currency exchange rates often influence how an actuary working with reinsurance aggregates losses in development triangles. Many global reinsurers who aggregate experience on a global basis review triangles at the prevailing exchange rates to prevent distortions in the age-to-age factors arising from fluctuations in currency exchange. This leads to differences in the values within the triangles from analysis to analysis.

An example helps demonstrate the effect of changes in currency exchange on age-to-age factors. Two reported loss development triangles are constructed based on the following assumptions:

- Cumulative reporting loss pattern of 20%, 60%, 90%, and 100% at 12, 24, 36, and 48 months, respectively.
- Ultimate losses of 1 million Euros for accident year 2014 with 20% each for the United States, Canada, Japan, U.K., and the rest of Europe.
- Annual growth in losses for each country of 5%.

The exchange rates at December 31 of each year are used to create the two triangles. In the first triangle, presented in Table 3. 1, reported loss are based on each country's reported losses restated at each maturity age at the currency exchange rate of December 31, 2019.

Table 3. 1. Global Reported Losses Based on Currency Exchange Rates at December 31, 2019

Accident Year	12	24	36	48	60	72
2014	206	618	927	1,030	1,030	1,030
2015	216	649	973	1,082	1,082	
2016	227	681	1,022	1,136		
2017	238	715	1,073			
2018	250	751				
2019	263					

In the second triangle, reported losses are based on the aggregation of reported losses from each country using the exchange rate at December 31 of each year. For example, the reported losses of the United States are adjusted by the triangle of US$-Euro exchange rates seen in Table 3. 2.

Table 3. 2. US$-Euro Exchange Rates

Accident Year	12	24	36	48	60	72
2014	1.21100	1.08660	1.05225	1.19990	1.14550	1.12270
2015	1.08660	1.05225	1.19990	1.14550	1.12270	
2016	1.05225	1.19990	1.14550	1.12270		
2017	1.19990	1.14550	1.12270			
2018	1.14550	1.12270				
2019	1.12270					

Reported losses for each of the other countries are similarly adjusted to produce the global reported loss triangle seen in Table 3. 3.

Table 3. 3. Global Reported Losses Based on Currency Exchange Rates at Each Year-End

Accident Year	12	24	36	48	60	72
2014	200	626	942	977	995	1,030
2015	219	659	924	1,045	1,082	
2016	231	647	987	1,136		
2017	226	691	1,073			
2018	242	751				
2019	263					

Not surprisingly, the age-to-age factors are noticeably different dependent on how losses are adjusted for currency exchange. Table 3. 4 compares the age-to-age factors of the first reported loss triangle with those of the second reported loss triangle.

Table 3. 4. Age-to-Age Factors for Global Reported Losses

Accident Year	12-24	24-36	36-48	48-60	60-72
	Reported Losses Adjusted by Dec 31, 2019 Exchange Rates				
2014	3.00	1.50	1.11	1.00	1.00
2015	3.00	1.50	1.11	1.00	
2016	3.00	1.50	1.11		
2017	3.00	1.50			
2018	3.00				
	Reported Losses by Exchange Rates at Each Year-end				
2014	3.13	1.51	1.04	1.02	1.04
2015	3.01	1.40	1.13	1.04	
2016	2.80	1.53	1.15		
2017	3.05	1.55			
2018	3.11				

Adjusting losses by a common currency exchange rate allows for the true reporting pattern to be seen without distortions from currency exchange. While the example is simplistic, in practice, the process can be complicated. Thus, adjustments to assumed losses for the effect of changes in currency can be extremely difficult and require approximations by the actuary.

Background About Examples

The examples included in this chapter are based primarily on the worldwide aggregated data of the largest reinsurers obtained from internet searches. The data are disguised through additive and multiplicative adjustments applied to reported and paid losses as well as earned premiums. The actual years in the experience period are not identified, in part so that the examples do not become dated with the passage of time. Similarly, the currency and units (i.e., thousands or millions) are not identified. It is not the purpose of this text to evaluate any specific reinsurer's experience but instead to explore common relationships between primary insurance and reinsurance and between different types of reinsurance.

Given that the examples in this chapter are constructed from the aggregated global experience of the world's largest reinsurers, the experience in these examples tends to have far greater stability than what an actuary actually sees when analyzing reinsurance experience by HRG. For financial reporting, reinsurers aggregate their experience into roughly 10 to 20 segments. In the commentary supporting the publicly available financial reports, one reinsurer notes that a single segment in their financial report includes the experience of 40 HRGs. One reinsurer reported that they maintain more than 500 HRGs, and another uses more than 1,000 HRGs for actuarial reserving analyses. Thus, the loss development triangle for a particular HRG for a reinsurer would be expected to have significantly less data with substantially more volatility than the examples of this chapter. It is not unusual for the loss development triangles for some HRGs to have values of nil.

Numeric examples are presented to examine the relationships in development experience for the following:

- Primary insurance and reinsurance for a similar type of business (professional lines, Exhibit I).
- Proportional and non-proportional reinsurance for the same line of business (liability, Exhibit II).
- Reinsurance excluding catastrophe and reinsurance catastrophe (property, Exhibit III).

For each of these examples, detailed exhibits are included at the end of the chapter and organized as follows:

- Sheets 1–4: Reported and paid loss development triangles including data and age-to-age factors, and cumulative development factors.
- Sheet 5: Reporting and payment patterns.
- Sheet 6: Development of expected loss ratios.
- Sheet 7: Projection of ultimate losses using expected method and Bornhuetter-Ferguson method.
- Sheet 8: Estimation of IBNR and total unpaid losses.

Data for the professional lines example are aggregated by accident year, and the data for the liability and property examples are aggregated by treaty year. For these latter two examples, the treaty year premium must be adjusted to reflect earnings at the end of the year when estimating unpaid losses, and details of these calculations are presented later in this chapter and in Sheet 8 of Exhibits II and III. An example of the development of written premium to ultimate is included for liability non-proportional and facultative reinsurance in Exhibit II, Sheet 9.

The development examples in this chapter incorporate several simplifying approaches that are described below.

Average Age-to-Age Factors

Three average age-to-age factors are calculated: simple three years, medial seven years (i.e., average of seven years excluding high and low values), and volume weighted five years. The intent is to present averages from different time periods to demonstrate potential volatility in these averages. In practice, the actuary would select the types of average and the experience periods for averages that reflect the specific circumstances of the insurer or reinsurer, its internal and external environments, and the credibility of the data.

Tail Factors

Tail factors for reported losses are selected based on the maximum of 1.00 and the latest observed factor (e.g., the reported tail factor from 120 months-to-ultimate is based on the maximum of 1.00 and the observed factor from 108-to-120 months). Tail factors for paid losses are derived from a review of the projected ultimate losses using the development method with reported losses for the most mature years. In practice, the actuary would use several approaches to select the tail factor. One approach is to rely on industry benchmark development factors. Another common approach is to fit a curve to the selected or observed development factors to extrapolate the tail factors. Many commercial reserving software programs as well as open-source code have routines for such extrapolation. A more in-depth discussion of tail factors is beyond the scope of this text. Actuaries seeking additional information are referred to actuarial literature available on the CAS web site and the CAS Tail Factors Working Party.

Expected Loss Ratios

The projected ultimate losses using the development method applied to paid and reported losses are shown on the exhibit for the development of expected loss ratios. For these examples, the initial estimates of ultimate losses are based solely on the projections using reported losses. In practice, the actuary would likely consider reported loss and paid loss development projections as well as expected loss ratios from pricing or financial planning and possibly also industry information.

In deriving expected loss ratios, there are no adjustments for loss or premium trend, changes in rate level, the effect of tort reform, or other changes in the claims environment, all of which could be significant. Four averages are calculated (latest three, five, and seven years and latest five years excluding high and low), and the selected expected loss ratio is based on the latest five years. The selected expected loss ratios are then used for the expected and Bornhuetter-Ferguson projections.

For the examples that rely on data aggregated by treaty year, an adjustment is required for premium to reflect earnings through the valuation date.

GL Captive Insurer

Data for the two final examples of this chapter use GL Captive Insurer, which is based on GL Self-Insurer from *Estimating Unpaid Claims Using Basic Techniques*. These examples present the perspective of a ceding company as opposed to the reinsurer.

Comparison of Age-to-Age Factors and Development Patterns

As noted previously, examples are presented to examine the relationships in development experience for the following:

- Primary insurance and reinsurance for a similar type of business.
- Proportional and non-proportional reinsurance for the same line of business.

- Property reinsurance excluding catastrophe and property reinsurance catastrophe.

Primary Insurance and Reinsurance for a Similar Type of Business

The first example, presented in Exhibit I at the end of this chapter, relies on the development data for professional lines of a global insurer that writes primary insurance and reinsurance. The focus is on the volatility of age-to-age factors and the differences in reporting and payment patterns. Greater volatility in age-to-age factors can lead to greater volatility in the indications of expected loss ratios for reinsurance when compared with primary insurance.

For professional lines of business, claim payment and reporting patterns are considered to be medium to long tail in nature for both primary insurance and reinsurance. For the primary insurance, the professional lines HRG includes the following:

- Directors & Officers (D&O) Liability.
- Employment Practices Liability (EPL).
- Fiduciary Liability.
- Crime.
- Errors & Omissions (E&O).
- Cyber Liability.
- Professional Indemnity.
- Other financial insurance related coverages for public and private commercial enterprises, financial institutions, non-profit organizations, and professional service providers.

Professional lines primary business is written predominantly on a claims-made basis.

For the reinsurance, the professional lines HRG includes:

- D&O liability
- EPL
- Medical malpractice
- Professional indemnity
- Environmental liability
- Miscellaneous E&O

D&O liability is a much greater proportion of the reinsurance business than the primary insurance business. For this example, the professional lines liability reinsurance HRG includes both non-proportional and proportional treaties, although the majority of exposures are excess policies. D&O exposures typically attach at higher levels than the rest of the portfolio. Like the primary insurance, the reinsurance is predominantly written on a claims-made basis, and most treaties are written on a risks-attaching basis.

Exhibit I, Sheets 1–4 present reported and paid loss development triangles, age-to-age and average age-to-age factors, and cumulative development factors. Reporting and payment patterns are summarized in Exhibit I, Sheet 5.

Comparison of Volatility in Age-to-Age Factors

The standard deviation and absolute differences of the age-to-age factors are calculated for each age-to-age interval from 12–24 months through 72–84 months as measures of the volatility in the reported and paid loss development. The standard deviation is a measure of the amount of variability (i.e., dispersion) in the age-to-age factors around the average. The absolute difference is equal to the highest age-to-age factor minus the lowest age-to-age factor. Table 3. 5 summarizes these results.

Table 3. 5. Professional Lines
Measures of Variability in the Age-to-Age Factors

	Age-to-Age Interval					
	12-24	24-36	36-48	48-60	60-72	72-84
	Standard Deviation - Reported Age-to-Age Factors					
Insurance	0.50	0.06	0.05	0.07	0.04	0.03
Reinsurance	0.84	0.16	0.14	0.10	0.08	0.12
	Standard Deviation - Paid Age-to-Age Factors					
Insurance	0.73	0.17	0.18	0.10	0.07	0.03
Reinsurance	2.91	0.46	0.19	0.12	0.07	0.04
	Absolute Difference - Reported Age-to-Age Factors					
Insurance	1.763	0.177	0.163	0.189	0.093	0.081
Reinsurance	2.181	0.528	0.379	0.257	0.214	0.263
	Absolute Difference - Paid Age-to-Age Factors					
Insurance	2.167	0.516	0.539	0.274	0.180	0.062
Reinsurance	7.643	1.179	0.568	0.331	0.179	0.080

As expected, there is more volatility seen at the earlier maturity ages with paid losses than with reported losses for both primary insurance and reinsurance due to the longer time frame for claims settlement and thus lower volume of paid loss data. One also readily observes much greater volatility in the age-to-age factors for the professional lines reinsurance when compared with the professional lines primary insurance. In this example, the differences are evident in both the reported loss and paid loss age-to-age factors and extend from 12–24 months through 72–84 months. Greater volatility in age-to-age factors can lead to greater uncertainty in the selection of age-to-age factors and resulting projections of ultimate losses.

Longer Reported and Payment Patterns for Reinsurance versus Primary Insurance

In Exhibit I, Sheet 5, reporting and payment patterns based on the three averages (i.e., simple three, medial seven, and volume weighted five) are shown for professional lines primary insurance and reinsurance. One readily observes longer (i.e., slower) reporting and payment patterns for the reinsurance than the primary insurance. The reasons for longer patterns are related to the lags in reporting that were previously discussed in Chapter 2 and include the need for the claims to first be recognized by the ceding company before they can be reported to the reinsurer, the time required for claims to develop beyond the ceding company's attachment point, and delays associated with bordereau reporting.

It is important to remember that these examples use a very simplistic approach for the selection of tail factors. In practice, the actuary would conduct a much more comprehensive analysis of the potential for losses beyond the experience period, and tail factors for reported and paid losses could be significantly different from the selections in this chapter's examples.

Proportional and Non-proportional Reinsurance for the Same Line of Business

While the previous example compared the volatility in losses for a similar type of business for primary insurance and reinsurance, this next example compares the loss experience for the same line of business. The development triangles included in this section are based on the experience of a global reinsurer for liability proportional treaty reinsurance and liability non-proportional treaty and facultative reinsurance. The focus of this example is on the volatility of age-to-age factors and the ratios of paid-to-reported losses as well as the length of the development patterns. Exhibit II, Sheets 1–4 present the reported and paid loss triangles. Exhibit II, Sheet 5 contains the reporting and payment patterns for liability proportional treaty reinsurance and liability non-proportional treaty and facultative reinsurance.

There are two notable differences in the loss development patterns of this example:

- There is significantly more volatility in the age-to-age factors for the non-proportional treaty and facultative reinsurance than for the proportional treaty reinsurance.
- The cumulative development factors are greater (i.e., longer development patterns) for the non-proportional treaty and facultative reinsurance than for the proportional treaty reinsurance.

Further details about these two observations follow.

Comparison of Volatility in the Age-to-Age Factors of Proportional versus Non-proportional Reinsurance

Table 3. 6 summarizes the standard deviations and absolute differences of the age-to-age factors from 12–24 months through 72–84 months. The greater volatility of the reported and paid losses is readily apparent when comparing the experience of proportional treaty and non-proportional treaty and facultative experience for the liability line of business.

Table 3. 6. Liability Reinsurance
Measures of Variability in the Age-to-Age Factors

	Age-to-Age Interval					
	12-24	24-36	36-48	48-60	60-72	72-84
	Standard Deviation - Reported Age-to-Age Factors					
Proportional	0.16	0.12	0.08	0.06	0.05	0.02
Non-Proportional and Facultative	1.53	0.30	0.15	0.40	0.05	0.07
	Standard Deviation - Paid Age-to-Age Factors					
Proportional	0.83	0.39	0.20	0.10	0.04	0.01
Non-Proportional and Facultative	37.77	0.35	0.39	0.15	0.20	0.11
	Absolute Difference - Reported Age-to-Age Factors					
Proportional	0.499	0.348	0.239	0.176	0.127	0.056
Non-Proportional and Facultative	4.837	0.953	0.420	1.117	0.140	0.163
	Absolute Difference - Paid Age-to-Age Factors					
Proportional	2.627	0.904	0.503	0.283	0.092	0.028
Non-Proportional and Facultative	116.571	1.179	1.110	0.380	0.502	0.250

Longer Reporting and Payment Patterns for Non-proportional versus Proportional Reinsurance

For this reinsurer, longer reporting and payment patterns are readily seen in Exhibit II, Sheet 5 when comparing proportional treaty to non-proportional treaty and facultative reinsurance for liability. This is not unexpected given the delays associated with non-proportional reinsurance and the long-tail nature of liability coverage. The reader is again cautioned about the simplistic process used for selecting tail factors in the examples of this chapter.

Variability in Ratios of Paid-to-Reported Losses

Many actuaries use development triangles for diagnostic purposes so that they can better understand how changes in operations and the external environment influence the loss data. Given the absence of data for claim counts and units of exposure for reinsurance, the ratio of paid-to-reported losses is one of the few triangle diagnostics that an actuary can review.

Examining the consistency of paid losses relative to reported losses is important for testing whether there might have been changes in case outstanding adequacy or in settlement patterns. Because this diagnostic is a ratio, further investigation is required if any changes are observed to determine if the change is occurring in paid losses (i.e., the numerator) or in the case outstanding, which are a critical component of the reported losses (i.e., the denominator). It is important to recognize that the absence of observed change in these ratios does not necessarily mean that changes are not occurring. There could be offsetting changes in both claim settlement practices and the adequacy of case outstanding that result in no change to the ratios of paid-to-reported losses.

Table 3. 7 presents the ratios of paid-to-reported losses for the liability reinsurance example. The two measures of variability are shown for these ratios below each triangle. There is significantly more variability seen at all maturity ages from 12 months through 72 months in the ratios for non-proportional and facultative reinsurance than for proportional reinsurance.

Table 3. 7. Liability Reinsurance
Ratios of Paid-to Reported Losses

Treaty	Ratios Paid-to-Reported Losses as of (months)									
Year	12	24	36	48	60	72	84	96	108	120
	Liability - Proportional									
1	0.22	0.28	0.50	0.54	0.61	0.71	0.78	0.84	0.86	0.85
2	0.20	0.34	0.46	0.51	0.65	0.72	0.77	0.83	0.87	
3	0.18	0.31	0.44	0.52	0.66	0.73	0.81	0.87		
4	0.20	0.34	0.45	0.55	0.66	0.74	0.81			
5	0.20	0.39	0.47	0.62	0.67	0.75				
6	0.20	0.30	0.48	0.60	0.68					
7	0.19	0.29	0.45	0.58						
8	0.20	0.28	0.44							
9	0.18	0.30								
10	0.20									
Std Dev	0.01	0.03	0.02	0.04	0.02	0.02				
Abs Diff	0.044	0.108	0.058	0.108	0.072	0.040				
	Liability - Non-Proportional and Facultative									
1	0.19	0.18	0.36	0.38	0.29	0.34	0.35	0.78	0.81	0.81
2	0.22	0.15	0.32	0.50	0.60	0.61	0.64	0.66	0.74	
3	0.14	0.23	0.36	0.52	0.53	0.54	0.60	0.66		
4	0.04	0.15	0.30	0.44	0.53	0.66	0.68			
5	0.13	0.19	0.32	0.44	0.51	0.71				
6	0.13	0.15	0.31	0.36	0.49					
7	0.18	0.19	0.33	0.49						
8	0.13	0.31	0.34							
9	0.02	0.30								
10	0.26									
Std Dev	0.07	0.06	0.02	0.06	0.11	0.15				
Abs Diff	0.245	0.160	0.061	0.156	0.315	0.371				

The same drivers of greater volatility in age-to-age factors for non-proportional and facultative reinsurance versus proportional reinsurance can drive the greater volatility in ratios of paid-to-reported losses. It is important to recognize that the volatility in the age-to-age factors and the diagnostics can contribute to overall greater uncertainty in the selection of age-to-age factors. This can then lead to uncertainty in the projected ultimate losses derived from the development method. In turn, this can lead to greater uncertainty in projections of ultimate losses from other methods, as they are often dependent on input from the development method.

Premium Development

A written premium development triangle was constructed to demonstrate how reinsurance premiums aggregated by treaty year can develop over time. Premium development is more pronounced for risks attaching reinsurance but also varies from one reinsurer to another depending on the distribution of renewal dates during the year. (See description of underwriting year in Chapter 1.) The ultimate losses for treaty years in which the premium is not fully earned require an adjustment to reflect only the portion of ultimate losses that are associated with occurrences prior to the valuation date. Exhibit II, Sheet 9 presents the premium development triangle, age-to-age factors, cumulative development factors, and projection of ultimate written premium by treaty year.

Concluding Remarks

The greater volatility and longer loss development patterns should not be surprising given that proportional reinsurance attaches on a ground-up basis, whereas non-proportional reinsurance is excess of loss coverage. Furthermore, there are many different types of non-proportional reinsurance, including excess per risk, excess per occurrence, catastrophe cover, and aggregate stop-loss. Each of these types of reinsurance could produce very different development patterns, none of which would be expected to be similar to or as stable as ground-up losses. While this example presents non-proportional treaty and facultative on a combined basis, the actuary would consider whether analysis with more segmented data would be appropriate.

Property Reinsurance excluding Catastrophe and Property Reinsurance Catastrophe

The next example compares the volatility in the age-to-age factors for property reinsurance excluding catastrophe and property reinsurance catastrophe. The property triangles include both treaty and facultative reinsurance, proportional and non-proportional, as well as personal and commercial lines of business. While in practice, these different types of risks would not be combined for detailed actuarial analyses of unpaid losses, the observed relationships are still important for understanding the volatility in this major line of business.

Catastrophe and Large Loss Events

Many actuaries exclude unusually large losses arising from catastrophe and other large loss events from development triangles, as such losses can significantly distort development factors and resulting estimates of unpaid losses. For reinsurers, carried reserves for these types of events tend not to be based on aggregated development analyses but instead on ground-up exposure-based assessments that reflect information provided by ceding companies on a contract-by-contract basis. Actuaries may supplement information from claims professionals with results from catastrophe models, particularly in the time period immediately following a catastrophe event when claims teams may not have access to the affected area.

In this example, losses associated with catastrophe events are included in the development triangle for property catastrophe reinsurance. Observe the tremendous volatility in losses down each column of the reported loss triangle, which is presented in Exhibit III, Sheet 2 and in Table 3. 8. The label "net reported losses" in this example refers to losses that are net of retrocessions.

Table 3. 8. Property Reinsurance Catastrophe – Reported Losses

Treaty	Net Reported Losses as of (months)									
Year	12	24	36	48	60	72	84	96	108	120
1	13,440	30,393	31,135	31,714	32,019	32,358	32,523	32,577	32,482	32,467
2	2,905	4,172	4,024	3,966	3,944	3,910	3,890	3,905	3,914	
3	4,240	6,040	6,416	6,282	6,343	6,715	6,645	6,600		
4	13,080	14,350	16,228	16,786	16,807	16,806	16,742			
5	4,892	9,050	9,448	9,066	8,963	8,912				
6	5,531	44,749	55,431	57,542	59,903					
7	10,150	13,806	14,332	16,540						
8	1,546	4,184	4,211							
9	15,554	18,677								
10	920									

The reported losses at 12 months range from a low of 920 to a high of 15,554; at 24 months, the reported losses range from a low of 4,172 to a high of 44,749. Great variability is seen down each column of the triangle.

The loss development seen in triangles can be distorted by the timing of catastrophe events as well as the wide swings in losses associated with such events. For example, one year may have a catastrophic ice storm in January that is almost fully developed by year-end (i.e., December 31), and the following year may have a late season hurricane that occurs the first week of December. The extent of claims reporting and settlement will be completely different for these two events as of December 31 (i.e., as of 12 months in a development triangle), and thus the loss development seen from 12-to-24 months will be completely different. The situation could be further exacerbated with treaties that are risks-attaching, where catastrophe events associated with a treaty year could occur within a time frame of up to three years. (See discussion of underwriting year in Chapter 2.) This could be a driver of the significant differences from 12-to-24 months for treaty years 6 and 7 (i.e., catastrophe events at significantly different times of the treaty year).

The fundamental assumption of the development method is that the relative change in a given year's losses from one evaluation point to the next is similar to the relative change in prior years' losses at similar evaluation points. This assumption may not always be appropriate for property reinsurance catastrophe.

Comparison of Volatility in Age-to-Age Factors

The reported and paid loss triangles (including age-to-age factors, average age-to-age factors, and cumulative development factors) are seen in Exhibit III, Sheets 1–4. Reporting and payment patterns are seen in Exhibit III, Sheet 5.

As with the prior examples, the standard deviations and absolute differences of age-to-age factors are calculated for each age interval from 12-to-24 months through 72-to-84 months. The measures of variability are shown in Table 3. 9.

Table 3. 9. Property Reinsurance
Measures of Variability in the Age-to-Age Factors

	Age-to-Age Interval					
	12-24	24-36	36-48	48-60	60-72	72-84
	Standard Deviation - Reported Age-to-Age Factors					
Property Reinsurance excluding Catastrophe	0.66	0.05	0.03	0.01	0.00	0.00
Property Reinsurance Catastrophe	2.20	0.09	0.06	0.02	0.03	0.01
	Standard Deviation - Paid Age-to-Age Factors					
Property Reinsurance excluding Catastrophe	2.23	0.09	0.04	0.01	0.02	0.00
Property Reinsurance Catastrophe	6.24	0.12	0.13	0.02	0.04	0.03
	Absolute Difference - Reported Age-to-Age Factors					
Property Reinsurance excluding Catastrophe	1.804	0.162	0.083	0.024	0.011	0.006
Property Reinsurance Catastrophe	6.993	0.274	0.194	0.052	0.067	0.016
	Absolute Difference - Paid Age-to-Age Factors					
Property Reinsurance excluding Catastrophe	7.476	0.233	0.111	0.020	0.040	0.008
Property Reinsurance Catastrophe	19.671	0.355	0.357	0.059	0.082	0.065

The volatility is substantially higher for catastrophe reinsurance than for property excluding catastrophe reinsurance for both reported and paid losses. This is not surprising given the nature of catastrophes, both natural and man-made. Greater variability is also seen in the ratios of paid-to-reported losses that are presented in Table 3. 10.

Table 3. 10. Property Reinsurance
Ratios of Paid-to Reported Losses

Treaty	Ratios Paid-to-Reported Losses as of (months)									
Year	12	24	36	48	60	72	84	96	108	120
	Property Reinsurance excluding Catastrophe									
1	0.28	0.61	0.84	0.91	0.94	0.98	0.99	0.99	0.99	1.00
2	0.30	0.60	0.82	0.90	0.95	0.97	0.97	0.98	0.98	
3	0.26	0.61	0.79	0.90	0.96	0.98	0.99	0.99		
4	0.21	0.65	0.83	0.93	0.96	0.97	0.99			
5	0.26	0.57	0.82	0.92	0.96	0.98				
6	0.33	0.54	0.78	0.91	0.95					
7	0.30	0.64	0.77	0.91						
8	0.28	0.57	0.77							
9	0.32	0.67								
10	0.39									
Std Dev	0.05	0.04	0.03	0.01	0.01	0.01				
Abs Diff	0.188	0.126	0.069	0.037	0.019	0.016				
	Property Reinsurance Catastrophe									
1	0.16	0.68	0.92	0.97	0.97	1.04	0.99	0.99	1.00	1.00
2	0.13	0.65	0.87	0.92	0.95	0.97	0.98	0.98	0.98	
3	0.51	0.74	0.88	0.94	0.95	0.94	0.97	0.98		
4	0.31	0.72	0.80	0.91	0.98	0.98	0.99			
5	0.16	0.65	0.81	0.92	0.96	0.97				
6	0.24	0.62	0.79	0.89	0.91					
7	0.22	0.45	0.63	0.76						
8	0.55	0.61	0.75							
9	0.73	0.83								
10	0.19									
Std Dev	0.21	0.11	0.09	0.07	0.02	0.04				
Abs Diff	0.599	0.388	0.295	0.209	0.065	0.104				

Given the significant volatility evident in the property reinsurance catastrophe loss development triangle, methods that rely on selected age-to-age factors are often not appropriate. Instead, actuaries can turn to catastrophe models and discussions with claims professionals. Catastrophe models can be particularly valuable for catastrophe events that occur close to a financial reporting date in circumstances where an insurer (or reinsurer) has not had time to process many claims. This assumes that the catastrophe event lends itself to reliable catastrophe modeling (such as hurricanes and earthquakes). As time progresses and the insurer (or reinsurer) has time to deploy claims adjusters on site and begin to process claims, the insight from the claims team will be invaluable to the actuary estimating unpaid losses.

Table 3. 11 presents an alternative for the projection of ultimate losses using the development method for property catastrophe reinsurance. In this approach, the losses associated with specific catastrophes

are excluded from the calculation and replaced with estimates derived from interaction with the claims team and review of indications from catastrophe models.

Table 3. 11. Alternative Projection with Adjustments for Large Catastrophes

Treaty Year	Losses at 12/31/10 Reported	Losses at 12/31/10 Paid	Catastrophe Losses at 12/31/10 Reported	Catastrophe Losses at 12/31/10 Paid	Catastrophe Losses at 12/31/10 Estimated Ultimate	Cum Dev Factor at 12/31/10 Reported	Cum Dev Factor at 12/31/10 Paid	Projected Ultimate Losses with Cat Adj Based on Reported	Projected Ultimate Losses with Cat Adj Based on Paid	Projected Ultimate Losses without Cat Adj Based on Reported	Projected Ultimate Losses without Cat Adj Based on Paid
1	32,467	32,438	28,500	28,500	28,500	1.000	1.010	32,465	32,477	32,452	32,762
2	3,914	3,817	-	-	-	0.999	1.010	3,910	3,856	3,910	3,856
3	6,600	6,443	-	-	-	0.997	1.012	6,578	6,520	6,578	6,520
4	16,742	16,563	-	-	-	0.997	1.016	16,696	16,835	16,696	16,835
5	8,912	8,647	-	-	-	0.997	0.994	8,889	8,596	8,889	8,596
6	59,903	54,576	50,000	49,000	50,500	1.007	1.042	60,469	56,309	60,299	56,853
7	16,540	12,558	-	-	-	1.032	1.100	17,062	13,811	17,062	13,811
8	4,211	3,167	-	-	-	1.076	1.297	4,530	4,108	4,530	4,108
9	18,677	15,577	13,000	8,900	20,000	1.244	1.898	27,065	32,670	23,242	29,558
10	920	179	-	-	-	2.988	6.626	2,749	1,186	2,749	1,186
Total	168,886	153,965	91,500	86,400	99,000			180,413	176,368	176,409	174,086

The mathematics of the projected ultimate losses with catastrophe adjustment are as follows:

- [(Reported losses – catastrophe reported losses) x reported cumulative development factor + estimated ultimate catastrophe losses].
- [(Paid losses – catastrophe paid losses) x paid cumulative development factor + estimated ultimate catastrophe losses].

The projected ultimate losses from the standard application of the development method are seen in the last two columns of Table 3. 11. There are notable differences in the indicated IBNR for treaty year 9 between the projections with and without adjustment for catastrophe. Another option that the actuary could consider is deriving separate development patterns from data inclusive and exclusive of years with unusually large catastrophe events.

Implications of Volatility in Loss Development Experience

Greater volatility in age-to-age factors can lead to greater uncertainty in the projections of ultimate losses and the resulting estimates of unpaid losses, not only for projections based on the development method but also projections based on other frequently used methods. Actuaries often use estimates of ultimate losses from the development method for mature years to determine the expected loss ratios used in the expected method. Thus, volatility in the age-to-age factors can result in uncertainty in the projections of the development method, which can lead to uncertainty in the selection of the expected loss ratio. The Bornhuetter-Ferguson method relies on the selected development patterns and the expected loss estimates. Thus, volatility and uncertainty in these can lead to uncertainty in the Bornhuetter-Ferguson projections of ultimate losses. Professional judgment is critically important for actuaries estimating unpaid losses for reinsurance.

The examples continue in Sheets 6–8 of the exhibits at the end of the chapter for:

- Professional lines – primary insurance and reinsurance.
- Liability – proportional treaty reinsurance and non-proportional treaty and facultative reinsurance.
- Property – reinsurance excluding catastrophe and reinsurance catastrophe.

Sheet 6 shows the development of the expected loss ratios. Sheet 7 presents the results of the expected method and the Bornhuetter-Ferguson method with reported and paid losses. Finally, Sheet 8 shows indicated IBNR and total unpaid losses.

Details of the calculations are assumed to be known and thus are not included. (For more information, see *Estimating Unpaid Claims Using Basic Techniques*.) One important difference with primary insurance and reinsurance is the need to earn the premium when analyses are conducted using treaty year data. For the liability and property examples, where data are aggregated by treaty year, the expected loss ratios are developed for the complete treaty year; similarly, ultimate losses are developed for the full treaty year for all years in the experience period. On Sheet 8 of Exhibits II and III, an adjustment is made for the most recent treaty years to reduce ultimate losses for the portion of premium unearned as of the valuation date (i.e., December 31, 10).

Observations

In Sheet 6, where expected loss ratios are selected, the standard deviation and absolute difference of the indicated ultimate loss ratios are calculated for each category of business. Similar to the greater volatility observed in age-to-age factors, greater volatility is also seen in the indicated ultimate loss ratios. Table 3. 12 summarizes the standard deviations and absolute differences for the above examples.

Table 3. 12. Measures of Variability in the Indicated Ultimate Loss Ratios

	Standard Deviation	Absolute Difference
Professional Lines - Primary Insurance	0.04	13%
Professional Lines - Reinsurance	0.14	41%
Liability Proportional Treaty Reinsurance	0.08	23%
Liability Nonproportional Treaty and Facultative Reinsurance	0.14	44%
Property excluding Catastrophe Reinsurance	0.17	51%
Property Catastrophe Reinsurance	0.64	157%

Range of Indicated IBNR and Total Unpaid

Calculations are extended to project ultimate losses with the development method (with reported and paid losses), the expected method, and the Bornhuetter-Ferguson method (also with reported and paid losses). The indicated IBNR and total unpaid losses are then calculated. Indicated IBNR is equal to the projected ultimate losses less total reported losses, and total unpaid losses are equal to the projected ultimate losses less total paid losses.

Sheet 8 presents the projected ultimate losses from each method by year (with adjustment for earning of the premium where losses are aggregated by treaty year) and the indicated IBNR and total unpaid losses resulting from each method on a total all years combined basis.

Not surprisingly, there is a greater range of indicated IBNR as measured by the maximum value minus the minimum value for reinsurance than for primary insurance in the professional lines example, for non-proportional treaty than proportional and facultative reinsurance than for proportional treaty reinsurance in the liability example, and for catastrophe than excluding catastrophe for the property reinsurance example.

Quota Share and Stop-Loss Reinsurance Examples

The final two examples in this chapter are from the perspective of the ceding company (i.e., the reinsured). They expand on the example of GL Self-Insurer found in *Estimating Unpaid Claims Using*

Basic Techniques.[51] For purposes of this reinsurance text, GL Self-Insurer is presented as GL Captive Insurer since captive insurers routinely purchase reinsurance.

Quota Share Reinsurance

Recall that with proportional reinsurance, the reinsurer shares the experience of the ceding company from the ground-up. For quota share, where premiums and losses are shared based on a specified percentage, the age-to-age factors are identical for losses gross of reinsurance, ceded losses, and losses net of reinsurance.[52]

With quota share reinsurance, the ceded losses are equal to gross losses multiplied by the percentage ceded. It is very important to understand the meaning of the percentage cited for quota share reinsurance, as the percentage can be used to refer to the percentage ceded or the percentage retained. The actuary should always seek clarification to ensure proper application of the percentage.

For a ceding company, the estimation of ultimate losses and unpaid losses for a line of business with a quota share reinsurance treaty is often a straightforward calculation. The percentage ceded is applied to the ultimate losses, case reserves, paid losses, and IBNR to determine the losses ceded to the reinsurer. If the percentage ceded remains constant for all years in the experience period, the calculation can be performed on a total basis for all years combined. Frequently, the percentage ceded changes over time, and the calculations are performed by year.

Table 3. 13 presents an example where the quota share reinsurance percentages are assumed to vary by year. (Note “QS” is used in a column heading as an abbreviation for quota share.) For GL Captive Insurer, accident year is equivalent to policy year as there is a single policy with a January 1 effective date. In this example, the quota share percentages are presented as the percentage ceded by GL Captive Insurer.

[51] The reported and paid losses are from Chapter 8 of *Estimating Unpaid Claims Using Basic Techniques, and t*he selected ultimate losses are assumed equal to the reported development projection.

[52] Surplus share reinsurance differs from quota share, and thus differences in age-to-age factors would exist due to the variable nature of the percentage of losses shared in surplus share reinsurance. However, the differences are likely not nearly as pronounced as they are between proportional and non-proportional reinsurance.

Table 3. 13. GL Captive Insurer – Example of the Application of Quota Share Reinsurance from the Ceding Company's Perspective Development of Losses ($000s) Ceded to Quota Share Reinsurance at December 31, 11

	Gross of Quota Share Reinsurance					Ceded to Quota Share Reinsurance					Retained
	Selected	At December 31, 11				At December 31, 11					Ultimate
Accident	Ultimate	Paid	Case	Indicated	QS %		Case		Total		Losses
Year	Losses	Losses	Oustanding	IBNR	Ceded	Paid	Oustanding	IBNR	Unpaid	Ultimate	After QS
1	914	890	10	14	50%	445	5	7	12	457	457
2	1,224	1,170	30	24	50%	585	15	12	27	612	612
3	1,339	1,265	35	39	50%	633	18	20	37	670	670
4	1,892	1,600	200	92	50%	800	100	46	146	946	946
5	1,562	1,200	250	112	40%	480	100	45	145	625	937
6	1,583	1,050	350	183	35%	368	123	64	187	554	1,029
7	2,986	900	1,500	586	30%	270	450	176	626	896	2,090
8	2,509	860	940	709	25%	215	235	177	412	627	1,882
9	2,424	525	975	924	20%	105	195	185	380	485	1,939
10	2,328	750	450	1,128	20%	150	90	226	316	466	1,862
11	1,862	170	430	1,262	15%	26	65	189	254	279	1,583
Total	20,623	10,380	5,170	5,073		4,076	1,395	1,146	2,541	6,616	14,007

The calculations above would likely not be the same for an actuary working with a primary insurer or a reinsurer. For a primary insurer, the calculations can become complicated if the quota share coverage is from a risks-attaching reinsurance treaty with a ceded percentage that changes over time and the reserving analysis of gross results is prepared on an accident year basis. In this situation, the change in the ceded percentage applies based on the policy year of the underlying risks not on the accident year of the insured event. For a reinsurer, there would be numerous quota share treaties in a single HRG with different ceding percentages and different terms and conditions, and thus the previous simple calculation would not be applicable.

Stop-Loss Reinsurance

The example with GL Captive Insurer continues with stop-loss coverage where the quota share arrangement inures to the benefit of the stop-loss coverage. Table 3. 14 presents the results, which are described after the table.

Table 3. 14. GL Captive Insurer – Example of the Application of Stop-Loss Limits from the Ceding Company's Perspective

Accident Year	Retained Ult Losses After QS	Stop-Loss Limit	Retained Ult Losses After QS and Stop Loss	Losses at December 31, 11 Net of Quota Share and Stop Loss: Reported	Paid	Case O/S	IBNR
1	457	750	457	450	445	5	7
2	612	750	612	600	585	15	12
3	670	750	670	650	633	18	20
4	946	750	750	750	750	-	-
5	937	750	750	750	720	30	-
6	1,029	1,500	1,029	910	683	228	119
7	2,090	1,500	1,500	1,500	630	870	-
8	1,882	3,000	1,882	1,350	645	705	532
9	1,939	3,000	1,939	1,200	420	780	739
10	1,862	3,000	1,862	960	600	360	902
11	1,583	3,000	1,583	510	145	366	1,073
Total	14,007		13,034	9,630	6,255	3,376	3,404

The retained ultimate losses after quota share are derived from Table 3. 13 and are equal to ultimate losses gross of quota share minus ultimate losses ceded to quota share. Ultimate losses after quota share can also be calculated as ultimate losses gross of quota share multiplied by 1.0 minus the quota share ceded percentage. Stop-loss limits are assumed for the purpose of this example.

Retained ultimate losses after quota share and stop-loss are calculated as:

Minimum [retained ultimate losses after quota share, stop-loss limit].

Reported and paid losses after quota share and stop-loss are calculated in a similar way. Observe that reported and paid losses for accident year 4 are both capped by the stop-loss limit of 750, and there is nil case outstanding and nil IBNR after quota share and stop-loss. For accident year 5, the reported losses are capped but the paid losses are not, and thus there is case outstanding of 30 net of quota share and stop-loss; however, there is no net IBNR for accident year 5. Similar observations are made for accident year 7, where reported losses are capped by the stop-loss of 1500 but the paid losses are not, and case outstanding are 870 with no IBNR.

In practice, once a ceding company breaches stop-loss coverage, it is not uncommon for the reinsurer to increase the price or the attachment point of stop-loss reinsurance (or both). Depending on market conditions, stop-loss reinsurance can be extremely challenging to secure after the ceding company exceeds its retention on more than one occasion.

In this example, the reported losses for accident year 7 of 2,400 (sum of paid losses of 900 and case outstanding of 1,500) are significantly greater than all other accident years. (See Table 3.13 for details by accident year.) Assume that there is an individual large loss for this accident year with an estimated ultimate value of 500. Further assume that GL Captive Insurer has excess per occurrence reinsurance with an attachment point of 100 that inures to the benefit of the quota share and stop-loss coverages. The ultimate loss gross and net of all reinsurance coverage is calculated as shown in Table 3. 15.

Table 3. 15. GL Self-Insurer – Accident Year Losses Net of Excess Per Occurrence, Quota Share, and Stop-Loss Reinsurance

(1) Selected ultimate loss gross of all reinsurance	2,986
(2) Single large loss	500
(3) Excess per occurrence reinsurance - attachment point	100
(4) Ceded losses to excess per occurrence reinsurer (4) = [(2) - (3)]	400
(5) Ultimate losses net of excess per occurrence reinsurance (5) = [(1) - (4)]	2,586
(6) Quota share ceded percentage	30%
(7) Ultimate losses net of excess per occurrend and quota share reinsurance (7) = [(5) x (1.0 - (6))]	1,810
(8) Stop loss limit	1,500
(9) Ultimate losses net of all reinsurance (9) = minimum [(7), (8)]	1,500

In this example, the loss ceded to the excess per occurrence reinsurance is first removed from the results before the application of the quota share ceded percentage. The ultimate losses net of quota share are then determined with the application of the stop-loss limit as the final step. Stop-loss limits typically apply after all other reinsurance. This form of reinsurance is used to protect the net result of the ceding company.

It is very important for the actuary to have complete details about the types of reinsurance (including attachment points, limits, participation percentages, and treatment of LAE) as well as the order in which different reinsurance contracts are applied. The determination of ceded losses can be a very complex process, and it is critical for the actuary to understand and clearly document the calculations and assumptions.

Conclusion

The estimation of ultimate losses and unpaid losses is a critical task of actuaries working with insurance and reinsurance. While the methods described in this chapter are used extensively, they should not be used mechanically without supplementing with professional judgment. Actuaries should meet regularly with underwriting teams and claims personnel to ensure that as much information as possible is considered before final decisions are made about the reserves to book in financial statements. Without incorporating critical insight from others, results derived from mechanical application of the development, expected, and Bornhuetter-Ferguson methods could produce inappropriate results.

P&C Insurance Company Valuation

Richard Goldfarb, FCAS, CFA, FRM

Original Draft: 2005 (Revised: 2010)

Abstract

This study note was prepared for use on the CAS Exam Syllabus. Its purpose is to describe various valuation approaches presented in introductory finance textbooks and to discuss practical implementation issues that arise when using these methods to value a Property & Casualty insurance company.

The methods described focus on those used by practitioners, including the dividend discount model, the discounted cash flow model using free cash flow, the abnormal earnings model and relative valuation using multiples. Applications of option pricing methods in equity valuation are briefly discussed, including the real options framework.

Acknowledgements

Several reviewers provided comments and suggestions that significantly enhanced the final version of the study note, including Walt Stewart, Wei Chuang, Victor Choi, Derek Jones, Mike Belfatti and Patrick Charles. I'd like to especially thank Emily Gilde for painstakingly verifying and correcting all of the calculations in an earlier draft and improving the content in several areas. Any remaining errors, of course, are my own.

Note Regarding 2010 Revision

The 2010 revision reflects a change to the title of the study note resulting from revisions to the numbering convention used for the CAS exam for which this study note was originally produced. In addition, some typographical errors have been corrected. All other content remains the same as in the 2005 version.

Table of Contents

1. Introduction

This study note was prepared for use on the CAS Exam Syllabus. Its purpose is to describe various valuation approaches presented in introductory finance textbooks and to discuss practical implementation issues that arise when using these methods to value a Property & Casualty insurance company.

2. Summary of Valuation Methods

This section provides a brief overview of several methods used to value the common shareholders' equity of financial and non-financial companies. Discussion of the various practical implementation issues for P&C insurance company valuation will be covered in subsequent sections.

2.1 Dividend Discount Model (DDM)

The DDM is the basic model presented in introductory finance textbooks. The method is based on the premise that the equity value of any firm is simply the present value of *all* future dividends. To apply this methodology, dividend payments are forecasted for all future periods and then discounted to present value using an appropriate (risk-adjusted) discount rate. Alternatively, dividends can be forecasted over a finite horizon and a terminal value can be used to reflect the value of all remaining dividends to be received beyond the explicit forecast horizon.

2.2 Discounted Cash Flow (DCF)

The DCF method is closely related to the DDM approach discussed above. However, rather than forecast and discount the actual dividends, the DCF method focuses on *free cash flow*.

The free cash flow is defined as all cash that could be paid as a dividend, regardless of whether or not it actually will be paid in the period it is generated. Free cash flow is measured net of any amounts required to be reinvested in the firm to maintain its operations and generate growth at the rate assumed in the forecasts.

The implicit assumption in this method is that the free cash flow not paid as a dividend is invested to earn an appropriate (risk-adjusted) return. When an investment earns a fair risk-adjusted rate of return, there is no positive or negative effect on the value of the firm from retaining rather than paying out the free cash flow.

There are two variations of this approach. These variations are referred to as the *Free Cash Flow to the Firm* (FCFF) approach and the *Free Cash Flow to Equity* (FCFE) approach.

- FCFF – In this variation, the focus is on the free cash flow to the entire firm, prior to taking into account any debt payments or tax consequences associated with the debt payments[1]. FCFF thus represents the cash that could be paid to all sources of capital, including both the debtholders and the equity holders. Discounting the FCFF produces a value for the entire firm. The value of the equity portion of the firm is then determined by subtracting the market value of the debt from the total firm value. The ease with which most debt instruments can be valued makes it relatively easy to value the equity portion of the firm using this indirect approach.
- FCFE – In this variation, the focus is on the free cash flows to the equity holders only, as opposed to the free cash flows to the entire firm. The free cash flow to equity, FCFE, therefore represents the cash generated by the firm, over and above its reinvestment and debt financing costs, which *could* be paid to the shareholders of the firm. This is estimated using the same approach used to estimate the FCFF, with the additional step of subtracting the debt payments, net of their associated tax consequences, from the free cash flow to the firm to

[1] Debt payments are deductible for corporate tax purposes.

derive the free cash flow to equity. The resulting valuation thus represents the equity valuation directly by determining the present value of these free cash flows.

An important distinction between the FCFF and FCFE methods is that they each use a different discount rate. The FCFF approach uses a discount rate that reflects the overall risk to both debtholders and equity holders (a so-called weighted average cost of capital); the FCFE approach uses a discount rate that reflects the risk to the equity holders only.

2.3 Abnormal Earnings (AE)

The AE method separates the book value of the firm from the value of the future earnings. The book value of a firm represents the value of the firm's equity assuming that the firm earns only the investors' required return on book value in all future periods. Valuations in excess of book value must therefore be the result of earnings in excess of the investors' required earnings. These earnings in excess of the investors' required earnings are referred to as the "abnormal earnings"[2]. The abnormal earnings in all future periods can be discounted and then added to the current book value to obtain the equity value of the firm.

An important distinction between this method and the DDM and DCF methods discussed earlier is that these latter methods both adjust the accounting-based net income measure into a cash flow measure, such as dividends paid or free cash flow. This translation is done to remove any potential distortions introduced by accounting rules designed to defer the recognition of revenues and expenses.

While it makes sense to unwind accounting distortions, some analysts point out that these distortions eventually unwind themselves. In some cases, using unadjusted accounting values may actually provide a more accurate valuation than would result using "cash flow" figures derived from unwinding certain accounting distortions, especially when applied over finite horizons[3].

Another important distinction between the abnormal earnings approach and the DCF or DDM approaches is that this method focuses on the *source* of value creation – the firm's ability to earn a return on equity in excess of investors' required returns. The DCF and DDM focus only on the effect of this value creation – the firm's ability to pay cash flows to its owners.

2.4 Relative Valuation Using Multiples

One common characteristic of the previously discussed methods is that they all require detailed assumptions regarding revenues, expenses, growth rates, etc. in perpetuity. These assumptions, when taken together, result in forecasts of key valuation variables such as dividends, free cash flows or earnings.

The net effect of all of these assumptions can often be summarized as a "multiple" to be applied to a selected financial measure, such as next-period's earnings, cash flow or book value, which will be demonstrated in more detail later in this study note. When these assumptions regarding revenues, expenses, growth rates, etc. are the same for comparable firms, then a shortcut valuation can be estimated using the multiples calculated from the valuation of these comparable firms. In other words, the firm's equity can be valued *relative* to other firms.

Valuation multiples of comparable firms play an important role in all valuations. Even when the multiples are not being used to perform the primary valuation, the valuation multiples of comparable firms often serve as a critical reasonableness check, indicating whether or not the assumptions driving the DDM, DCF or Abnormal Earnings approaches make sense in the aggregate and whether they differ materially from the assumptions inherent in the valuations of other comparable firms.

---

[2] This method of valuation often appears under a variety of other names, including the "residual income" method or the "economic value added" method. The latter terminology was popularized by consulting firm Stern Stewart in the 1990s as "EVA™" and is a registered trademark of that firm. The more generic term "abnormal earnings" is used in this study note.

[3] See Sougiannis, Theodore and Penman, Stephen H., "A Comparison of Dividend, Cash Flow, and Earnings Approaches to Equity Valuation".

2.5 Option Pricing Theory

In a 1974 paper[4], Robert Merton showed that the equity of a firm could be viewed as a call option on the assets of the firm with a strike price equal to the (undiscounted) value of the liabilities. The equity owners can be thought of as having sold the assets of the firm to the debtholders but have the right to buy back the assets by repaying the face value of the debt on the maturity date.

Using this perspective of equity as a call option, some analysts have attempted to use option pricing formulas such as the Black-Scholes formula, or more typically variations of this formula, to value the equity of a firm.

Although theoretically sound, this approach is difficult to implement. There are numerous practical limitations associated with determining the necessary inputs, accurately reflecting the real-world complexity of many firms' capital structure (e.g. there are often multiple classes of debt with multiple maturity dates), and other issues.

Nonetheless, the theoretical foundation of option pricing has recently proven to be useful in thinking about specific sources of value from so-called *real options*. Some examples of real options include options to expand current operations, options to make follow-on investments, options to abandon projects and other forms of managerial flexibility.

Given this overview of the various valuation approaches, the next section of this study note will discuss their specific application to the valuation of P&C insurance companies.

[4] See Merton, Robert C. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates".

3. Dividend Discount Model (DDM)

3.1 Overview of the DDM

The DDM in many ways serves as the foundation of the other methods that will be covered in this study note. As a result, a relatively detailed explanation is warranted. But given the coverage of this approach in introductory finance textbooks[5], it should be sufficient to simply summarize the key points here.

To begin, one can think of the value of a share of stock as the discounted (present) value of the expected future dividends. Since this definition includes all dividends paid, there is no need to adjust this definition in the case of firms that do not *currently* pay dividends – eventually some dividends will have to be paid, even if they merely represent a liquidating dividend at some distant date.

In symbols,

$$V_0 = \frac{E(Div_1)}{(1+k)} + \frac{E(Div_2)}{(1+k)^2} + \frac{E(Div_3)}{(1+k)^3} + \cdots$$

where, $E(Div_i)$ reflects the expected dividends to be paid at the end of period i and k is the appropriate discount rate (see below).

In the case where dividends are expected to grow (in perpetuity) at a constant rate, g, this can be simplified as:

$$V_0 = \frac{E(Div_1)}{k-g}$$

In the more general case, dividends may be projected over a finite horizon and then assumed to grow at a constant rate in perpetuity beyond that horizon. For example, if a three-year horizon is used, the formula can be written as the present value of each of the next three dividends plus the present value of the *remaining* future dividends beginning in year four. Since the dividends are assumed to grow at a constant rate in perpetuity beginning in year four, the previous formula can be used to represent this value at the end of the third year, which is referred to as the *terminal value*.

The resulting formula in the case of a three year horizon is therefore,

$$V_0 = \frac{E(Div_1)}{(1+k)} + \frac{E(Div_2)}{(1+k)^2} + \frac{E(Div_3)}{(1+k)^3} + \frac{\text{Terminal Value}}{(1+k)^3}$$

$$\text{where, Terminal Value} = \frac{E(Div_4)}{k-g}$$

Before getting into the details of how to estimate the dividends, the growth rates and the appropriate discount rate, consider the following example.

Example 1 – Application of DDM

Assume that as of the end of 2004, the expected dividends for an insurance company are estimated as follows:

Table 1: Estimated Dividends

Year	Expected Dividend
2005	100
2006	120
2007	135
2008	150
2009	165

---

[5] See Bodie, Kane and Marcus (6th Edition), Chapter 18.

From 2009 on, the dividends are expected to grow at a constant rate of 5% per year and the appropriate risk adjusted discount rate is 15%.

The DDM can be used to value of the equity of this firm as of the end of 2004.

The first step is to calculate the PV of each of the first five dividends using the discount rate of 15%. This gives a value of the dividends to be earned during the next five years (excluding the dividends beyond that point) as follows:

$$V_{2005-2009} = \frac{100}{1.15} + \frac{120}{1.15^2} + \frac{135}{1.15^3} + \frac{150}{1.15^4} + \frac{165}{1.15^5} = 434$$

To value the remaining dividends beyond 2009, note that the dividends are expected to grow at a rate of 5% from year 2010 on. This suggests that the 2010 dividend is 165*1.05 = 173.25 and the value *as of the end of 2009* is:

$$V_{2009} = \frac{E(Div_{2010})}{k-g} = \frac{173.25}{.15-.05} = 1{,}732.5$$

This value of 1,732.5 represents the *terminal value* as of the end of the explicit dividend forecast horizon. The present value of this amount as of the end of 2004 is $1{,}732.5/1.15^5 = 861$.

Adding the present value of this terminal value to the present value of the dividends for years 2005 through 2009, the total value of all future dividends is $V_{2004} = 434 + 861 = \$1{,}295$.

3.2 Terminal Value

In the previous example, the dividends from year 2010 on were worth a total of $1,732.5 as of the end of 2009 and had a present value of $861 as of the end of 2004. This *terminal value* beyond the explicit dividend forecast horizon is driven largely by the assumption of 5% perpetual dividend growth beyond 2009. Given the fact that the terminal value represents 66.5% of the total value of the firm's equity, it is important to consider these terminal value assumptions carefully.

For convenience, the terminal value as of the end of 2009 can be expressed as:

$$\text{Terminal Value} = \frac{173.25}{.15-.05} = 165 * \frac{1.05}{.15-.05} = 165 * 10.5 = Div_{2009} * 10.5 = 1{,}732.5$$

In other words, the terminal value at the end of 2009 is worth "10.5 times the 2009 dividend". This suggests treating 10.5 as a *multiple* to be applied to the current dividend amount as of the terminal date. This multiple effectively summarizes in one number the net effect of the following assumptions:

i) Dividends will grow at a constant rate forever;

ii) The growth rate is 5%;

iii) The appropriate discount rate is 15%.

3.3 Application of the DDM

The following three key assumptions are required to implement the DDM:

- Expected Dividends During Forecast Horizon
- Dividend Growth Rates Beyond Forecast Horizon
- Appropriate Risk-Adjusted Discount Rate

Each of these assumptions will be discussed in more detail in this section.

3.3.1 Expected Dividends During Forecast Horizon

Forecasting expected future dividends is a complex exercise with a substantial degree of uncertainty. Fundamentally, this will involve forecasts of revenues, expenses, investment needs, cash flow needs and other values for several future periods. These forecasts will require careful consideration of prior business written, expected renewals and new business written.

For the sake of brevity, this study note will assume that such forecasts have already been performed. The models used for these forecasts will not be discussed here. For a detailed discussion of the process one might follow to prepare these forecasts for a generic firm, see *Business Analysis & Valuation*, by Palepu, Bernard and Healey. For a more focused discussion of how this could be done for a P&C insurance company, see *The Application of Fundamental Valuation Principles to Property/Casualty Insurance Companies*, by Blackburn, Jones, Schwartzman and Siegman or *Using the Public Access DFA Model: A Case Study* by D'Arcy, Gorvett, Hettinger and Walling.

3.3.2 Dividend Growth Rates Beyond Forecast Horizon

Estimates of growth rates for revenues, expenses and other variables are inherently part of the process of estimating dividends during the forecast horizon.

Beyond the explicit forecast horizon though, growth rates used in the terminal value calculation are more difficult to determine. One simple approach is to use the growth rates during the explicit forecast horizon to extrapolate the future growth rates.

Another approach is to base the growth rate on the dividend payout ratio, representing the portion of earnings paid as dividends[6], and the return on equity, which represents the profit per dollar of reinvested earnings. This reflects the fact that growth in earnings, and hence dividends, is driven by the retention of some portion of the current period's earnings so that they can be reinvested to generate additional future period income.

Typically, the term *plowback ratio* is used to refer to that portion of earnings retained and reinvested in the firm and the firm's return on equity (ROE) is often used to indicate the income generated from such reinvestment. Combining these, the growth rate, *g*, is estimated as:

$$g = plowback * ROE$$

The assumed growth rate plays a significant role in the ultimate valuation, particularly due to its impact on the terminal value estimate. When estimating the terminal value, the growth rate should reflect the steady-state <u>perpetual</u> growth rate and should not reflect any bias resulting from higher than normal short-term growth estimates. For instance, a growth rate in excess of the growth rate for the entire economy should be assessed carefully, as this implies the firm's share of the total economy will eventually rise to unreasonable levels.

It is important to recognize that high growth rates do not necessarily increase the value of the firm. If all other assumptions were held constant, then mathematically this would be the case. However, assumptions about growth rates, dividend payout rates and the risk-adjusted discount rate cannot be made independently of each other. For instance, simultaneously high growth rates and high dividend payout rates are unlikely to be sustainable and so the effects of high growth rates are likely to be offset by lower dividend amounts.

Additionally, the dividend payments for firms with high growth rates are likely to be riskier (in a systematic risk sense) than those of firms with low growth rates. The high growth firms often depend upon a favorable economic climate for their growth, which introduces more systematic risk. As a result, the effects of high growth rates are likely to be offset by discounting the dividends to present value using higher risk-adjusted discount rates.

3.3.3 Appropriate Risk-Adjusted Discount Rate

A key element of the previous example is the appropriate discount rate to use in the calculation of the present value of the expected cash flows. An entire study note could be devoted to this topic alone. Some of the most important issues associated with the choice of discount rates will be discussed here; additional details are available from various sources contained in the References section of the paper[7].

---

[6] Since stock buybacks are economically equivalent to large cash dividends, these should be included in any reference to "dividends" in the text.

[7] See, in particular, Bodie, Kane and Marcus and Cornell, Bradford, 1993, Corporate Valuation: Tools for Effective Appraisal and Decision Making, Business One Irwin, New York, NY.

3.3.3.1 Risk-Adjusted Discount Rates vs. Risk-Adjusted Cash Flows

When valuing uncertain or risky cash flows, it is important to reflect this risk in the value that is calculated. The most common approach to making this risk adjustment is to discount the cash flows at a risk-adjusted discount rate that is higher than the risk free rate, thereby producing a value that is lower than it otherwise would be in the absence of this risk.

However, reflecting this risk in the discount rate is not the only way to accomplish this objective. Alternative approaches that incorporate the risk adjustment directly in the cash flows may even be preferred. Halliwell[8], for instance, presents compelling arguments for reflecting risk adjustments in the cash flows, using utility theory to produce *certainty equivalent* cash flows that can be discounted at risk free discount rates. This approach is closely related to the *risk neutral* valuation approach widely used to value derivative securities, as well as other probability transform methods advocated for pricing insurance risks, such as the Proportional Hazard Transform or the Wang Transform[9,10].

While the certainty equivalent, risk-neutral and probability transform approaches are appealing on theoretical grounds, the use of risk-adjusted discount rates is currently more common in practice. No clear consensus yet exists on how to apply these alternative approaches consistently in many real-world applications. Therefore, this study note will follow the more common approach using risk-adjusted discount rates and will focus on some of the principal issues involved in this process.

3.3.3.2 Private vs. Equilibrium Market Valuation

Before addressing specific methods of determining discount rates, it is important to make a distinction between a *private valuation* and an *equilibrium market valuation*.

In a private valuation, individual investors are assumed to have their own view of "risk" and to hold different existing portfolios. Any potential investment is assessed relative to the investor’s existing portfolio. As a result, the value of any stream of risky or uncertain cash flows may have a different value to different investors.

In an equilibrium market valuation, it is often assumed that all investors hold the same portfolio, assess the risk associated with a new investment in an identical fashion and also have the same estimates of future cash flows. Alternatively, it can be recognized that investors will not have identical risk and cash flows assessments, but only the marginal investor’s risk and cash flow assumptions will determine the “market” price of the investment. In this case, it is not necessary to assume that every investor will place the same value on a given investment, but if an investor’s private valuation differs from others’ valuations they simply will not trade at the market price.

Theoretical rate of return models often used to determine risk-adjusted discount rates tend to focus on market equilibrium rates of return. As a result, they serve as a useful starting point for determining any one investor’s appropriate discount rate for a given opportunity, but may not reflect all factors that need to be considered by any specific investor.

3.3.3.3 Determining the Discount Rate

The most popular model used to estimate (equilibrium) shareholder return expectations is the Capital Asset Pricing Model (CAPM)[11]. The CAPM attempts to describe the relationship between the “risk” of an equity investment and the return investors *expect* to earn on that investment. In this model, risk is defined in terms of the investment’s *beta*, a measure of systematic risk (risk that cannot be diversified away in a large portfolio). The beta reflects the degree to which the percentage changes in *market value* (the rates of return) co-vary with the rates of return on a hypothetical portfolio

---

[8] See Halliwell, Leigh J., "A Critique of Risk-Adjusted Discounting".

[9] See Wang, Shaun, "Insurance Pricing and Increased Limits Ratemaking by Proportional Hazards Transforms".

[10] See Appendix C of Halliwell.

[11] The discussion of only the CAPM as the source of discount rates in this study note is not intended to suggest a particular preference for this model. Other models, including Arbitrage Pricing Theory (APT), a Multi-factor CAPM or the Fama-French 3-Factor Model could certainly be used in place of the CAPM throughout.

consisting of all risky assets that an investor may choose to invest in. This portfolio of all risky assets is referred to as the *market portfolio.*

Mathematically, the CAPM can be expressed as follows:

$$k = r_f + \beta\,(E[r_m] - r_f)$$

where,

k = expected or required equity return

r_f = risk free rate

$E[r_m]$ = expected market return

$E[r_m] - r_f$ = expected equity market risk premium

β = Beta, a measure of the systematic market risk

This model is mechanically trivial to implement. However, there are important considerations to note when estimating beta, the risk-free rate and the expected equity market risk premium.

3.3.3.3(a) Estimating Beta

There are two common methods used to determine the beta for the purposes of valuation – measuring the target firm's beta directly or using an industry-wide beta.

- *Firm Beta* - Historical stock price data of the firm can be used to directly measure the CAPM Beta. The estimation is performed using a linear regression of the company's returns against the market returns. The company's historical beta can then be assumed to remain constant for the prospective period. Betas measured in this way are commonly reported by Bloomberg and other sources, sometimes inclusive of various statistical adjustments to improve the estimates, as discussed in Bodie, Kane and Marcus[12].
- *Industry Beta* - Beta estimates for individual firms are often unreliable due to statistical issues affecting individual firm data and changes in firm risk over time. Somewhat more reliable and stable are industrywide mean or median values. For example, Cummins and Phillips[13] estimate an industry-wide CPM beta for P&C insurers of approximately 0.843. This estimate reflects an average across all P&C insurers, each with different mixes of business and different degrees of financial leverage (debt). Therefore, the industry average should be interpreted carefully and adjustments may be required to reflect factors such as:

    a. Mix of Business – With respect to adjustments for different mixes of business, ideally only those firms with a comparable mix to the firm being valued should be used. However, as the definition of "comparable firms" gets more precise, the number of eligible firms drops significantly and the result becomes less reliable. Ultimately, judgment is needed.

    b. Financial Leverage – When firms raise capital by issuing debt, the leverage that is introduced impacts the degree of risk to the equity holders, making cash flows to equity holders riskier and the betas higher. This effect will show up in any estimates of the betas of firms with debt outstanding and therefore may make the betas of different firms difficult to compare.

    To make the various betas easier to compare and to allow for the use of an industrywide mean or median beta, the beta is often defined to reflect solely

[12] See Bodie, Kane and Marcus, Chapter 10.

[13] See Cummins and Phillips, "Estimating the Cost of Equity Capital for a Property-Liability Insurer", March 2004

the business risk of the firm and not the effect of debt leverage. This is the beta that would exist had the firm been capitalized entirely with equity and is often referred to as the *all-equity beta*.

Introductory finance texts provide a full description of how one could *de-lever* the equity betas to estimate the beta for an all-equity firm, so that material will not be reviewed here[14]. However, once the average all-equity beta for the industry is obtained, the equity beta for any particular firm would be found by readjusting the beta to reflect the amount of debt leverage for that particular firm[15].

While this approach to de-levering and then re-levering industry betas is often covered in the introductory finance textbooks, its application to insurance company valuation is somewhat limited, and perhaps unnecessary. This is because policyholder liabilities also result in leverage effects that are not fully accounted for when the beta is adjusted solely for debt leverage. Therefore, it may be reasonable to assume that the *total leverage* of all firms in the insurance industry is similar and that the appropriate leveraged equity return for any particular firm is based on the industry average equity beta, without any further adjustments.

In the above discussion, the focus was on the beta for the equity of the firm so that the expected returns to the equity holders can be measured. The equity holders' returns expectations are relevant because the intent of the DDM is to value the dividends *to the equity holders*. These expected returns to the equity holders will differ from the firm's *weighted average cost of capital* (WACC), which reflects the returns to both debt[16] and equity providers. The WACC is a commonly referenced estimate of the "cost of capital" but is not directly used in the DDM. An alternative valuation model that does use the WACC will be discussed in a subsequent section.

Below are some representative estimates of equity betas for various publicly traded insurers and reinsurers as of October 2004[17]:

Table 2: P&C Insurer and Reinsurer Equity Betas (Oct. 2004)

Company	Beta
American International Group, Inc	0.89
The Allstate Corporation	0.38
The Progressive Corp.	0.83
Chubb Corporation	0.72
ACE Limited	0.72
XL Capital Ltd.	0.59
CNA Financial Corporation	0.64
Market Value Weighted Average	0.79

[14] See Brealey & Meyers, Principles of Corporate Finance.

[15] The so-called Miles-Ezzel formula reflects the relationship between the levered equity return and the all-equity return. The levered return, r_e, is related to the unlevered equity return (r), the pre-tax debt return (r_d), the effective corporate tax rate (T) and the market values of the debt (D) and equity (E) according to the formula:

$$r_e = r + (1-T)(D/E)(r-r_d).$$

[16] The debt return used in the WACC formula is usually the after-tax yield on the debt.

[17] Source: Yahoo! Finance

3.3.3.3(b) Estimating the Risk Free Rate

The risk-free rate plays an important role in the standard CAPM. It should be based on current yields on risk free securities, which are often represented using zero-coupon U.S. Treasury yields.

To properly reflect the shape of the term structure, it is also appropriate to discount each cash flow at a rate that reflects the time to payment. Therefore, one would want to use a different required return for each time period, k_t, to discount each cash flow at time period t, rather than a single discount rate k for all time periods. This will also involve estimating a different equity risk premium (see below) for each time period.

In practice, it is common to avoid this complexity and instead use a single risk free rate and a single equity risk premium for all maturities. One still has a choice of which maturity to use for the risk free rate. The options include:

- 90-Day T-Bills – These are the purest "risk free" instruments as they are free of both credit and reinvestment risk. In textbook applications these are the securities most often used.
- Maturity Matched T-Notes – Some practitioners prefer to use a Treasury security with a term that matches the average maturity of the cash flows being valued.
- T-Bonds – Yields on 20-year Treasury bonds likely represent the most reasonable current estimate of the long run average *short-term* yields. These are also the most stable and the most logical choice for corporate decision-making because they come closest to matching the duration of the market portfolio and of the cash flows being valued.

  However, long-term yields also reflect a liquidity or term premium. As a result, the historical term premium between long-term and short-term yields should be netted out of the long-term yields. Bradford Cornell estimates that this term premium has historically been approximately 1.2%[18].

For the remainder of this study note, the risk free rate will be based on the 20-year T-bond yield, adjusted to reflect a 1.2% term premium, as a proxy for the long term average short-term yield.

3.3.3.3(c) Estimating the Equity Market Risk Premium

The actual spread between the market return and the short-term risk free rate has historically averaged approximately 6% to 8%. As a result, some authors recommend using this as a forecast of the future equity risk premium.

However, many authors have noted a so-called *equity premium puzzle* in that the historical premiums seem too high relative to any commonly proposed theories of investor behavior. Many attribute the historical return premium over risk free investments to be the result of good luck on the part of equity investors and/or bad luck on the part of bond investors. A 2004 CAS paper by Derrig and Orr[19] surveys the literature on the equity risk premium and documents estimates of the expected equity risk premium ranging from 4% to 8%, somewhat lower than the historical average.

The key considerations in determining the appropriate equity risk premium include the following:

- Short-term vs. Long-term Risk Free Rates as Benchmark – The market risk premium reflects the spread between the expected market return and the risk free rate. Since the risk free rate appears twice in the CAPM formula, it is important to use a consistent definition of the risk free rate in both the CAPM formula and in the measurement of the market risk premium. If a short-term yield is used in the CAPM, the market risk premium should be measured relative to short-term yields. Alternatively, if long-terms yields are

---

[18] See Cornell, *Corporate Valuation*, Chapter 7.

[19] Derrig, Richard A. and Elisha D. Orr, "Equity Risk Premium: Expectations Great and Small".

used as the risk free rate, the market risk premium should reflect the spread between the market returns and the long-term risk free yields.

- Arithmetic vs. Geometric Averages – When calculating average risk premiums, a choice must be made between arithmetic and geometric averages. Generally, arithmetic averages are preferred for single period forecasts. However for multiple period forecasts or long-term averages, geometric averages are preferred[20].
- Historical vs. Implied Risk Premiums – As noted in the Derrig and Orr study, risk premiums can be estimated based on either historical averages or by estimating the risk premium that is implied by current market prices.

  For the historical risk premiums, a choice has to be made with respect to the time period over which to measure the average returns, as the equity risk premium has fluctuated significantly over the past 75 or so years.

  The table below demonstrates the effect of using different time periods (as well as different choices for the risk free asset and arithmetic vs. geometric averages):

Table 3: Historical U.S. Risk Premiums[21]

	Stocks vs. T-Bills		Stocks vs. T-Bonds	
Period	Arithmetic	Geometric	Arithmetic	Geometric
1928-2000	8.41%	7.17%	6.53%	5.51%
1962-2000	6.41%	5.25%	5.30%	4.52%
1990-2000	11.42%	7.64%	12.67%	7.09%

  Note that the use of historical data, as shown in the above table, is not the only approach used to estimate risk premiums. An alternative method is to infer the equity risk premium from current market prices. For instance, one could use the DDM on an aggregate market index and solve for the risk premium given assumptions about the risk free rate, aggregate dividends and aggregate growth rates.

Taking these considerations into account, it is difficult to recommend any single value to be used for the equity risk premium. Any analysis should consider a range of possible values and the impact of different assumptions should be reviewed. A baseline risk premium of 5.5% will be used throughout the remainder of this study note and sensitivity analysis will be performed.

3.4 P&C Insurance Company Example

In this section, a simplified example of the DDM will be used to demonstrate the valuation of a P&C insurance company. To keep the discussion focused on the valuation methodology and not the detailed accounting issues, the example will rely upon simplified extracts from forecasted financial statements prepared in accordance with U.S. GAAP accounting rules.

---

[20] See Damodaran, *Investment Valuation*

[21] Source: Damodaran, *Investment Valuation*

Example 2 – DDM for Sample Insurance Company

Consider the following 5-year forecasts of the financial results for Sample Insurance Company. The data below shows actual (2004) and 5 years of forecasted (2005 – 2009) income statement and balance sheet items, each according to U.S. GAAP.

Table 4: U.S. GAAP Income Statement ($000's)

	2004	2005	2006	2007	2008	2009
Selected US GAAP Income Statement Items						
Net Income Before Tax	14,598	15,366	16,134	16,941	17,788	18,678
Corporate Income Tax	5,109	5,378	5,647	5,929	6,226	6,537
Net Income After Tax	9,489	9,988	10,487	11,012	11,562	12,141
Selected US GAAP Balance Sheet Items						
Total Assets	471,550	493,359	523,125	558,165	598,112	642,413
Total Liabilities	371,550	388,365	412,887	442,421	476,588	514,818
US GAAP Equity	100,000	104,994	110,238	115,744	121,525	127,595
Total Liabilities and Equity	471,550	493,359	523,125	558,165	598,112	642,413
Dividends Paid (50% of NI)	4,744	4,994	5,244	5,506	5,781	6,070

The following additional information is available for Sample Insurance Company:

- Dividend Payout Ratio – The firm has a current dividend payout ratio equal to 50% of its after-tax net income and intends to maintain this payout ratio indefinitely.
- Risk Free Rate – The current yield[22] of the 20-year U.S. Treasury Bond is approximately 4.33% with annual compounding. This rate will be used as the risk free rate.
- Company's Equity Beta – The company's actual equity beta cannot be estimated directly because it is a relatively new company with limited historical equity price data.
- Equity Betas for Peer Companies – The industry beta for this company's closest peers is estimated to be 0.84. The companies in the peer group have comparable levels of financial leverage (debt outstanding as a percentage of the firm value) and operating leverage (premiums as a percentage of GAAP equity).

The following steps are used to implement the DDM to value this company:

Step 1: Determine Dividend During Forecast Period

These amounts were provided in the table above and are summarized here for convenience:

Table 5: U.S. GAAP Income Statement ($000's)

	2005	2006	2007	2008	2009
Dividends Paid	4,994	5,244	5,506	5,781	6,070

[22] As of June 2, 2004, the 20-year CMT yield with semi-annual compounding is 5.47%. Subtracting the 1.2% term premium and converting to an annually compounding basis results in the 4.33% risk free rate.

Step 2: Estimate Dividend Growth Rate Beyond Year 2009

Refer to the selected financial data shown below:

Table 6: Selected Financial Data

	2005	2006	2007	2008	2009
GAAP Equity Beginning of Period)	100,000	104,994	110,238	115,744	121,525
GAAP Equity (End of Period)	104,994	110,238	115,744	121,525	127,595
Net Income	9,988	10,487	11,012	11,562	12,141
Dividend	4,994	5,244	5,506	5,781	6,070

Based on these values, the following values needed to estimate the growth rate in dividends beyond the 2009 forecast horizon are obtained:

Table 7: Growth Rate Data

	2005	2006	2007	2008	2009
Dividend Payout Ratio	50.0%	50.0%	50.0%	50.0%	50.0%
Plowback Ratio	50.0%	50.0%	50.0%	50.0%	50.0%
ROE = NI / Beginning GAAP Equity	10.0%	10.0%	10.0%	10.0%	10.0%
Dividend Growth Rate					
Expected Plowback Ratio	50.0%				
Expected Average ROE	10.0%				
Growth Rate	5.0%				

As shown in the table, the formula expressing the growth rate as the plowback ratio multiplied by the ROE is used to obtain a growth rate of 5.0% beyond the forecast horizon. This is consistent with the dividend growth rate during the forecast horizon. This may not always be the case, for instance, if the long-term average ROE or dividend payout ratios are expected to differ from the short-term values during the forecast horizon.

Step 3: Estimate Required Equity Return

The CAPM equity beta, based on the equity betas of peer companies, was stated earlier and assumed to equal 0.84. Using CAPM with the following parameters, the appropriate discount rate is estimated to be 8.95%, as shown below.

Table 8: Discount Rate

Risk Free Rate	4.33%
Equity Risk Premium	5.50%
Equity Beta	0.84
Discount Rate	8.95%

Step 4: Determine Value

The dividends and terminal value amounts can now be combined to estimate the total equity value by discounting each amount at the 8.95% discount rate:

Table 9: Valuation Using DDM

	2005	2006	2007	2008	2009	Terminal Value
Dividend	4,994	5,244	5,506	5,781	6,070	161,354
PV Factor	0.918	0.842	0.773	0.710	0.651	0.651
PV	4,584	4,418	4,257	4,103	3,954	105,110
Value	126,426					

The terminal value was determined based on an assumption of constant growth beyond 2009 of 5.0%, the discount rate of 8.95% and the year 2009 dividends of 6,070.

$$\text{Terminal Value} = \frac{6{,}070 * (1.05)}{.0895 - .05} = 161{,}354$$

The present value of this terminal value estimate is then $161{,}354/1.0895^5 = 105{,}110$.

The total estimated value of the equity is then the sum of the present values of the five dividend payments and the terminal value, which totals $126.4 million.

Step 4: Sensitivity Analysis

Notice that the present value of the terminal value component is approximately $105 million. This means that 83% of the total value of the firm is reflected in the terminal value, which assumes perpetual growth in dividends of 5%. The magnitude of the terminal value relative to the total value of the firm suggests the need to be very careful about the sensitivity of the result to this growth assumption.

Below is a table that shows the sensitivity of the terminal value and the total equity value to estimates of the growth rates. The different rates shown represent the results of alternative assumptions regarding the ROE beyond the forecast horizon, with the dividend payout rate remaining constant. For example, if the ROE were to decline to the level of the investor's required return (8.95%) the growth rate would decline to 4.475%. The resulting total valuation would decrease from $126.4 million to $114.2 million. This represents a reduction of 9.7%.

Table 10: Sensitivity to Alternative Growth Rate Assumptions

Growth Rate	Nominal Terminal Value	PV Terminal Value	Equity Value
4.000%	127,531	83,077	104,393
4.475%	142,543	92,856	114,172
5.000%	161,354	105,110	126,426
6.000%	218,108	142,081	163,397

The sensitivity of the firm value to the estimated discount rate can also be tested. For instance, using alternative assumptions about the equity risk premium would result in the following alternative estimates of the CAPM discount rate and equity value:

Table 11: Sensitivity to Alternative Equity Risk Premium

Equity Risk Premium	CAPM Discount Rate	Equity Value
4.0%	7.69%	185,644
5.5%	8.95%	126,426
6.0%	9.37%	114,276
8.0%	11.06%	82,407

Combining these ranges of discount rates and ranges of growth rates beyond the forecast horizon, the following estimates of total equity value would be obtained:

Table 12: Sensitivity to Growth and Discount Assumptions

	Growth Rate Beyond Forecast Horizon			
Discount Rate	4.000%	4.475%	5.000%	6.000%
7.69%	140,176	159,347	185,644	284,921
8.95%	104,393	114,172	126,426	163,397
9.37%	96,198	104,309	114,276	143,082
11.06%	73,081	77,389	82,407	95,419

Notice that the valuation in this table ranges from a low of $73 million to a high of $285 million. This is a rather large range. But recall that the growth rates and discount rates are not independent of each other. Rapid growth is unlikely to be possible without assuming more risk; stable, low growth businesses are unlikely to exhibit high systematic risk. In the case of the previous table, the equity risk premium was varied but the estimated CAPM betas were not altered to ensure consistency with the assumed growth rates. This suggests that the more extreme values in the table are less realistic than many of the other entries in the table.

4. Discounted Cash Flow

The valuation approach based on the present value of future dividends is easy to understand. A fair amount of detail for that model was presented because many of the elements of the application to a real valuation exercise apply equally well to other methods. However, the dividend discount model (DDM) has some important limitations. Actual dividend payments are highly discretionary and can be difficult to forecast. In addition, the increased use of stock buybacks as an efficient vehicle for returning funds to shareholders requires that, at a minimum, a fairly liberal definition of "dividend" be adopted.

An alternative, though very closely related, approach is to focus on *free cash flows* rather than dividends. The free cash flows represent all of the cash that could be paid out as dividends or other payments to the capital providers, after making appropriate adjustments to reflect amounts needed to support current operations and the expected growth. The key difference between this approach, referred to here as the discounted cash flow (DCF) method, and the DDM is simply the recognition that free cash flow not paid as a dividend immediately would be invested to earn a fair risk adjusted return (i.e. it would not be stuffed in a drawer). As long as this can be assumed to be the case, there is no impact on value, positive or negative, from not paying the funds out immediately. For the purpose of valuation, it is acceptable to assume that the entire free cash flow is in fact paid as a dividend.

The DCF approach abstracts away from actual dividend policy and focuses on the cash that could be paid in each future period. This is not meant to suggest that "cash flow" is measured exactly as it might be defined under Generally Accepted Accounting Principles (GAAP). This is because free cash flow also reflects the capital expenditures needed to maintain the firm's operations and generate the earnings growth inherent in the forecasts.

When applying the free cash flow approach, there are two alternative methods used. One approach is to focus on the free cash flows to the entire firm and the other approach is to focus on the free cash flow to the equity holders only.

4.1 Free Cash Flow to the Firm

The Free Cash Flow to the Firm (FCFF) approach values the entire firm and then subtracts off the market value of the debt to value the equity indirectly. This valuation methodology is discussed in some detail in Chapter 18 of Bodie, Kane and Marcus as well as other introductory finance texts. While this approach has many advantages when applied to most industries, it is problematic when applied to financial services firms such as insurance companies.

Damodaran discusses the difficulties applying the FCFF method to banks and insurance companies. His key points can be summarized as follows:

- Policyholder Liabilities vs. Debt - The FCFF method values the entire firm and then subtracts off the value of the debt to value the equity. This approach treats the debt as a source of capital that is more like the equity of the firm rather than a part of the firm's normal business activities. As noted earlier with respect to the levered equity beta, the distinction between debt and policyholders liabilities for a P&C insurance company is rather arbitrary and there is no economic rationale for different treatment of these two sources of liability.

- WACC and APV – The FCFF approach is applied by first using the firm's weighted average cost of capital (WACC) as the discount rate for the free cash flows to determine the value of the entire firm. The market value of the debt is then subtracted from this amount to determine the value of the equity.

  Alternatively, the free cash flows could be discounted using the unlevered, all-equity discount rate (assuming that there is no debt) to derive the value of the firm without consideration of the debtholders' claims, the tax consequences of the debt or the impact of debt on the riskiness of the equity holders' claims. The equity value is determined by subtracting the market value of the debt from the firm value and then making two adjustments. The first adjustment reflects the debt's tax consequences by adding the value of the debt's tax shields. The second adjustment reflects the debt's effect on equity risk by incorporating an estimate of the potential cost of financial distress. This alternative approach is often referred to as an Adjusted Present Value (APV) approach.

In either case, the existence of policyholder liabilities makes it difficult to precisely define either the WACC or the unlevered, all-equity discount rate needed for the APV approach.

Since this study note focuses on valuation for P&C insurance companies, the FCFF approach will not be presented in any detail here[23].

4.2 Free Cash Flow to Equity

When valuing insurance companies, it is preferable to focus on the Free Cash Flow to Equity (FCFE) method. FCFE is very similar to FCFF but it reflects free cash flows *after* deductions for interest payments, net of any tax consequences of these interest payments, and any net change in borrowings (i.e. repayment of debt and new debt issued). This focus on the cash flows to the equity holders also means that the discount rate reflects only the risk to the equity holders rather than the WACC mentioned above. This allows the use of the levered equity return as the discount rate, which is useful given the difficulties identified earlier with the estimation of the unlevered equity return for P&C insurance companies.

The typical textbook definition of FCFE is summarized as shown in the following table:

Table 13: Definition of Free Cash Flow to Equity

	Net Income
plus	Non-Cash Charges (Expenses)
less	Net Working Capital Investment
less	Capital Expenditures
plus	Net Borrowing
	Free Cash Flow to Equity (FCFE)

Typically, expenses that are deducted under U.S. GAAP accounting but do not represent actual cash expenditures are added back to the reported net income to determine the cash flow available to be paid to equity holders. These amounts are referred to in the table above as *Non-Cash Charges.* For a P&C insurer, the most significant of these "non-cash" expense items on the income statement are the increases in the loss and expense reserves. These increases in reserves have a large impact on the reported income but not on the actual cash flow. This would seem to suggest that changes in reserves could be added back to net income, but this is not the case, as will be explained below.

Notice that two other components of the free cash flow to equity calculation include changes in net working capital and capital expenditures. Both of these amounts represent uses of cash flow needed to maintain the firm's operations and support the growth that is planned. *Working Capital Investment* shown in the above table reflects net short term (non-cash) assets held to facilitate company operations, such as inventory or accounts receivable. *Capital Expenditures* typically refer to investment in property, plant, equipment and other physical items. For P&C insurance companies, net working capital is not typically significant and will not be discussed in detail here[24].

The definition of capital expenditures for P&C insurance companies is more complicated because it must be adjusted to include changes in loss and expense reserve balances as well as increases in capital held ("invested") to meet regulatory and/or rating agency capital requirements consistent with the company's business plan.

---

[23] The interested reader should refer to Damodaran's *Investment Valuation* for a thorough treatment of this valuation approach.

[24] Refer to Damadoran and Stowe, et. al. for extensive discussion of the other components of Non-Cash Charges and Net Working

- Treatment of Increase in Loss and Expense Reserves – Recall that the FCFE represents the cash flow that *could* be paid to shareholders in any particular period. In the simple case of a two year insurance policy where the firm collects the premium net of expenses up front and then pays claims at the end of the second period, it would not be sufficient to treat the net premiums as the (positive) free cash flow in the first period and the claim payments as the (negative) free cash flow in the second period. This is because some of the premium collected in the first period is not *free* to be paid to shareholders. Instead, some portion of the premium must be held in claim reserves, usually on an undiscounted basis.

  The implication of this is that when calculating FCFE, changes in loss and expense reserves can be included in the definition of capital expenditures. Since these changes in reserves reflect the most significant *Non-Cash Charges*, which according to the usual definition of FCFE would be added back to Net Income, and also reflect a significant portion of *Capital Expenditures*, which would be subtracted from Net Income, these two adjustments will cancel each other out. The result is that the increases in loss and expense reserves, which have already been reflected in the net income figures, can be ignored in the steps used to estimate FCFE through adjustments to net income.

- Treatment of Increase in Required Capital – In addition to reserve requirements, insurers are subject to regulatory and/or rating agency capital requirements. Just as a widget manufacturer must invest capital in machinery to make widgets, an insurance company must invest capital before it can sell an insurance policy. Such regulatory minimum capital requirements should be treated as "capital expenditures" for the purposes of determining free cash flow. Furthermore, the ability of an insurer to meet its growth targets and profitability targets is tied closely to public perception of its financial strength and credit standing. Therefore, capital required to maintain the firm's target credit rating implied by the business plan should also be treated as equivalent to a capital expenditure. In both of these cases, the regulatory and rating agency capital requirements serve to reduce the free cash flow relative to U.S. GAAP definitions of net income.

  To focus attention on the valuation methodology as opposed to accounting and regulatory issues in this study note, specific regulatory or rating agency capital requirements will not be addressed here. In the numerical examples shown, the minimum capital requirements are approximated using simplified capital standards that are meant to mirror Standard & Poor's guidelines applicable to AA-rated insurers. The interested reader should refer to Standard & Poor's "Property/Casualty Insurance Ratings Criteria" for more information on this important aspect of valuation.

  In a real-world application, there are likely to be multiple constraints on free cash flow resulting from the need to hold capital in the firm. The most binding constraint could be the result of regulatory restrictions, rating agency restrictions or perhaps management's own assessment of the capital needed to support the risk-taking activities of the firm without negatively impacting the firm's ability to achieve its growth plans. In this case it would be necessary to determine the most binding constraint on capital and assess how it impacts free cash flow.

The resulting definition of FCFE that can be used for P&C insurers is therefore adjusted as follows:

Table 14: Simplified Definition of Free Cash Flow to Equity for P&C Insurer

	Net Income
Plus	Non-Cash Charges – Excluding Changes in Reserves
Less	Net Working Capital Investment
Less	Increase in Required Capital
Plus	<u>Net Borrowing</u>
	Free Cash Flow to Equity (FCFE)

Example 3 – Free Cash Flow to Equity Calculation for ABC Insurance Company

Consider a hypothetical P&C insurer, ABC Insurance Company. In the current period the company had beginning U.S. GAAP Equity equal to $103.500 million and U.S. GAAP Net Income equal to $17.193 million. Based on their internal financial model that reflects their growth plans for the coming year, they have determined that the capital needed (at the start of their next accounting period) to maintain their AA-rating is $108.624 million.

For simplicity, assume that there are no non-cash charges included in the net income figure other than changes in reserves, there are no net working capital investments and there are no increases in borrowings.

The Free Cash Flow to Equity for this firm in the current period can be calculated as follows:

Table 15: Calculation of Free Cash Flow to Equity for ABC Insurance Company ($ Millions)

Beginning US GAAP Equity	103.500
Net Income	17.193
Ending US GAAP Equity - Before Dividends	120.693
Minimum Capital - Based on Target S&P AA Rating	108.624
Beginning US GAAP Equity	103.500
Increase in Required Capital	5.124
Net Income	17.193
Non Cash Charges (Excluding Change in Reserves)	0
Net Working Capital Investment	0
Capital Expenditures = Increase in Required Capital	5.124
Net Borrowing	0
Free Cash Flow to Equity	12.069

Notice that the FCFE could also be calculated as the difference between the ending GAAP equity and the minimum required capital, as shown here:

Table 16: Alternative Calculation of Free Cash Flow to Equity for ABC Insurance Company

Ending US GAAP Equity - Before Dividends	120.693
Minimum Capital - Based on Target S&P AA Rating	108.624
Free Cash Flow to Equity	12.069

4.3 Applying the FCFE Method

Once the FCFE values are determined, much of the remainder of the valuation exercise is similar to what was done using the DDM. The free cash flows during the forecast horizon are valued using an appropriate risk-adjusted discount rate and the terminal value is estimated by assuming a constant growth rate in free cash flow and an appropriate discount rate.

Below, several details regarding this methodology will be addressed. The financial model for ABC Insurance Company used in *Example 3* above will be used as a reference. The Net Income, Equity and Free Cash Flow to Equity amounts for the years 2005 – 2009 were calculated using the same methodology and the key elements are summarized as follows:

Table 17: Free Cash Flow to Equity for ABC Insurance Company 2005 – 2009 ($000's)

	2005	2006	2007	2008	2009
Beginning US GAAP Equity	103,500	108,624	113,274	117,648	122,422
Net Income	17,193	17,236	17,446	18,376	18,967
Ending US GAAP Equity - Before Dividends	120,693	125,860	130,720	136,024	141,388
Minimum Capital - Based on Target S&P AA Rating	108,624	113,274	117,648	122,422	127,250
Beginning US GAAP Equity	103,500	108,624	113,274	117,648	122,422
Increase in Required Capital	5,124	4,650	4,374	4,774	4,828
Free Cash Flow to Equity	12,069	12,586	13,072	13,602	14,139

4.3.1 Growth Rates

Earlier in the discussion of the DDM approach, growth rates were estimated using historical averages or by relying on the fundamental principle that growth is the result of income that is reinvested in the firm and that subsequently earns a positive return (ROE).

When using the FCFE method, it is important to note the implicit assumption that *all* free cash flow to equity is paid to shareholders. Therefore, the definition of reinvestment for purposes of determining growth rates is slightly different than it was in the DDM. In that case it was sufficient to simply compare the dividends paid to the firm's net income.

For a P&C insurance company, the best determinant of growth is the portion of net income that is used to increase the capital base of the firm, since the capital base of the firm determines the *maximum* growth that can be achieved given the regulatory and rating agency constraints[25].

Combining this with the return on equity provides an estimate of the growth rate beyond the forecast horizon, as shown below using the ABC Insurance Company example data.

Table 18: Estimated Growth Rate Beyond Forecast Horizon ($000's)

	2005	2006	2007	2008	2009
Net Income	17,193	17,236	17,446	18,376	18,967
Free Cash Flow to Equity	12,069	12,586	13,072	13,602	14,139
Reinvested Capital	5,124	4,650	4,374	4,774	4,828
Reinvestment Rate	29.8%	27.0%	25.1%	26.0%	25.5%
Beginning Capital	103,500	108,624	113,274	117,648	122,422
ROE	16.6%	15.9%	15.4%	15.6%	15.5%
Free Cash Flow Growth Rate					
During Forecast Horizon		4.3%	3.9%	4.1%	3.9%
Beyond Forecast Horizon - Estimated					3.9%

[25] It can be argued that growth is also constrained by the firm's investment in quality personnel. See Damodaran, Investment Valuation, for a more detailed discussion of this issue.

In the above table, the following calculations are shown:

- Reinvested Capital = Net Income – Free Cash Flow
- Reinvestment Rate = Reinvested Capital / Net Income
- ROE = Net Income / Beginning Capital
- Forecast Horizon Growth Rate = $FCFE_t$ / $FCFE_{t-1}$
- Horizon Growth Rate = Reinvestment $Rate_{2009}$ * ROE_{2009} = 3.9%

4.3.2 Discount Rate

The appropriate discount rate for this method is determined in essentially the same manner as in the DDM. It is, however, important to ensure that the assumption regarding the riskiness of the cash flows is consistent with the assumption regarding the distribution of the free cash flow to shareholders.

Compared to the DDM, the FCFE model assumes that more cash is distributed to shareholders in each period because *all* cash that could be paid as a dividend is assumed to be paid. The values used in the calculation are not impacted by the firm's actual dividend policy. This does not affect the overall valuation because of the implicit assumption that any cash that was not distributed in the form of dividends and was not needed to support growth in the insurance operations would be invested in marketable securities and would earn an appropriate risk-adjusted return. Investments in marketable securities should generally be a zero net present value activity and so value is neither created nor destroyed from this activity.

The riskiness of the dividend cash flows can be thought of as representing an average of the riskiness of the insurance operations and the investment operations. As a result, it is likely to be the case that the appropriate discount rate in the FCFE model is different than the discount rate in the DDM model. The two models assume different proportions of investment income and underwriting income because the FCFE method pays out all free cash flow while the DDM model pays out only the assumed dividends and reinvests the balance in marketable securities. The DDM model's measure of risk is therefore impacted by a larger proportion of the risk coming from marketable securities than from underwriting risk.

Specifically quantifying this difference in risk is a challenge. When the CAPM is used as the basis for the risk-adjusted discount rate, what matters is systematic risk and not total risk. For most practical purposes the precision of the discount rate calculation is low enough that this distinction is often ignored. Therefore, for simplicity the example below will assume the same discount rates can be used in the DDM and FCFE models.

4.3.3 Example of FCFE Method Using ABC Insurance Company Data

The following example uses the data referenced above in Table 17 for the ABC Insurance Company to demonstrate the FCFE method and to perform sensitivity analysis of the results.

Example 4 – Valuation of ABC Insurance Company using FCFE Method

Using the estimated FCFE for ABC Insurance Company, the 3.9% growth rate assumption discussed in the text and the same 8.95% discount rate assumption used earlier, the calculations using the FCFE method are as shown below.

Table 19: Valuation Using Free Cash Flow to Equity Method ($000's)

	2005	2006	2007	2008	2009	Terminal Value
FCFE	12,069	12,586	13,072	13,602	14,139	290,899
PV Factor	0.918	0.842	0.773	0.710	0.651	0.651
PV	11,078	10,603	10,108	9,654	9,210	189,499
Value	240,152					

The terminal value shown above was determined based on an assumption of constant growth beyond 2009 of 3.9%, the discount rate of 8.95% and the year 2009 FCFE of 14,139.

$$\text{Terminal Value} = \frac{14{,}139 * (1.039)}{.0895 - .039} = 290{,}899$$

The total estimated value of the equity is the sum of the present values of the five FCFE amounts and the present value of the terminal value. The total equity value is $240.2 million.

Sensitivity Analysis

Notice that the discounted terminal value is 290,899/(1.0895^5) = 189,499. This means that 79% of the total value of the firm is reflected in the terminal value, which assumes perpetual growth in FCFE of 3.9%. This suggests the need to be very careful about the sensitivity of the results to this growth assumption.

Below is a table that shows the sensitivity of the terminal value and the total equity value to estimates of the growth rates. The different rates shown represent the results of alternative assumptions regarding the ROE beyond the forecast horizon. For example, if the ROE were to decline to the level of the investor's required return (8.95%) then the growth rate would decline to 2.3%. The resulting equity valuation would decrease from $240.2 million to $192.3 million, a reduction of 20%.

Table 20: Sensitivity to Alternative Growth Rate Assumptions

Growth Rate	Terminal Value	PV of Terminal Value	Total Equity Value
2.3%	217,507	141,689	192,342
3.1%	249,185	162,325	212,978
3.9%	290,899	189,499	240,152
4.8%	357,052	232,593	283,246

The sensitivity of the firm value to the estimated discount rate can also be tested. For instance, using alternative assumptions about the equity risk premium would result in the following alternative estimates of the CAPM discount rate and equity value:

Table 21: Sensitivity to Alternative Equity Risk Premium

ERP	Discount Rate	Equity Value
4.0%	7.69%	320,023
5.5%	8.95%	240,152
6.0%	9.37%	221,706
8.0%	11.06%	169,355

Combining these ranges of discount rates and ranges of growth rates beyond the forecast horizon, the following estimates of total equity value would be obtained:

Table 22: Sensitivity to Growth and Discount Assumptions

	Growth Rate Beyond Forecast Horizon			
Discount Rate	2.3%	3.1%	3.9%	4.8%
7.69%	237,683	266,794	320,023	419,443
8.95%	185,829	212,978	240,152	283,246
9.37%	180,823	201,211	221,706	252,652
11.06%	146,872	160,882	169,355	181,227

Notice that the valuation in this table ranges from a low of approximately $147 million to a high of $419 million. It may be unrealistic to assume that the highest growth rates and the lowest discount rates would apply simultaneously, making the most extreme values potential less reliable. Nonetheless, this highlights the wide range of results that can be obtained and the need to carefully consider all of the assumptions made.

4.3.4 Observations Regarding Example 4

Before proceeding further, some important observations with respect to the application of the FCFE method are noted.

- Terminal Value – The terminal value calculated in the previous example ($290,899) was based on assumptions of the Year 2009 FCFE, the growth rate beyond that point and the discount rate. This terminal value is 290,899/14,139 = 20.6 times the Year 2009 FCFE. In other words, the impact of the growth rate and discount rate assumptions could have been combined into a single multiple of the FCFE and expressed the terminal value as "20.6 times" FCFE.
- Average Discount Rates – Most firms' overall earnings and cash flows represent the total amounts across a variety of businesses, each with their own risk profile. The discount rate therefore represents an average discount rate reflecting the average risk from all of these separate businesses and activities. To the extent that the mix of business or degree of financial leverage is changing, these changes should be reflected in different discount rates for different time periods or cash flows.
- Market Value of Net Cash Flows – The use of a single discount rate for the *net* free cash flow to equity implicitly discounts each of its components at the same rate. Therefore, cash flows from investment returns and cash flows from liability payments, as well as other cash flows, are discounted at the same weighted average rate, even though the risk characteristics of the component cash flows likely vary considerably. It is worth considering whether this is appropriate.

  Most textbook presentations of the FCFE approach focus on the valuation of industrial firms in which investments in cash and marketable securities are usually minimal. In these cases, the definition of FCFE does *not* include investment income on currently held marketable securities. These non-operating assets are excluded from the valuation and added back in at their current market values at the end. For insurance companies, this distinction between operating and non-operating assets is considerably more difficult to make. As a result, it is typical to include investment income cash flows in the definition of FCFE.

  Including investment returns in the definition of free cash flow and then calculating their present value at an average rate for all cash flows is unlikely to reproduce a present value equal to the market value of the investment at inception. When investments are restricted to marketable securities, especially those most often found in P&C insurance investment portfolios, $1,000 invested in stocks is worth the same on the date of the investment as $1,000 invested in corporate bonds or $1,000 invested in risk-free bonds. It is true that their income and cash flow profiles differ and so their *future* value will differ. However, their present values at the date the investment is made should be identical. This result will only occur though if the discount rates used to determine the present values differ and reflect the riskiness of the respective investments. The use of an average rate for all cash flows will not produce the correct value for any particular investment.

When future investment cash flows are included in the aggregate cash flows, it can appear to be the case that value is either created or destroyed based on different assumptions about the asset portfolio composition. This misleading result occurs because the discount rate used reflects the average risk for the entire firm's net cash flows rather than the appropriate risk-adjusted rate for the investment asset cash flows themselves.

Similarly, using an average discount rate to calculate the present value of liability cash flows is unlikely to produce an accurate risk-adjusted value for this liability, as the appropriate risk-adjusted discount rate for liability cash flows is a rate *below* the risk free rate[26]. This would reflect the positive risk premium that would have to be paid in order to transfer this uncertain liability to a third party.

For this reason, some analysts argue that the assets and liabilities should be valued separately to ensure *market consistent* valuation of each. But separately valuing each component of the free cash flow may not be practical. This is because the cash flow specific risk-adjusted discount rates may be extremely difficult to quantify. This is particularly true for assets and liabilities that are not currently reflected on the firm's balance sheet.

As a result, this study note will follow the common practice of discounting net cash flows at an average rate. Sensitivity testing can be used to ensure that assumptions regarding investment policy have reasonable and appropriate impacts on the value of the firm. Further discussion of this issue in the context of the valuation of life insurance companies can be found in Girard.

[26] See Butsic, "Determining the Proper Interest Rate for Loss Reserve Discounting: An Economic Approach".

5. Abnormal Earnings Valuation Method

The DCF approach to valuation just described is relatively simple to understand and focuses attention directly on the net cash flow generating capacity of the firm. Furthermore, the process of thinking through the cash flow generating activities of the firm, quantifying the firm's capital needs and contemplating the risk factors is an important and worthwhile part of any valuation exercise.

However, the DCF method suffers from some practical weaknesses. To estimate free cash flows, the analyst must first forecast financial statements (income statements and balance sheets) according to a specific set of accounting standards (U.S. GAAP, U.S. Statutory or International Accounting Standards). Then, a variety of adjustments are made to the forecasts of net income to estimate the free cash flow. The resulting values for free cash flow (to equity) may then bear little resemblance to the forecasts that management is familiar with, such as the values used within the firm's internal planning process, the financial results of peer companies or the forecasts of external analysts. This might make it difficult to assess the reasonableness of the forecasted free cash flows or estimate their future growth rates.

An alternative method that relies more directly on accounting measures of net income rather than cash flows is referred to here as the Abnormal Earnings (AE) approach. Using this method, the accounting net income is not adjusted to reflect cash flows. Instead, reported book value and forecasted net income under the applicable accounting framework are used directly.

Before presenting this approach, it is useful to note that finance textbooks have long advocated a preference for cash flow models as opposed to accounting-based earnings models in order to accurately reflect the timing of the cash flows and to avoid problems associated with arbitrary methodology choices that may not represent real effects on firm value. More recently, several academics and practitioners have demonstrated that a discounted accounting-based earnings approach often produces more accurate valuation estimates and may offer additional benefits by framing the problem differently than the traditional cash flow models[27].

5.1 Background on Abnormal Earnings Method

Recall from the pricing of bonds that the value of a default free bond merely represents the present value of its coupon and principal payments, discounted at the appropriate (maturity matched) zero-coupon yields. In the event that the coupon rate and the yields are equal, the bond's market value will equal its face value (principal amount). This is because the periodic interest paid on the bond, based on its coupon rate, is exactly equal to the periodic interest that investors demand. Similarly, if the coupon rate exceeds the yields, the bond will be have a *higher* value than the face value; if the coupon rate is below the yields then its market value will be *below* the face value.

This same concept can be extended to the valuation of a firm based on its accounting values. The book value of the firm reflects the value of the firm's equity capital, at least according to a specific accounting standard (e.g. U.S. GAAP). If the firm can earn a return on this capital exactly equal to a "normal" return demanded by its shareholders, then the market value of the firm's equity should exactly equal its book value[28]. This is similar to the notion that the market value and face value of a bond are equal if the coupon rate and yield are equal.

This suggests that positive (negative) deviations from book value must be due to the firm's ability to earn more (less) than this "normal" rate demanded by shareholders. By focusing attention solely on these "abnormal" earnings, the present value of all future abnormal earnings can be calculated and added to the book value to determine the total value of the firm's equity.

---

[27] See Sougiannis and Penham.

[28] For simplicity, I will assume that the assets and liabilities are both fairly stated on the balance sheet according to the appropriate accounting methods and that there is no systematic bias in the reported book value.

In mathematical terms, the abnormal earnings (AE) in any given period, t, are equal to:

$$
\begin{aligned}
AE_t &= \text{Net Income}_t - [\text{Required Equity Return}_t * \text{Book Value of Equity}_{t-1}] \\
&= NI_t - k * BV_{t-1} \\
&= (ROE_t - k)\, BV_{t-1}
\end{aligned}
$$

where, NI_t is the net income for period t, BV_{t-1} is the beginning book value for period t, ROE_t is the return on equity in period t and k is the required return.

Of course, the actual abnormal earnings for future periods at the time of the valuation are not known. The expected values of these abnormal earnings, denoted $E[AE_t]$, are used.

Then the value of the equity of the firm is simply:

$$
\begin{aligned}
\text{Value of Equity} &= \text{Beginning Book Value of Equity} + \text{PV(Expected Abnormal Earnings)} \\
&= BV_0 + \sum_{t=1}^{\infty} \frac{E[AE_t]}{(1+k)^t}
\end{aligned}
$$

Just as with the DDM and DCF approaches, the abnormal earnings approach is typically implemented by forecasting abnormal earnings for several periods (the forecast horizon). Then, a terminal value must be calculated that reflects abnormal earnings beyond this forecast horizon.

In the DDM and DCF valuation approaches, the terminal value calculation usually assumes that the dividends or free cash flows will continue in perpetuity and often the amounts are assumed to grow at a constant rate. In the case of the AE method, these terminal valuation assumptions are often different. Abnormal earnings are less likely to continue in perpetuity and are more likely to decline to zero as new competition is attracted to businesses with positive abnormal earnings.

The difficulty of achieving sustained growth in abnormal earnings is one reason why practitioners often favor the AE approach. This method forces the analyst to explicitly consider the limits of growth from a value perspective. Growth in earnings may be easy to achieve by simply increasing the book value of the firm, but this growth adds value only if the earnings exceed the shareholders' expected returns. Normal earnings growth does not add value; only abnormal earnings add value.

5.2 Accounting Distortions

It may be surprising that the arbitrary nature of certain accounting rules does not necessarily limit the usefulness of unadjusted earnings for valuation purposes. How, for instance, can one ignore the reality that P&C insurance reserves must be carried at their nominal value rather than their discounted value?

To reconcile this apparent weakness, note that the abnormal earnings approach includes both the current book value and the discounted value of future abnormal earnings in the value of the equity. As a result, accounting rules that distort estimates of earnings will also distort the estimates of book value[29] and will eventually reverse themselves. This is an important point and is worth demonstrating. An example used by Palepu, Bernard and Healy, in their textbook, *Business Analysis and Valuation*, will be used here.

Assume a manufacturing firm could have capitalized $100 of expenditures and included them in the value of its inventory, but instead decided to treat these costs as a current period expense. Both their income and end-of-period book value will be reduced by $100 in the current period. For instance, assume that their book value would have been $1,000 had they capitalized these costs but is only $900

---

[29] Technically, for this to be true the forecasts must satisfy what is referred to as the "clean surplus condition". The clean surplus condition assumes that changes in book value solely reflect earnings, dividends and capital contributions. It precludes accounting entries that impact book value without flowing through earnings, such as in the case of foreign currency translations under U.S. GAAP accounting. U.S. and international accounting standards do not always adhere to the clean surplus condition, so adjustments may be required. See Ohlson, *Earnings, Book Values and Dividends in Equity Valuation* for more details.

as the result of expensing these costs. Further assume that they will sell the inventory for $200 in two years and that the required rate of return is 13%.

As shown in the table below, the two approaches will begin with different book values. In the first period, there are no earnings. In the second period, the goods are sold for $200, causing one method to report income of $100 and one method to report income of $200. But the use of different starting book values causes the resulting equity values, found by adding the present value of the Period 1 and Period 2 abnormal earnings to the book value, to be identical.

Table 23: Demonstration of Self-Correcting Accounting

	Method 1 Capitalize Cost	Method 2 Expense Cost
Beginning Book Value	1,000.00	900.00
Period 1		
Sales	0.00	0.00
less Inventory Cost	0.00	0.00
Earnings	0.00	0.00
less Required Return * Book Value	130.00	117.00
Abnormal Earnings	-130.00	-117.00
PV(Abnormal Earnings) = AE/1.13	-115.04	-103.54
Period 2		
Sales	200.00	200.00
less Inventory Cost	100.00	0.00
Earnings	100.00	200.00
less Required Return * Book Value	130.00	117.00
Abnormal Earnings	-30.00	83.00
PV(Abnormal Earnings) = AE/1.13^2	-23.49	65.00
Value	861.46	861.46

It is important to not take too much comfort from the self-correcting nature of the accounting entries. The example above seems to suggest that the choice of accounting methods is irrelevant. However, there are many reasons to prefer an accounting system that reflects the economic reality as accurately as possible. The accounting values will influence the perception of the business' performance by those performing the valuation and could affect the choice of assumptions. So while the DCF and AE approaches will produce the *same* value, they may produce an incorrect value if the accounting system severely distorts the perception of value creation.

More importantly, as will be shown in the detailed discussion below, the DCF and AE approaches result in a significantly different split between the value within the forecast horizon and the value attributed to the terminal value. A more accurate accounting system will result in more of the value being accurately reflected in the book value (or within the forecast horizon) and less of it attributed to the terminal value. Given the healthy skepticism needed to assess terminal value estimates, this could be an important consideration in some valuations.

5.3 Application to P&C Insurance Companies

5.3.1 Example

To see how the abnormal earnings approach could be used to value a P&C insurance company, the example used earlier will be continued. The following components of the AE method are highlighted for clarity:

- Book Value – The beginning book value is perhaps the easiest component to estimate, since it will in most cases be the reported book value of the equity of the firm. Nonetheless, two adjustments may need to be made. First, any systematic bias in the reported asset and liability values should be eliminated. For P&C insurers, this may involve restating the reported loss reserves. Second, it is common to make an adjustment to reflect the *tangible book value* rather than the reported book value. The tangible book value of the firm is simply the reported book value adjusted to remove the impact of intangible assets such as goodwill. In subsequent periods, the (tangible) book value is adjusted to reflect the net income less dividends and share repurchases plus any capital contributions[30].

- Net Income During Forecast Horizon – The net income estimates for the forecast horizon are determined using the same forecasting models used earlier. Here, no adjustments are made to reflect free cash flows. In this process it is acceptable, though not necessary, to adjust the accounting basis to remove any biases that may exist in the accounting system and develop net income estimates that more closely reflect economic reality.

  For example, under U.S. GAAP accounting P&C loss reserves generally are not discounted[31]. Some analysts would therefore argue that the book value should be adjusted to reflect the discounted loss reserves as this might more closely reflect the economic value of these liabilities. If this is done, then there should be a corresponding adjustment to the assumed ROE, since the same earnings will be generated from a larger capital base.

  If reserves are discounted, it is also important to consider what rate is appropriate to discount the loss reserves. Some would use a risk-free rate. However, this would not truly reflect the economic value of the liabilities unless the liabilities were adjusted to also include a risk margin[32].

- Required Rate of Return[33] – As in the DDM and DCF approaches, the required return used in an AE valuation should reflect the equity investors' appropriate discount rate. The CAPM can be used for this purpose.

- Abnormal Earnings – Abnormal earnings equal the amount by which net income exceeds the required income. Required income is the product of the required rate of return and the beginning of period book value.

- Growth Rate Beyond Forecast Horizon – In this model growth in abnormal earnings reflects both the growth rate in the book value of the firm as well as the amount by which the ROE exceeds the required return. Even in cases where the book value is growing significantly, as in the case where dividends are not paid and the invested asset portfolio grows, abnormal earnings could be declining and could even be zero. For this reason, terminal value growth rates under this method will quite often be very low (or negative).

  Recalling the clean surplus condition discussed in Footnote 29, it is also important to ensure that the growth in book value that is assumed does not require additional capital contributions. Otherwise, the valuation will not accurately reflect the value to the current equity holders.

[30] This follows the "clean surplus condition" discussed in Footnote 29.

[31] One notable exception is certain tabular workers' compensation reserves.

[32] See Butsic or the CAS Fair Value White Paper.

[33] The terms "cost of capital" or "hurdle rate" are quite commonly used to refer to this required return in this context.

Example 5 – Abnormal Earnings Valuation for ABC Insurance Company

Using the same financial model results for ABC Insurance Company as in the previous example, key financial statement variables are summarized below and used to estimate the Abnormal Earnings in each period of the forecast.

Table 24: Calculation of Abnormal Earnings

	2005	2006	2007	2008	2009
GAAP Equity - Beginning of Year	103,500	108,624	113,274	117,648	122,422
Required Return	8.95%	8.95%	8.95%	8.95%	8.95%
Normal Earnings	9,263	9,722	10,138	10,529	10,957
Net Income	17,193	17,236	17,446	18,376	18,967
Abnormal Earnings	7,930	7,514	7,308	7,847	8,010

To estimate the equity value, it is important to estimate the growth rate of the abnormal earnings. One fairly optimistic approach would be to estimate the rate of growth in the book value of the firm and assume that the difference between the ROE and the required return is constant in perpetuity.

Table 25: Calculation of Abnormal Earnings Growth Rates

	2005	2006	2007	2008	2009
GAAP Equity - Beginning of Year	103,500	108,624	113,274	117,648	122,422
GAAP Equity - End of Year	108,624	113,274	117,648	122,422	127,250
Growth in Book Value	5,124	4,650	4,374	4,774	4,828
Book Value Growth Rate	5.0%	4.3%	3.9%	4.1%	3.9%

These book value growth rates and constant abnormal earnings as a percentage of book value would result in an abnormal earnings growth rate of roughly 4.0%. Using that assumption in perpetuity would be very optimistic. It is more likely that the difference between ROE and the required return will decline to zero over a finite time horizon. For simplicity here, abnormal earnings will be assumed to be constant (growth rate equal to zero) and the valuation will be done using different assumptions with regard to the time horizon over which the abnormal earnings will persist.

The simplest case to show first is the case where abnormal earnings continue in perpetuity.

Table 26: Valuation Using Abnormal Earnings Method – Constant AE in Perpetuity

	2005	2006	2007	2008	2009	Terminal Value
Abnormal Earnings	7,930	7,514	7,308	7,847	8,010	89,494
PV Factor	0.918	0.842	0.773	0.710	0.651	0.651
PV	7,279	6,330	5,651	5,569	5,218	58,299
Sum of PV(AE)	88,345					
Beginning Book Value	103,500					
Total Equity Value	191,845					

To calculate the Terminal Value in the table above, the 2009 abnormal earnings of $8,010 are assumed to be constant and continue in perpetuity. When discounted to the valuation date, the terminal value represents 30% of the total equity value.

Sensitivity Analysis

In any valuation exercise, it is important to test the sensitivity of the results to many of the key assumptions. For example, the terminal value assumed abnormal earnings in perpetuity. As noted, abnormal earnings should often be assumed to decline to zero over some finite horizon. In the long run, abnormal earnings require that the firm earn an ROE in excess of the shareholders' required return. These will be sustainable only if there is a competitive advantage that will not ultimately be competed away.

In the numerical example above, the abnormal earnings were assumed to continue in perpetuity. A more realistic assumption is that the firm is able to earn abnormal returns (i.e. achieve an ROE in excess of the shareholders' required return) for only *n*-years after the forecast horizon. The following table shows what would happen if the abnormal earnings declined linearly over a 5-, 10- or 15-year period[34]. In this case, the terminal value estimates and the resulting total equity values would be as shown below:

Table 27: Sensitivity of Equity Value to Abnormal Earnings Horizon

	Version A – 5 Years		Version B – 10 Years		Version C – 15 Years	
Year	AE	PV of AE	AE	PV of AE	AE	PV of AE
2010	6,675	6,126	7,282	6,683	7,509	6,892
2011	5,340	4,499	6,553	5,521	7,009	5,904
2012	4,005	3,097	5,825	4,504	6,508	5,032
2013	2,670	1,895	5,097	3,618	6,007	4,264
2014	1,335	870	4,369	2,846	5,507	3,587
2015	0	0	3,641	2,177	5,006	2,993
2016	0	0	2,913	1,598	4,505	2,473
2017	0	0	2,184	1,100	4,005	2,017
2018	0	0	1,456	673	3,504	1,620
2019	0	0	728	309	3,004	1,275
2020	0	0	0	0	2,503	975
2021	0	0	0	0	2,002	716
2022	0	0	0	0	1,502	493
2023	0	0	0	0	1,001	302
2024	0	0	0	0	501	138
Terminal Value		16,486		29,030		38,681
PV of Terminal Value		10,740		18,911		25,198
PV of AE 2005-2009		30,047		30,047		30,047
Beginning Book Value		103,500		103,500		103,500
Total Equity Value		144,287		152,458		158,745

The assumption of constant abnormal earnings in perpetuity resulted in $58,299 of terminal value. This value declines substantially (to $10,740; $18,911; or $25,198), if the abnormal earnings eventually decline to zero over a 5-, 10- or 15-year horizon. This emphasis on the ability of the firm to generate abnormal earnings, which is the real source of value creation, is one of the key advantages of this method as compared to the DDM and DCF methods.

[34] For this analysis, the assumption is that there are n more years of potential abnormal earnings and that the amount decreases by $1/(n+1)$ times the 2009 estimated abnormal earnings each year. This ensures n additional years of positive abnormal earnings.

5.3.2 Observations Regarding Example 5

As demonstrated in the previous example, the AE approach takes a different perspective than the DDM and DCF methods. Neither dividends nor free cash flows are really *sources* of value creation. Instead, these measures are more accurately the *consequences* of value creation. By emphasizing the firm's ability to earn abnormal profits, the abnormal earnings approach makes use of assumptions that are more directly tied to value creation.

An additional benefit of the approach is that it de-emphasizes the importance of the terminal value estimates and the assumptions that drive those. In the examples demonstrating the DDM and DCF methods, the terminal values represented 83% and 79% of the total equity value. In the AE estimate, the terminal value represented only 30% of the total equity value even when the abnormal earnings were expected to continue in perpetuity.

These points are emphasized here to remind the reader that the AE method is not simply an algebraic recharacterization of the free cash flow method. Blackburn, et. al. demonstrate that under consistent assumptions these approaches are, in fact, mathematically equivalent. However, the two methods may not necessarily produce the same answers in practice. The use of one method or the other may cause the analyst to focus on different aspects of the business and could result in different assumptions being made.

6. Relative Valuation Using Multiples

The DDM, DCF and AE methods discussed so far share as a critical starting point the availability of long-term forecasts of key financial statement variables. Given the popularity of dynamic financial models in recent years and the simplistic nature of the presentation here, this may not have seemed like a daunting exercise. This is misleading. In reality, reliable forecasts of publicly traded insurers are extremely difficult for outsiders to build.

First, an outsider or minority investor may not have access to data in sufficient detail to properly parameterize the model. Second, without the kind of market knowledge and specific planning data used by company executives, growth and rate adequacy estimates may be difficult to obtain. And third, even a relatively short horizon such as 5 years may stretch the limits of one's forecasting ability.

In this section, a methodology for valuation that appears to avoid the need to deal with these forecasts is presented. In reality, this approach requires the same assumptions needed to prepare the detailed forecasts in the DDM, DCF and AE models are used, though not as explicitly. As a result, this approach tends to appear to be easier to implement.

6.1 Price-Earnings Ratio

6.1.1 P-E Ratio Based on Fundamentals

In various earlier discussions of the terminal value it was noted that one could collapse all of the assumptions underlying a DDM, DCF or Abnormal Earnings into a single multiple.

For instance, in the DDM model a constant dividend payout rate and constant growth rate in perpetuity result in the following formula for the price (per share) of the equity:

$$P_0 = \frac{E(\text{Earnings Per Share}_1) * \text{Dividend Payout Rate}}{k - g}$$

Dividing both sides by the expected earnings per share (EPS) and dropping, for convenience, the expected value operator, this can be written as:

$$\frac{P_0}{EPS_1} = \frac{\text{Dividend Payout Rate}}{k - g}$$

This indicates that the "Price-Earnings Ratio" (P-E ratio) is tied directly to the DDM and can be used to summarize, in a single number, the combined effect of the constant dividend payout rate, the constant growth rate and the appropriate discount rate. The price is then simply this P-E ratio times the expected earnings per share next period.

To see what "typical" P-E ratios might be, assume that the ROE is fixed at 15% but that the dividend payout ratios and discount rates are allowed to vary. The ROE, dividend payout rates and growth rate are linked through the formula,

$$g = (1 - \textit{Dividend Payout Rate}) * \textit{ROE}$$

As a result, the following range of P-E ratios could be obtained using different discount rates and dividend payout rates:

Table 28: Illustrative P-E Ratios (ROE = 15%)

	Dividend Payout Ratio		
Discount Rate	40%	50%	60%
10.0%	40.0	20.0	15.0
12.5%	11.4	10.0	9.2
15.0%	6.7	6.7	6.7

Notice that when the discount rate and the ROE are both 15%, the P-E ratio is constant across different dividend payout rates. This demonstrates a point made previously that the dividend payout ratio, and hence the growth rate, does not affect the value of the firm if the firm's ROE is equal to the discount rate.

6.1.2 Representative P&C Industry P-E Ratios

In the basic formula for the P-E ratio shown above, the estimated future period's earnings were used as the basis for determining the ratio of price to "earnings". The P-E ratio could also be presented in terms of the <u>prior</u> period's earnings; often both approaches are used in practice. To avoid confusion, the former approach using expected future earnings is referred to as the *forward* or *leading* P-E ratio; the latter approach using prior period's earnings is referred to as the *trailing* P-E ratio.

The following table indicates the trailing and forward P-E ratios of several P&C insurers as of June 6, 2005:

Table 29: P&C Insurance Trailing and Forward P-E Ratios[35]

Company	Market Capitalization ($ B)	Trailing P-E Ratio	Forward P-E
American International Group	142.17	14.85	9.89
Hartford Financial Services	22.13	10.07	9.12
Chubb Corporation	16.47	9.92	10.07
ACE Limited	12.55	11.57	7.14
XL Capital Ltd.	10.44	9.24	7.33
Sample Average	203.76	13.44	9.52
P&C Insurance Industry[36]	517.18	13.07	NA

In this table, the trailing P-E ratios are based upon current market prices and 2004 GAAP earnings. It is important to recognize that these trailing P-E ratios for any individual company can be distorted by unusually positive or negative earnings surprises in the past year. For this reason, analysts will often favor the use of *core earnings* that smooth the effects of unusual, non-recurring events or the use of forward P-E ratios that reflect analyst estimates of prospective earnings. The forward P-E ratios shown reflect consensus analyst estimates of prospective earnings.

6.1.3 Alternative Uses for P-E Ratios

The P-E ratio can be used for several purposes:

- Validation of Assumptions – The number of assumptions required to forecast financial results and estimate terminal values can be daunting. In many cases, it may be difficult to verify each assumption against objective benchmarks. However, once the valuation is performed it may be possible to recharacterize the value as a ratio to forward or trailing earnings and compare the resulting P-E ratio to the P-E ratios implied by the market values of peer companies.

  This is instructive because if two firms are expected to have comparable growth rates, dividend payout rates, discount rates, etc. then they should have comparable P-E ratios. If differences in P-E ratios cannot be explained as a result of differences in one or more of these key variables, this might indicate that one or more of the assumptions are inappropriate.

[35] Source: Yahoo! Finance, June 6, 2005.

[36] The industry average trailing P-E is weighted by market value. The universe includes all firms included in the Yahoo! Finance P&C Insurance Industry sector but excludes Berkshire Hathaway (an outlier with significant non-insurance operations) as well as Renaissance Re (due to an apparent data error) and any firm with negative earnings in the most recent period. Industry-wide forward P-E ratios were not available and are not shown.

- Shortcut to Valuation – Aside from the validation of an otherwise full-fledged forecast and valuation, the P-E ratio of peer companies might serve as a useful shortcut to valuation in cases where industry average performance is expected. In this case, a group of peer companies would be selected and their mean or median P-E ratios could be used. Of course, given the skewed nature of such ratios, the median industry P-E may be preferred.
- Terminal Value – Even in instances where a full valuation based on separate forecasts is performed, it may be useful to rely on peer P-E ratios to help guide the terminal value calculation.

  In this case, the one additional point to note is that a reasonable terminal value should be based on assumptions appropriate as of the end of the forecast horizon. If, for instance, the industry is expected to experience excessive short-term growth and then slow down to a low-growth steady state, the current valuations of peer companies will reflect this short-term high growth rate to some extent. The current P-E ratios may therefore overstate the appropriate P-E ratio at the forecast horizon.

6.2 Price to Book Value Ratio

The P-E ratio described above is just one of numerous "multiples" that can be used in this way. As another example, consider the Price-Book Value multiple (or equivalently the Price to Tangible Book Value). The P-BV ratio is commonly preferred over the P-E ratio when valuing banks, insurance companies and other financial services firms with substantial holdings in marketable securities.

6.2.1 P-BV Ratio Based on Fundamentals

As before, the P-BV ratio is tied directly to the other methods discussed.

For instance, consider the abnormal earnings approach, which can be written as:

$$\begin{aligned}\text{Price} &= \text{BV}_0 + \sum \frac{\text{AE}_\text{i}}{(1+k)^\text{i}}\\ &= \text{BV}_0 + \frac{[\text{BV}_0 * \text{ROE}_1 - \text{BV}_0 * k]}{(1+k)} + \frac{[\text{BV}_1 * \text{ROE}_2 - \text{BV}_1 * k]}{(1+k)^2} + \frac{[\text{BV}_2 * \text{ROE}_3 - \text{BV}_2 * k]}{(1+k)^3} + \cdots\end{aligned}$$

If the book value is assumed to grow at a constant rate, *g*, and the ROE is assumed to be constant, then this can be written as:

$$\begin{aligned}\text{Price} &= \text{BV}_0 + \frac{\text{BV}_0[\text{ROE} - k]}{(1+k)} + \frac{\text{BV}_0(1+g)[\text{ROE} - k]}{(1+k)^2} + \frac{\text{BV}_0(1+g)^2[\text{ROE} - k]}{(1+k)^3} + \cdots\\ &= \text{BV}_0 + \frac{\text{BV}_0[\text{ROE} - k]}{(k\text{ - g})}\end{aligned}$$

Finally, dividing both sides by the beginning book value, the P-BV ratio is given as:

$$\frac{\text{Price}}{\text{BV}} = 1 + \frac{\text{ROE} - k}{k\text{ - g}}$$

Note that this derivation assumed that the growth rate in book value and the excess return per period (ROE – k) would persist in perpetuity. This will rarely be the case. The excess returns would eventually invite competition that will put pressure on the ROE, the growth rate or both. Alternate formulas that reflect a period after which the excess returns decline to zero can be easily derived[37]. Nonetheless, the previous formula demonstrates the important link between the P-BV multiple and fundamental firm characteristics such as the ROE, the growth rate and the discount rate.

---

[37] For example, if after 5 years the ROE is assumed to decline to the level of the cost of capital, the P-BV ratio would be: $\frac{\text{Price}}{\text{BV}} = 1 + \frac{ROE - k}{k\text{ - }g}\left(1 - \left(\frac{1+g}{1+k}\right)^5\right)$.

If a constant ROE of 15% is assumed, the growth rate and the discount rate can be varied to derive the following range of P-BV ratios:

Table 30: Illustrative P-BV Ratios (ROE = 15%)

	Growth Rates		
Discount Rate	0%	2%	4%
10.0%	1.50	1.63	1.83
12.5%	1.20	1.24	1.29
15.0%	1.00	1.00	1.00

6.2.3 Representative P&C Industry P-BV Ratios

The P-BV ratios for several P&C insurers are shown below:

Table 31: P&C Insurance Trailing P-BV Ratios[38]

Company	Market Capitalization ($ B)	Trailing P-BV
American International Group	142.17	1.77
Hartford Financial Services	22.13	1.54
Chubb Corporation	16.47	1.57
ACE Limited	12.55	1.25
XL Capital Ltd.	10.44	1.34
Sample Average	203.76	1.67
P&C Insurance Industry	517.18	1.54

6.2.3 Alternative Uses for P-BV Ratios

Just as in the case of the P-E ratios, the P-BV ratio can be used to validate other forecasts, serve as a shortcut or be used as a terminal value estimate in other approaches. Because it is linked directly to these other methods, industry peer P-BV multiples can serve as a useful benchmark.

6.3 Firm vs. Equity Multiples

Recall the two alternative methods of applying the DCF approach. The FCFF method values the entire firm and subtracts the value of debt to obtain the equity value; the FCFE method values the equity directly. The two examples shown above, the P-E and the P-BV, both focus on per share equity measures in the denominator. These multiples could just as readily have used a firmwide measure, such as firmwide revenue or total asset value as the basis for a multiple. However, for the same reasons that valuing the equity directly using free cash flows to equity (FCFE) is preferred when valuing P&C insurers, it is preferable to avoid firmwide valuation multiples and limit the use of multiples to equity measures.

6.4 Market vs. Transaction Multiples

The P-E and P-BV ratios shown above were based on the market price of the companies' shares on a particular day, their most recent financial statement values and current analyst estimates for next year's earnings and book value. Of course the market value and forecasted financial statement values fluctuate, sometimes significantly, from day to day and so it may often be useful to observe these ratios over a number of time periods.

[38] Source: Yahoo! Finance, June 6, 2005.

Some practitioners prefer to avoid these fluctuations of market multiples and focus instead on *transaction multiples* based on actual merger or acquisition prices or initial public offerings (IPOs). For example, below is a table of recent transaction multiples for several P&C insurance companies:

Table 32: Transaction Multiples[39]

Company	P-E	P-BV	Transaction	Year
Aspen Insurance Holdings Ltd.	13.10	1.10	IPO	2003
AXIS Capital Holdings Ltd.	28.60	1.40	IPO	2003
Endurance Specialty Holdings Ltd.	13.20	1.00	IPO	2003
Infinity Property and Casualty Corp.	13.90	0.90	IPO	2003
Mercer Insurance Group, Inc.	20.00	0.70	IPO	2003
United National Group, Ltd.	18.50	1.30	IPO	2003
Safety Insurance Group, Inc.	24.50	1.30	IPO	2002
Montpelier Re Holdings, Ltd.	20.20	1.00	IPO	2002
Travelers Property Casualty	17.50	1.00	IPO	2002

One advantage of transaction multiples is that typically the price in these transactions is based on a complex negotiation with sophisticated parties on both sides. As a result, some practitioners consider these prices to be more meaningful than multiples based solely on current market prices. However, there are several reasons to be cautious:

- Control Premiums – M&A transaction prices typically contain what might be considered "control premiums" that reflect the buyer's willingness to pay more for a company in order to gain control of its operations and make different strategic and managerial decisions than the current management. In these cases, the multiples based on current operations and/or current analyst forecasts might be misleading.
- Overpricing in M&A Transactions – Academic studies of M&A transactions[40] show that when mergers and acquisitions increase total shareholder value, most of these gains accrue to the target firm's shareholders and not the acquiring firm. This suggests that acquiring firms have a tendency to overpay. There are multiple causes for this, including managerial hubris, the difficulties of integrating management structures and the failure of planned synergies to fully materialize. But regardless of the reason, it would be prudent to consider this when using M&A transaction multiples.
- Underpricing in IPO Transactions – When firms undertake an initial public offering (IPO) there is a great deal of disclosure and thorough analyses conducted by the firm's bankers as well as investors. This analysis conducted during the IPO process ought to suggest a greater degree of reliability for IPO prices than general market prices. However, the underpricing of IPOs, reflected in the downward bias in initial offering prices, has been widely recognized and documented in numerous academic studies[41]. In recent years, particularly during the technology bubble of the late 1990s, a misalignment of the investment bankers' and managers' interest with those of the shareholders greatly exacerbated this problem[42]. IPO pricing multiples should therefore be interpreted carefully.
- Reported Financial Variables – Even in cases where the prices in M&A and IPO transactions are more reliable, it may not be the case that the reported multiples are as accurate. This is because the reported multiples will be based on either the prior period's

[39] Source: Conning & Company

[40] See Damodaran, Investment Fables

[41] See Ritter, "Initial Public Offerings"

[42] See Partnoy, *Infectious Greed*

financial statements or some published analysts' estimates of next period's financial statements. The prices themselves may have been based on different forecasts. As a result, the multiples may not accurately reflect the buyer's underlying assumptions about growth rates, ROE assumptions and discount rates.

- Underlying Economic Assumptions – By definition, transaction multiples will typically come from *past* transactions that may have been carried out in a different economic environment. Key valuation variables that are imbedded in these multiples, such as interest rates, industry growth rates and industry profitability outlooks, may no longer be appropriate.

To understand the potential variation in valuation multiples over time, consider the following table of P&C insurance multiples over a 10-year period:

Table 33: P&C Insurance Industry Mean Market Multiples[43]

Year	Price to Earnings	Price to Book Value
1985	21.0	1.5
1986	10.0	1.6
1987	19.0	1.2
1988	12.0	1.5
1989	10.0	1.3
1990	11.0	1.5
1991	15.0	1.3
1992	15.0	1.1
1993	18.0	1.4
1994	9.0	1.3
Average	14.0	1.4

Even during this short time period, P&C valuation multiples exhibit variation that would be significant in practice, with high and low multiples as much as 50% above and 36% below the mean multiples.

[43] Source: Conning & Company

Example 6 – Relative Valuation

Consider a P&C insurer with projected 2005 Earnings of $1.5 billion and a beginning book value of $10 billion. Using the average forward P-E ratio for the five firms shown in Table 29 and the average trailing P-BV ratio for the five firms shown in Table 31, the following three estimates of the value of this firm can be produced:

Table 34: Valuation Based on Earnings and Book Value Multiples

Method 1: Forward P-E Ratio	
Forward Earnings	$1.50 B
P-E	9.52
Equity Value	$14.28 B
Method 2: Trailing P-BV Ratio	
Trailing Book Value	$10.00 B
Trailing P-BV	1.67
Equity Value	$16.70 B
Average	$15.49 B

It is important to recognize that this example utilized the average forward P-E and trailing P-BV ratios for five selected companies that did not necessarily have identical operations. In an actual application, it would be important to assess the appropriateness of each of the peer companies used in this average. Companies with different underlying fundamentals (growth rates, risk profiles, leverage ratios, etc.) would not be expected to have identical P-E or P-BV ratios and therefore the peer group has to be carefully constructed.

6.5 Application of Relative Valuation for Multi-Line Firms

Among the key issues to assess in the selection of peer companies is the comparability of the underlying businesses. This becomes particularly difficult in a realistic application because most insurers operate in a variety of markets, each with their own growth rates and risk profiles. The universe of closely comparable firms is actually quite small.

This issue is best illustrated by deviating for a moment from the focus on P&C insurers only and considering how relative valuation might be applied to a multi-line insurer with P&C, Life, and Financial Services businesses. In each case, relative valuation can be used with the segment-specific financial measures and multiples based on firms that operate in only the specific segment of interest. Alternatively, peer companies with comparably diverse operations can be used along with the firmwide financial measures. In either case, the peer groups are likely to be quite limited and considerable effort will be required to assess the results.

6.5.1 Use of Pure Play Peers

Consider the case of a hypothetical diversified insurer, referred to here as Study Note Insurer (SNI). SNI is assumed to represent a diversified financial services firm with operations that include P&C insurance, life insurance and other financial services businesses such as trading, premium financing, etc.

The valuation of SNI would proceed in the following fashion:

- Collect Financial Data by Segment

  Separate the firm into its distinct business segments, each with its own growth rate, profitability and risk level. The three business segments used include:

- P&C Insurance
- Life Insurance
- Financial Services

Use either published financial reports (for trailing values) or independent forecasts (for forward looking values) to obtain key financial variables for each of SNI's segments. In practice, this could prove to be more of a challenge than it appears, depending on the degree of segment detail provided in the firm's financial statements.

In the table below, segment-specific trailing earnings for the most recent fiscal year and an allocation of the total book value of the firm to each business segment are shown. The book values might reflect adjustments for reserve adequacy, the removal of goodwill or similar adjustments to ensure comparability with other firms.

Table 35: SNI P&C Segment Financial Data ($ Millions) – Actual Amounts from Latest Fiscal Year

Current Year	P&C	Life	Fin Services
Earnings	561	839	478
Book Value	3,058	6,160	2,137

Also of interest might be a *smoothed* estimate of earnings that reflects a forward-looking best-estimate of next period's earnings. These smoothed earnings will remove any unusual results from the most recent period and reflect amounts that might reflect a more useful base from which to project future earnings. In practice, it is common to use current actual book value and an average ROE to derive the smoothed earnings. For simplicity, the analysis is limited to the use of trailing earnings in this example.

- Peer Company Selections (Pure Play Companies)

The next step is to identify peer companies in each of the business segments. Ideally, one would want to identify publicly traded firms whose operations consist solely of either P&C insurance, life insurance or financial services businesses. The reliance on single-business entities, known as "pure play" firms, is intended to ensure that the underlying financial characteristics of each business are reflected.

To ensure that the selected companies are appropriate peers for each of SNI's segments, it would be necessary to compare the firms' respective businesses (products offered, markets served, etc.). The ROE, financial leverage and growth rates of the firms would be reviewed to ensure that the firms were comparable on all of these bases.

To highlight the limitations one might encounter, only two peers are identified for the P&C segment and one of them is assumed to have negative trailing earnings that make its trailing P-E ratio meaningless. Four life insurance and two financial services two peers are also identified.

- Choice of Multiples

To avoid relying on a single multiple, several valuation multiples would be used, such as Price/Earnings (trailing) and Price/Book Value (trailing).

The following table shows the peer multiples for the P&C segment:

Table 36: P&C Insurance Segment Peer Multiples

Multiple	P&C Peer 1	P&C Peer 2	Simple Average
P-E	17.07	N/A	17.07
P-BV	1.75	2.27	2.01

The Life Insurance segment multiples are as follows:

Table 37: Life Insurance Segment Peer Multiples

Multiple	Life Peer 1	Life Peer 2	Life Peer 3	Life Peer 4	Simple Average
P-E	20.10	19.06	13.77	25.78	19.68
P-BV	2.41	2.33	3.00	4.25	3.00

And the Financial Services segment multiples are as follows:

Table 38: Financial Services Segment Peer Multiples

Multiple	Asset Mgt Peer 1	Asset Mgt Peer 2	Simple Average
P-E	29.75	19.89	24.82
P-BV	6.10	2.78	4.44

- Application of Multiples for Segment Valuation

  The P&C segment financial data is then combined with the P&C peer multiples to obtain the following estimates of the value of the P&C segment.

Table 39: P&C Segment Valuation ($ Millions)

Valuation Basis	SNI Amount	Peer Multiple	Segment Value
Earnings	561	17.07	9,576
Book Value	3,058	2.01	6,147
Average			**7,862**

Similar analyses are done for the other two segments, as shown in the following two tables.

Table 40: Life Segment Valuation ($ Millions)

Valuation Basis	SNI Amount	Peer Multiple	Segment Value
Earnings	839	19.68	16,512
Book Value	6,160	3.00	18,480
Average			**17,496**

Table 41: Financial Services Segment Valuation ($ Millions)

Valuation Basis	SNI Amount	Peer Multiple	Segment Value
Earnings	478	24.82	11,864
Book Value	2,137	4.44	9,488
Average			**10,676**

- Total Firm Value

  The total value of SNI's equity would reflect the sum of the segment values, as shown in the table below:

Table 42: SNI Valuation Summary ($ Millions)

Segment	Value
P&C Insurance	7,862
Life Insurance	17,496
Financial Services	10,676
Total	36,034

- Validation Against Other Diversified Insurers

  Since the universe of possible peer companies by segment is very limited, it may be difficult to select more than a few firms in each segment. If these selected peer companies are not truly comparable, the results could be biased.

  As an alternative to the segment valuation, other diversified insurance/financial services firms could also be used as the source of peer multiples. These diversified firms would be selected so that they are similar to SNI in many respects – similar businesses, similar ROE, similar S&P claims paying rating, similar CAPM betas, etc.

  Peer multiples for three diversified insurers are summarized as follows:

Table 43: Peer Multiples – Diversified Insurance/Financial Services

Multiple	Diversified Peer 1	Diversified Peer 2	Diversified Peer 3	Average
P-E	17.53	16.89	11.48	15.30
P-BV	2.34	2.25	1.35	1.98

  When the average multiples are applied to SNI's total earnings and book value across all segments, the following results are obtained:

Table 44: SNI Valuation – Diversified Insurance/Financial Services Peers ($ Millions)

Valuation Basis	SNI Amount	Peer Multiple	Equity Value
Earnings	1,878	15.30	28,733
Book Value	11,355	1.98	22,483
Average			**25,608**

Additional Considerations

The following additional observations are made with respect to the above example:

- Choice of Peer Companies – The valuation relied heavily on the assumption that the average multiples for the selected peer companies are appropriate for SNI. The validity of the chosen peer companies depends on whether the ROE, growth rate and discount rate assumptions are comparable for these firms (or at least the net effect is comparable). This is ultimately a matter of informed judgment.

Consider, for instance, the peer companies selected for the Life Insurance segment:

Table 45: Comparison of Life Segment Peer P-E Multiples

Life Peer 1	20.10
Life Peer 2	19.06
Life Peer 3	13.77
Life Peer 4	25.78
Simple Average	19.68

The first two firms' multiples are approximately equal to the average multiple. However, one firm's P-E is approximately 30% lower than this average and another firm's P-E is approximately 30% higher than this average. As a result, which of these four firms are included in the average multiple calculation can have a material impact. Determining which of the firms has operations most like SNI's operations is important.

Notice also that the valuation used trailing P-E ratios in the analysis. The large differences in P-E ratios could merely reflect special circumstances in the latest reporting year for one or more of these firms that caused their earnings to be artificially lower or higher than expected. This may not truly reflect differences in expected ROEs, growth rates or discount rates and therefore should not be used to proxy for the appropriate ROE, growth and discount rate assumptions that would be used in an explicit DCF valuation.

Growth rates and discount rates for SNI and their peers could very well differ substantially due to underlying fundamental differences in their operations.

- Simple Average vs. Weighted Average Multiples – Notice that when valuing the various segments, the peer companies' respective multiples were averaged using a simple average. If the peer firms are not roughly the same size, a weighted average might be more appropriate.

7. Option Pricing Methods

Many recently published valuation textbooks now include extensive discussion of the use of option pricing theory in the valuation of the equity of a firm. This section briefly discusses the rationale behind this approach and its potential applicability to insurance company valuation.

Two related approaches are presented: (a) valuing the equity as a call option rather than as a discounted stream of future dividends, cash flow or abnormal earnings and (b) the valuation of *real options* as an additional source of value to be added to the DCF, AE or relative valuation results.

7.1 Valuing Equity as a Call Option

7.1.1 Background

This method is based on Merton's characterization of equity as a call option on the company's assets, with a strike price equal to the face value of the debt.

When a firm is owned entirely by equity holders, they own all of the assets of the firm – the physical assets plus the income that those assets produce over the life of the company. If the equity holders issue debt (i.e. borrow money), then the equity holders no longer own all of the value of the firm, *V*. Instead, they own the excess of the value of the firm over the debt that they have to repay at time *T*, denoted *D*. In other words $E_T = \max(V_T - D, 0)$, which looks like a call on the value of V_T with a strike price of *D*.

When the equity holders borrowed the present value of *D*, they gave all of the assets of the firm to the bondholders, who will keep them if the debt is not repaid. However, by repaying the debt at time *T*, the equity holders have the right to buy back the assets of the firm by paying *D*. If $V_T < D$ on that date, they will not buy the assets back and will let the bondholders keep the assets. In other words, they will default.

To value the equity of a firm as a call option on the assets, the Black-Scholes option pricing formula can be used, with some modifications. For instance, instead of using the value of the stock and its volatility as inputs, the value and volatility of <u>all</u> of the firm's assets are the critical inputs. In addition, the strike price is set equal to the face value of the debt and the expiration date for the option is set equal to the (single) expiration date of the debt.

7.1.2 Application to P&C Insurers

For many years after Merton's original presentation, this approach remained a purely theoretical discussion and was not commonly used as a valuation framework because of its many practical limitations. In recent years, as option pricing methods have become more widely understood, the use of this approach has grown. For instance, a variation of this approach is now used to estimate probabilities of default for publicly traded firms[44].

However, when it comes to the valuation of P&C insurance companies, this is still largely a theoretical model. The reason for this is similar to why equity valuation methods rather than firm valuation methods are generally preferred for insurance company valuations – the notion of "debt" for an insurance company is not well defined. An insurer's policyholder liabilities are essentially indistinguishable from other debt from the perspective of the equity holder. Due to the complexity of the policyholder liabilities, a single expiration date for all of an insurer's "debt" cannot be readily approximated.

Given the limitations of this approach in a practical valuation analysis, this approach will not be explored further in this study note.

---

[44] The most widely known application is the Moody's/KMV Credit Default Model.

7.2 Real Options Valuation

7.2.1 Background on Real Options

Another use of option pricing theory of relevance to valuation is the real options framework. The real options approach attempts to value various sources of managerial flexibility that can often be thought of as put and call options. Some of the most common real options include the following[45]:

- Abandonment Option – Many projects can be terminated early and the investment sold for its liquidation value less closing-down costs. This option is valued as an American put on the value of the project with a strike equal to the net liquidation proceeds.
- Expansion Option – Projects that are successful often contain an option to expand the scope of the project and capture more profits. This is valued as an American call option on the (gross) value of the additional capacity with a strike price equal to the cost of creating the capacity.
- Contraction Option – This is the opposite of the expansion option. It is valued as an American put on the (gross) value of the lost capacity with a strike equal to the cost savings.
- Option to Defer – Otherwise known as the option to wait, this is an American call on the value of a project. It essentially measures the value of being able to hold off on a project until more information is known – hence, preventing the bad outcomes at the expense of maybe giving up some interim revenue in the good outcomes.
- Option to Extend – This is an option to extend the life of a project by paying a fixed amount. It is valued as a European call option on the asset's future value.

The argument that managerial flexibility has value that should be included within the equity valuation is appealing. However, care must be taken to distinguish between managerial choices that have value and managerial choices that do not. For instance, all firms have the "flexibility" to buy assets at their market prices, but this does not in itself create value. Value is created only when assets can be purchased at less than their fair value or when the firm has exclusive access to opportunities.

7.2.2 Example of Real Option Analysis

The valuation of real options is considerably more complex than the valuation of options on financial instruments. Practices vary widely with respect to implementation of standard option pricing models for these sorts of options. For the sake of clarity, this section will provide a brief demonstration of just one particular method used by some insurance company equity analysts. The example will be intentionally simplified to highlight the rationale behind this methodology. The specific formulas used here have certain limitations and may not be applicable in all situations.

Assume an insurer has a *new* business opportunity that it has not yet exploited due to uncertainty with regard to its value. Based on current assumptions, the opportunity will require an initial investment of $500 million and will generate an expected ROE (in perpetuity) of 8.95%, exactly equal to its cost of capital. There is uncertainty with respect to the ROE that will be achieved, but this uncertainty will diminish over a three year period.

Using the Abnormal Earnings valuation methodology, it is easy to see that the *gross* value of the opportunity equals the initial book value of $500 million because the expected abnormal earnings are equal to zero in every future period. Given the required investment of $500 million, the *net* value of this opportunity is zero and there would be no incentive for the firm to enter into this business.

Nonetheless, there may be a real option value to consider here. Assume that the firm's flexibility allows it to essentially lock in the required investment for a set period, say 3 years for the sake of the example. During this time the uncertainty with respect to the ROE that can be achieved will be

[45] This list is taken from Hull. Other sources for more information on real options valuation include Damodaran and Trigeorgis.

resolved. If it turns out that the ROE on this business exceeds the current expected value of 8.95% in perpetuity and the firm can still invest only $500 million in book value to enter the business, then there may be a real option value associated with this flexibility.

The value of their flexibility to delay making the investment may be estimated using the Black-Scholes option pricing formula and an assumption regarding the volatility of the value of the business' cash flows. The volatility assumption would be based upon the volatility of the ROE and would be impacted by other valuation factors such as whether the abnormal earnings continue in perpetuity. For the sake of simplicity, the volatility is arbitrarily set at 20% for this example.

The specific formula is summarized as follows:

$$\text{Real Option Value} = AN(d_1) - Ie^{-rT}N(d_2)$$

where A = Current Value of Cash Flows ($500), I = Required Investment ($500), r = continuously compounded risk-free interest rate (4.55%), T = Time to Expiration (3), and σ = Volatility of Current Value (20%). As in the standard Black-Scholes model, N( ) is the standard normal CDF, and d_1 and d_2 are defined as follows:

$$d_1 = \frac{\ln(A/I) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(A/I) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

Table 46: Real Option Value of New P&C Insurance Opportunities

Asset Value (A)	500
Strike Price (I)	500
Volatility (σ)	20.0%
Time to Expiration in years (T)	3.00
Risk Free Rate (r)	4.55%
d_1	0.567
d_2	0.221
$N(d_1)$	0.715
$N(d_2)$	0.587
Option Value ($ Millions)	101.1

As a result of these calculations, it would be appropriate to include an additional $101.1 million to the valuation of the firm. The underlying new business opportunity does not have any value to the firm now, even if the investment were made to enter the business. However, the firm's ability to wait for three years before committing to the investment provides it with a real option. The value of this option, as opposed to the value of the underlying business, should be added to the estimates produced by valuing all of the firm's existing businesses.

7.2.3 Practical Considerations

The calculations described in the previous example were intended to demonstrate the concepts underlying attempts to include the value of managerial flexibility in the value of a firm. In practice, it may be substantially more difficult to a) identify the new businesses for which some real option value may exist, b) assess the current value of these businesses and c) determine whether the firm actually has the ability to enter these businesses at a fixed price or at a price that otherwise differs from the businesses' market value. It is appropriate to contemplate the potential for firms to have exclusive rights or exclusive abilities to capitalize on new business opportunities, but placing a dollar value on these opportunities requires considerably more judgment and insight than the simplified example here might suggest.

7.2.4 Key Valuation Considerations

In addition to the practical considerations raised in the previous section, there are also a variety of technical issues that must be considered in the actual valuation formula. The following is a sample of some of these considerations:

- Valuing the Underlying Business Cash Flows – In this example the gross value of the business was valued using the AE method but the abnormal earnings were assumed to continue in perpetuity. This assumption made the value of the underlying cash flows change each period primarily as the result of the volatility of the ROE.

  In practice, abnormal earnings periods usually have a finite life. As a result, after each period passes with the option not exercised, the gross value of the cash flows will decline. This effect is comparable to the effect on the stock price after cash dividends are paid and adjustments to the option valuation formula similar to those made when valuing options on stocks that pay dividends may be appropriate.

- Time to Option Maturity – In this example the time to maturity was assumed to be known and had a finite value. In practice, real options are likely to have uncertain maturities – or possibly no maturity date at all.

- Exercise Type – The example was simplified by assuming that the option could be exercised only at maturity. In practice, real options are more likely to be American-style options that can be exercised any time until maturity. Appropriate adjustments to the option pricing formula would therefore be made in these cases[46].

- Appropriate Valuation Formula – This example used the Black-Scholes formula to value the option. For certain real options, the implicit assumption of a lognormal underlying asset price distribution may be inappropriate and other valuation formulas may be appropriate.

7.2.5 Assessing the Reasonableness of Real Option Values

To assess the reasonableness of the real option valuation results, it is helpful to consider the following characteristics that make real options more valuable:

- Options are more valuable when new information will be discovered prior to their expiration date that will allow for a more informed decision. If no new information exists, then waiting to make a decision might be convenient but it won't necessarily add significant value to the firm.

- Expansion options are valuable only if there is some exclusive right or ability to exercise them. It is not sufficient to say that new business opportunities might come along in the future. If there is competition, other firms might also attempt to capitalize on these opportunities, driving up the exercise cost and eroding any net value impact to the firm upon exercise.

- The exercise price must be fixed in order for the option to have value. As an extreme example, an "option" to purchase an asset at some future date at the <u>then current</u> market price does not have any value.

---

[46] See Hull.

8. Additional Considerations

Given the limited scope of this study note, a variety of complicating factors have been ignored. This section will include a partial list of these factors, but readers are encouraged to review the sources included in the References section for more complete details.

Topics of particular interest may include the following:

- Complex Capital Structures – The valuation methods discussed here reflect the value to all of the stakeholders who have a claim on the equity value of the firm. These stakeholders may include a broader group than just the current shareholders of the firm. Determining that value of the common shareholders' interests therefore might require more than just dividing the total equity value by the number of outstanding common shares.

  One adjustment may include special consideration for preferred shareholders. Another more complicated adjustment is to reflect the value of any outstanding warrants or employee stock options. These are call options issued by the firm to investors, management or other employees. The value of the publicly traded shares must take into account the effect on firm value and the number of shares outstanding if and when these options are exercised.

- Valuation of Non-Operating Assets – The methods discussed here assumed that the assets of the firm were used to generate the earnings and cash flows depicted in the valuation formulas. Other assets may require special considerations.

- International Considerations – A variety of issues associated with international operations have been ignored, including methods needed to assess the consolidated financial statements for globally diversified firms and methods used to reflect currency risk in the valuation methods.

The text by Damodaran and the text by Stowe, Robinson, Pinto and McLeavey each provide complete discussions of these and other related valuation topics.

9. References

Blackburn, Jones, Schwartzman and Siegman, 2003, "The Application of Fundamental Valuation Principles to Property/Casualty Insurance Companies". (Available on the CAS Website at: http://casact.org/research/Valpap2.pdf)

Bodie, Kane and Marcus, 2004, *Investments*, 6th Edition, McGraw Hill.

Butsic, Robert P., 1988, "Determining the Proper Interest Rate for Loss Reserve Discounting: An Economic Approach", Casualty Actuarial Society Discussion Paper Program, Vol: May Page(s): 147. (Available on the CAS Website at: http://www.casact.org/pubs/dpp/dpp88/88dpp147.pdf)

CAS Task Force on Fair Value Liabilities, 2000, *White Paper on* Fair Valuing Property/Casualty Insurance Liabilities, Casualty Actuarial Society. (Available on the CAS Website at: http://www.casact.org/research/tffvl/index.htm)

Cornell, Bradford, 1993, *Corporate Valuation: Tools for Effective Appraisal and Decision Making*, New York: Business One Irwin.

Cornell, Bradford, 1999, *The Equity Risk Premium*, New York: John Wiley & Sons, Inc.

Cummins, J. David and Richard D. Phillips, 2004 (Revised), "Estimating the Cost of Equity Capital for Property-Liability Insurers", Presented at 2003 CAS Enterprise Risk Management Seminar.

Damodaran, Aswath, 2002, *Investment Valuation*, 2nd Edition, John Wiley & Sons.

Damodaran, Aswath, 2004, *Investment Fables*, Financial Times Prentice Hall.

Derrig, Richard A and Elisha D. Orr, 2004, "Equity Risk Premium: Expectations Great and Small", Casualty Actuarial Society Forum, Vol: Winter Page(s): 1-44. (Available on the CAS Website at: http://www.casact.org/pubs/forum/04wforum/04wf001.pdf)

D'Arcy, Stephen P., Gorvett, Richard W.; Hettinger, Thomas E.; Walling III, Robert J., 1998, "Using the Public Access DFA Model: A Case Study", Casualty Actuarial Society Forum, Vol: Summer Page(s): 53-118. (Available on the CAS Website at: http://www.casact.org/pubs/forum/98sforum/98sf053.pdf)

Girard, Luke, 2000, "Market Value of Insurance Liabilities: Reconciling the Actuarial Appraisal and Option Pricing Methods," *North American Actuarial Journal*, Volume 4, Number 1.

Girard, Luke, 2002, "An Approach to Fair Valuation of Insurance Liabilities Using the Firm's Cost of Capital," *North American Actuarial Journal*, Volume 6, Number 2.

Halliwell, Leigh J., 2001, "A Critique of Risk-Adjusted Discounting", ASTIN Colloquium, (Available on the CAS Website at:http://www.casact.org/coneduc/reinsure/astin/2000/halliwell1.doc)

Hull, John C. 2006, *Options, Futures and Other Derivative Securities*, 6th Edition, Prentice Hall.

Merton, Robert C. 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", Journal of Finance, Vol. 29, pp. 449-470.

Ohlson, 1995, "Earnings, Book Values and Dividends in Equity Valuation", Contemporary Accounting Research, Volume 11, No. 2 (Spring 1995).

Palepu, Bernard & Healey, 1996, *Business Analysis & Valuation*, Southwestern Publishing.

Partnoy, Frank, 2003, *Infectious Greed: How Deceit and Risk Corrupted the Financial Markets*, Times Books.

Ritter, J.R., 1998, "Initial Public Offerings", Contemporary Finance Digest, v2, 5-31.

Sougiannis, Theodore and Penman, Stephen H., "A Comparison of Dividend, Cash Flow, and Earnings Approaches to Equity Valuation". (Internet address: http://ssrn.com/abstract=15043)

Standard & Poor's, "Property/Casualty Insurance Ratings Criteria". Available on the S&P website at www.standardandpoors.com/ratings.

Stowe, Robinson, Pinto and McLeavey, 2002, *Analysis of Equity Investments: Valuation*. AIMR.

Stewart, G.B., 1991, *The Quest for Value*, New York: Harper Business.

Trigeorgis, Lenos, 1996, *Real Options*, MIT Press.

Wang, Shaun, 1995, "Insurance Pricing and Increased Limits Ratemaking by Proportional Hazards Transforms", Insurance Mathematics and Economics, 17, pp. 43–54.

"Credible Loss Ratio Claims Reserves: The Benktander, Neuhaus and Mack Methods Revisited"

Due to copyright restrictions, the text is not included in this complete PDF.

To access the Werner Hurlimann text, please use the following link.

Please note, candidates are not responsible for mathematical proofs.

http://www.actuaries.org/LIBRARY/ASTIN/vol39no1/81.pdf

Errata to

Credible Loss Ratio Claims Reserves: The Benktander, Neuhaus and Mack Methods Revisited

By Hürlimann, W. in *ASTIN Bulletin* 39(1), 2009

Casualty Actuarial Society[1]

Version 1.0, January 31, 2020

This note presents errata to various tables and formulas in Hürlimann's paper on "Credible Loss Ratio Claims Reserves." Items printed in **red** indicate an update, clarification, or change.

1. Errata

- Table 7.4 of Hürlimann (page 95) should be amended from:

Origin Period	Method				
	collective	individual	Neuhaus	Benktander	optimal
all periods	86,752	87,810	86,751	86,837	86,486
1	14,307	14,307	14,307	14,307	14,307
2	9,964	9,882	9,906	9,891	9,966
3	12,772	12,660	12,706	12,686	12,779
4	11,443	11,112	11,313	11,266	11,484
5	20,826	22,947	21,022	21,219	20,364
6	17,440	16,902	17,498	17,469	17,586

to:

Origin Period	Method				
	collective	individual	Neuhaus	Benktander	optimal
all periods	85,992	87,810	86,751	86,837	86,752
1	14,307	14,307	14,307	14,307	14,307
2	10,043	9,882	9,906	9,891	9,964
3	12,878	12,660	12,706	12,686	12,772
4	11,731	11,112	11,313	11,266	11,443
5	19,284	22,947	21,022	21,219	20,826
6	17,749	16,902	17,498	17,469	17,440

[1] This note was prepared by the Exam 7 Syllabus Committee.

© 2020, Casualty Actuarial Society, All Rights Reserved

- Table 7.5 of Hürlimann (page 95) should be amended from:

Origin Period	Method				
	collective	individual	Neuhaus	Benktander	optimal
2	1.027133	1.028713	1.014146	1.022818	1
3	1.058036	1.065943	1.003002	1.038856	1
4	1.115378	1.153525	1.002692	1.044128	1
5	1.198612	1.376096	1.120972	1.012892	1
6	1.244417	1.740080	1.409648	1.002206	1

to:

Origin Period	Method				
	collective	individual	Neuhaus	Benktander	optimal
2	1.027133	1.028713	1.014146	1.022818	1
3	1.058037	1.065943	1.023277	1.038856	1
4	1.115379	1.153525	1.023764	1.044128	1
5	1.198610	1.376096	1.003211	1.012892	1
6	1.244422	1.740080	1.008555	1.002208	1

- Table 7.10 of Hürlimann (page 97) should be amended from:

Origin Period	Method				
	collective	individual	Neuhaus	Benktander	optimal
2	1.001629	1.001634	1.001405	1.001615	1
3	1.029900	1.031834	1.021187	1.024673	1
4	1.046368	1.051243	1.029035	1.033985	1
5	1.111731	1.146991	1.036943	1.044625	1
6	1.228790	1.548854	1.000149	1.000894	1

to:

Origin Period	Method				
	collective	individual	Neuhaus	Benktander	optimal
2	1.001566	1.001571	1.001342	1.001551	1
3	1.029900	1.031833	1.021187	1.024673	1
4	1.046373	1.051248	1.029039	1.033989	1
5	1.111729	1.146989	1.036940	1.044623	1
6	1.228789	1.548852	1.000148	1.000893	1

© 2020, Casualty Actuarial Society, All Rights Reserved

- The following formula from Hürlimann (page 88, formula 4.14) should be amended from:

$$mse\left(R_i^{ind}\right) = E\left[(R_i^{ind} - R_i)^2\right] = Var\left[R_i^{ind} - R_i\right] = Var\left[R_i^{coll}\right] - 2Cov\left[R_i^{ind}, R_i\right] + Var[R_i]$$

to:

$$mse\left(R_i^{ind}\right) = E\left[(R_i^{ind} - R_i)^2\right] = Var\left[R_i^{ind} - R_i\right] = Var\left[R_i^{ind}\right] - 2Cov\left[R_i^{ind}, R_i\right] + Var[R_i]$$

- The following formula from Hürlimann (page 92) should be amended from:

$$\widehat{Var}[R_i^c] = \left(\hat{Z}_i{}^2 \cdot (1 + f_i) \cdot \left[1 + \frac{1 - p_i}{p_i}\frac{\hat{t}_i}{1 - \hat{t}_i}\right] - 2\hat{Z}_i + 1\right) \cdot Var\left[R_i^{coll}\right]$$

to:

$$\widehat{Var}[R_i^c] = \left(\hat{Z}_i{}^2 \cdot (1 + f_i) \cdot \left[1 + \frac{1 - p_i}{p_i}\frac{\hat{t}_i}{1 + \hat{t}_i}\right] - 2\hat{Z}_i + 1\right) \cdot Var\left[R_i^{coll}\right]$$

© 2020, Casualty Actuarial Society, All Rights Reserved

Measuring the Variability of Chain Ladder Reserve Estimates

by Thomas Mack

MEASURING THE VARIABILITY

OF CHAIN LADDER RESERVE ESTIMATES

Thomas Mack, Munich Re

Abstract:

The variability of chain ladder reserve estimates is quantified without assuming any specific claims amount distribution function. This is done by establishing a formula for the so-called standard error which is an estimate for the standard deviation of the outstanding claims reserve. The information necessary for this purpose is extracted only from the usual chain ladder formulae. With the standard error as decisive tool it is shown how a confidence interval for the outstanding claims reserve and for the ultimate claims amount can be constructed. Moreover, the analysis of the information extracted and of its implications shows when it is appropriate to apply the chain ladder method and when not.

Submitted to the 1993 CAS Prize Paper Competition on 'Variability of Loss Reserves'

Presented at the May, 1993 meeting of the Casualty Actuarial Society.

Reproduction in whole or in part without acknowledgement to the Casualty Actuarial Society is specifically prohibited.

1. Introduction and Overview

The chain ladder method is probably the most popular method for estimating outstanding claims reserves. The main reason for this is its simplicity and the fact that it is distribution-free, i.e. that it seems to be based on almost no assumptions. In this paper, it will be seen that this impression is wrong and that the chain ladder algorithm rather has far-reaching implications. These implications also allow it to measure the variability of chain ladder reserve estimates. With the help of this measure it is possible to construct a confidence interval for the estimated ultimate claims amount and for the estimated reserves.

Such a confidence interval is of great interest for the practitioner because the estimated ultimate claims amount can never be an exact forecast of the true ultimate claims amount and therefore a confidence interval is of much greater information value. A confidence interval also automatically allows the inclusion of business policy into the claims reserving process by using a specific confidence probability. Moreover, there are many other claims reserving procedures and the results of all these procedures can vary widely. But with the help of a confidence interval it can be seen whether the difference between the results of the chain ladder method and any other method is significant or not.

The paper is organized as follows: In Chapter 2 a first basic assumption underlying the chain ladder method is derived from the formula used to estimate the ultimate claims amount. In Chapter 3, the comparison of the age-to-age factor formula used by the chain ladder method with other possibilities leads to a second underlying assumption regarding the variance of the claims amounts. Using both of these derived assumptions and a third assumption on the independence of the accident years, it is possible to calculate the so-called standard error of the estimated ultimate claims amount. This is done in Chapter 4 where it is also shown that this standard error is the appropriate measure of variability for the construction of a confidence interval. Chapter 5 illustrates how any given run-off triangle can be checked using some plots to ascertain whether the assumptions mentioned can be considered to be met. If these plots show that the assumptions do not seem to be met, the chain ladder method should not be applied. In Chapter 6 all formulae and instruments established including two statistical tests set out in Appendices G and H are applied to a numerical example. For the sake of comparison, the reserves and standard errors according to a well-known claims reserving software package are also quoted. Complete and detailed proofs of all results and formulae are given in the Appendices A - F.

The proofs are not very short and take up about one fifth of the paper. But the resulting formula (7) for the standard error is very simple and can be applied directly after reading the basic notations (1) and (2) in the first two paragraphs of the next chapter. In the numerical example, too, we could have applied formula (7) for the standard error immediately after the completion of the run-off triangle. But we prefer to first carry through the analysis of whether the chain ladder assumptions are met in this particular case as this analysis generally should be made first. Because this analysis comprises many tables and plots, the example takes up another two fifths of the paper (including the tests in Appendices G and H).

2. Notations and First Analysis of the Chain Ladder Method

Let C_{ik} denote the accumulated total claims amount of accident year i, $1 \le i \le I$, either paid or incurred up to development year k, $1 \le k \le I$. The values of C_{ik} for $i+k \le I+1$ are known to us (run-off triangle) and we want to estimate the values of C_{ik} for $i+k > I+1$, in particular the ultimate claims amount C_{iI} of each accident year i = 2, ..., I. Then,

$$R_i = C_{iI} - C_{i,I+1-i}$$

is the outstanding claims reserve of accident year i as $C_{i,I+1-i}$ has already been paid or incurred up to now.

The chain ladder method consists of estimating the ultimate claims amounts C_{iI} by

(1) $$\hat{C}_{iI} = C_{i,I+1-i} \cdot \hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{I-1}, \qquad 2 \le i \le I,$$

where

$$(2) \qquad \mathbf{f_k} = \sum_{j=1}^{I-k} C_{j,k+1} \Big/ \sum_{j=1}^{I-k} C_{jk}\,, \qquad 1 \le k \le I-1,$$

are the so-called age-to-age factors.

This manner of projecting the known claims amount $C_{i,I+1-i}$ to the ultimate claims amount $\mathbf{C_{iI}}$ uses for all accident years $i \ge I+1-k$ the same factor $\mathbf{f_k}$ for the increase of the claims amount from development year k to development year k+1 although the observed individual development factors $C_{i,k+1}/C_{ik}$ of the accident years $i \le I-k$ are usually different from one another and from $\mathbf{f_k}$. This means that each increase from C_{ik} to $C_{i,k+1}$ is considered a random disturbance of an expected increase from C_{ik} to $C_{ik}f_k$ where f_k is an unknown 'true' factor of increase which is the same for all accident years and which is estimated from the available data by $\mathbf{f_k}$.

Consequently, if we imagine to be at the end of development year k we have to consider $C_{i,k+1}, \ldots, C_{iI}$ as random variables whereas the realizations of $C_{i1}, \ldots, C_{ik}$ are known to us and are therefore no longer random variables but scalars. This means that for the purposes of analysis every C_{ik} can be a random variable or a scalar, depending on the development year at the end of which we imagine to be but independently of whether C_{ik} belongs to the known part $i+k \le I+1$ of the run-off triangle or not. When taking expected values or variances we therefore must always also state the development year at the end of which we imagine to be. This will be done by explicitly indicating those variables C_{ik} whose values are assumed to be known. If nothing is indicated all C_{ik} are assumed to be unknown.

What we said above regarding the increase from C_{ik} to $C_{i,k+1}$ can now be formulated in stochastic terms as follows: The chain ladder method assumes the existence of accident-year-independent factors $f_1, \ldots, f_{I-1}$ such that, given the development $C_{i1}, \ldots, C_{ik}$, the realization of $C_{i,k+1}$ is 'close' to $C_{ik}f_k$, the latter being the expected value of $C_{i,k+1}$ in its mathematical meaning, i.e.

(3) $E(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}f_k \,, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

Here to the right of the '|' those C_{ik} are listed which are assumed to be known. Mathematically speaking, (3) is a conditional expected value which is just the exact mathematical formulation of the fact that we already know $C_{i1}, \ldots, C_{ik}$, but do not know $C_{i,k+1}$. The same notation is also used for variances since they are specific expectations. The reader who is not familiar with conditional expectations should not refrain from further reading because this terminology is easily understandable and the usual rules for the calculation with expected values also apply to conditional expected values. Any special rule will be indicated wherever it is used.

We want to point out again that the equations (3) constitute an assumption which is not imposed by us but rather implicitly underlyies the chain ladder method. This is based on two aspects of the basic chain ladder equation (1): One is the fact that (1) uses the same age-to-age factor f_k for different accident years $i = I+1-k, \ldots, I$. Therefore equations (3) also postulate age-to-age parameters f_k which are the same for all accident years. The other is the fact that (1) uses only the most recent observed value $C_{i,I+1-i}$ as basis for the projection to ultimate ignoring on the one hand all amounts $C_{i1}, \ldots, C_{i,I-i}$ observed earlier and on the other hand the fact that $C_{i,I+1-i}$ could substantially deviate from its expected value. Note that it would easily be possible to also project to ultimate the amounts $C_{i1}, \ldots, C_{i,I-i}$ of the earlier development years with the help of the age-to-age factors $f_1, \ldots, f_{I-1}$ and to combine all these projected amounts together with $C_{i,I+1-i}f_{I+1-i}\cdot\ldots\cdot f_{I-1}$ into a common estimator for C_{iI}. Moreover, it would also easily be possible to use the values $C_{j,I+1-i}$ of the earlier accident years $j < i$ as additional estimators for $E(C_{i,I+1-i})$ by translating them into accident year i with the help of a measure of volume for each accident year. These possibilities are all ignored by the chain ladder method which uses $C_{i,I+1-i}$ as the only basis for the projection to ultimate. This means that the chain ladder method implicitly must use an assumption which states that the information contained in $C_{i,I+1-i}$ cannot be augmented by additionally using $C_{i1}, \ldots, C_{i,I-i}$ or $C_{1,I+1-i}, \ldots, C_{i-1,I+1-i}$. This is very well reflected by the equations (3).

Having now formulated this first assumption underlying the chain ladder method we want to emphasize that this is a rather strong assumption which has important consequences and which cannot be taken as met for every run-off triangle. Thus the widespread impression the chain ladder method would work with almost no assumptions is not justified. In Chapter 5 we will elaborate on the linearity constraint contained in assumption (3). But here we want to point out another consequence of formula (3). We can rewrite (3) into the form

$$E(C_{i,k+1}/C_{ik} | C_{i1},\ldots,C_{ik}) = f_k$$

because C_{ik} is a scalar under the condition that we know C_{i1}, ..., C_{ik}. This form of (3) shows that the expected value of the individual development factor $C_{i,k+1}/C_{ik}$ equals f_k irrespective of the prior development C_{i1}, ..., C_{ik} and especially of the foregoing development factor $C_{ik}/C_{i,k-1}$. As is shown in Appendix G, this implies that subsequent development factors $C_{ik}/C_{i,k-1}$ and $C_{i,k+1}/C_{ik}$ are uncorrelated. This means that after a rather high value of $C_{ik}/C_{i,k-1}$ the expected size of the next development factor $C_{i,k+1}/C_{ik}$ is the same as after a rather low value of $C_{ik}/C_{i,k-1}$. We therefore should not apply the chain ladder method to a business where we usually observe a rather small increase $C_{i,k+1}/C_{ik}$ if $C_{ik}/C_{i,k-1}$ is higher than in most other accident years, and vice versa. Appendix G also contains a test procedure to check this for a given run-off triangle.

3. Analysis of the Age-to-Age Factor Formula: the Key to Measuring the Variability

Because of the randomness of all realizations C_{ik} we can not infer the true values of the increase factors $f_1, \ldots, f_{I-1}$ from the data. They only can be estimated and the chain ladder method calculates estimators $\mathbf{f}_1, \ldots, \mathbf{f}_{I-1}$ according to formula (2). Among the properties which a good estimator should have, one prominent property is that the estimator should be unbiased, i.e. that its expected value $E(\mathbf{f}_k)$ (under the assumption that the whole run-off triangle is not yet known) is equal to the true value f_k, i.e. that $E(\mathbf{f}_k) = f_k$. Indeed, this is the case here as is shown in Appendix A under the additional assumption that

(4) the variables $\{C_{i1}, \ldots, C_{iI}\}$ and $\{C_{j1}, \ldots, C_{jI}\}$ of different accident years $i \neq j$ are independent.

Because the chain ladder method neither in (1) nor in (2) takes into account any dependency between the accident years we can conclude that the independence of the accident years is also an implicit assumption of the chain ladder method. We will therefore assume (4) for all further calculations. Assumption (4), too, cannot be taken as being met for every run-off triangle because certain calendar year effects (such as a major change in claims handling or in case reserving or greater changes in the inflation rate) can affect several accident years in the same way and can thus distort the independence. How such a situation can be recognized is shown in Appendix H.

A closer look at formula (2) reveals that

$$\mathbf{f_k} = \frac{\sum_{j=1}^{I-k} C_{j,k+1}}{\sum_{j=1}^{I-k} C_{jk}} = \sum_{j=1}^{I-k} \frac{C_{jk}}{\sum_{j=1}^{I-k} C_{jk}} \cdot \frac{C_{j,k+1}}{C_{jk}}$$

is a weighted average of the observed individual development factors $C_{j,k+1}/C_{jk}$, $1 \le j \le I-k$, where the weights are proportional to C_{jk}. Like $\mathbf{f_k}$ every individual development factor $C_{j,k+1}/C_{jk}$, $1 \le j \le I-k$, is also an unbiased estimator of f_k because

$$
\begin{aligned}
E(C_{j,k+1}/C_{jk}) &= E(E(C_{j,k+1}/C_{jk} \mid C_{j1},\ldots,C_{jk})) && \text{(a)}\\
&= E(E(C_{j,k+1} \mid C_{j1},\ldots,C_{jk})/C_{jk}) && \text{(b)}\\
&= E(C_{jk}f_k/C_{jk}) && \text{(c)}\\
&= E(f_k) && \\
&= f_k\ . && \text{(d)}
\end{aligned}
$$

Here equality (a) holds due to the iterative rule $E(X) = E(E(X|Y))$ for expectations, (b) holds because, given C_{j1} to C_{jk}, C_{jk} is a scalar, (c) holds due to assumption (3) and (d) holds because f_k is a scalar. (When applying expectations iteratively, e.g. $E(E(X|Y))$, one first takes the conditional expectation $E(X|Y)$ assuming Y being known and then averages over all possible realizations of Y.)

Therefore the question arises as to why the chain ladder method uses just $\mathbf{f_k}$ as estimator for f_k and not the simple average

$$\frac{1}{I-k} \sum_{j=1}^{I-k} C_{j,k+1}/C_{jk}$$

of the observed development factors which also would be an unbiased estimator as is the case with any weighted average

$$g_k = \sum_{j=1}^{I-k} w_{jk}\, C_{j,k+1}/C_{jk} \qquad \text{with} \qquad \sum_{j=1}^{I-k} w_{jk} = 1$$

of the observed development factors. (Here, w_{jk} must be a scalar if $C_{j1}, \ldots, C_{jk}$ are known.)

Here we recall one of the principles of the theory of point estimation which states that among several unbiased estimators preference should be given to the one with the smallest variance, a principle which is easy to understand. We therefore should choose the weights w_{jk} in such a way that the variance of g_k is minimal. In Appendix B it is shown that this is the case if and only if (for fixed k and all j)

w_{jk} is inversely proportional to $\mathrm{Var}(C_{j,k+1}/C_{jk}|C_{j1},\ldots,C_{jk})$.

The fact that the chain ladder estimator $\mathbf{f_k}$ uses weights which are proportional to C_{jk} therefore means that C_{jk} is assumed to be inversely proportional to $\mathrm{Var}(C_{j,k+1}/C_{jk}|C_{j1},\ldots,C_{jk})$, or stated the other way around, that

$$\mathrm{Var}(C_{j,k+1}/C_{jk}|C_{j1},\ldots,C_{jk}) = \alpha_k^2/C_{jk}$$

with a proportionality constant α_k^2 which may depend on k but not on j and which must be non-negative because variances are always non-negative. Since here C_{jk} is a scalar and because generally $Var(X/c) = Var(X)/c^2$ for any scalar c, we can state the above proportionality condition also in the form

(5) $\quad Var(C_{j,k+1}|C_{j1},\ldots,C_{jk}) = C_{jk}\alpha_k^2, \quad 1 \le j \le I, \ 1 \le k \le I-1,$

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

As it was the case with assumptions (3) and (4), assumption (5) also has to be considered a basic condition implicitly underlying the chain ladder method. Again, condition (5) cannot a priori be assumed to be met for every run-off triangle. In Chapter 5 we will show how to check a given triangle to see whether (5) can be considered met or not. But before we turn to the most important consequence of (5): Together with (3) and (4) it namely enables us to quantify the uncertainty in the estimation of C_{iI} by $\mathbf{C_{iI}}$.

4. Quantifying the Variability of the Ultimate Claims Amount

The aim of the chain ladder method and of every claims reserving method is the estimation of the ultimate claims amount C_{iI} for the accident years i = 2, ..., I. The chain ladder method does this by formula (1), i.e. by

$$\mathbf{C_{iI}} = C_{i,I+1-i}\cdot\mathbf{f_{I+1-i}}\cdot\ldots\cdot\mathbf{f_{I-1}} \ .$$

This formula yields only a point estimate for C_{iI} which will normally turn out to be more or less wrong, i.e. there is only a very small probability for C_{iI} being equal to $\mathbf{C}_{iI}$. This probability is even zero if C_{iI} is considered to be a continuous variable. We therefore want to know in addition if the estimator $\mathbf{C}_{iI}$ is at least on average equal to the mean of C_{iI} and how large on average the error is. Precisely speaking we first would like to have the expected values $E(\mathbf{C}_{iI})$ and $E(C_{iI})$, $2 \le i \le I$, being equal. In Appendix C it is shown that this is indeed the case as a consequence of assumptions (3) and (4).

The second thing we want to know is the average distance between the forecast $\mathbf{C}_{iI}$ and the future realization C_{iI}. In Mathematical Statistics it is common to measure such distances by the square of the ordinary Euclidean distance ('quadratic loss function'). This means that one is interested in the size of the so-called mean squared error

$$\mathrm{mse}(\mathbf{C}_{iI}) = E((C_{iI} - \mathbf{C}_{iI})^2|D)$$

where $D = \{ C_{ik} \mid i+k \le I+1 \}$ is the set of all data observed so far. It is important to realize that we have to calculate the mean squared error on the condition of knowing all data observed so far because we want to know the error due to <u>future</u> randomness only. If we calculated the unconditional error $E(C_{iI}-\mathbf{C}_{iI})^2$, which due to the iterative rule for expectations is equal to the mean value $E(E((C_{iI} - \mathbf{C}_{iI})^2|D))$ of the conditional mse over all possible data sets D, we also would include all deviations from the data observed so far which obviously makes no sense if we want to establish a confidence interval for C_{iI} on the basis of the given particular run-off triangle D.

The mean squared error is exactly the same concept which also underlyies the notion of the variance

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

of any random variable X. Var(X) measures the average distance of X from its mean value E(X).

Due to the general rule $E(X-c)^2 = \mathrm{Var}(X) + (E(X)-c)^2$ for any scalar c we have

$$\mathrm{mse}(\mathbf{C}_{iI}) = \mathrm{Var}(C_{iI}|D) + (E(C_{iI}|D) - \mathbf{C}_{iI})^2$$

because $\mathbf{C}_{iI}$ is a scalar under the condition that all data D are known. This equation shows that the mse is the sum of the pure future random error $\mathrm{Var}(C_{iI}|D)$ and of the estimation error which is measured by the squared deviation of the estimate $\mathbf{C}_{iI}$ from its target $E(C_{iI}|D)$. On the other hand, the mse does not take into account any future changes in the underlying model, i.e. future deviations from the assumptions (3), (4) and (5), an extreme example of which was the emergence of asbestos. Modelling such deviations is beyond the scope of this paper.

As is to be expected and can be seen in Appendix D, $\mathrm{mse}(\mathbf{C}_{iI})$ depends on the unknown model parameters f_k and $\alpha_k{}^2$. We therefore must develop an estimator for $\mathrm{mse}(\mathbf{C}_{iI})$ which can be calculated from the known data D only. The square root of such an estimator is usually called '<u>standard error</u>' because it is an estimate of the standard deviation of C_{iI} in cases in which we have to estimate the mean value, too. The standard error $\mathrm{s.e.}(\mathbf{C}_{iI})$ of

$\mathbf{C}_{iI}$ is at the same time the standard error s.e.($\mathbf{R}_i$) of the reserve estimate

$$\mathbf{R}_i = \mathbf{C}_{iI} - C_{i,I+1-i}$$

of the outstanding claims reserve

$$R_i = C_{iI} - C_{i,I+1-i}$$

because

$$\mathrm{mse}(\mathbf{R}_i) = E((\mathbf{R}_i - R_i)^2|D) = E((\mathbf{C}_{iI} - C_{iI})^2|D) = \\ = \mathrm{mse}(\mathbf{C}_{iI})$$

and because the equality of the mean squared errors also implies the equality of the standard errors. This means that

$$\text{(6)} \qquad \text{s.e.}(\mathbf{R}_i) = \text{s.e.}(\mathbf{C}_{iI}) \; .$$

The derivation of a formula for the standard error s.e.($\mathbf{C}_{iI}$) of $\mathbf{C}_{iI}$ turns out to be the most difficult part of this paper; it is done in Appendix D. Fortunately, the resulting formula is simple:

$$\text{(7)} \qquad (\text{s.e.}(\mathbf{C}_{iI}))^2 = \mathbf{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\boldsymbol{\alpha}_k^2}{\mathbf{f}_k^2} \left(\frac{1}{\mathbf{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right)$$

where

$$\text{(8)} \qquad \boldsymbol{\alpha}_k^2 = \frac{1}{I-k-1} \sum_{j=1}^{I-k} C_{jk} \left(\frac{C_{j,k+1}}{C_{jk}} - \mathbf{f}_k \right)^2 , \qquad 1 \le k \le I-2.$$

is an unbiased estimator of α_k^2 (the unbiasedness being shown in Appendix E) and

$$\mathbf{C}_{ik} = C_{i,I+1-i} \cdot \mathbf{f}_{I+1-i} \cdot \ldots \cdot \mathbf{f}_{k-1} \, , \quad k > I+1-i,$$

are the amounts which are automatically obtained if the run-off triangle is completed step by step according to the chain ladder method. In (7), for notational convenience we have also set

$$\mathbf{C}_{i,I+1-i} = C_{i,I+1-i}.$$

Formula (8) does not yield an estimator for α_{I-1} because it is not possible to estimate the two parameters f_{I-1} and α_{I-1} from the single observation $C_{1,I}/C_{1,I-1}$ between development years I-1 and I. If $\mathbf{f}_{I-1} = 1$ and if the claims development is believed to be finished after I-1 years we can put $\boldsymbol{\alpha}_{I-1} = 0$. If not, we extrapolate the usually decreasing series $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_{I-3}, \boldsymbol{\alpha}_{I-2}$ by one additional member, for instance by means of loglinear regression (cf. the example in Chapter 6) or more simply by requiring that

$$\boldsymbol{\alpha}_{I-3} / \boldsymbol{\alpha}_{I-2} = \boldsymbol{\alpha}_{I-2} / \boldsymbol{\alpha}_{I-1}$$

holds at least as long as $\boldsymbol{\alpha}_{I-3} > \boldsymbol{\alpha}_{I-2}$. This last possibility leads to

$$(9) \qquad \boldsymbol{\alpha}^2_{I-1} = \min\left(\boldsymbol{\alpha}^4_{I-2}/\boldsymbol{\alpha}^2_{I-3},\ \min(\boldsymbol{\alpha}^2_{I-3},\ \boldsymbol{\alpha}^2_{I-2}) \right) .$$

We now want to establish a confidence interval for our target variables C_{iI} and R_i. Because of the equation

$$C_{iI} = C_{i,I+1-i} + R_i$$

the ultimate claims amount C_{iI} consists of a known part $C_{i,I+1-i}$ and an unknown part R_i. This means that the probability distribution function of C_{iI} (given the observations D which include $C_{i,I+1-i}$) is completely determined by that of R_i. We therefore need to establish a confidence interval for R_i only and can then simply shift it to a confidence interval for C_{iI}.

For this purpose we need to know the distribution function of R_i. Up to now we only have estimates $\mathbf{R_i}$ and s.e.($\mathbf{R_i}$) for the mean and the standard deviation of this distribution. If the volume of the outstanding claims is large enough we can, due to the central limit theorem, assume that this distribution function is a Normal distribution with an expected value equal to the point estimate given by $\mathbf{R_i}$ and a standard deviation equal to the standard error s.e.($\mathbf{R_i}$). A symmetric 95%-confidence interval for R_i is then given by

$$(\mathbf{R_i} - 2\cdot \text{s.e.}(\mathbf{R_i}) \; , \; \mathbf{R_i} + 2\cdot \text{s.e.}(\mathbf{R_i})).$$

But the symmetric Normal distribution may not be a good approximation to the true distribution of R_i if this latter distribution is rather skewed. This will especially be the case if s.e.($\mathbf{R_i}$) is greater than 50 % of $\mathbf{R_i}$. This can also be seen at the above Normal distribution confidence interval whose lower limit then becomes negative even if a negative reserve is not possible.

In this case it is recommended to use an approach based on the Lognormal distribution. For this purpose we approximate the unknown distribution of R_i by a Lognormal distribution with parameters μ_i and $\sigma_i{}^2$ such that mean values as well as variances of both distributions are equal, i.e. such that

$$\exp(\mu_i + \sigma_i{}^2/2) = \mathbf{R_i} \; ,$$
$$\exp(2\mu_i + \sigma_i{}^2)(\exp(\sigma_i{}^2)-1) = (\text{s.e.}(\mathbf{R_i}))^2 \; .$$

This leads to

$$\text{(10)}\quad \begin{aligned} \sigma_i^2 &= \ln(1 + (\text{s.e.}(\mathbf{R_i}))^2/\mathbf{R_i}^2) \ , \\ \mu_i &= \ln(\mathbf{R_i}) - \sigma_i^2/2 \ . \end{aligned}$$

Now, if we want to estimate the 90th percentile of R_i, for example, we proceed as follows. First we take the 90th percentile of the Standard Normal distribution which is 1.28. Then $\exp(\mu_i+1.28\sigma_i)$ with μ_i and σ_i^2 according to (10) is the 90th percentile of the Lognormal distribution and therefore also approximately of the distribution of R_i. For instance, if $\text{s.e.}(\mathbf{R_i})/\mathbf{R_i} = 1$, then $\sigma_i^2 = \ln(2)$ and the 90th percentile is $\exp(\mu_i + 1.28\sigma_i) = \mathbf{R_i}\exp(1.28\sigma_i - \sigma_i^2/2) = \mathbf{R_i}\exp(.719) = 2.05\cdot\mathbf{R_i}$. If we had assumed that R_i has approximately a Normal distribution, we would have obtained in this case $\mathbf{R_i} + 1.28\cdot\text{s.e.}(\mathbf{R_i}) = 2.28\cdot\mathbf{R_i}$ as 90th percentile.

This may come as a surprise since we might have expected that the 90th percentile of a Lognormal distribution always must be higher than that of a Normal distribution with same mean and variance. But there is no general rule, it depends on the percentile chosen and on the size of the ratio $\text{s.e.}(\mathbf{R_i})/\mathbf{R_i}$. The Lognormal approximation only prevents a negative lower confidence limit. In order to set a specific lower confidence limit we choose a suitable percentile, for instance 10%, and proceed analogously as with the 90% before. The question of which confidence probability to choose has to be decided from a business policy point of view. The value of 80% = 90% - 10% taken here must be regarded merely as an example.

We have now shown how to establish confidence limits for every R_i and therefore also for every $C_{iI} = C_{i,I+1-i} + R_i$. We may also be interested in having confidence limits for the overall reserve

$$R = R_2 + \ldots + R_I ,$$

and the question is whether, in order to estimate the variance of R, we can simply add the squares $(\text{s.e.}(\mathbf{R_i}))^2$ of the individual standard errors as would be the case with standard deviations of independent variables. But unfortunately, whereas the R_i's itself are independent, the estimators $\mathbf{R_i}$ are not because they are all influenced by the same age-to-age factors $\mathbf{f_k}$, i.e. the $\mathbf{R_i}$'s are positively correlated. In Appendix F it is shown that the square of the standard error of the overall reserve estimator

$$\mathbf{R} = \mathbf{R_2} + \ldots + \mathbf{R_I}$$

is given by

(11) $$(\text{s.e.}(\mathbf{R}))^2 =$$

$$= \sum_{i=2}^{I} \left\{ (\text{s.e.}(\mathbf{R_i}))^2 + \mathbf{C_{iI}}\left(\sum_{j=i+1}^{I} \mathbf{C_{jI}}\right) \sum_{k=I+1-i}^{I-1} \frac{2\alpha_k^2/\mathbf{f_k}^2}{\sum_{n=1}^{I-k} C_{nk}} \right\}$$

Formula (11) can be used to establish a confidence interval for the overall reserve amount R in quite the same way as it was done before for R_i. Before giving a full example of the calculation of the standard error, we will deal in the next chapter with the problem of how to decide for a given run-off triangle whether the chain ladder assumptions (3) and (5) are met or not.

5. Checking the Chain Ladder Assumptions Against the Data

As has been pointed out before, the three basic implicit chain ladder assumptions

(3) $E(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}f_k$,

(4) Independence of accident years ,

(5) $\mathrm{Var}(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}\alpha_k^2$,

are not met in every case. In this chapter we will indicate how these assumptions can be checked for a given run-off triangle. We have already mentioned in Chapter 3 that Appendix H develops a test for calendar year influences which may violate (4). We therefore can concentrate in the following on assumptions (3) and (5).

First, we look at the equations (3) for an arbitrary but fixed k and for i = 1, ..., I. There, the values of C_{ik}, $1 \le i \le I$, are to be considered as given non-random values and equations (3) can be interpreted as an ordinary regression model of the type

$$Y_i = c + x_ib + \epsilon_i \ , \qquad 1 \le i \le I,$$

where c and b are the regression coefficients and ϵ_i the error term with $E(\epsilon_i) = 0$, i.e. $E(Y_i) = c + x_ib$. In our special case, we have $c = 0$, $b = f_k$ and we have observations of the independent variable $Y_i = C_{i,k+1}$ at the points $x_i = C_{ik}$ for i =

1, ..., I-k. Therefore, we can estimate the regression coefficient $b = f_k$ by the usual least squares method

$$\sum_{i=1}^{I-k} (C_{i,k+1} - C_{ik}f_k)^2 \stackrel{!}{=} \text{minimum} .$$

If the derivative of the left hand side with respect to f_k is set to 0 we obtain for the minimizing parameter f_k the solution

$$(12) \qquad f_{k0} = \sum_{i=1}^{I-k} C_{ik}C_{i,k+1} \Big/ \sum_{i=1}^{I-k} C_{ik}^2 .$$

This is not the same estimator for f_k as according to the chain ladder formula (2). We therefore have used an additional index '0' at this new estimator for f_k. We can rewrite f_{k0} as

$$f_{k0} = \sum_{i=1}^{I-k} \frac{C_{ik}^2}{\sum_{i=1}^{I-k} C_{ik}^2} \cdot \frac{C_{i,k+1}}{C_{ik}}$$

which shows that f_{k0} is the C_{ik}^2-weighted average of the individual development factors $C_{i,k+1}/C_{ik}$, whereas the chain ladder estimator $\hat{f}_k$ is the C_{ik}-weighted average. In Chapter 3 we saw that these weights are inversely proportional to the underlying variances $\mathrm{Var}(C_{i,k+1}/C_{ik}|C_{i1},\ldots,C_{ik})$. Correspondingly, the estimator f_{k0} assumes

$\mathrm{Var}(C_{i,k+1}/C_{ik}|C_{i1},\ldots,C_{ik})$ being proportional to $1/C_{ik}^2$,

or equivalently

$\mathrm{Var}(C_{i,k+1}|C_{i1},\ldots,C_{ik})$ being proportional to 1

which means that $\mathrm{Var}(C_{i,k+1}|C_{i1},\ldots,C_{ik})$ is the same for all observations i = 1, ..., I-k. This is not in agreement with the chain ladder assumption (5).

Here we remember that indeed the least squares method implicitly assumes equal variances $Var(Y_i) = Var(\epsilon_i) = \sigma^2$ for all i. If this assumption is not met, i.e. if the variances $Var(Y_i) = Var(\epsilon_i)$ depend on i, one should use a weighted least squares approach which consists of minimizing the weighted sum of squares

$$\sum_{i=1}^{I} w_i(Y_i - c - x_i b)^2$$

where the weights w_i are in inverse proportion to $Var(Y_i)$.

Therefore, in order to be in agreement with the chain ladder variance assumption (5), we should use regression weights w_i which are proportional to $1/C_{ik}$ (more precisely to $1/(C_{ik}\alpha_k^2)$, but α_k^2 can be amalgamated with the proportionality constant because k is fixed). Then minimizing

$$\sum_{i=1}^{I-k} (C_{i,k+1} - C_{ik}f_k)^2 / C_{ik}$$

with respect to f_k yields indeed

$$f_{k1} = \sum_{i=1}^{I-k} C_{i,k+1} / \sum_{i=1}^{I-k} C_{ik}$$

which is identical to the usual chain ladder age-to-age factor $\mathbf{f_k}$.

It is tempting to try another set of weights, namely $1/C_{ik}^2$ because then the weighted sum of squares becomes

$$\sum_{i=1}^{I-k} (C_{i,k+1} - C_{ik}f_k)^2 / C_{ik}^2 = \sum_{i=1}^{I-k} \left(\frac{C_{i,k+1}}{C_{ik}} - f_k \right)^2 .$$

Here the minimizing procedure yields

$$(13) \qquad f_{k2} = \frac{1}{I-k} \sum_{i=1}^{I-k} \frac{C_{i,k+1}}{C_{ik}} ,$$

which is the ordinary unweighted average of the development factors. The variance assumption corresponding to the weights used is

$Var(C_{i,k+1}|C_{i1},\ldots,C_{ik})$ being proportional to C_{ik}^2

or equivalently

$Var(C_{i,k+1}/C_{ik}|C_{i1},\ldots,C_{ik})$ being proportional to 1.

The benefit of transforming the estimation of the age-to-age factors into the regression framework is the fact that the usual regression analysis instruments are now available to check the underlying assumptions, especially the linearity and the variance assumption. This check is usually done by carefully inspecting plots of the data and of the residuals:

First, we plot $C_{i,k+1}$ against C_{ik}, $i = 1, \ldots, I-k$, in order to see if we really have an approximately linear relationship around a straight line through the origin with slope $\mathbf{f_k} = f_{k1}$. Second, if linearity seems acceptable, we plot the weighted residuals

$(C_{i,k+1} - C_{ik}\mathbf{f_k}) / \sqrt{C_{ik}}$, $1 \le i \le I-k$,

(whose squares have been minimized) against C_{ik} in order to see if the employed variance assumption really leads to a plot in which the residuals do not show any specific trend but appear purely random. It is recommended to compare all three residual plots (for $i = 1, \ldots, I-k$)

Plot 0: $C_{i,k+1} - C_{ik}f_{k0}$ against C_{ik} ,

Plot 1: $(C_{i,k+1} - C_{ik}f_{k1})/\sqrt{C_{ik}}$ against C_{ik} ,

Plot 2: $(C_{i,k+1} - C_{ik}f_{k2})/C_{ik}$ against C_{ik} ,

and to find out which one shows the most random behaviour. All this should be done for every development year k for which we have sufficient data points, say at least 6, i.e. for $k \leq I-6$.

Some experience with least squares residual plots is useful, especially because in our case we have only very few data points. Consequently, it is not always easy to decide whether a pattern in the residuals is systematic or random. However, if Plot 1 exhibits a nonrandom pattern, and either Plot 0 or Plot 2 does not, and if this holds true for several values of k, we should seriously consider replacing the chain ladder age-to-age factors $f_{k1} = \mathbf{f_k}$ with f_{k0} or f_{k2} respectively. The following numerical example will clarify the situation a bit more.

6. Numerical Example

The data for the following example are taken from the 'Historical Loss Development Study', 1991 Edition, published by the Reinsurance Association of America (RAA). There, we find on page 96 the following run-off triangle of Automatic Facultative business in General Liability (excluding Asbestos & Environmental):

	c_{i1}	c_{i2}	c_{i3}	c_{i4}	c_{i5}	c_{i6}	c_{i7}	c_{i8}	c_{i9}	c_{i10}
i=1	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834
i=2	106	4285	5396	10666	13782	15599	15496	16169	16704	
i=3	3410	8992	13873	16141	18735	22214	22863	23466		
i=4	5655	11555	15766	21266	23425	26083	27067			
i=5	1092	9565	15836	22169	25955	26180				
i=6	1513	6445	11702	12935	15852					
i=7	557	4020	10946	12314						
i=8	1351	6947	13112							
i=9	3133	5395								
i=10	2063									

The above figures are cumulative incurred case losses in \$ 1000. We have taken the accident years from 1981 (i=1) to 1990 (i=10) which is enough for the sake of example but does not mean that we believe to have reached the ultimate claims amount after 10 years of development.

We first calculate the age-to-age factors $\mathbf{f_k} = f_{k,1}$ according to formula (2). The result is shown in the following table together with the alternative factors f_{k0} according to (12) and f_{k2} according to (13):

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
f_{k0}	2.217	1.569	1.261	1.162	1.100	1.041	1.032	1.016	1.009
f_{k1}	2.999	1.624	1.271	1.172	1.113	1.042	1.033	1.017	1.009
f_{k2}	8.206	1.696	1.315	1.183	1.127	1.043	1.034	1.018	1.009

If one has the run-off triangle on a personal computer it is very easy to produce the plots recommended in Chapter 5 because most spreadsheet programs have the facility of plotting X-Y graphs. For every $k = 1, \ldots, 8$ we make a plot of the amounts $C_{i,k+1}$ (y-axis) of development year k+1 against the amounts C_{ik} (x-axis) of development year k for $i = 1, \ldots, 10-k$, and draw a straight line through the origin with slope f_{k1}. The plots for k = 1 to 8 are shown in the upper graphs of Figures 1 to 8, respectively. (All figures are to be found at the end of the paper after the appendices.) The number above each point mark indicates the corresponding accident year. (Note that the point mark at the upper or right hand border line of each graph does not belong to the plotted points $(C_{ik}, C_{i,k+1})$, it has only been used to draw the regression line.) In the lower graph of each of the Figures 1 to 8 the corresponding weighted residuals $(C_{i,k+1} - C_{ik})/\sqrt{C_{ik}}$ are plotted against C_{ik} for $i = 1, \ldots, 10-k$.

The two plots for k = 1 (Figure 1) clearly show that the regression line does not capture the direction of the data points very well. The line should preferably have a positive intercept on the y-axis and a flatter slope. However, even then we would have a high dispersion. Using the line through the origin we will probably underestimate any future C_{i2} if C_{i1} is less than 2000 and will overestimate it if C_{i1} is more than 4000. Fortunately, in the one relevant case i = 10 we have $C_{10,1}$ = 2063 which means that the resulting forecast $C_{10,2} = C_{10,1}f_2$ =

$2063 \cdot 2.999 = 6187$ is within the bulk of the data points plotted. In any case, Figure 1 shows that any forecast of $C_{10,2}$ is associated with a high uncertainty of about ± 3000 or almost $\pm 50\%$ of an average-sized $C_{i,2}$ which subsequently is even enlarged when extrapolating to ultimate. If in a future accident year we have a value C_{i1} outside the interval (2000, 4000) it is reasonable to introduce an additional parameter by fitting a regression line with positive intercept to the data and using it for the projection to C_{i2}. Such a procedure of employing an additional parameter is acceptable between the first two development years in which we have the highest number of data points of all years.

The two plots for $k = 2$ (Figure 2) are more satisfactory. The data show a clear trend along the regression line and quite random residuals. The same holds for the two plots for $k = 4$ (Figure 4). In addition, for both $k = 2$ and $k = 4$ a weighted linear regression including a parameter for intercept would yield a value of the intercept which is not significantly different from zero. The plots for $k = 3$ (Figure 3) seem to show a curvature to the left but because of the few data points we can hope that this is incidental. Moreover, the plots for $k = 5$ have a certain curvature to the right such that we can hope that the two curvatures offset each other. The plots for $k = 6$, 7 and 8 are quite satisfactory. The trends in the residuals for $k = 7$ and 8 have no significance in view of the very few data points.

We need not to look at the regression lines with slopes f_{k0} or f_{k2} as these slopes are very close to $\mathbf{f_k}$ (except for k=1). But we should look at the corresponding plots of weighted residuals in order to see whether they appear more satisfactory than the previous ones. (Note that due to the different weights the residuals will be different even if the slopes are equal.) The residual plots for f_{k0} and k = 1 to 4 are shown in Figures 9 and 10. Those for f_{k2} and k = 1 to 4 are shown in Figures 11 and 12. In the residual plot for $f_{1,0}$ (Figure 9, upper graph) the point furthest to the left is not an outlier as it is in the plots for $f_{1,1} = \mathbf{f_1}$ (Figur 1, lower graph) and $f_{1,2}$ (Figure 11, upper graph). But with all three residual plots for k=1 the main problem is the missing intercept of the regression line which leads to a decreasing trend in the residuals. Therefore the improvement of the outlier is of secondary importance. For k = 2 the three residuals plots do not show any major differences between each other. The same holds for k = 3 and 4. The residual plots for k = 5 to 8 are not important because of the small number of data points. Altogether, we decide to keep the usual chain ladder method, i.e. the age-to-age factors $\mathbf{f_k} = f_{k,1}$, because the alternatives $f_{k,0}$ or $f_{k,2}$ do not lead to a clear improvement.

Next, we can carry through the tests for calendar year influences (see Appendix H) and for correlations between subsequent development factors (see Appendix G). For our example neither test leads to a rejection of the underlying assumption as is shown in the appendices mentioned.

Having now finished all preliminary analyses we calculate the estimated ultimate claims amounts C_{iI} according to formula (1), the reserves $R_i = C_{iI} - C_{i,I+1-i}$ and its standard errors (7). For the standard errors we need the estimated values of α_k^2 which according to formula (8) are given by

k	1	2	3	4	5	6	7	8	9
α_k^2	27883	1109	691	61.2	119	40.8	1.34	7.88	

A plot of $\ln(\alpha_k^2)$ against k is given in Figure 13 and shows that there indeed seems to be a linear relationship which can be used to extrapolate $\ln(\alpha_9^2)$. This yields $\alpha_9^2 = \exp(-.44) = .64$. But we use formula (9) which is more easily programmable and in the present case is a bit more on the safe side: it leads to $\alpha_9^2 = 1.34$. Using formula (11) for s.e.(R) as well we finally obtain

	$C_{i,10}$	R_i	$s.e(C_{i,10}) = s.e.(R_i)$	$s.e.(R_i)/R_i$
i=2	16858	154	206	134 %
i=3	24083	617	623	101 %
i=4	28703	1636	747	46 %
i=5	28927	2747	1469	53 %
i=6	19501	3649	2002	55 %
i=7	17749	5435	2209	41 %
i=8	24019	10907	5358	49 %
i=9	16045	10650	6333	59 %
i=10	18402	16339	24566	150 %
Overall		52135	26909	52 %

(The numbers in the 'Overall'-row are **R**, s.e.(**R**) and s.e.(**R**)/**R**.) For i = 2, 3 and 10 the percentage standard error (last column) is more than 100% of the estimated reserve R_i. For i = 2 and 3 this is due to the small amount of the corresponding reserve and is not important because the absolute amounts of the standard errors are rather small. But the standard error of 150 % for the most recent accident year i = 10 might lead to some concern in practice. The main reason for this high standard error is the high uncertainty of forecasting next year's value $C_{10,2}$ as was seen when examining the plot of C_{i2} against C_{i1}. Thus, one year later we will very likely be able to give a much more precise forecast of $C_{10,10}$.

Because all standard errors are close to or above 50 % we use the Lognormal distribution in all years for the calculation of confidence intervals. We first calculate the upper 90%-confidence limit (or with any other chosen percentage) for the overall outstanding claims reserve R. Denoting by μ and σ^2 the parameters of the Lognormal distribution approximating the distribution of R and using s.e.(**R**)/**R** = .52 we have σ^2 = .236 (cf. (10)) and, in the same way as in Chapter 4, the 90th percentile is $\exp(\mu + 1.28\sigma) = \mathbf{R}\cdot\exp(1.28\sigma-\sigma^2/2) = 1.655\cdot\mathbf{R}$ = 86298. Now we allocate this overall amount to the accident years i = 2,..., 10 in such a way that we reach the same level of confidence for every accident year. Each level of confidence corresponds to a certain percentile t of the Standard Normal distribution and - according to Chapter 4 - the corresponding percentile of the distribution of R_i is $\mathbf{R_i}\exp(t\sigma_i - \sigma_i^2/2)$ with $\sigma_i^2 = \ln(1 + (\text{s.e.}(\mathbf{R_i}))^2/\mathbf{R_i}^2)$. We therefore only have to choose t in such a way that

$$\sum_{i=2}^{I} \mathbf{R_i}\cdot\exp(t\sigma_i - \sigma_i^2/2) = 86298 .$$

This can easily be solved with the help of spreadsheet software (e.g. by trial and error) and yields t = 1.13208 which corresponds to the 87th percentile per accident year and leads to the following distribution of the overall amount 86298:

	$\mathbf{R_i}$	s.e.$(\mathbf{R_i})/\mathbf{R_i}$	σ_i^2	upper confidence limit $\mathbf{R_i}\exp(t\sigma_i-\sigma_i^2/2)$
i=2	154	1.34	1.028	290
i=3	617	1.01	.703	1122
i=4	1636	.46	.189	2436
i=5	2747	.53	.252	4274
i=6	3649	.55	.263	5718
i=7	5435	.41	.153	7839
i=8	10907	.49	.216	16571
i=9	10650	.59	.303	17066
i=10	16339	1.50	1.182	30981
Total	52135			86298

In order to arrive at the lower confidence limits we proceed completely analogously. The 10th percentile, for instance, of the total outstanding claims amount is $\mathbf{R}\cdot\exp(-1.28\sigma-\sigma^2/2) = .477\cdot\mathbf{R} = 24871$. The distribution of this amount over the individual accident years is made as before and leads to a value of $t = -.8211$ which corresponds to the 21st percentile. This means that a 87% - 21% = 66% confidence interval for each accident year leads to a 90% - 10% = 80% confidence interval for the overall reserve amount. In the following table, the confidence intervals thus obtained for R_i are already shifted (by adding $C_{i,I+1-i}$) to confidence intervals for the ultimate claims amounts C_{iI} (for instance, the upper limit 16994 for i=2 has been obtained by adding $C_{2,9} = 16704$ and 290 from the preceding table):

	$C_{i,10}$	confidence intervals for 80% prob. overall	empirical limits
i=2	16858	(16744 , 16994)	(16858 , 16858)
i=3	24083	(23684 , 24588)	(23751 , 24466)
i=4	28703	(28108 , 29503)	(28118 , 29446)
i=5	28927	(27784 , 30454)	(27017 , 31699)
i=6	19501	(17952 , 21570)	(16501 , 22939)
i=7	17749	(15966 , 20153)	(14119 , 23025)
i=8	24019	(19795 , 29683)	(16272 , 48462)
i=9	16045	(11221 , 22461)	(8431 , 54294)
i=10	18402	(5769 , 33044)	(5319 , 839271)

The column "empirical limits" contains the minimum and maximum size of the ultimate claims amount resulting if in formula (1) each age-to-age factor f_k is replaced with the minimum (or maximum) individual development factor observed so far. These factors are defined by

$$f_{k,min} = \min \{ C_{i,k+1}/C_{ik} \mid 1 \le i \le I-k \}$$
$$f_{k,max} = \max \{ C_{i,k+1}/C_{ik} \mid 1 \le i \le I-k \}$$

and can be taken from the table of all development factors which can be found in Appendices G and H. They are

	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9
$f_{k,min}$	1.650	1.259	1.082	1.102	1.009	.993	1.026	1.003	1.009
$f_{k,max}$	40.425	2.723	1.977	1.292	1.195	1.113	1.043	1.033	1.009

In comparison with the confidence intervals, these empirical limits are narrower in the earlier accident years $i \le 4$ and wider in the more recent accident years $i \ge 5$. This was to be expected because the small number of development factors observed between the late development years only leads to a rather small variation between the minimum and maximum factors. Therefore these empirical limits correspond to a confidence probability which is rather small in the early accident years and becomes larger and larger towards the recent accident years. Thus, this empirical approach to establishing confidence limits does not seem to be reasonable.

If we used the Normal distribution instead of the Lognormal we had obtained a 90th percentile of $\mathbf{R} + 1.28 \cdot \mathbf{R} \cdot (\text{s.e.}(\mathbf{R})/\mathbf{R}) = 1.661 \cdot \mathbf{R}$ (which is almost the same as the $1.655 \cdot \mathbf{R}$ with the Lognormal) and a 10th percentile of $\mathbf{R} - 1.28 \cdot \mathbf{R} \cdot (\text{s.e.}(\mathbf{R})/\mathbf{R}) = .34 \cdot \mathbf{R}$ (which is lower than the $.477 \cdot \mathbf{R}$ with the Lognormal). Also, the allocation to the accident years would be different.

Finally, we compare the standard errors obtained to the output of the claims reserving software package ICRFS by Ben Zehnwirth.

This package is a modelling framework in which the user can specify his own model within a large class of models. But it also contains some predefined models, inter alia also a 'chain ladder model'. But this is not the usual chain ladder method, instead, it is a loglinearized approximation of it. Therefore, the estimates of the oustanding claims amounts differ from those obtained here with the usual chain ladder method. Moreover, it works with the logarithms of the incremental amounts $C_{i,k+1}-C_{ik}$ and one must therefore eliminate the negative increment $C_{2,7}-C_{2,6}$. In addition, $C_{2,1}$ was identified as an outlier and was eliminated. Then the ICRFS results were quite similar to the chain ladder results as can be seen in the following table:

	est. outst. claims amount R_i		standard error	
	chain ladder	ICRFS	chain ladder	ICRFS
i=2	154	394	206	572
i=3	617	825	623	786
i=4	1636	2211	747	1523
i=5	2747	2743	1469	1724
i=6	3649	4092	2002	2383
i=7	5435	5071	2209	2972
i=8	10907	11899	5358	6892
i=9	10650	14569	6333	9689
i=10	16339	25424	24566	23160
Overall	52135	67228	26909	28414

Even though the reserves R_i for i=9 and i=10 as well as the overall reserve R differ considerably they are all within one standard error and therefore not significantly different. But it should be remarked that this manner of using ICRFS is not intended by Zehnwirth because any initial model should be further adjusted according to the indications and plots given by the program. In this particular case there were strong indications for developing the model further but then one would have to give up the 'chain ladder model'.

7. Final Remark

This paper develops a rather complete methodology of how to attack the claims reserving task in a statistically sound manner on the basis of the well-known and simple chain ladder method. However, the well-known weak points of the chain ladder method should not be concealed: These are the fact that the estimators of the last two or three factors f_I, f_{I-1}, f_{I-2} rely on very few observations and the fact that the known claims amount C_{I1} of the last accident year (sometimes $C_{I-1,2}$, too) forms a very uncertain basis for the projection to ultimate. This is most clearly seen if C_{I1} happens to be 0: Then we have $\mathbf{C_{iI}} = 0$, $\mathbf{R_I} = 0$ and s.e.$(\mathbf{R_I}) = 0$ which obiously makes no sense. (Note that this weakness often can be overcome by translating and mixing the amounts C_{i1} of earlier accident years i < I into accident year I with the help of a measure of volume for each accident year.)

Thus, even if the statistical instruments developed do not reject the applicability of the chain ladder method, the result must be judged by an actuary and/or underwriter who knows the business under consideration. Even then, unexpected future changes can make all estimations obsolete. But for the many normal cases it is good to have a sound and simple method. Simple methods have the disadvantage of not capturing all aspects of reality but have the advantage that the user is in a position to know exactly how the method works and where its weaknesses are. Moreover, a simple method can be explained to non-actuaries in more detail. These are invaluable advantages of simple models over more sophisticated ones.

Appendix A: Unbiasedness of Age-to-Age Factors

Proposition: Under the assumptions

(3) There are unknown constants $f_1, \ldots, f_{I-1}$ with

$E(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}f_k, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

(4) The variables $\{C_{i1}, \ldots, C_{iI}\}$ and $\{C_{j1}, \ldots, C_{jI}\}$ of different accident years $i \neq j$ are independent.

the age-to-age factors $\mathbf{f}_1, \ldots, \mathbf{f}_{I-1}$ defined by

$$(2) \quad \mathbf{f}_k = \sum_{j=1}^{I-k} C_{j,k+1} \Big/ \sum_{j=1}^{I-k} C_{jk}, \quad 1 \le k \le I-1,$$

are unbiased, i.e. we have $E(\mathbf{f}_k) = f_k$, $1 \le k \le I-1$.

Proof: Because of the iterative rule for expectations we have

$$(A1) \quad E(\mathbf{f}_k) = E(E(\mathbf{f}_k|B_k))$$

for any set B_k of variables C_{ij} assumed to be known. We take

$$B_k = \{ C_{ij} \mid i+j \le I+1, \ j \le k \}, \quad 1 \le k \le I-1.$$

According to the definition (2) of $\mathbf{f}_k$ and because C_{jk}, $1 \le j \le I-k$, is contained in B_k and therefore has to be treated as scalar, we have

$$(A2) \quad E(\mathbf{f}_k|B_k) = \sum_{j=1}^{I-k} E(C_{j,k+1}|B_k) \Big/ \sum_{j=1}^{I-k} C_{jk}.$$

Because of the independence assumption (4) conditions relating to accident years other than that of $C_{j,k+1}$ can be omitted, i.e. we get

$$(A3) \quad E(C_{j,k+1}|B_k) = E(C_{j,k+1}|C_{j1},\ldots,C_{jk}) = C_{jk}f_k$$

using assumption (3) as well. Inserting (A3) into (A2) yields

(A4) $$E(\mathbf{f_k}|B_k) = \sum_{j=1}^{I-k} C_{jk} f_k \Big/ \sum_{j=1}^{I-k} C_{jk} = f_k \,.$$

Finally, (A1) and (A4) yield

$$E(\mathbf{f_k}) = E(f_k) = f_k$$

because f_k is a scalar.

Appendix B: Minimizing the Variance of Independent Estimators

Proposition: Let $T_1, \ldots, T_I$ be independent unbiased estimators of a parameter t, i.e. with

$$E(T_i) = t \;, \quad 1 \le i \le I,$$

then the variance of a linear combination

$$T = \sum_{i=1}^{I} w_i T_i$$

under the constraint

$$\text{(B1)} \qquad \sum_{i=1}^{I} w_i = 1$$

(which guarantees $E(T) = t$) is minimal iff the coefficients w_i are inversely proportional to $Var(T_i)$, i.e. iff

$$w_i = c/Var(T_i) \;, \quad 1 \le i \le I.$$

Proof: We have to minimize

$$Var(T) = \sum_{i=1}^{I} w_i^2 Var(T_i)$$

(due to the independence of $T_1, \ldots, T_I$) with respect to w_i under the constraint (B1). A necessary condition for an extremum is that the derivatives of the Lagrangian are zero, i.e. that

$$\text{(B2)} \qquad \frac{\partial}{\partial w_i} \left(\sum_{i=1}^{I} w_i^2 Var(T_i) + \lambda\Big(1 - \sum_{i=1}^{I} w_i\Big) \right) = 0 \;, \quad 1 \le i \le I,$$

with a constant multiplier λ whose value can be determined by additionally using (B1). (B2) yields

$$2w_i Var(T_i) - \lambda = 0$$

or

$$w_i = \lambda \,/\, (2\cdot \mathrm{Var}(T_i)) \ .$$

These weights w_i indeed lead to a minimum as can be seen by calculating the extremal value of Var(T) and applying Schwarz's inequality.

Corrollary: In the chain ladder case we have estimators $T_i = C_{i,k+1}/C_{ik}$, $1 \le i \le I-k$, for f_k where the variables of the set

$$A_k = \bigcup_{i=1}^{I-k} \{ C_{i1}, \ldots, C_{ik} \}$$

of the corresponding accident years $i = 1, \ldots, I-k$ up to development year k are considered to be given. We therefore want to minimize the conditional variance

$$\mathrm{Var}\Big(\sum_{i=1}^{I-k} w_i T_i \,|\, A_k\Big) \ .$$

From the above proof it is clear that the minimizing weights should be inversely proportional to $\mathrm{Var}(T_i|A_k)$. Because of the independence (4) of the accident years, conditions relating to accident years other than that of $T_i = C_{i,k+1}/C_{ik}$ can be omitted. We therefore have

$$\mathrm{Var}(T_i|A_k) = \mathrm{Var}(C_{i,k+1}/C_{ik}|C_{i1},\ldots,C_{ik})$$

and arrive at the result that

the minimizing weights should be

inversely proportional to $\mathrm{Var}(C_{i,k+1}/C_{ik}|C_{i1},\ldots,C_{ik})$.

Appendix C: Unbiasedness of the Estimated Ultimate Claims Amount

Proposition: Under the assumptions

(3) There are unknown constants $f_1, \ldots, f_{I-1}$ with

$E(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}f_k, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

(4) The variables $\{C_{i1}, \ldots, C_{iI}\}$ and $\{C_{j1}, \ldots, C_{jI}\}$ of different accident years $i \neq j$ are independent.

the expected values of the estimator

(1) $\mathbf{C}_{iI} = C_{i,I+1-i}\mathbf{f}_{I+1-i}\cdot\ldots\cdot\mathbf{f}_{I-1}$

for the ultimate claims amount and of the true ultimate claims amount C_{iI} are equal, i.e. we have $E(\mathbf{C}_{iI}) = E(C_{iI})$, $2 \le i \le I$.

Proof: We first show that the age-to-age factors $\mathbf{f}_k$ are uncorrelated. With the same set

$$B_k = \{ C_{ij} \mid i+j \le I+1,\ j \le k \}, \quad 1 \le k \le I-1,$$

of variables assumed to be known as in Appendix A we have for $j < k$

$$
\begin{aligned}
E(\mathbf{f}_j\mathbf{f}_k) &= E(E(\mathbf{f}_j\mathbf{f}_k|B_k)) && \text{(a)}\\
&= E(\mathbf{f}_jE(\mathbf{f}_k|B_k)) && \text{(b)}\\
&= E(\mathbf{f}_jf_k) && \text{(c)}\\
&= E(\mathbf{f}_j)f_k && \text{(d)}\\
&= f_jf_k\ . && \text{(e)}
\end{aligned}
$$

Here (a) holds because of the iterative rule for expectations, (b) holds because $\mathbf{f}_j$ is a scalar for B_k given and for $j < k$, (c) holds due to (A4), (d) holds because f_k is a scalar and (e) was shown in Appendix A.

This result can easily be extended to arbitrary products of different f_k's, i.e. we have

(C1) $\quad E(f_{I+1-i}\cdot\ldots\cdot f_{I-1}) = f_{i+1-i}\cdot\ldots\cdot f_{I-1}$.

This yields

$$
\begin{aligned}
E(C_{iI}) &= E(E(C_{iI}|C_{i1},\ldots,C_{i,I+1-i})) && \text{(a)}\\
&= E(E(C_{i,I+1-i}f_{I+1-i}\cdot\ldots\cdot f_{I-1}|C_{i1},\ldots,C_{i,I+1-i})) && \text{(b)}\\
&= E(C_{i,I+1-i}E(f_{I+1-i}\cdot\ldots\cdot f_{I-1}|C_{i1},\ldots,C_{i,I+1-i})) && \text{(c)}\\
&= E(C_{i,I+1-i}E(f_{I+1-i}\cdot\ldots\cdot f_{I-1})) && \text{(d)}\\
&= E(C_{i,I+1-i})\cdot E(f_{I+1-i}\cdot\ldots\cdot f_{I-1}) && \text{(e)}\\
&= E(C_{i,I+1-i})\cdot f_{I+1-i}\cdot\ldots\cdot f_{I-1}\ . && \text{(f)}
\end{aligned}
$$

Here (a) holds because of the iterative rule for expectations, (b) holds because of the definition (1) of C_{iI}, (c) holds because $C_{i,I+1-i}$ is a scalar under the stated condition, (d) holds because conditions which are independent from the conditioned variable $f_{I+1-i}\cdot\ldots\cdot f_{I-1}$ can be omitted (observe assumption (4) and the fact that f_{I+1-i}, ..., f_{I-1} only depend on variables of accident years < i), (e) holds because $E(f_{I+1-i}\cdot\ldots\cdot f_{I-1})$ is a scalar and (f) holds because of (C1).

Finally, repeated application of the iterative rule for expectations and of assumption (3) yields for the expected value of the true reserve C_{iI}

$$
\begin{aligned}
E(C_{iI}) &= E(E(C_{iI}|C_{i1},\ldots,C_{i,I-1}))\\
&= E(C_{i,I-1}f_{I-1})\\
&= E(C_{i,I-1})f_{I-1}\\
&= E(E(C_{i,I-1}|C_{i1},\ldots,C_{I-2}))f_{I-1}
\end{aligned}
$$

$$= E(C_{i,I-2} f_{I-2}) f_{I-1}$$

$$= E(C_{i,I-2}) f_{I-2} f_{I-1}$$

$$= \text{etc.}$$

$$= E(C_{i,I+1-i}) f_{I+1-i} \cdot \ldots \cdot f_{I-1}$$

$$= E(C_{iI}) \ .$$

Appendix D: Calculation of the Standard Error of $\hat{C}_{iI}$

Proposition: Under the assumptions

(3) There are unknown constants $f_1, \ldots, f_{I-1}$ with

$E(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}f_k, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

(4) The variables $\{C_{i1}, \ldots, C_{iI}\}$ and $\{C_{j1}, \ldots, C_{jI}\}$ of different accident years $i \neq j$ are independent.

(5) There are unknown constants $\alpha_1, \ldots, \alpha_{I-1}$ with

$\mathrm{Var}(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}\alpha_k^2, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

the standard error s.e.$(\hat{C}_{iI})$ of the estimated ultimate claims amount $\hat{C}_{iI} = C_{i,I+1-i}\hat{f}_{I+1-i}\cdot\ldots\cdot\hat{f}_{I-1}$ is given by the formula

$$(\mathrm{s.e.}(\hat{C}_{iI}))^2 = \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\alpha}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{ik}} + \frac{1}{\sum_{j=1}^{I-k} C_{jk}} \right)$$

where $\hat{C}_{ik} = C_{i,I+1-i}\hat{f}_{I+1-i}\cdots\hat{f}_{k-1}$, $k > I+1-i$, are the estimated values of the future C_{ik} and $\hat{C}_{i,I+1-i} = C_{i,I+1-i}$.

Proof: As stated in Chapter 4, the standard error is the square root of an estimator of $\mathrm{mse}(\hat{C}_{iI})$ and we have also seen that

(D1) $\quad \mathrm{mse}(\hat{C}_{iI}) = \mathrm{Var}(C_{iI}|D) + (E(C_{iI}|D) - \hat{C}_{iI})^2$.

In the following, we use the abbreviations

$E_i(X) = E(X|C_{i1}, \ldots, C_{i,I+1-i})$,

$\mathrm{Var}_i(X) = \mathrm{Var}(X|C_{i1}, \ldots, C_{i,I+1-i})$.

Because of the independence of the accident years we can omit in (D1) that part of the condition $D = \{ C_{ik} \mid i+k \le I+1 \}$ which is independent from C_{iI}, i.e. we can write

(D2) $\quad \mathrm{mse}(\hat{C}_{iI}) = \mathrm{Var}_i(C_{iI}) + (E_i(C_{iI}) - \hat{C}_{iI})^2$.

We first consider $Var_i(C_{iI})$. Because of the general rule $Var(X) = E(X^2) - (E(X))^2$ we have

(D3) $\quad Var_i(C_{iI}) = E_i(C_{iI}^2) - (E_i(C_{iI}))^2$.

For the calculation of $E_i(C_{iI})$ we use the fact that for $k \geq I+1-i$

(D4)
$$\begin{aligned} E_i(C_{i,k+1}) &= E_i(E(C_{i,k+1}|C_{i1}, \ldots, C_{ik})) \\ &= E_i(C_{ik}f_k) \\ &= E_i(C_{ik})f_k \ . \end{aligned}$$

Here, we have used the iterative rule for expectations in its general form $E(X|Z) = E(E(X|Y)|Z)$ for $\{Y\} \supset \{Z\}$ (mostly we have $\{Z\} = \emptyset$). By successively applying (D4) we obtain for $k \geq I+1-i$

(D5)
$$\begin{aligned} E_i(C_{i,k+1}) &= E_i(C_{i,I+1-i})f_{I+1-i}\cdot\ldots\cdot f_k \\ &= C_{i,I+1-i}f_{I+1-i}\cdot\ldots\cdot f_k \end{aligned}$$

because $C_{i,I+1-i}$ is a scalar under the condition 'i'.

For the calculation of the first term $E_i(C_{iI}^2)$ of (D3) we use the fact that for $k \geq I+1-i$

(D6)
$$\begin{aligned} E_i(C_{i,k+1}^2) &= E_i(E(C_{i,k+1}^2|C_{i1}, \ldots, C_{ik}) & \text{(a)} \\ &= E_i(\ Var(C_{i,k+1}|C_{i1}, \ldots, C_{ik}) + & \text{(b)} \\ &\qquad + (E(C_{i,k+1}|C_{i1}, \ldots, C_{ik}))^2\ ) \\ &= E_i(\ C_{ik}\alpha_k^2 + (C_{ik}f_k)^2\ ) & \text{(c)} \\ &= E_i(C_{ik})\alpha_k^2 + E_i(C_{ik}^2)f_k^2\ . \end{aligned}$$

Here, (a) holds due to the iterative rule for expectations, (b) due to the rule $E(X^2) = Var(X) + (E(X))^2$ and (c) holds due to (3) and (5).

Now, we apply (D6) and (D5) successively to get

(D7) $$E_i(C_{iI}^2) = E_i(C_{i,I-1})\alpha_{I-1}^2 + E_i(C_{i,I-1}^2)f_{I-1}^2 \quad \text{(D6)}$$

$$= C_{i,I+1-1}f_{I+1-1}\cdots f_{I-2}\alpha_{I-1}^2 + \quad \text{(D5)}$$

$$+ E_i(C_{i,I-2})\alpha_{I-2}^2 f_{I-1}^2 + \quad \text{(D6)}$$

$$+ E_i(C_{i,I-2}^2)f_{I-2}^2 f_{I-1}^2$$

$$= C_{i,I+1-1}f_{I+1-1}\cdots f_{I-2}\alpha_{I-1}^2 +$$

$$+ C_{i,I+1-1}f_{I+1-1}\cdots f_{I-3}\alpha_{I-2}^2 f_{I-1}^2 + \quad \text{(D5)}$$

$$+ E_i(C_{i,I-3})\alpha_{I-3}^2 f_{I-2}^2 f_{I-1}^2 + \quad \text{(D6)}$$

$$+ E_i(C_{i,I-3}^2)f_{I-3}^2 f_{I-2}^2 f_{I-1}^2$$

$$= \text{etc.}$$

$$= C_{i,I+1-i} \sum_{k=I+1-i}^{I-1} f_{I+1-i}\cdots f_{k-1}\alpha_k^2 f_{k+1}^2\cdots f_{I-1}^2$$

$$+ C_{i,I+1-i}^2 f_{I+1-i}^2\cdot\ldots\cdot f_{I-1}^2$$

where in the last step we have used $E_i(C_{i,I+1-i}) = C_{i,I+1-i}$ and $E_i(C_{i,I+1-i}^2) = C_{i,I+1-i}^2$ because under the condition 'i' $C_{i,I+1-i}$ is a scalar.

Due to (D5) we have

(D8) $$(E_i(C_{iI}))^2 = C_{i,I+1-i}^2 f_{I+1-i}^2\cdot\ldots\cdot f_{I-1}^2 .$$

Inserting (D7) and (D8) into (D3) yields

(D9) $$\mathrm{Var}_i(C_{iI}) = C_{i,I+1-i} \sum_{k=I+1-i}^{I-1} f_{I+1-i}\cdots f_{k-1}\alpha_k^2 f_{k+1}^2\cdots f_{I-1}^2$$

We estimate this first summand of $\mathrm{mse}(\mathbf{C}_{iI})$ by replacing the unknown parameters f_k, α_k^2 with their unbiased estimators $\mathbf{f}_k$ and $\boldsymbol{\alpha}_k^2$, i.e. by

(D10) $$C_{i,I+1-i} \sum_{k=I+1-i}^{I-1} \mathbf{f}_{I+1-i}\cdots \mathbf{f}_{k-1}\cdot\boldsymbol{\alpha}_k^2\cdot\mathbf{f}_{k+1}^2\cdots\mathbf{f}_{I-1}^2 =$$

$$= C_{i,I+1-i}^2 \hat{f}_{I+1-i}^2 \cdots \hat{f}_{I-1}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\alpha}_k^2/\hat{f}_k^2}{C_{i,I+1-i}\hat{f}_{I+1-i}\cdots\hat{f}_{k-1}}$$

$$= \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\alpha}_k^2/\hat{f}_k^2}{\hat{C}_{ik}}$$

where we have used the notation $\hat{C}_{ik}$ introduced in the proposition for the estimated amounts of the future C_{ik}, $k > I+1-i$, including $\hat{C}_{i,I+1-i} = C_{i,I+1-i}$.

We now turn to the second summand of the expression (D2) for $\text{mse}(\hat{C}_{iI})$. Because of (D5) we have

$$E_i(C_{iI}) = C_{i,I+1-i} f_{I+1-i} \cdot \ldots \cdot f_{I-1}$$

and therefore

(D11) $(E_i(C_{iI}) - \hat{C}_{iI})^2 =$

$$= C_{i,I+1-i}^2 (f_{I+1-i} \cdot \ldots \cdot f_{I-1} - \hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{I-1})^2 .$$

This expression cannot simply be estimated by replacing f_k with $\hat{f}_k$ because this would yield 0 which is not a good estimator because $\hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{I-1}$ generally will be different from $f_{I+1-i} \cdot \ldots \cdot f_{I-1}$ and therefore the squared difference will be positive. We therefore must take a different approach. We use the algebraic identity

$$F = f_{I+1-i} \cdot \ldots \cdot f_{I-1} - \hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{I-1}$$
$$= S_{I+1-i} + \ldots + S_{I-1}$$

with

$$S_k = \hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{k-1} f_k f_{k+1} \cdot \ldots \cdot f_{I-1} -$$
$$- \hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{k-1} \hat{f}_k f_{k+1} \cdot \ldots \cdot f_{I-1}$$
$$= \hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{k-1} (f_k - \hat{f}_k) f_{k+1} \cdot \ldots \cdot f_{I-1} .$$

This yields

$$F^2 = (S_{I+1-i} + \ldots + S_{I-1})^2$$

$$= \sum_{k=I+1-i}^{I-1} S_k^2 + 2 \sum_{j<k} S_j S_k \,.$$

where in the last summation j and k run from I+1-i to I-1. Now we replace S_k^2 with $E(S_k^2|B_k)$ and S_jS_k, $j < k$, with $E(S_jS_k|B_k)$. This means that we approximate S_k^2 and S_jS_k by varying and averaging as little data as possible so that as many values C_{ik} as possible from data observed are kept fixed. Due to (A4) we have $E(f_k-\mathbf{f_k}|B_k) = 0$ and therefore $E(S_jS_k|B_k) = 0$ for $j < k$ because all $\mathbf{f_r}$, $r < k$, are scalars under B_k. Because of

(D12) $$E((f_k-\mathbf{f_k})^2|B_k) = \mathrm{Var}(\mathbf{f_k}|B_k)$$

$$= \sum_{j=1}^{I-k} \mathrm{Var}(C_{j,k+1}|B_k)/(\sum_{j=1}^{I-k} C_{jk})^2$$

$$= \sum_{j=1}^{I-k} \mathrm{Var}(C_{j,k+1}|C_{j1},\ldots,C_{jk})/(\sum_{j=1}^{I-k} C_{jk})^2$$

$$= \sum_{j=1}^{I-k} C_{jk}\alpha_k^2 \;/\; (\sum_{j=1}^{I-k} C_{jk})^2$$

$$= \alpha_k^2 \;/\; \sum_{j=1}^{I-k} C_{jk}$$

we obtain

$$E(S_k^2|B_k) = \mathbf{f}_{I+1-i}^2 \cdots \mathbf{f}_{k-1}^2 \alpha_k^2 f_{k+1}^2 \cdots f_{I-1}^2 \;/\; \sum_{j=1}^{I-k} C_{jk} \,.$$

Taken together, we have replaced $F^2 = (\Sigma\, S_k)^2$ with $\Sigma_k E(S_k^2|B_k)$ and because all terms of this sum are positive we can replace all unknown parameters f_k, α_k^2 with their unbiased estimators

f_k, α_k^2. Altogether, we estimate $F^2 = (f_{I+1-i} \cdot \ldots \cdot f_{I-1} - \hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{I-1})^2$ by

$$\sum_{k=I+1-i}^{I-1} \left(\hat{f}_{I+1-i}^2 \cdots \hat{f}_{k-1}^2 \cdot \hat{\alpha}_k^2 \cdot \hat{f}_{k+1}^2 \cdots \hat{f}_{I-1}^2 \Big/ \sum_{j=1}^{I-k} C_{jk} \right) =$$

$$= \hat{f}_{I+1-i}^2 \cdot \ldots \cdot \hat{f}_{I-1}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\alpha}_k^2 / \hat{f}_k^2}{\sum_{j=1}^{I-k} C_{jk}} .$$

Using (D11), this means that we estimate $(E_i(C_{iI}) - \hat{C}_{iI})^2$ by

$$\text{(D13)} \quad C_{i,I+1-i}^2 \hat{f}_{I+1-i}^2 \cdot \ldots \cdot \hat{f}_{I-1}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\alpha}_k^2 / \hat{f}_k^2}{\sum_{j=1}^{I-k} C_{jk}} =$$

$$= \hat{C}_{iI}^2 \sum_{k=I+1-i}^{I-1} \frac{\hat{\alpha}_k^2 / \hat{f}_k^2}{\sum_{j=1}^{I-k} C_{jk}} .$$

From (D2), (D10) and (D13) we finally obtain the estimator $(\text{s.e.}(\hat{C}_{iI}))^2$ for $\text{mse}(\hat{C}_{iI})$ as stated in the proposition.

Appendix E: Unbiasedness of the Estimator α_k^2

Proposition: Under the assumptions

(3) There are unknown constants $f_1, \ldots, f_{I-1}$ with

$E(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}f_k, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

(4) The variables $\{C_{i1}, \ldots, C_{iI}\}$ and $\{C_{j1}, \ldots, C_{jI}\}$ of different accident years $i \neq j$ are independent.

(5) There are unknown constants $\alpha_1, \ldots, \alpha_{I-1}$ with

$Var(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}\alpha_k^2, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

the estimators

$$\alpha_k^2 = \frac{1}{I-k-1} \sum_{j=1}^{I-k} C_{jk} \left(\frac{C_{j,k+1}}{C_{jk}} - f_k \right)^2, \quad 1 \le k \le I-2,$$

of α_k^2 are unbiased, i.e. we have

$$E(\alpha_k^2) = \alpha_k^2, \qquad 1 \le k \le I-2.$$

Proof: In this proof all summations are over the index j from j=1 to j=I-k. The definition of α_k^2 can be rewritten as

(E1) $\quad (I-k-1)\alpha_k^2 = \Sigma\ (C_{j,k+1}^2/C_{jk} - 2\cdot C_{j,k+1}f_k + C_{jk}f_k^2)$

$\qquad\qquad = \Sigma\ (C_{j,k+1}^2/C_{jk}) - \Sigma\ (C_{jk}f_k^2)$

using $\Sigma C_{j,k+1} = f_k \Sigma C_{jk}$ according to the definition of f_k. Using again the set

$$B_k = \{ C_{ij} \mid i+j \le I+1, \ j \le k \}$$

of variables C_{ij} assumed to be known, (E1) yields

(E2) $\quad E((I-k-1)\alpha_k^2|B_k) = \Sigma\ E(C_{j,k+1}^2|B_k)/C_{jk} - \Sigma\ C_{jk}E(f_k^2|B_k)$

because C_{jk} is a scalar under the condition of B_k being known. Due to the independence (4) of the accident years, conditions which are independent from the conditioned variable can be omitted in $E(C_{j,k+1}{}^2|B_k)$, i.e.

$$(E3) \quad E(C_{j,k+1}{}^2|B_k) = E(C_{j,k+1}{}^2|C_{j1},\ldots,C_{jk})$$
$$= Var(C_{j,k+1}|C_{j1},\ldots,C_{jk}) + (E(C_{j,k+1}|C_{j1},\ldots,C_{jk}))^2$$
$$= C_{jk}\alpha_k{}^2 + (C_{jk}f_k)^2$$

where the rule $E(X^2) = Var(X) + (E(X))^2$ and the assumptions (5) and (3) have also been used.

From (D12) and (A4) we gather

$$(E4) \quad E(\mathbf{f_k}{}^2|B_k) = Var(\mathbf{f_k}|B_k) + (E(\mathbf{f_k}|B_k))^2$$
$$= \alpha_k{}^2 / \Sigma C_{jk} + f_k{}^2 .$$

Inserting (E3) and (E4) into (E2) we obtain

$$E((I-k-1)\boldsymbol{\alpha_k}{}^2|B_k) =$$
$$= \sum_{j=1}^{I-k} (\alpha_k{}^2 + C_{jk}f_k{}^2) - \sum_{j=1}^{I-k} (C_{jk}\alpha_k{}^2/ \sum_{j=1}^{I-k} C_{jk} + C_{jk}f_k{}^2)$$
$$= (I-k)\alpha_k{}^2 - \alpha_k{}^2$$
$$= (I-k-1)\alpha_k{}^2 .$$

From this we immediately obtain $E(\boldsymbol{\alpha_k}{}^2|B_k) = \alpha_k{}^2$.

Finally, the iterative rule for expectations yields

$$E(\boldsymbol{\alpha_k}{}^2) = E(E(\boldsymbol{\alpha_k}{}^2|B_k)) = E(\alpha_k{}^2) = \alpha_k{}^2 .$$

Appendix F: The Standard Error of the Overall Reserve Estimate

Proposition: Under the assumptions

(3) There are unknown constants $f_1, \ldots, f_{I-1}$ with

$E(C_{i,k+1} | C_{i1}, \ldots, C_{ik}) = C_{ik} f_k, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

(4) The variables $\{C_{i1}, \ldots, C_{iI}\}$ and $\{C_{j1}, \ldots, C_{jI}\}$ of different accident years $i \neq j$ are independent.

(5) There are unknown constants $\alpha_1, \ldots, \alpha_{I-1}$ with

$\mathrm{Var}(C_{i,k+1} | C_{i1}, \ldots, C_{ik}) = C_{ik} \alpha_k^2, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$

the standard error s.e.(**R**) of the overall reserve estimate

$$\mathbf{R} = \mathbf{R}_2 + \ldots + \mathbf{R}_I$$

is given by

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

Proof: This proof is analogous to that in Appendix D. The comments will therefore be brief.

We first must determine the mean squared error mse(**R**) of **R**.

Using again $D = \{ C_{ik} \mid i+k \le I+1 \}$ we have

$$\text{(F1)} \quad \mathrm{mse}\left(\sum_{i=2}^{I} \mathbf{R}_i\right) = E\left(\left(\sum_{i=2}^{I} \mathbf{R}_i - \sum_{i=2}^{I} R_i\right)^2 \Big| D\right)$$

$$= E\left(\left(\sum_{i=2}^{I} \mathbf{C}_{iI} - \sum_{i=2}^{I} C_{iI}\right)^2 \Big| D\right)$$

$$= \mathrm{Var}\left(\sum_{i=2}^{I} C_{iI} \Big| D\right) + \left(E\left(\sum_{i=2}^{I} C_{iI} \Big| D\right) - \sum_{i=2}^{I} \mathbf{C}_{iI} \right)^2 .$$

The independence of the accident years yields

(F2) $\quad \mathrm{Var}\left(\sum_{i=2}^{I} C_{iI} \mid D\right) = \sum_{i=2}^{I} \mathrm{Var}(C_{iI} \mid C_{i1}, \ldots, C_{i,I+1-i})$,

whose summands have been calculated in Appendix D, see (D9). Furthermore

(F3) $$\left(E\left(\sum_{i=2}^{I} C_{iI} \mid D\right) - \sum_{i=2}^{I} \hat{C}_{iI} \right)^2 = \left(\sum_{i=2}^{I} \left(E(C_{iI} \mid D) - \hat{C}_{iI} \right) \right)^2 =$$

$$= \sum_{2 \le i,j \le I} (E(C_{iI} \mid D) - \hat{C}_{iI}) \cdot (E(C_{jI} \mid D) - \hat{C}_{jI})$$

$$= \sum_{2 \le i,j \le I} C_{i,I+1-i} C_{j,I+1-j} F_i F_j$$

$$= \sum_{i=2}^{I} (C_{i,I+1-i} F_i)^2 + 2 \sum_{i<j} C_{i,I+1-i} C_{j,I+1-j} F_i F_j$$

with (like in (D11))

$$F_i = f_{I+1-i} \cdot \ldots \cdot f_{I-1} - \hat{f}_{I+1-i} \cdot \ldots \cdot \hat{f}_{I-1}$$

which is identical to F of Appendix D but here we have to carry the index i, too. In Appendix D we have shown (cf. (D2) and (D11)) that

$$\mathrm{mse}(\hat{R}_i) = \mathrm{Var}(C_{iI} \mid C_{i1}, \ldots, C_{i,I+1-i}) + (C_{i,I+1-i} F_i)^2 .$$

Comparing this with (F1), (F2) and (F3) we see that

(F4) $$\mathrm{mse}\left(\sum_{i=2}^{I} \hat{R}_i\right) = \sum_{i=2}^{I} \mathrm{mse}(\hat{R}_i) + \sum_{2 \le i < j \le I} 2 \cdot C_{i,I+1-i} C_{j,I+1-j} F_i F_j .$$

We therefore need only develop an estimator for $F_i F_j$. A procedure completely analogous to that for F^2 in the proof of Appendix D yields for $F_i F_j$, $i<j$, the estimator

$$\sum_{k=I+1-i}^{I-1} \hat{f}_{I+1-j} \cdot \ldots \cdot \hat{f}_{I-i} \hat{f}^2_{I+1-i} \cdot \ldots \cdot \hat{f}^2_{k-1} \hat{\alpha}^2_k \hat{f}^2_{k+1} \cdot \ldots \cdot \hat{f}^2_{I-1} \Big/ \sum_{n=1}^{I-k} C_{nk} ,$$

which immediately leads to the result stated in the proposition.

Appendix G: Testing for Correlations between Subsequent Development Factors

In this appendix we first prove that the basic assumption (3) of the chain ladder method implies that subsequent development factors $C_{ik}/C_{i,k-1}$ and $C_{i,k+1}/C_{ik}$ are not correlated. Then we show how we can test if this uncorrelatedness is met for a given run-off triangle. Finally, we apply this test procedure to the numerical example of Chapter 6.

Proposition: Under the assumption

(3) There are unknown constants $f_1, \ldots, f_{I-1}$ with

$$E(C_{i,k+1}|C_{i1},\ldots,C_{ik}) = C_{ik}f_k, \quad 1 \le i \le I, \quad 1 \le k \le I-1.$$

subsequent development factors $C_{ik}/C_{i,k-1}$ and $C_{i,k+1}/C_{ik}$ are uncorrelated, i.e. we have (for $1 \le i \le I$, $2 \le k \le I-1$)

$$E\left(\frac{C_{ik}}{C_{i,k-1}}\cdot\frac{C_{i,k+1}}{C_{ik}}\right) = E\left(\frac{C_{ik}}{C_{i,k-1}}\right)\cdot E\left(\frac{C_{i,k+1}}{C_{ik}}\right).$$

Proof: For $j \le k$ we have

$$\text{(G1)} \quad E(C_{i,k+1}/C_{ij}) = E(E(C_{i,k+1}/C_{ij}|C_{i1},\ldots,C_{ik})) \quad \text{(a)}$$

$$= E(E(C_{i,k+1}|C_{i1},\ldots,C_{ik})/C_{ij}) \quad \text{(b)}$$

$$= E(C_{ik}f_k/C_{ij}) \quad \text{(c)}$$

$$= E(C_{ik}/C_{ij})f_k\,. \quad \text{(d)}$$

Here equation (a) holds due to the iterative rule $E(X) = E(E(X|Y))$ for expectations, (b) holds because, given $C_{i1}, \ldots, C_{ik}$, C_{ij} is a scalar for $j \le k$, (c) holds due to (3) and (d) holds because f_k is a scalar.

From (G1) we obtain through the specialization $j = k$

(G2) $\quad E(C_{i,k+1}/C_{ik}) = E(C_{ik}/C_{ik})f_k = f_k$

and through $j = k-1$

(G3) $$E\left(\frac{C_{ik}}{C_{i,k-1}}\cdot\frac{C_{i,k+1}}{C_{ik}}\right) = E\left(\frac{C_{i,k+1}}{C_{i,k-1}}\right) \overset{(G1)}{=} E\left(\frac{C_{ik}}{C_{i,k-1}}\right)f_k .$$

Inserting (G2) into (G3) completes the proof.

Designing the test procedure:

The usual test for uncorrelatedness requires that we have identically distributed pairs of observations which come from a Normal distribution. Both conditions are usually not fulfilled for adjacent columns of development factors. (Note that due to (G2) the development factors $C_{i,k+1}/C_{ik}$, $1 \le i \le I-k$, have the same expectation but assumption (5) implies that they have different variances.) We therefore use the test with Spearman's rank correlation coefficient because this test is distribution-free and because by using ranks the differences in the variances of $C_{i,k+1}/C_{ik}$, $1 \le i \le I-k$, become less important. Even if these differences are negligeable the test will only be of an approximate nature because, strictly speaking, it is a test for independence rather than for uncorrelatedness. But we will take this into account when fixing the critical value of the test statistic.

For the application of Spearman's test we consider a fixed development year k and rank the development factors $C_{i,k+1}/C_{ik}$ observed so far according to their size starting with the smallest one on rank one and so on. Let r_{ik}, $1 \le i \le I-k$, denote the rank of $C_{i,k+1}/C_{ik}$ obtained in this way, $1 \le r_{ik} \le I-k$. Then we do the same with the preceding development factors $C_{ik}/C_{i,k-1}$, $1 \le i \le I-k$, leaving out $C_{I+1-k,k}/C_{I+1-k,k-1}$ for which the subsequent development factor has not yet been observed. Let s_{ik}, $1 \le i \le I-k$, be the ranks obtained in this way, $1 \le s_{ik} \le I-k$. Now, Spearman's rank correlation coefficient T_k is defined to be

$$\text{(G4)} \qquad T_k = 1 - 6 \sum_{i=1}^{I-k} (r_{ik} - s_{ik})^2 \,/\, ((I-k)^3 - I + k) \;.$$

From a textbook of Mathematical Statistics it can be seen that

$$-1 \le T_k \le +1 \;,$$

and, under the null-hypothesis,

$$E(T_k) = 0 \;,$$

$$\mathrm{Var}(T_k) = 1/(I-k-1) \;.$$

A value of T_k close to 0 indicates that the development factors between development years k-1 and k and those between years k and k+1 are not correlated. Any other value of T_k indicates that the factors are (positively or negatively) correlated.

For a formal test we do not want to consider every pair of columns of adjacent development years separately in order to avoid an accumulation of the error probabilities. We therefore consider the triangle as a whole. This also is preferable from a practical point of view because it is more important to know whether correlations globally prevail than to find a small part of the triangle with correlations. We therefore combine all values T_2, T_3, ..., T_{I-2} obtained in the same way like T_k. (There is no T_1 because there are no development factors before development year k=1 and similarly there is also no T_I; even T_{I-1} is not included because there is only one rank and therefore no randomness.) According to Appendix B we should not form an unweighted average of T_2, ..., T_{I-2} but rather use weights which are inversely proportional to $Var(T_k) = 1/(I-k-1)$. This leads to weights which are just equal to one less than the number of pairs (r_{ik}, s_{ik}) taken into account by T_k which seems very reasonable.

We thus calculate

$$\text{(G5)} \quad T = \sum_{k=2}^{I-2} (I-k-1)T_k \Big/ \sum_{k=2}^{I-2} (I-k-1)$$

$$= \sum_{k=2}^{I-2} \frac{I-k-1}{(I-2)(I-3)/2} T_k ,$$

$$E(T) = \sum_{k=2}^{I-2} E(T_k) = 0 ,$$

$$\text{(G6)} \quad Var(T) = \sum_{k=2}^{I-2} (I-k-1)^2 \, Var(T_k) \Big/ \Big(\sum_{k=2}^{I-2} (I-k-1) \Big)^2$$

$$= \sum_{k=2}^{I-2} (I-k-1) \Big/ \Big(\sum_{k=2}^{I-2} (I-k-1) \Big)^2$$

$$= \frac{1}{(I-2)(I-3)/2}$$

where for the calculation of Var(T) we used the fact that under the null-hypothesis subsequent development factors and therefore also different T_k's are uncorrelated.

Because the distribution of a single T_k with $I-k \geq 10$ is Normal in good approximation and because T is the aggregation of several uncorrelated T_k's (which all are symmetrically distributed around their mean 0) we can assume that T has approximately a Normal distribution and use this to design a significance test. Usually, when applying a significance test one rejects the null-hypothesis if it is very unlikely to hold, e.g. if the value of the test statistic is outside its 95% confidence interval. But in our case we propose to use only a 50% confidence interval because the test is only of an approximate nature and because we want to detect correlations already in a substantial part of the run-off triangle. Therefore, as the probability for a Standard Normal variate lying in the interval (-.67, .67) is 50% we do not reject the null-hypothesis of having uncorrelated development factors if

$$-\frac{.67}{\sqrt{((I-2)(I-3)/2)}} \leq T \leq +\frac{.67}{\sqrt{((I-2)(I-3)/2)}} .$$

If T is outside this interval we should be reluctant with the application of the chain ladder method and analyze the correlations in more detail.

<u>Application to the example of Chapter 6</u>:

We start with the table of all development factors:

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9
i=1	1.6	1.32	1.08	1.15	1.20	1.11	1.033	1.00	1.01
i=2	40.4	1.26	1.98	1.29	1.13	0.99	1.043	1.03	
i=3	2.6	1.54	1.16	1.16	1.19	1.03	1.026		
i=4	2.0	1.36	1.35	1.10	1.11	1.04			
i=5	8.8	1.66	1.40	1.17	1.01				
i=6	4.3	1.82	1.11	1.23					
i=7	7.2	2.72	1.12						
i=8	5.1	1.89							
i=9	1.7								

As described above we first rank column F_1 according to the size of the factors, then leave out the last element and rank the column again. Then we do the same with columns F_2 to F_8. This yields the following table:

r_{i1}	s_{i2}	r_{i2}	s_{i3}	r_{i3}	s_{i4}	r_{i4}	s_{i5}	r_{i5}	s_{i6}	r_{i6}	s_{i7}	r_{i7}	s_{i8}	r_{i8}
1	1	2	2	1	1	2	2	5	4	4	3	2	1	1
9	8	1	1	7	6	6	5	3	2	1	1	3	2	2
4	3	4	4	4	3	3	3	4	3	2	2	1		
3	2	3	3	5	4	1	1	2	1	3				
8	7	5	5	6	5	4	4	1						
5	4	6	6	2	2	5								
7	6	8	7	3										
6	5	7												
2														

We now add the squared differences between adjacent rank columns of equal length, i.e. we add $(s_{ik} - r_{ik})^2$ over i for every k, $2 \le k \le 8$. This yields 68, 74, 20, 24, 6, 6 and 0. (Remember that we have to leave out k = 1 because there is no s_{i1}, and k = 9 because there is only one pair of ranks and therefore no randomness.) From these figures we obtain Spearman's rank correlation coefficients T_k according to formula (G4):

k	2	3	4	5	6	7	8
T_k	4/21	-9/28	3/7	-1/5	2/5	-1/2	1
I-k-1	7	6	5	4	3	2	1

The (I-k-1)-weighted average of the T_k's is T = .070 (see formula (G5)). Because of Var(T) = 1/28 (see (G6)) the 50% confidence limits for T are $\pm.67/\sqrt{28} = \pm.127$. Thus, T is within its 50%-interval and the hypothesis of having uncorrelated development factors is not rejected.

Appendix H: Testing for Calendar Year Effects

One of the three basic assumptions underlying the chain ladder method was seen to be assumption (4) of the independence of the accident years. The main reason why this independence can be violated in practice is the fact that we can have certain calendar year effects such as major changes in claims handling or in case reserving or external influences such as substantial changes in court decisions or inflation. Note that a constant rate of inflation which has not been removed from the data is extrapolated into the future by the chain ladder method. In the following, we first generally describe a procedure to test for such calendar year influences and then apply it to our example.

Designing the test procedure:

A calendar year influence affects one of the diagonals

$$D_j = \{ C_{j1}, C_{j-1,2}, \ldots, C_{2,j-1}, C_{1j} \}, \quad 1 \le j \le I,$$

and therefore also influences the adjacent development factors

$$A_j = \{ C_{j2}/C_{j1}, C_{j-1,3}/C_{j-1,2}, \ldots, C_{1,j+1}/C_{1j} \}$$

and

$$A_{j-1} = \{ C_{j-1,2}/C_{j-1,1}, C_{j-2,3}/C_{j-2,2}, \ldots, C_{1j}/C_{1,j-1} \}$$

where the elements of D_j form either the denominator or the numerator. Thus, if due to a calendar year influence the elements of D_j are larger (smaller) than usual, then the elements of A_{j-1} are also larger (smaller) than usual and the elements of A_j are smaller (larger) than usual.

Therefore, in order to check for such calendar year influences we only have to subdivide all development factors into 'smaller' and 'larger' ones and then to examine whether there are diagonals where the small development factors or the large ones clearly prevail. For this purpose, we order for every k, $1 \le k \le I-1$, the elements of the set

$$F_k = \{ C_{i,k+1}/C_{ik} \mid 1 \le i \le I-k \} ,$$

i.e. of the column of all development factors observed between development years k and k+1, according to their size and subdivide them into one part LF_k of larger factors being greater than the median of F_k and into a second part SF_k of smaller factors below the median of F_k. (The median of a set of real numbers is defined to be a number which divides the set into two parts with the same number of elements.) If the number I-k of elements of F_k is odd there is one element of F_k which is equal to the median and therefore assigned to neither of the sets LF_k and SF_k; this element is eliminated from all further considerations.

Having done this procedure for each set F_k, $1 \le k \le I-1$, every development factor observed is

- either eliminated (like e.g. the only element of F_{I-1})
- or assigned to the set $L = LF_1 + \ldots + LF_{I-2}$ of larger factors
- or assigned to the set $S = SF_1 + \ldots + SF_{I-2}$ of smaller factors. In this way, every development factor which is not eliminated has a 50% chance of belonging to either L or S.

Now we count for every diagonal A_j, $1 \le j \le I-1$, of development factors the number L_j of large factors, i.e. elements of L, and the number S_j of small factors, i.e. elements of S. Intuitively, if there is no specific change from calendar year j to calendar year j+1, A_j should have about the same number of small factors as of large factors, i.e. L_j and S_j should be of approximately the same size apart from pure random fluctuations. But if L_j is significantly larger or smaller than S_j or, equivalently, if

$$Z_j = \min(L_j, S_j) \ ,$$

i.e. the smaller of the two figures, is significantly smaller than $(L_j+S_j)/2$, then there is some reason for a specific calendar year influence.

In order to design a formal test we need the first two moments of the probability distribution of Z_j under the hypothesis that each development factor has a 50 % probability of belonging to either L or S. This distribution can easily be established. We give an example for the case where $L_j+S_j = 5$, i.e. where the set A_j contains 5 development factors without counting any eliminated factor. Then the number L_j has a Binomial distribution with n = 5 and p = .5, i.e.

$$\text{prob}(L_j = m) = \binom{n}{m} \frac{1}{2^n} = \binom{5}{m} \frac{1}{2^5} \ , \quad m = 0, 1, \ldots, 5.$$

Therefore

$$\text{prob}(S_j = 5) = \text{prob}(L_j = 0) = 1/32 \ ,$$

$$\text{prob}(S_j = 4) = \text{prob}(L_j = 1) = 5/32 \ ,$$

$$\mathrm{prob}(S_j = 3) = \mathrm{prob}(L_j = 2) = 10/32 ,$$

$$\mathrm{prob}(S_j = 2) = \mathrm{prob}(L_j = 3) = 10/32 ,$$

$$\mathrm{prob}(S_j = 1) = \mathrm{prob}(L_j = 4) = 5/32 ,$$

$$\mathrm{prob}(S_j = 0) = \mathrm{prob}(L_j = 5) = 1/32 .$$

This yields

$$\mathrm{prob}(Z_j = 0) = \mathrm{prob}(L_j = 0) + \mathrm{prob}(S_j = 0) = 2/32 ,$$

$$\mathrm{prob}(Z_j = 1) = \mathrm{prob}(L_j = 1) + \mathrm{prob}(S_j = 1) = 10/32 ,$$

$$\mathrm{prob}(Z_j = 2) = \mathrm{prob}(L_j = 2) + \mathrm{prob}(S_j = 2) = 20/32 ,$$

$$E(Z_j) = (0\cdot 2 + 1\cdot 10 + 2\cdot 20)/32 = 50/32 ,$$

$$E(Z_j^2) = (0\cdot 2 + 1\cdot 10 + 4\cdot 20)/32 = 90/32 ,$$

$$\mathrm{Var}(Z_j) = E(Z_j^2) - (E(Z_j))^2 = 95/256 .$$

The derivation of the general formula is straightforward but tedious. We therefore give only its result. If $n = L_j+S_j$ and $m = [(n-1)/2]$ denotes the largest integer $\leq (n-1)/2$ then

$$\text{(H1)} \qquad E(Z_j) = \frac{n}{2} - \binom{n-1}{m} \frac{n}{2^n} ,$$

$$\text{(H2)} \qquad \mathrm{Var}(Z_j) = \frac{n(n-1)}{4} - \binom{n-1}{m} \frac{n(n-1)}{2^n} + E(Z_j) - (E(Z_j))^2 .$$

It is not advisable to test each Z_j separately in order to avoid an accumulation of the error probabilities. Instead, we consider

$$Z = Z_2 + \ldots + Z_{I-1}$$

where we have left out Z_1 because A_1 contains at most one element which is not eliminated and therefore Z_1 is not a random variable but always $= 0$. Similarly, we have to leave out any other Z_j if $L_j+S_j \leq 1$. Because under the null-hypothesis different Z_j's are (almost) uncorrelated we have

$$E(Z) = E(Z_2) + \ldots + E(Z_{I-1}) ,$$

$$Var(Z) = Var(Z_2) + \ldots + Var(Z_{I-1})$$

and we can assume that Z approximately has a Normal distribution. This means that we reject (with an error probability of 5 %) the hypothesis of having no significant calendar year effects only if not

$$E(Z) - 2\cdot\sqrt{Var(Z)} \le Z \le E(Z) + 2\cdot\sqrt{Var(Z)} .$$

Application to the example of Chapter 6:

We start with the triangle of all development factors observed:

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9
i=1	1.6	1.32	1.08	1.15	1.20	1.11	1.033	1.00	1.01
i=2	40.4	1.26	1.98	1.29	1.13	0.99	1.043	1.03	
i=3	2.6	1.54	1.16	1.16	1.19	1.03	1.026		
i=4	2.0	1.36	1.35	1.10	1.11	1.04			
i=5	8.8	1.66	1.40	1.17	1.01				
i=6	4.3	1.82	1.11	1.23					
i=7	7.2	2.72	1.12						
i=8	5.1	1.89							
i=9	1.7								

We have to subdivide each column F_k into the subset SF_k of 'smaller' factors below the median of F_k and into the subset LF_k of 'larger' factors above the median. This can be done very easily with the help of the rank columns r_{ik} established in Appendix G: The half of factors with small ranks belongs to SF_k, those with large ranks to LF_k and if the total number is odd we have to eliminate the mean rank. Replacing a small rank with

'S', a large rank with 'L' and a mean rank with '*' we obtain the following picture:

	j=1	j=2	j=3	j=4	j=5	j=6	j=7	j=8	j=9
j=1	S	S	S	S	L	L	*	S	*
j=2	L	S	L	L	*	S	L	L	
j=3	S	S	*	S	L	S	S		
j=4	S	S	L	S	S	L			
j=5	L	L	L	L	S				
j=6	*	L	S	L					
j=7	L	L	S						
j=8	L	L							
j=9	S								

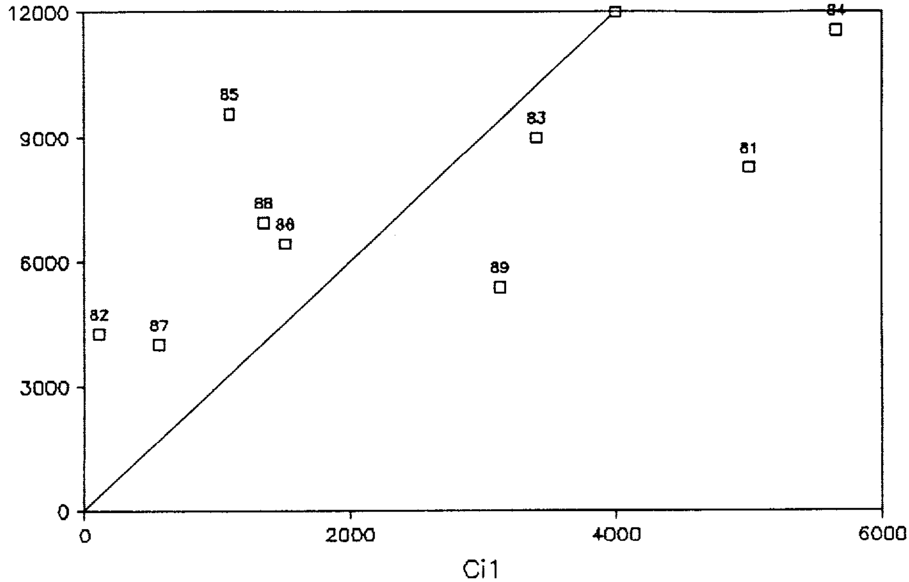
We now count for every diagonal A_j, $2 \le j \le 9$, the number L_j of L's and the number S_j of S's. With the notations $Z_j = \min(L_j, S_j)$, $n = S_j + L_j$, $m = [(n-1)/2]$ as above and using the formulae (H1), (H2) for $E(Z_j)$ and $Var(Z_j)$ we obtain the following table:

j	S_j	L_j	Z_j	n	m	$E(Z_j)$	$Var(Z_j)$
2	1	1	1	2	0	.5	.25
3	3	0	0	3	1	.75	.1875
4	3	1	1	4	1	1.25	.4375
5	1	3	1	4	1	1.25	.4375
6	1	3	1	4	1	1.25	.4375
7	2	4	2	6	2	2.0625	.6211
8	4	4	4	8	3	2.90625	.8037
9	4	4	4	8	3	2.90625	.8037
Total			14			12.875	$3.9785 = (1.9946)^2$

The test statistic $Z = \Sigma Z_j = 14$ is not outside its 95%-range $(12.875 - 2 \cdot 1.9946,\ 12.875 + 2 \cdot 1.9946) = (8.886,\ 16.864)$ and therefore the null-hypothesis of not having significant calendar year influences is not rejected so that we can continue to apply the chain ladder method.

Figure 1: Regression and Residuals

Ci2 against Ci1



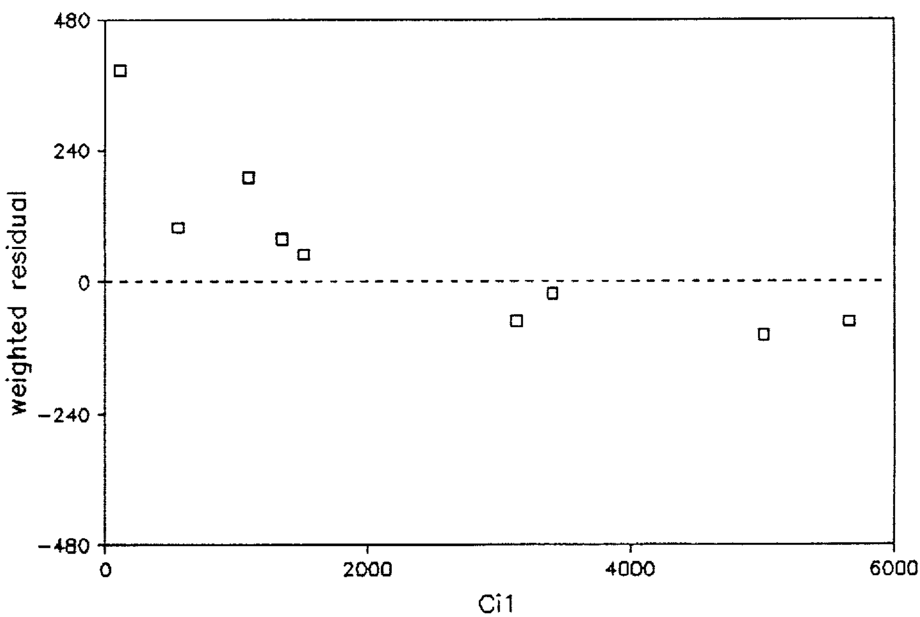
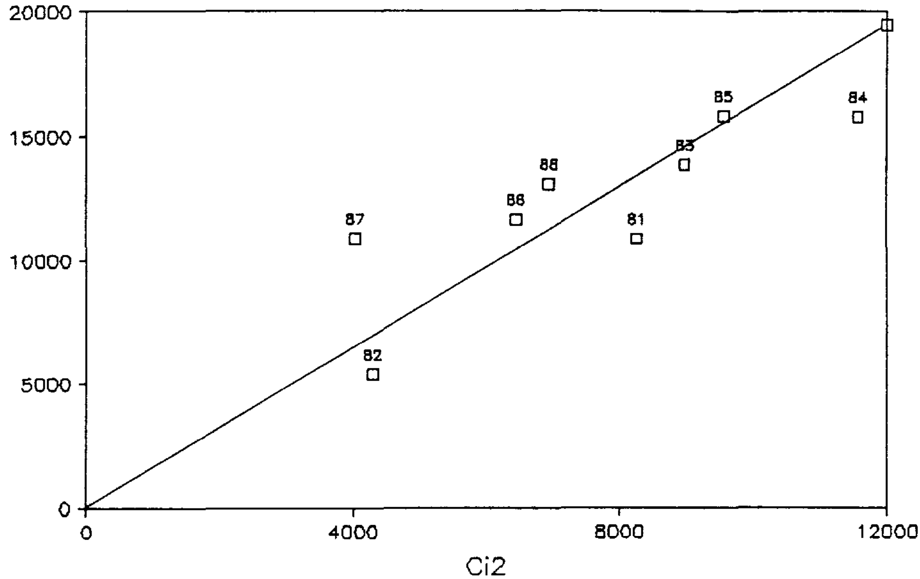


Figure 2: Regression and Residuals
Ci3 against Ci2



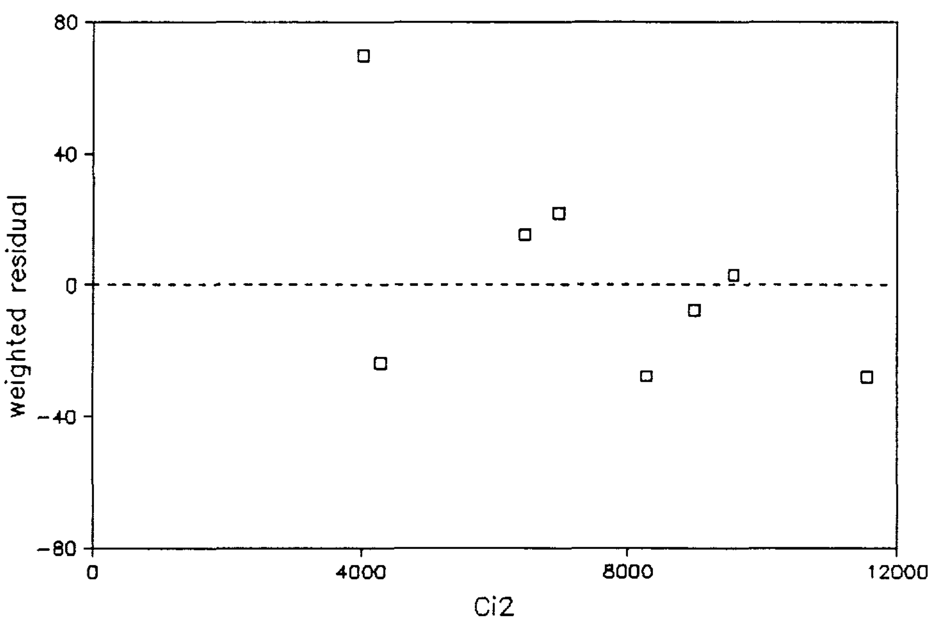


Figure 3: Regression and Residuals

Ci4 against Ci3

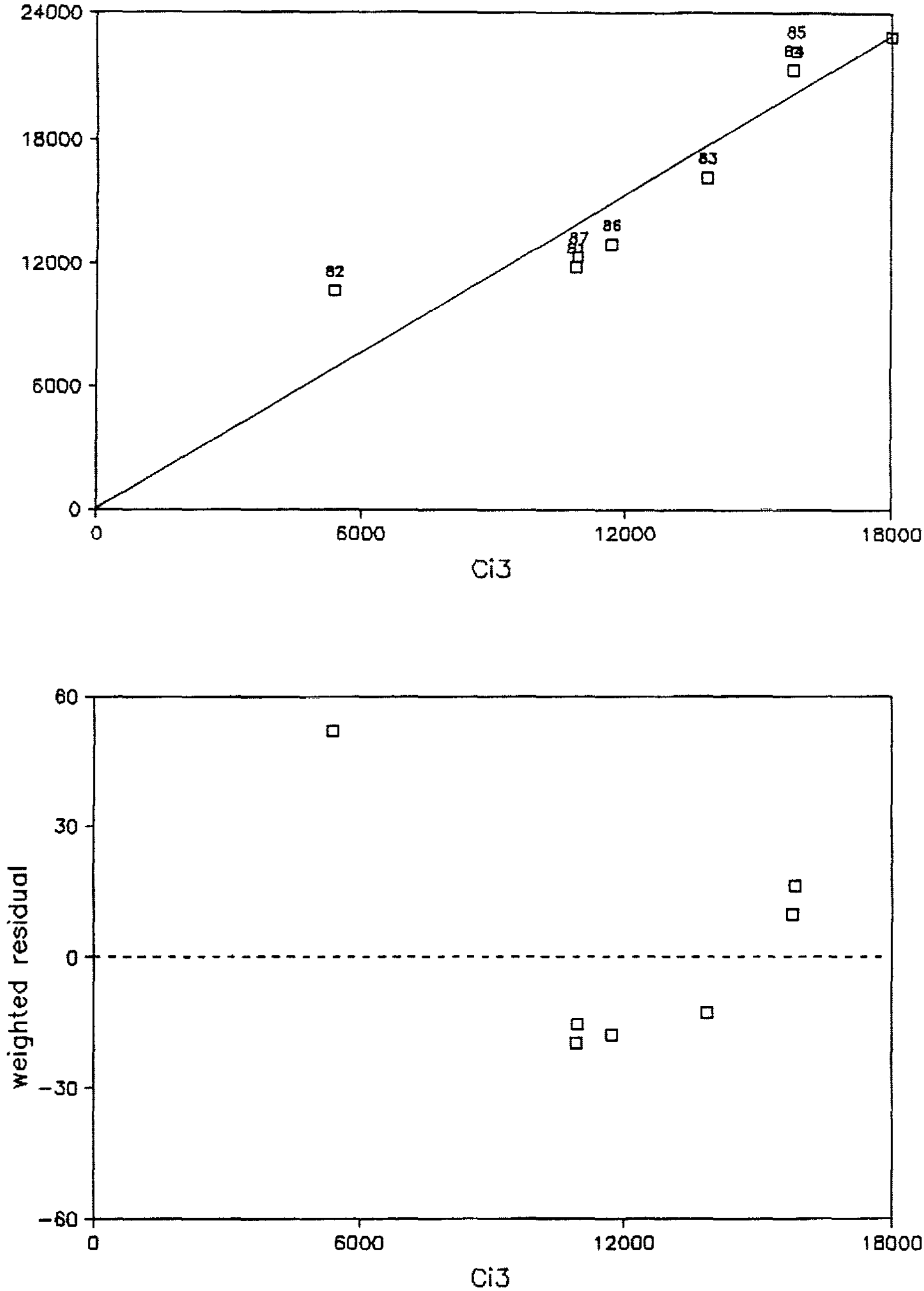
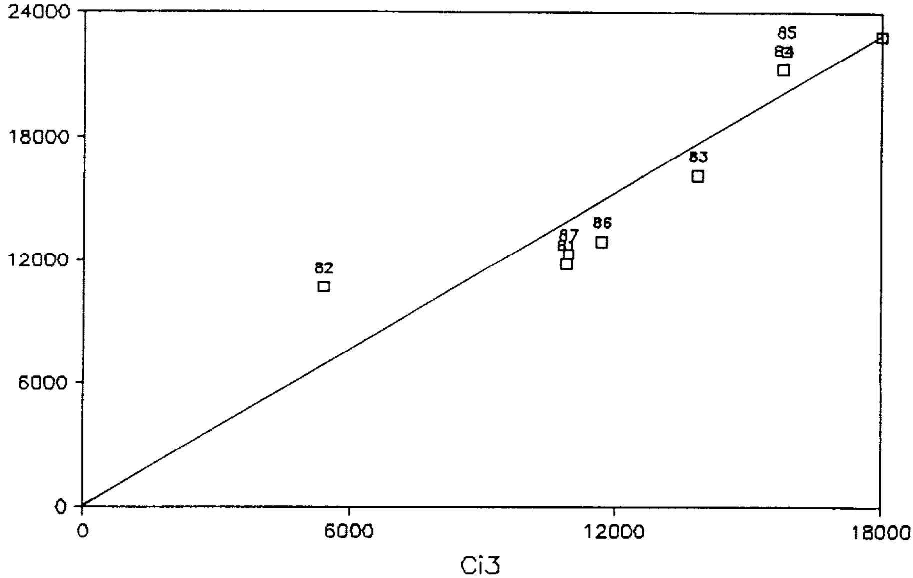


Figure 3: Regression and Residuals
Ci4 against Ci3



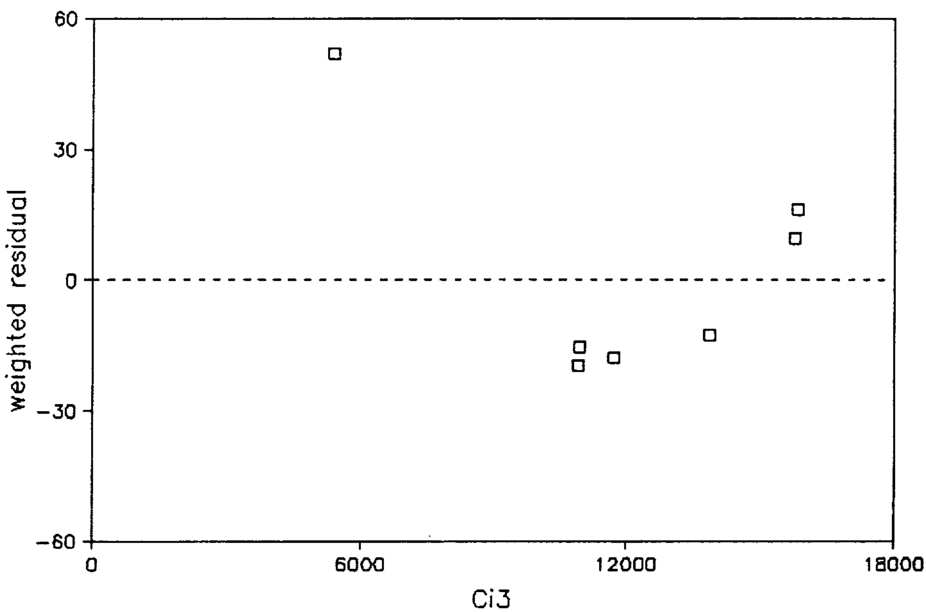


Figure 4: Regression and Residuals
Ci5 against Ci4

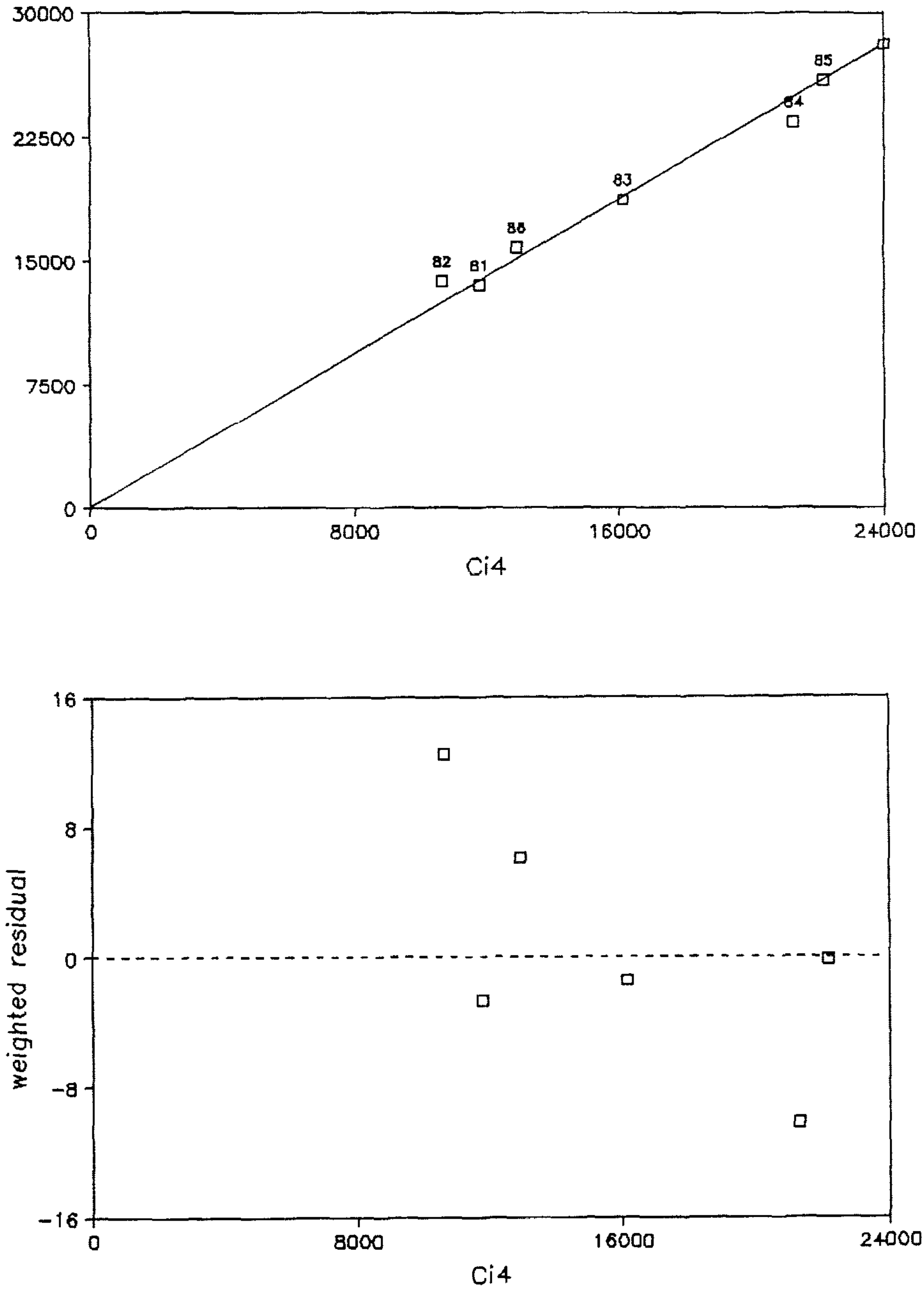


Figure 5: Regression and Residuals

Ci6 against Ci5

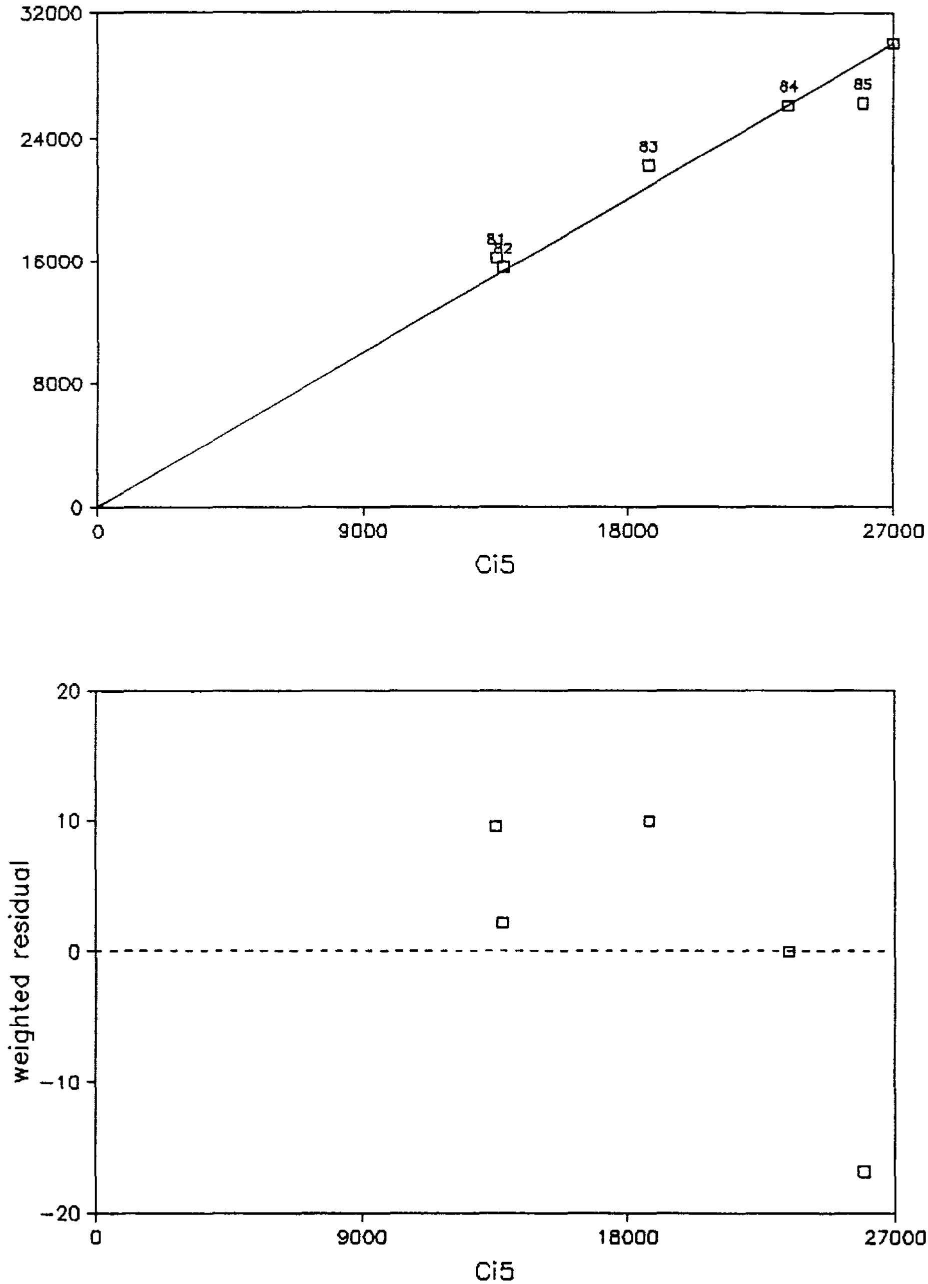


Figure 6: Regression and Residuals
Ci7 against Ci6

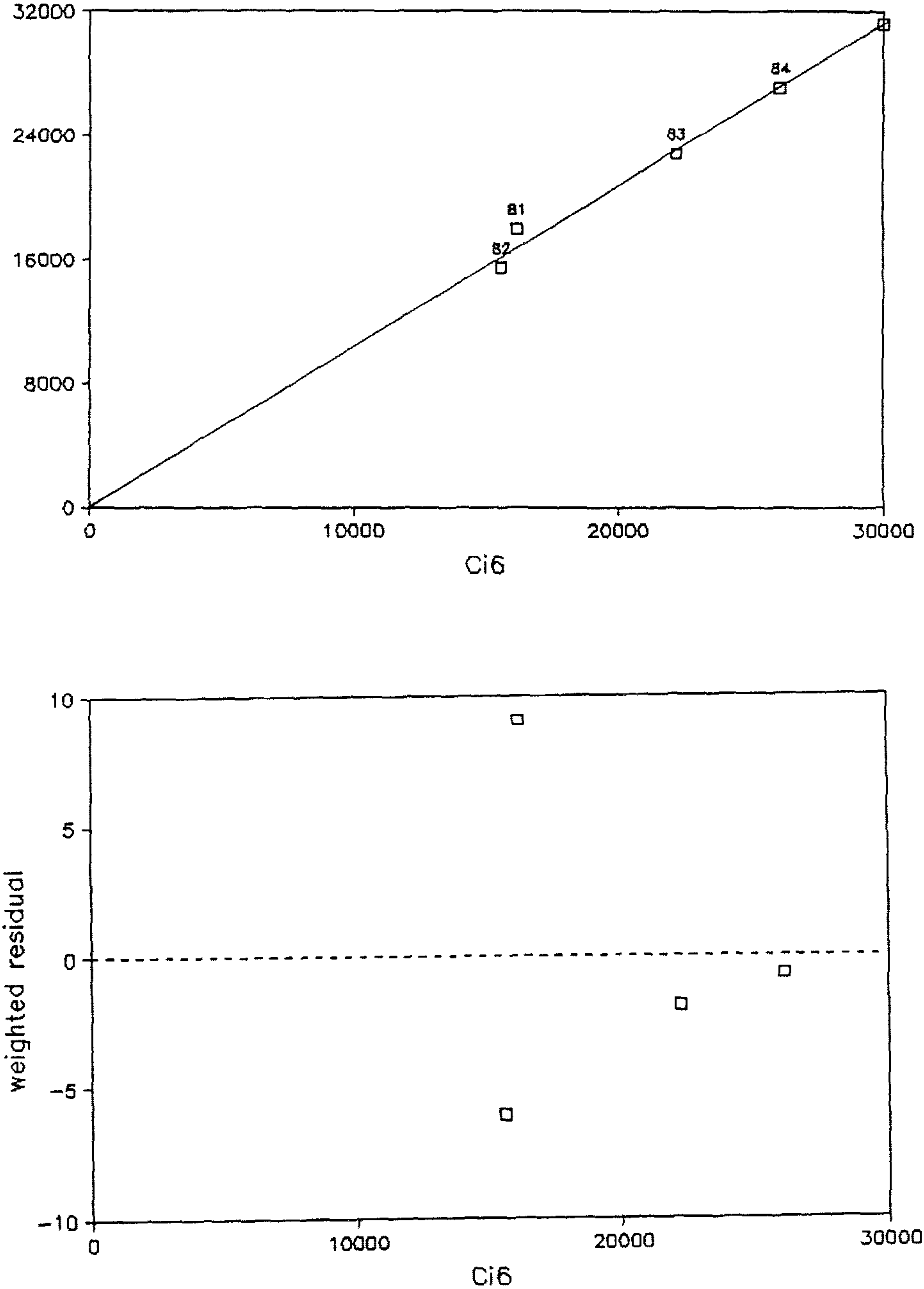


Figure 7: Regression and Residuals

Ci8 against Ci7

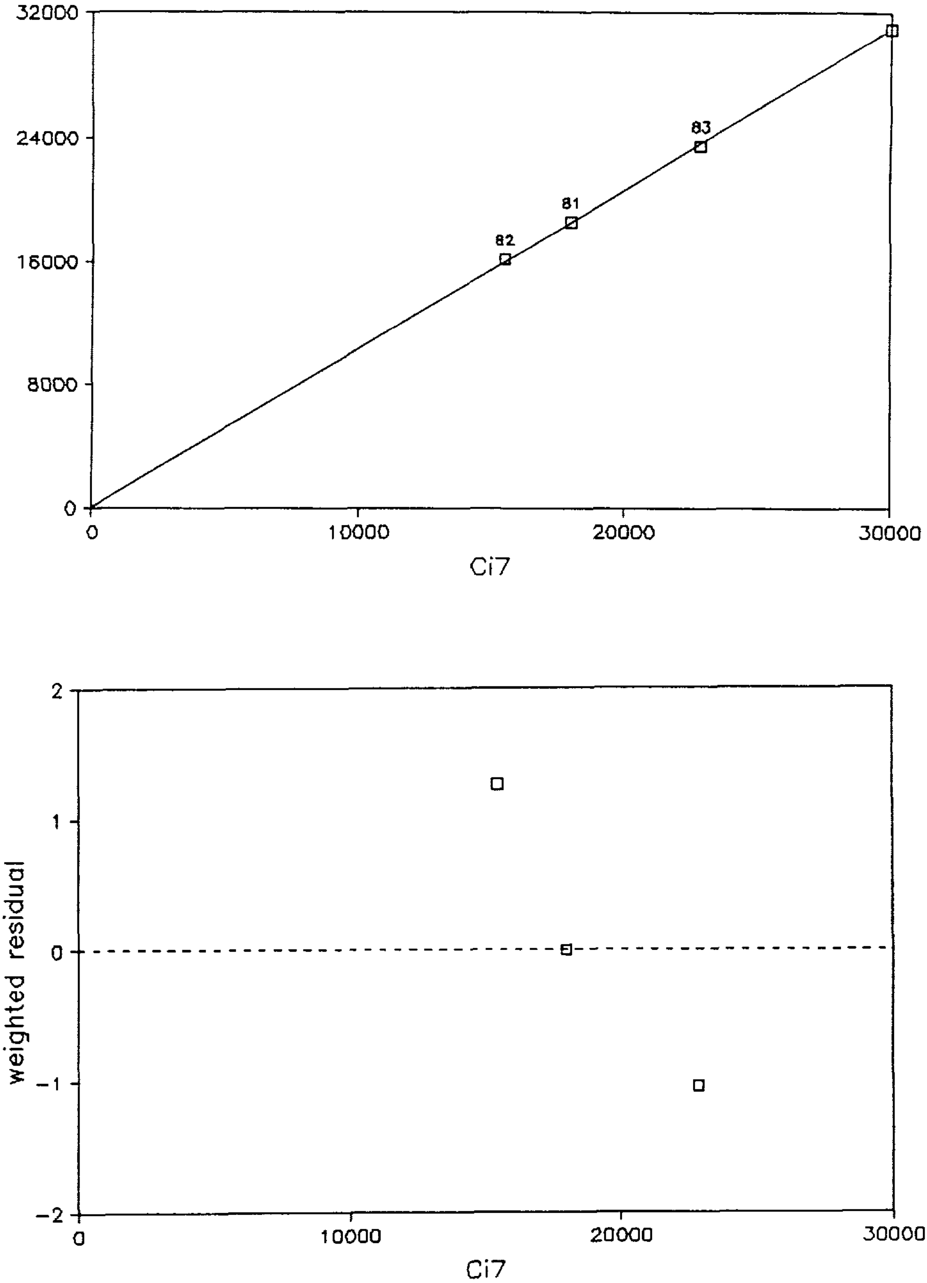
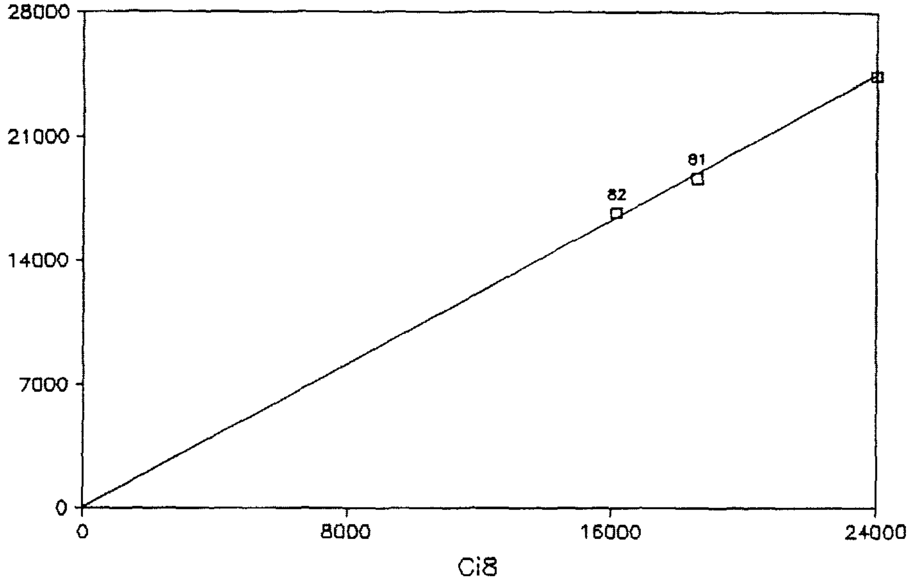


Figure 8: Regression and Residuals

Ci9 against Ci8



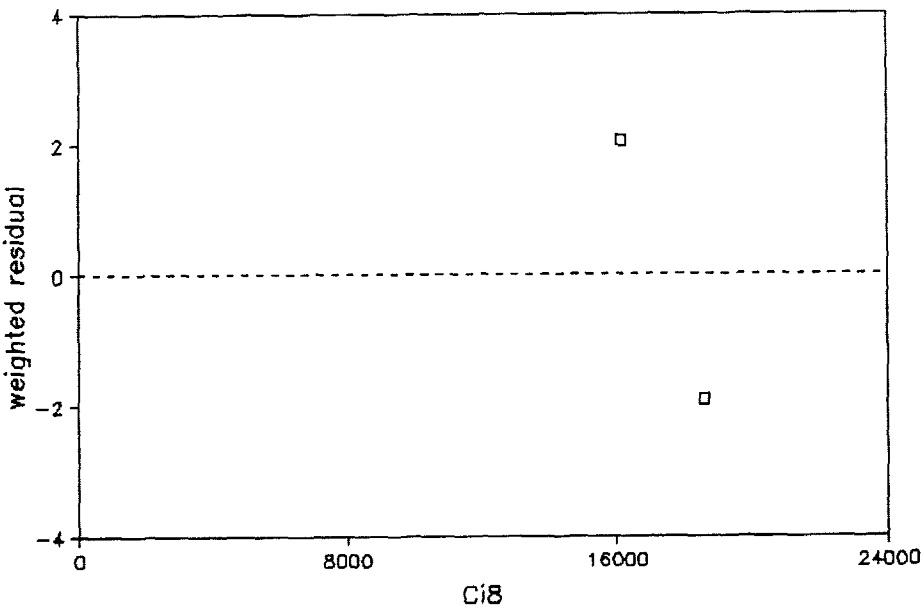
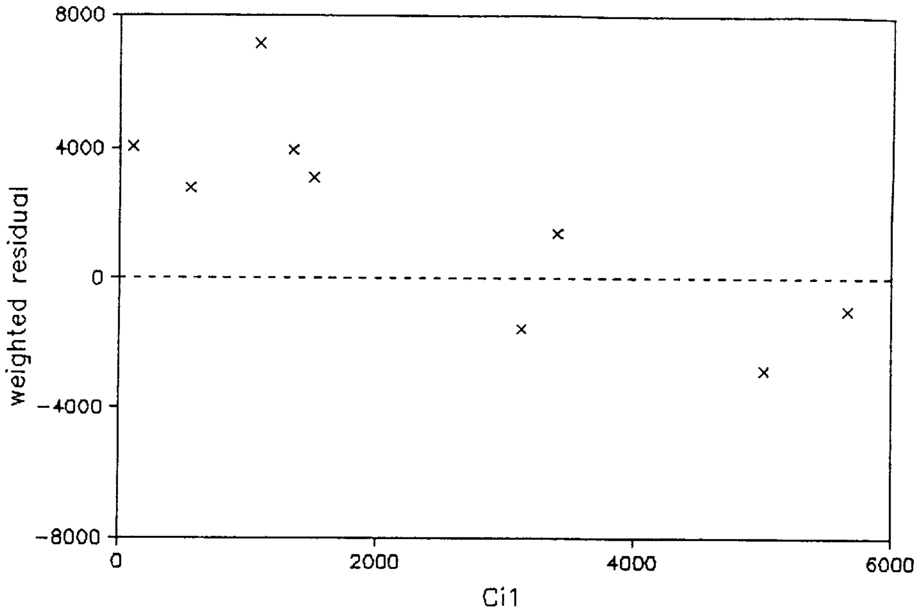


Figure 9: Residual Plots for fk0



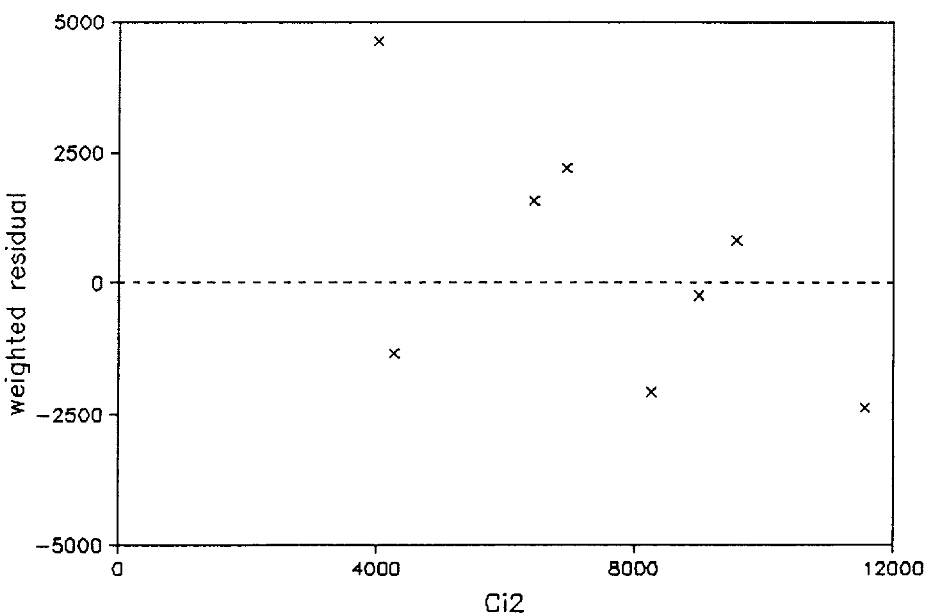


Figure 10: Residual Plots for fk0

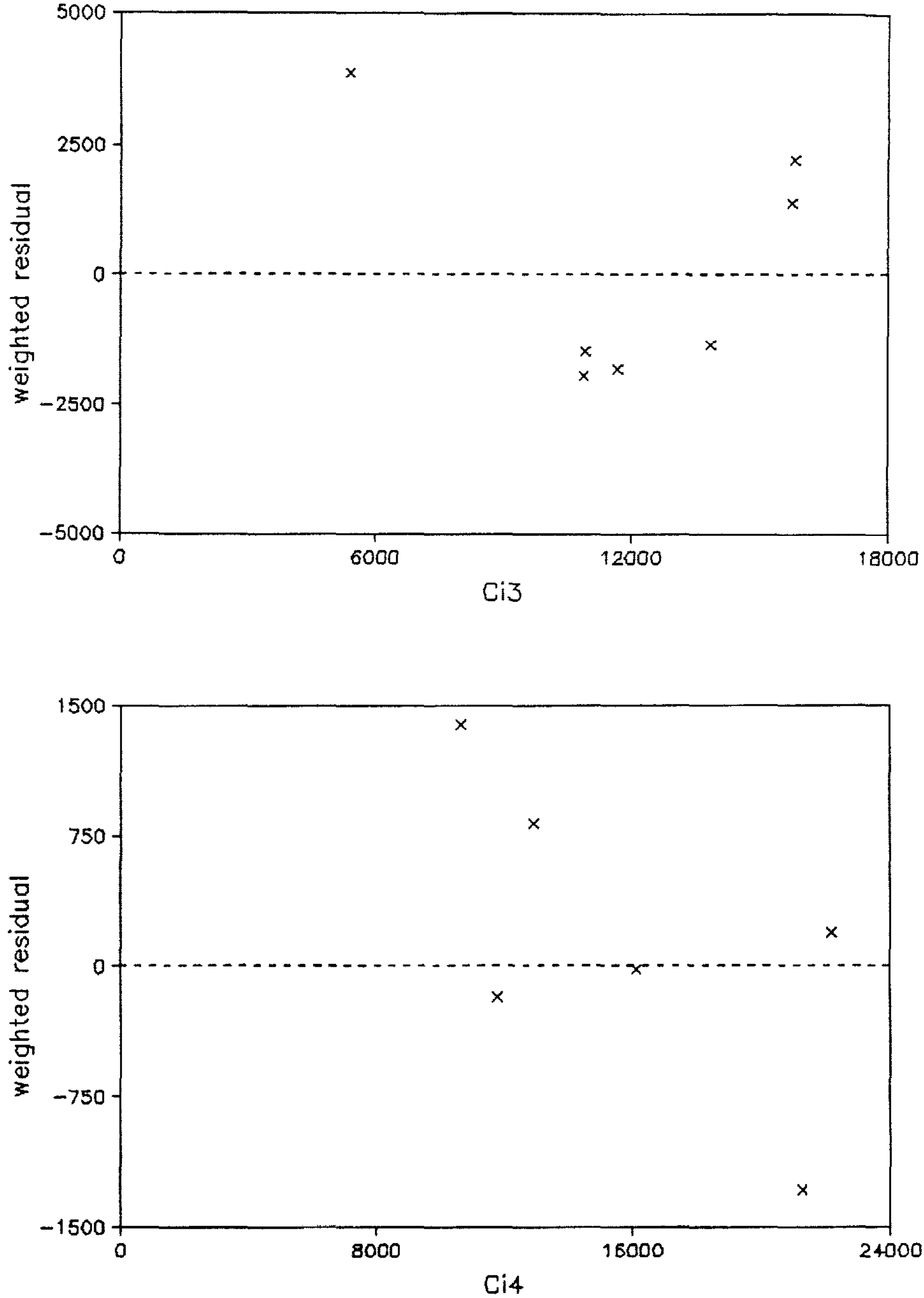


Figure 11: Residual Plots for fk2

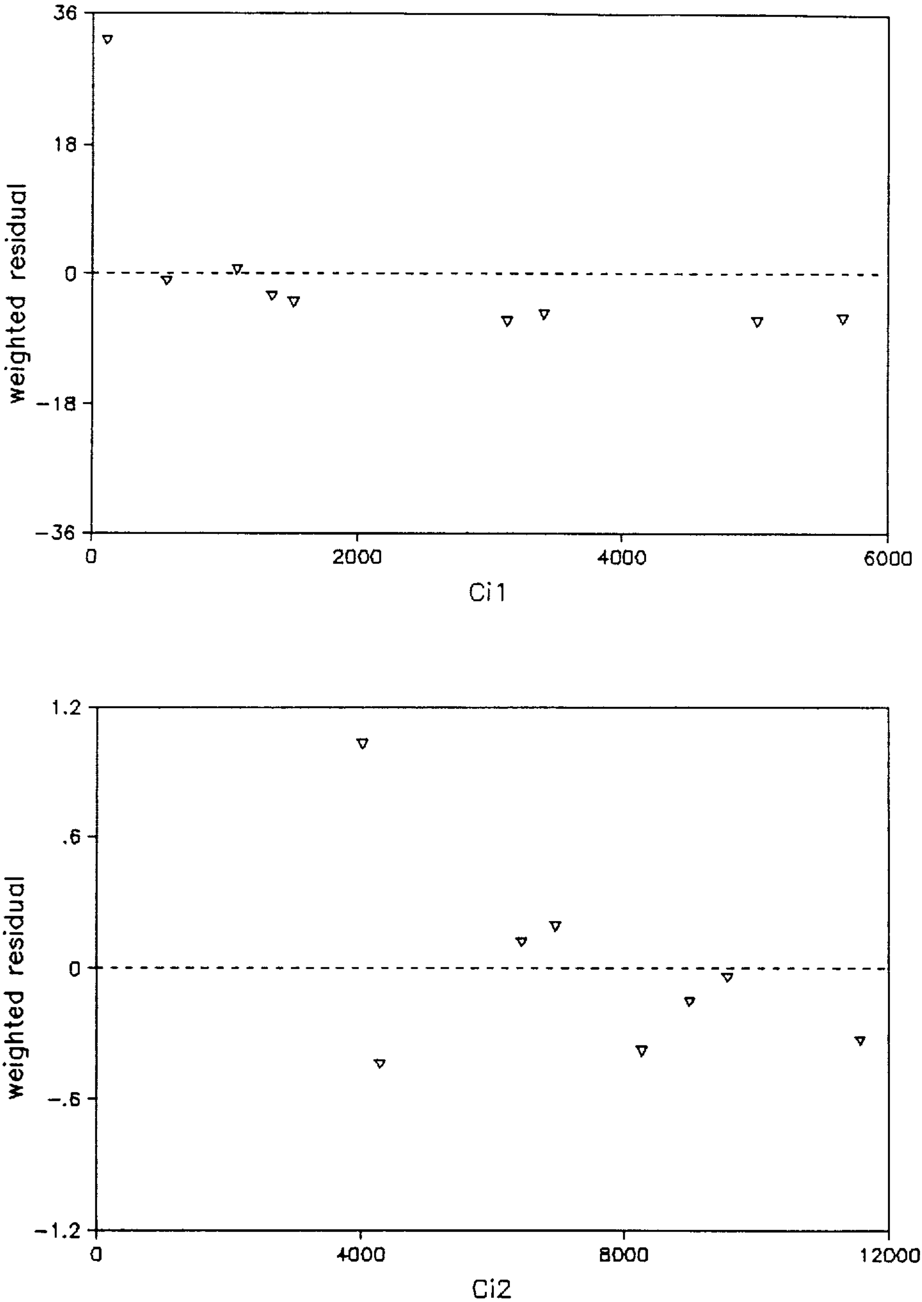
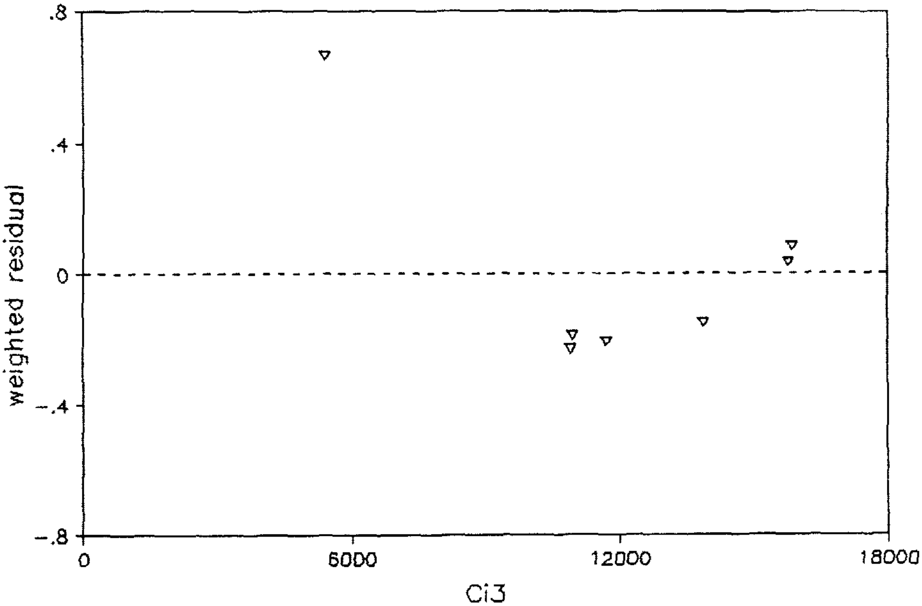


Figure 12: Residual Plots for fk2



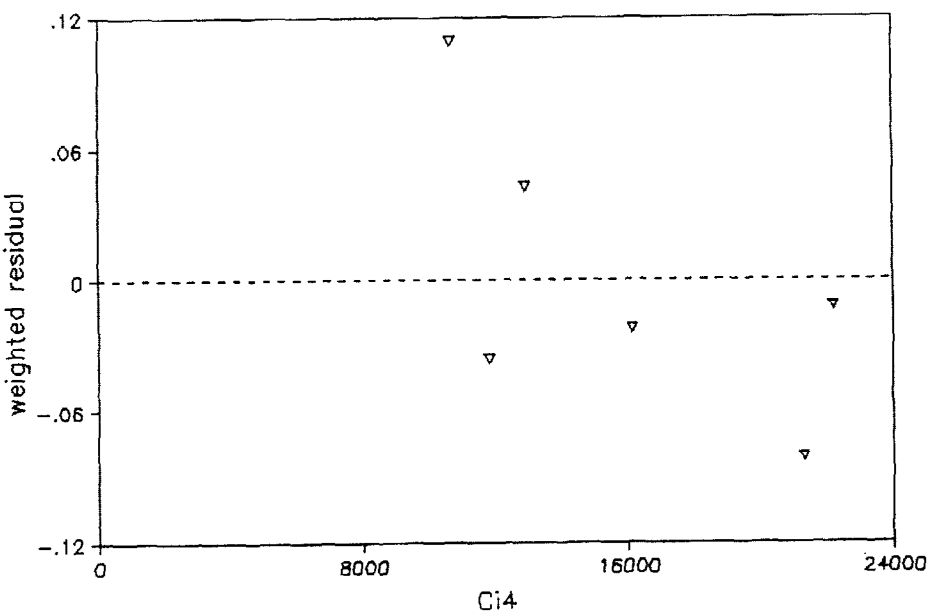
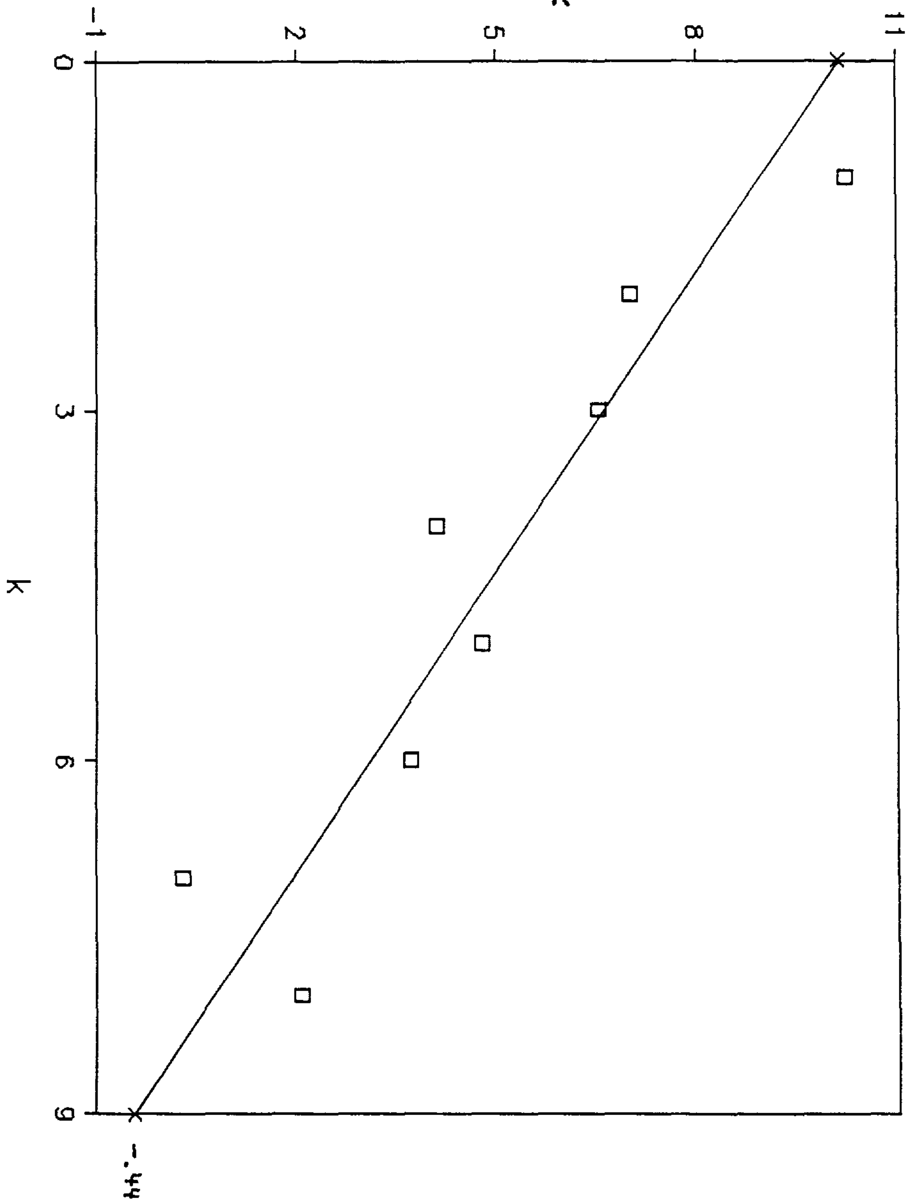


Figure 13: Plot of $\ln(\alpha_k^2)$ against k



CREDIBLE CLAIMS RESERVES: THE BENKTANDER METHOD

BY

THOMAS MACK

Munich Re, Munich

ABSTRACT

A claims reserving method is reviewed which was introduced by Gunnar Benktander in 1976. It is a very intuitive credibility mixture of Bornhuetter/Ferguson and Chain Ladder. In this paper, the mean squared errors of all 3 methods are calculated and compared on the basis of a very simple stochastic model. The Benktander method is found to have almost always a smaller mean squared error than the other two methods and to be almost as precise as an exact Bayesian procedure.

KEYWORDS

Claims Reserves, Chain Ladder, Bornhuetter/Ferguson, Credibility, Standard Error

1. INTRODUCTION

This note on the occasion of the 80st anniversary of Gunnar Benktander focusses on a claims reserving method which was published by him in 1976 in "The Actuarial Review" of the Casualty Actuarial Society (CAS) under the title "An Approach to Credibility in Calculating IBNR for Casualty Excess Reinsurance". The Actuarial Review is the quarterly newsletter of the CAS and is normally not subscribed outside of North America. This might be the reason why Gunnar's article did not become known in Europe. Therefore, the method has been proposed a second time by the Finnish actuary Esa Hovinen in his paper "Additive and Continuous IBNR", submitted to the ASTIN Colloquium 1981 in Loen/Norway. During that colloquium, Gunnar Benktander referred to his former article and Hovinen's paper was not published further. Therefore it was not unlikely that the method was invented a third time. Indeed, Walter Neuhaus published it in 1992 in the Scandinavian Actuarial Journal under the title "Another Pragmatic Loss Reserving Method or Bornhuetter/Ferguson Revisited". He mentioned neither Benktander nor Horvinen because he did not know about their articles. In recent years, the method has been used occasionally in actuarial reports under the name "Iterated Bornhuetter/Ferguson Method". The present article gives a short review of the method and connects it with the name of its first publisher. Furthermore, evidence is given that the method is very useful which should already be clear from the fact that it has been invented so many times. Using a simple stochastic model it is shown that the Benktander method outperformes the Bornhuetter/Ferguson method and the chain ladder method in many situations. Moreover, simple formulae for the mean squared error of all three methods are derived. Finally, a numerical example is given and a comparison with a credibility model and a Bayesian model is made.

2. Review of the method

To keep notation simple we concentrate on one single accident year and on paid claims. Furthermore, we assume the payout pattern to be given, i.e. we denote with p_j, $0 < p_1 < p_2 < \ldots < p_n = 1$, the proportion of the ultimate claims amount which is expected to be paid after j years of development. After n years of development, all claims are assumed to be paid. Let U_0 be the estimated ultimate claims amount, as it is expected prior to taking the own claims experience into account. For instance, U_0 can be taken from premium calculation. Then, being at the end of a fixed development year $k < n$,

$$R_{BF} = q_k U_0 \quad \text{with} \quad q_k = 1 - p_k$$

is the well-known Bornhuetter/Ferguson (BF) reserve (Bornhuetter/Ferguson 1972). The claims amount C_k paid up to now does not enter the formula for R_{BF}, i.e. this reserving method ignores completely the current claims experience of the portfolio under consideration. Note that the axiomatic relationship between any reserve estimate $\hat{R}$ and the corresponding ultimate claims estimate $\hat{U}$ is always

$$\hat{U} = C_k + \hat{R} \quad \text{and} \quad \hat{R} = \hat{U} - C_k$$

because the same relationship also holds for the true reserve $R = C_n - C_k$ and the corresponding ultimate claims $U = C_n$, i.e. we have

$$U = C_k + R \quad \text{and} \quad R = U - C_k.$$

For the Bornhuetter/Ferguson method this implies that the final estimate of the ultimate claims is the posterior estimate

$$U_{BF} = C_k + R_{BF}$$

whereas the prior estimate U_0 is only used to arrive at an estimate of the reserve. Note further that the payout pattern $\{p_j\}$ is defined by $p_j = E(C_j)/E(U)$.

Another well-known claims reserving method is the chain ladder (CL) method. This method grosses up the current claims amount C_k, i.e. uses

$$U_{CL} = C_k / p_k$$

as estimated ultimate claims amount and

$$R_{CL} = U_{CL} - C_k$$

as claims reserve. Note that there

$$R_{CL} = q_k U_{CL}$$

holds. This reserving method considers the current claims amount C_k to be fully credibly predictive for the future claims and ignores the prior expectation U_0 completely. One advantage of CL over BF is the fact that with CL different actuaries come always to similar results which is not the case with BF because there may be some dissent regarding U_0.

BF and CL represent extreme positions. Therefore Benktander (1976) proposed to replace the prior U_0 with a credibility mixture

$$U_c = cU_{CL} + (1 - c)U_0.$$

As the credibility factor c should increase similarly as the claims C_k develop, he proposed to take $c = p_k$ and to estimate the claims reserve by

$$R_{GB} = R_{BF} \cdot \frac{U_{p_k}}{U_0}.$$

This is the method as proposed by Gunnar Benktander (GB). Observe that we have

$$R_{GB} = q_k U_{p_k}$$

and

$$U_{p_k} = p_k U_{CL} + q_k U_0 = C_k + R_{BF} = U_{BF},$$

i.e.

$$R_{GB} = q_k U_{BF}.$$

This last equation means that the Benktander reserve R_{GB} is obtained by applying the BF procedure in an additional step to the posterior ultimate claims amount U_{BF} which was arrived at by the normal BF procedure. This way has been taken in some recent actuarial reports and has there been called "iterated Bornhuetter/Ferguson method".

Note again that the resulting posterior estimate

$$U_{GB} = C_k + R_{GB} = (1 - q_k^2)U_{CL} + q_k^2 U_0 = U_{1-q_k^2}$$

for the ultimate claims is different from U_{p_k} which was used as prior.

Esa Hovinen (1981) applied the credibility mixture directly to the reserves instead of the ultimates, i.e. proposed the reserve estimate

$$R_{EH} = cR_{CL} + (1 - c)R_{BF},$$

again with $c = p_k$. But the Hovinen reserve

$$R_{EH} = p_k q_k U_{CL} + (1 - p_k) q_k U_0 = q_k U_{p_k} = R_{GB}$$

is identical to the Benktander reserve.

We have already seen that the functions $R(U) = q_k U$ and $U(R) = C_k + R$ are not inverse to each other except for $U = U_{CL}$. In addition, Table 1 shows that the further iteration of the methods of BF and GB for an arbitrary starting point U_0 finally leads to the chain ladder method.

We want to state this as a theorem:

Theorem 1. For an arbitrary starting point $U^{(0)} = U_0$, the iteration rule

$$R^{(m)} = q_k U^{(m)} \quad \text{and} \quad U^{(m+1)} = C_k + R^{(m)}, \quad m = 0, 1, 2, \ldots,$$

gives credibility mixtures

$$U^{(m)} = (1 - q_k^m) U_{CL} + q_k^m U_0,$$
$$R^{(m)} = (1 - q_k^m) R_{CL} + q_k^m R_{BF}$$

between BF and CL which start at BF and lead via GB finally to CL for $m = \infty$.

TABLE 1

ITERATION OF BORNHUETTER/FERGUSON

Ultimate $U(R) = C_k + R$	*Connection*	*Reserve* $R(U) = q_k U$
U_0		
	$\searrow$	
		$R_{BF} = q_k U_0$
	$\swarrow$	
$U^{(1)} = U_{BF} = C_k + R_{BF}$		
$= (1 - q_k) U_{CL} + q_k U_0$	$\searrow$	
		$R^{(1)} = R_{GB} = q_k U_{BF}$
		$= (1 - q_k) R_{CL} + q_k R_{BF}$
$U^{(2)} = U_{GB} = C_k + R_{GB}$	$\swarrow$	
$= (1 - q_k^2) U_{CL} + q_k^2 U_0$		
.....	$\searrow$	
		.....
$U^{(m)} = (1 - q_k^m) U_{CL} + q_k^m U_0$		
	$\searrow$	
		$R^{(m)} = q_k U^{(m)}$
		$= (1 - q_k^m) R_{CL} + q_k^m R_{BF}$
$U^{(m+1)} = C_k + R^{(m)}$	$\swarrow$	
$= (1 - q_k^{m+1}) U_{CL} + q_k^{m+1} U_0$		
.....	$\searrow$	.....
$U^{(\infty)} = U_{CL}$	$\longleftrightarrow$	$R^{(\infty)} = R_{CL}$

Walter Neuhaus (1992) analyzed the situation in a full Bühlmann/Straub credibility framework (see section 6 for details) and compared the size of the mean squared error $mse(R_c) = E(R_c - R)^2$ of

$$R_c = cR_{CL} + (1 - c)R_{BF}$$

and the true reserve $R = U - C_k = C_n - C_k$ especially for

$$\begin{array}{ll} c = 0 & (BF) \\ c = p_k & \text{(GB, called PC-predictor by Neuhaus)} \\ c = c^* & \text{(optimal credibility reserve),} \end{array}$$

where $c^* \in [0; 1]$ can be defined to be that c which minimizes $mse(R_c)$. Neuhaus did not include $c = 1$ (CL) explicitely into his analysis.

Neuhaus showed that the mean squared error of the Benktander reserve R_{GB} is almost as small as of the optimal credibility reserve R_{c^*} except if p_k is small and c^* is large at the same time (cf. Figures 1 and 2 in Neuhaus (1992)). Moreover, he showed that the Benktander reserve R_{GB} has a smaller mean squared error than R_{BF} whenever $c^* > p_k/2$ holds. This result is very plausible because then c^* is closer to $c = p_k$ than to $c = 0$.

In the following we include the CL into the analysis and consider the case where U_0 is not necessarily equal to $E(U)$, i.e. consider the estimation error, too. This seems to be more realistic as in Neuhaus (1992) where $U_0 = E(U)$ was assumed. Instead of the credibility model used by Neuhaus, we introduce a less demanding stochastic model in order to compare the precision of R_{BF}, R_{CL} and R_{GB}. We derive a formula for the standard error of R_{BF} and R_{GB} (and R_{CL}) and show how the parameters required can be estimated. A numerical example is given in section 4. Moreover, there is a close connection to a paper by Gogol (1993) which will be dealt with in section 5. Finally, the connection to the credibility model is analyzed in section 6.

3. CALCULATION OF THE OPTIMAL CREDIBILITY FACTOR c^* AND OF THE MEAN SQUARED ERROR OF R_c

In order to compare R_{BF}, R_{CL} and R_{GB}, we use the mean squared error

$$mse(R_c) = E(R_c - R)^2$$

as criterion for the precision of the reserve estimate R_c (for a discussion see section 5). Because

$$R_c = cR_{CL} + (1 - c)R_{BF} = c(R_{CL} - R_{BF}) + R_{BF}$$

is linear in c, the mean squared error $mse(R_c)$ is a quadratic function of c and will therefore have a minimum.

In the following, we consider U_0 to be an estimation function which is independent from C_k, R, U and has expectation $E(U_0) = E(U)$ and variance $Var(U_0)$. Then we have

Theorem 2. The optimal credibility factor c^* which minimizes the mean squared error $mse(R_c) = E(R_c - R)^2$ is given by

$$c^* = \frac{p_k}{q_k} \cdot \frac{Cov(C_k, R) + p_k q_k Var(U_0)}{Var(C_k) + p_k^2 Var(U_0)}. \tag{1}$$

Proof

$$E(R_c - R)^2 = E[c(R_{CL} - R_{BF}) + R_{BF} - R]^2$$
$$= c^2 E(R_{CL} - R_{BF})^2 - 2cE[(R_{CL} - R_{BF})(R - R_{BF})] + E(R_{BF} - R)^2.$$

$$0 = \frac{\partial}{\partial c} E(R_c - R)^2 = 2cE(R_{CL} - R_{BF})^2 - 2E[(R_{CL} - R_{BF})(R - R_{BF})]$$

yields

$$c^* = \frac{E[(R_{CL} - R_{BF})(R - R_{BF})]}{E(R_{CL} - R_{BF})^2} = \frac{p_k}{q_k} \cdot \frac{E[(C_k - p_k U_0)(R - q_k U_0)]}{E(C_k - p_k U_0)^2}$$
$$= \frac{p_k}{q_k} \cdot \frac{Cov(C_k - p_k U_0,\ R - q_k U_0)}{Var(C_k - p_k U_0)} = \frac{p_k}{q_k} \cdot \frac{Cov(C_k, R) + p_k q_k Var(U_0)}{Var(C_k) + p_k^2 Var(U_0)}.$$

Here, we have used that $E(C_k) = p_k E(U_0)$ according to the definition of the payout pattern (and therefore $E(R) = q_k E(U_0)$). Q.E.D.

In order to estimate c^*, we need a model for $Var(C_k)$ and $Cov(C_k, R)$. The following model is not more than a slightly refined definition of the payout pattern:

$$E(C_k/U|U) = p_k, \tag{2}$$
$$Var(C_k/U|U) = p_k q_k \beta^2(U). \tag{3}$$

The factor q_k in (3) is necessary in order to secure that $Var(C_k|U) \to 0$ as k approaches n. A similar argument holds for p_k in case of very small values. A parametric example is obtained if the ratio C_k/U, given U, has a Beta(ap_k, aq_k)-distribution with $a > 0$; in this case $\beta^2(U) = (a+1)^{-1}$. Thus, in the simple cases, $\beta^2(U)$ depends neither on U nor on k. If the variability of C_k/U for high values of U is higher, then $\beta^2(U) = (U/U_0) \cdot \beta^2$ is a reasonable assumption.

From assumptions (2) and (3) and with $\alpha^2(U) := U^2\beta^2(U)$ we gather

$$E(C_k|U) = p_k U,$$
$$Var(C_k|U) = p_k q_k \alpha^2(U),$$
$$E(C_k) = p_k E(U),$$
$$Var(C_k) = p_k q_k E(\alpha^2(U)) + p_k^2 Var(U)$$
$$= p_k E(\alpha^2(U)) + p_k^2 (Var(U) - E(\alpha^2(U))), \tag{4}$$

$$
\begin{aligned}
Cov(C_k, U) &= Cov(E(C_k|U), U) = p_k Var(U),\\
Cov(C_k, R) &= Cov(C_k, U) - Var(C_k) = p_k q_k\big(Var(U) - E(\alpha^2(U))\big), \qquad (5)\\
E(R) &= E(U) - E(C_k) = q_k E(U),\\
Var(R) &= Var(U) - 2Cov(C_k, U) + Var(C_k)\\
&= Var(U)(1 - 2p_k + p_k^2) + p_k q_k E(\alpha^2(U))\\
&= q_k^2 Var(U) + p_k q_k E(\alpha^2(U))\\
&= q_k E(\alpha^2(U)) + q_k^2\big(Var(U) - E(\alpha^2(U))\big).
\end{aligned}
$$

By inserting (4) and (5) into (1), we immediately obtain

Theorem 3. Under the assumptions of model (2)-(3), the optimal credibility factor c^* which minimizes $mse(R_c)$ is given by

$$c^* = \frac{p_k}{p_k + t} \quad \text{with} \quad t = \frac{E(\alpha^2(U))}{Var(U_0) + Var(U) - E(\alpha^2(U))}. \qquad (6)$$

Some further straightforward calculations lead to

Theorem 4. Under the assumptions of model (2)-(3), we have the following formulae for the mean squared error:

$$
\begin{aligned}
mse(R_{BF}) &= E(\alpha^2(U))q_k(1 + q_k/t),\\
mse(R_{CL}) &= E(\alpha^2(U))q_k/p_k,\\
mse(R_c) &= E(\alpha^2(U))\left(\frac{c^2}{p_k} + \frac{1}{q_k} + \frac{(1-c)^2}{t}\right)q_k^2.
\end{aligned}
$$

Proof

$$
\begin{aligned}
mse(R_{BF}) &= E(R_{BF} - R)^2 = Var(R_{BF} - R) = Var(R_{BF}) + Var(R)\\
&= q_k^2 Var(U_0) + q_k^2\big(Var(U) - E(\alpha^2(U))\big) + q_k E(\alpha^2(U))\\
&= E(\alpha^2(U))\big(q_k + q_k^2/t\big),\\
mse(R_{CL}) &= E(R_{CL} - R)^2 = Var(R_{CL} - R)\\
&= Var(R_{CL}) - 2Cov(R_{CL}, R) + Var(R)\\
&= q_k^2 Var(C_k)/p_k^2 - 2q_k Cov(C_k, R)/p_k + Var(R)\\
&= E(\alpha^2(U))q_k/p_k,\\
mse(R_c) &= E(cR_{CL} + (1-c)R_{BF} - R)^2\\
&= E[c(R_{CL} - R) + (1-c)(R_{BF} - R)]^2\\
&= c^2 mse(R_{CL}) + 2c(1-c)E[(R_{CL} - R)(R_{BF} - R)] + (1-c)^2 mse(R_{BF}),
\end{aligned}
$$

$$\begin{aligned} E[(R_{CL}-R)(R_{BF}-R)] &= Cov(R_{CL}-R,\ R_{BF}-R) \\ &= -Cov(R_{CL},R)+Var(R) \\ &= Var(R)-q_k Cov(C_k,R)/p_k \\ &= q_k E\big(\alpha^2(U)\big). \end{aligned}$$

and putting all pieces together leads to the formula stated. Q.E.D.

An actuary who is able to assess $p_k = E(C_k/U|U)$ and U_0 (i.e. $E(U_0)$) should also be able to estimate $Var(U_0)$ and $Var(C_k/U|U)$ or $E(Var(C_k|U))$ as well as $Var(U)$. Therefrom, he can deduce $E\big(\alpha^2(U)\big) = E(Var(C_k|U))/(p_k q_k)$ – or $E\big(\alpha^2(U)\big) = Var(C_k/U|U)E(U^2)/(p_k q_k)$ if $Var(C_k/U|U)$ does not depend on U – and finally the parameter t. Then he has now a formula for the mean squared error of the BF method and a very simple formula for the CL method (where t is not needed) and can calculate the best estimate R_{c^*} including its mean squared error as well as the one of R_{GB}.

Regarding the very simple formula for $mse(R_{CL})$ we should note that this formula deviates from the corresponding one (i.e. for the unconditional mean squared error with known payout pattern) of the distribution-free chain ladder model of Mack (1993). The reason is that the models underlying are slightly different: Here we have

$$E\left(\frac{C_k}{U}|U\right) = p_k$$

and the model of Mack (1993) can be written as

$$E\left(\frac{U}{C_k}|C_k\right) = \frac{1}{p_k}.$$

Using theorem 4, we now compare the mean squared errors of the different methods in terms of p_k and t. First, we have

$$mse(R_{BF}) < mse(R_{CL}) \iff p_k < t,$$

i.e. we should use BF for the green years ($p_k < t$) and CL for the rather mature years ($p_k > t$). This is very plausible and the author is aware that some companies use this rule with $t = 0.5$. But the volatility measure t varies from one business to the other and therefore the actuary should try to estimate t in every single case as is shown in the next section.

Furthermore, we have

$$\begin{aligned} mse(R_{GB}) < mse(R_{BF}) &\iff t < 2 - p_k, \\ mse(R_{GB}) < mse(R_{CL}) &\iff t > p_k q_k/(1+p_k), \end{aligned}$$

i.e. GB is better than BF except t is very large and is better than CL except t is very small, see Figure 1 where for each of the three areas it is indicated which of BF, GB, CL is best. In the numerical example below, it will become clear that t is almost always in the GB area.

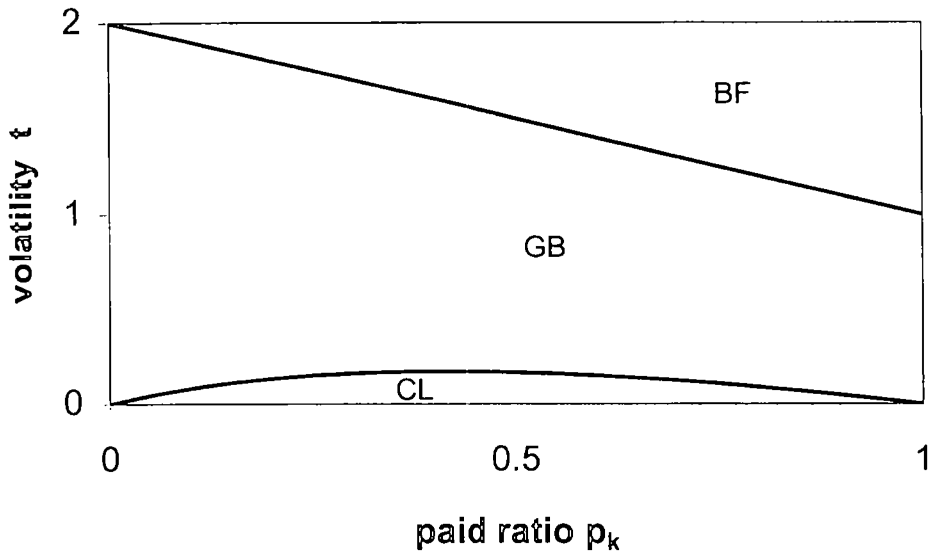


FIGURE 1: Areas of smallest mean squared error.

4. NUMERICAL EXAMPLE

Assume that the a priori expected ultimate claims ratio is 90% of the premium, i.e. $U_0 = 90\%$. Assuming further $p_k = 0.50$ for $k = 3$, we have $R_{BF} = 45\%$ (all % ages relate to the premium). Let the paid claims ratio be $C_k = 55\%$, then $U_{CL} = 110\%$ and $R_{CL} = 55\%$. Taken together, we have $R_{GB} = 50\%$.

In order to calculate the standard errors, we have to assess $Var(U)$, $Var(U_0)$ and $E(\alpha^2(U))$. For $Var(U)$, we can use a consideration of the following type: We assume that the ultimate claims ratio will never be below 60% and only once every 20 years above 150%. Then, assuming a shifted lognormal distribution with expectation 90%, we get $Var(U) = (35\%)^2$. This rather high variance is typical for a reinsurance business or a small direct portfolio.

Regarding $E(\alpha^2(U))$, we consider here the special case where $\beta^2(U) = \beta^2$ does not depend on U (e.g. using a Beta distribution), i.e. $E(\alpha^2(U)) = E(U^2)\beta^2 = E(U^2)Var(C_k/U|U)/(p_k q_k)$. Therefore, we have to assess $Var(C_k/U|U)$, i.e. the variability of the payment ratio C_k/U around its mean p_k. If we assume – e.g. by looking at the ratios C_k/U of past accident years – that C_k/U will be almost always between 0.30 and 0.70, then – using the two-sigma rule from the normal distribution – we have a standard deviation of 0.10, i.e. $Var(C_k/U|U) = 0.10^2$, which leads to $\beta^2 = Var(C_k/U|U)/(p_k q_k) = 0.20^2$ and $E(\alpha^2(U)) = E(U^2)\beta^2 = 0.193^2$.

Finally, the most difficult task is to assess $Var(U_0)$ but this has much less influence on t than $Var(U)$ (which is always larger) and $E(\alpha^2(U))$. Moreover, an actuary who is able to establish a point estimate U_0 should also be able to estimate the uncertainty $Var(U_0)$ of his point estimate. Thus, there will be a certain interval or range of values where the actuary takes his choice of U_0 from. Then he can take this interval and use the two-sigma rule to produce the standard deviation $\sqrt{Var(U_0)}$. Let us assume that in our example $Var(U_0) = (15\%)^2$.

Now we can calculate $t = 0.346$ and all standard errors (= square root of the estimated mean squared error) as well as the optimal credibility reserve R_{c^*}:

$$\begin{aligned} R_{BF} &= 45\% \pm 21.3\% \\ R_{CL} &= 55\% \pm 19.3\% \\ R_{GB} &= 50\% \pm 17.3\% \\ R_{c^*} &= 50.9\% \pm 17.2\% \qquad \text{with} \quad c^* = 0.591. \end{aligned}$$

Note that these standard errors are based on the *un*conditional mean squared error (cf. discussion in the next section) and on a known pattern $\{p_j\}$. Including the uncertainty of the p_j will increase the standard error.

For the purpose of comparison, we look at a more stable business, too: Assume that $Var(U) = (10\%)^2$, $Var(U_0) = (5\%)^2$ and $Var(C_k/U|U) = (0.03)^2$. Then, everything else being equal, we obtain $\beta^2 = 0.06^2$, $E(\alpha^2(U)) = 0.054^2$, $t = 0.309$ and

$$\begin{aligned} R_{BF} &= 45\% \pm 6.2\% \\ R_{CL} &= 55\% \pm 5.4\% \\ R_{GB} &= 50\% \pm 4.9\% \\ R_{c^*} &= 51.2\% \pm 4.9\% \qquad \text{with} \quad c^* = 0.618. \end{aligned}$$

In both cases, GB has a smaller mean squared error than BF and CL, and the size of t has not changed much, because the relative sizes of the three variances $Var(U)$, $Var(U_0)$, $Var(C_k/U|U)$ are similar. A closer look at formula (6) shows that the size of t is changed more if $E(\alpha^2(U))$ (i.e. $Var(C_k/U|U)$) is changed than if $Var(U)$ or $Var(U_0)$ are changed. In the first example, for instance, we had $Var(C_k/U|U) = 0.10^2$ and GB was better than CL and BF. If we change the variability of the paid ratio to $Var(C_k/U|U) \geq 0.153^2$, then $t \geq 1.51$ and BF is better than GB and CL. If we change it to $Var(C_k/U|U) \leq 0.074^2$, then $t \leq 0.164$ and CL is better than GB and BF, see Figure 1. But in the large range of normal values of $Var(C_k/U|U)$, GB is better than CL and BF. Because $Var(U_0)$ is always smaller than $Var(U)$, the size of t is essentially determined by the ratio $Var(C_k/U|U)/Var(U)$.

5. Application of an exact Bayesian model to the numerical example

If we make distributional assumptions for U and $C_k|U$, we can determine the exact distribution of $U|C_k$ according to Bayes' theorem. This was done by Gogol (1993) who assumed that U and $C_k|U$ have lognormal distributions because then $U|C_k$ has a lognormal distribution, too.

Applied to our first numerical example, this model is:

$$U \sim \text{Lognormal } (\mu, \sigma^2) \quad \text{with} \quad E(U) = 90\%,\ Var(U) = (35\%)^2,$$
$$C_k|U \sim \text{Lognormal } (\nu, \tau^2) \quad \text{with} \quad E(C_k|U) = p_k U,\ Var(C_k|U) = p_k q_k \beta^2 U^2$$

where $\beta^2 = 0.20^2$ is as before, i.e. such that $Var(C_k/U|U) = 0.10^2$.
This yields

$$\sigma^2 = \ln(1 + Var(U)/(E(U))^2) = 0.375^2,$$
$$\mu = \ln(E(U)) - \sigma^2/2 = -0.176,$$
$$\tau^2 = \ln\left(1 + \beta^2 q_k/p_k\right) = 0.198^2.$$

Then (see Gogol (1993)),

$$U|C_k \sim \text{Lognormal } (\mu_1, \sigma_1^2)$$

with

$$\mu_1 = z\left(\tau^2 + \ln(C_k/p_k)\right) + (1 - z)\mu = 0.067,$$
$$\sigma_1^2 = z\tau^2 = 0.175^2,$$
$$z = \sigma^2/(\sigma^2 + \tau^2) = 0.782.$$

This yields (at $C_k = 55\%$)

$$E(U|C_k) = \exp(\mu_1 + \sigma_1^2/2) = 108.6\%,$$
$$E(R|C_k) = E(U|C_k) - C_k = 53.6\%,$$
$$Var(R|C_k) = Var(U|C_k) = (E(U|C_k))^2\left(\exp(\sigma_1^2) - 1\right) = (19.2\%)^2.$$

If we compare this last result with the mean squared errors obtained in section 4, we should recall that $E(R|C_k)$ minimizes the *conditional* mean squared error

$$E\left((\hat{R} - R)^2|C_k\right) = Var(R|C_k) + \left(\hat{R} - E(R|C_k)\right)^2$$

among all estimators $\hat{R}$ which are a square integrable function of C_k as well as it minimizes the *un*conditional mean squared error

$$E\left(\hat{R} - R\right)^2 = E(Var(R|C_k)) + E\left(\hat{R} - E(R|C_k)\right)^2$$

because the first term of the r.h.s. does not depend on $\hat{R}$. But the resulting minimum values $Var(R|C_k)$ and $E(Var(R|C_k))$ are different.

Basically, in claims reserving we should minimize the *conditional* mean squared error, given C_k, because we are only interested in the future variability and because C_k remains a fixed part of the ultimate claims U. But if $E(R|C_k)$ is a linear function of C_k (like R_c), this function can be found by minimizing the unconditional (average) mean squared error. Moreover, the latter can often be calculated easier than the conditional mean squared error as it is the case in model (2)-(3). The unconditional mean squared error is the appropriate measure to compare the precision of different reserving methods.

Altogether, it is clear that the mean squared errors calculated in section 4 are average (unconditional) mean squared errors, averaged over all possible values of C_k. Therefore, in order to make a fair comparison of the various methods in our numerical example, we must calculate the unconditional mean squared error $E(Var(R|C_k))$ in the Bayesian model, too.

For this purpose, we have to integrate $Var(R|C_k)$ over C_k and therefore need the distribution of C_k which we obtain by mixing the distributions of $C_k|U$ and U:

$$C_k/p_k \sim \text{Lognormal}\ (\mu - \tau^2/2,\ \sigma^2 + \tau^2),$$
$$\exp(2z\ \ln(C_k/p_k)) \sim \text{Lognormal}\ \left(2z\mu - z\tau^2,\ 4z^2(\sigma^2 + \tau^2)\right).$$

This yields

$$\begin{aligned} E(Var(R|C_k)) &= E\left(\exp\left(2\mu_1 + \sigma_1^2\right)\left(\exp(\sigma_1^2) - 1\right)\right) \\ &= E(\exp(2z\ \ln(C_k/p_k)))\exp\left(3z\tau^2 + 2(1-z)\mu\right)\left(\exp(z\tau^2) - 1\right) \\ &= \exp\left(2\mu + 2\sigma^2\right)\left(\exp(z\tau^2) - 1\right) \\ &= (17.0\%)^2. \end{aligned}$$

This shows finally, that the exact Bayesian model on average has only a slightly smaller mean squared error than the optimal credibility reserve R_{c^*} and the Benktander reserve R_{GB}. But if we recall that, with the exact Bayesian procedure, we assume to exactly know the distributional laws without any estimation error, then the slight improvement in the mean squared error does not pay for the strong assumptions made.

6. Connection to the credibility model

Finally, we establish an interesting connection between the model (2)-(3) and the credibility model used in Neuhaus (1992). There, the Bühlmann/Straub credibility model was applied to the incremental losses and payouts: For $j = 1, \ldots, n$ (where n is such that $p_n = 1$) let

$$m_j = p_j - p_{j-1}$$

be the incremental payout pattern and

$$S_j = C_j - C_{j-1}$$

be the incremental claims (with the convention $p_0 = 0$ and $C_0 = 0$). Then the Bühlmann/Straub credibility model makes the following assumptions:

$$S_1|\Theta, \ldots, S_n|\Theta \quad \text{are independent,} \tag{7}$$

$$E(S_j/m_j|\Theta) = \mu(\Theta), \qquad 1 \le j \le n, \tag{8}$$

$$Var(S_j/m_j|\Theta) = \sigma^2(\Theta)/m_j \qquad 1 \le j \le n, \tag{9}$$

where Θ is the unknown distribution quality of the accident year. Assumption (7) can be crucial in practise. Model (7)-(9) can be set up without refering to p_j by just requiring $m_j > 0$ and $m_1 + \ldots + m_n = 1$. Then the following formulae still hold using $p_k := m_1 + \ldots + m_k$.

From (7)-(9) we obtain

$$E(C_k|\Theta) = p_k\mu(\Theta),$$
$$Var(C_k|\Theta) = p_k\sigma^2(\Theta).$$

The latter formula shows, that the credibility model is different from model (2)-(3) where we have $Var(C_k|U) = p_k q_k \alpha^2(U)$, i.e. we do not have $\Theta = U$.

In the credibility model (7)-(9) we obtain further

$$\begin{aligned}
E(C_k) &= p_k E(\mu(\Theta)) = p_k E(C_n) = p_k E(U),\\
Var(C_k) &= p_k E\big(\sigma^2(\Theta)\big) + p_k^2 Var(\mu(\Theta)),\\
Cov(C_k, U) &= E(Cov(C_k,\ C_k|\Theta)) + Cov(p_k\mu(\Theta),\ \mu(\Theta))\\
&= p_k\big(E\big(\sigma^2(\Theta)\big) + Var(\mu(\Theta))\big),\\
Cov(C_k, R) &= p_k q_k Var(\mu(\Theta)),\\
E(R) &= q_k E(\mu(\Theta)) = q_k E(U),\\
Var(R) &= q_k E\big(\sigma^2(\Theta)\big) + q_k^2 Var(\mu(\Theta)).
\end{aligned} \tag{10}$$

If we compare these formulae with the corresponding formulae of model (2)-(3) and take into account that here

$$Var(\mu(\Theta)) = Var(U) - E\big(\sigma^2(\Theta)\big)$$

holds (from (10) with $k = n$), then we see that these formulae are completely identical if $E\big(\alpha^2(U)\big) = E\big(\sigma^2(\Theta)\big)$. This leads immediately to

Theorem 5. The formulae of theorems 3 and 4 hold for model (7)-(9), too, after having replaced $E\big(\alpha^2(U)\big)$ with $E\big(\sigma^2(\Theta)\big)$.

In the credibility model, a natural estimate of $E(\sigma^2(\Theta))$ can be established: From

$$Var(S_j/m_j|\Theta) = \sigma^2(\Theta)/m_j$$

and

$$\sum_{j=1}^{k} m_j \frac{S_j}{m_j} \Big/ \sum_{j=1}^{k} m_j = C_k/p_k = U_{CL}$$

it follows that

$$\sigma^2 = \frac{1}{k-1}\sum_{j=1}^{k} m_j\left(\frac{S_j}{m_j} - U_{CL}\right)^2$$

is an unbiased estimator of $E(\sigma^2(\Theta))$. We can write

$$\sigma^2 = p_k s^2/(k-1)$$

where

$$s^2 = \sum_{j=1}^{k} m_j\left(\frac{S_j}{m_j} - U_{CL}\right)^2 \Big/ \sum_{j=1}^{k} m_j$$

can be calculated easily as the m_j-weighted average of the squared deviations of the observed ratios S_j/m_j from their weighted mean U_{CL}. Note that each S_j/m_j is an unbiased estimate of the expected ultimate claims $E(U)$.

If in our numerical example in addition to $p_3 = 0.50$ and $C_3 = 55\%$ we have $p_1 = 0.10$, $p_2 = 0.30$, $C_1 = 15\%$, $C_2 = 27\%$, then $m_1 = 0.10$, $m_2 = 0.20$, $m_3 = 0.20$, $S_1 = 15\%$, $S_2 = 12\%$, $S_3 = 28\%$, and the ratios $S_1/m_1 = 1.5$, $S_2/m_2 = 0.6$, $S_3/m_3 = 1.4$ have a variance $s^2 = 0.41^2$. Then the estimate for $E(\sigma^2(\Theta))$ is $\sigma^2 = 0.205^2$. With $C_1 = 10\%$ and $C_2 = 30\%$ we would get $\sigma^2 = 0.061^2$ indicating a more stable case.

Note that for the estimation of $E(\alpha^2(U))$ the observation of several accident years is necessary. Anyhow, model (2)-(3) is less demanding than model (7)-(9).

7. Conclusion

In claims reserving, the actuary has usually two independent estimators R_{BF} and R_{CL}, at his disposal: One is based on prior knowledge (U_0), the other is based on the claims already paid (C_k). It is a well-known lemma of Statistics that from several independent and unbiased estimators one can form a better estimator (i.e. with smaller variance) by putting them together via a linear combination. From this general perspective, too, it is clear that the *GB* reserve should be superior to *BF* or *CL*.

More precisely, the foregoing analysis has shown that *GB* has a smaller mean squared error than *BF* and *CL* if the payout pattern is neither extremely volatile nor extremely stable. This conclusion is derived within a model whose assumptions are nothing more than a precise definition of the term 'payout pattern'. Therefore, actuaries should include the Benktander method in their standard reserving methods.

Finally, we want to emphasize that all formulae derived rely on the assumption that the prior estimate U_0 and the observed claims C_k are independent. This means that these formulae probably will not hold any more for a 'prior' U_0 which has been adjusted during the development period as it is often done in practise. Such an adjustment is like choosing an U_c with an unknown c and gives a procedure which is much less objective than the Benktander method.

ACKNOWLEDGEMENT

This paper has benefitted from the discussions at and after the RESTIN meeting 1999, especially with Ole Hesselager.

REFERENCES

BENKTANDER, G. (1976) An Approach to Credibility in Calculating IBNR for Casualty Excess Reinsurance. In *The Actuarial Review*, April 1976, p. 7.

BORNHUETTER, R.L. and FERGUSON, R.E. (1972) The Actuary and IBNR. In *Proceedings of the Casualty Actuarial Society*, Vol. LIX, 181-195.

GOGOL, D. (1993) Using Expected Loss Ratios in Reserving. In *Insurance: Mathematics and Economics* **12**, 297-299.

HOVINEN, E. (1981) Additive and Continuous IBNR. *ASTIN Colloquium Loen/Norway*.

MACK, Th. (1993) Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin* **23** (1993), 213-225.

NEUHAUS, W. (1992) Another Pragmatic Loss Reserving Method or Bornhuetter/Ferguson Revisited. In *Scand. Actuarial J.* 1992, 151-162.

DR. THOMAS MACK
Munich Re
80791 München
Germany
E-mail: tmack@munichre.com

Correction Note

to the paper
„Credible Claims Reserves: The Benktander Method“
by Thomas Mack

In Chapter 5 („Application …“), there is a mistake.
The equation for μ_1 should be as follows:

$$\mu_1 = z\,(\,\tau^2/2 + \ln(C_k/p_k)\,) + (1 - z)\,\mu = 0.05155\;,$$

i. e. $\tau^2/2$ instead of τ^2 and a slightly different numerical result. This mistake entails the following further alterations later on in the same chapter:

$E(U|C_k) = \ldots = 106.9\%$ (instead of 108.6%),

$E(R|C_k) = \ldots = 51.9\%$ (instead of 53.6%),

$Var(R|C_k) = \ldots = \ldots = (18.9\%)^2$ (instead of 19.2%).

Finally, the last equations of Chapter 5 change as follows:

$$\begin{aligned} E(Var(R|C_k)) &= E(\,\exp(2\mu_1 + \sigma_1^2)\,(\exp(\sigma_1^2) - 1)\,) \\ &= E(\exp(2z\,\ln(C_k/p_k)))\,\exp(2z\tau^2 + 2(1-z)\mu)\,(\exp(z\tau^2) - 1) \\ &= \exp(2\mu + (1+z)\sigma^2)\,(\exp(z\tau^2) - 1) \\ &= (16.8\%)^2\;. \end{aligned}$$

(i. e. $2z\tau^2$ instead of $3z\tau^2$ in the second line, $(1+z)\sigma^2$ instead of $2\sigma^2$ in the third line and 16.8% instead of 17.0% in the forth line.) This concludes the list of corrections.

A Framework for Assessing Risk Margins

Prepared by the Risk Margins Taskforce
(Karl Marshall, Scott Collings, Matt Hodson & Conor O'Dowd)

This is the final version of a draft paper that was presented to the Institute of Actuaries of Australia 16th General Insurance Seminar 9-12 November 2008, Coolum, Australia. The changes between this and the draft are minimal and reflect our view that the fundamental principles and techniques discussed in the draft remain appropriate.

Contents

Abstract

The main purpose of this paper is to propose a comprehensive framework for assessing insurance liability risk margins and to provide practical advice on how to implement it. The key sources of uncertainty are examined and the main quantitative approaches to analysing uncertainty discussed, including commentary on the advantages and disadvantages of each approach. The framework recognises, however, that quantitative analysis of historical data cannot alone capture adequately all aspects of future uncertainty. There will always be a need for judgement to be applied and in many situations such considerations will dominate the risk margin assessment. The application of judgement, however, is arguably the most difficult aspect of any attempt to estimate future uncertainty and assess appropriate risk margins. Our paper examines the key judgmental aspects and introduces a structured approach to combining these qualitative considerations with the results of any available quantitative analysis.

Keywords: framework, risk margins, uncertainty, APRA, independent risk, systemic risk.

1. Introduction

1.1. Preamble

General Insurance actuaries in Australia have, for many years, been analysing the uncertainty involved in the claim process with a view to assessing appropriate risk margins for inclusion in insurance liabilities. The approaches adopted to date range from those that involve little analysis of the underlying claim portfolio to those that involve significant analysis of the uncertainty using a wide range of information and techniques, including stochastic modelling.

The Risk Margins Taskforce was created to provide GI actuaries in Australia with support and guidance in the assessment of risk margins. In particular, it was felt that actuaries would benefit greatly from a stronger awareness of the key considerations when analysing uncertainty and the tools at their disposal when undertaking such analysis. A better equipped actuarial profession could feel more confident that key stakeholders, including APRA, insurance company boards, senior management and auditors, better understand the nature of and feel more comfortable with the quality and consistency of actuarial advice in this area.

The main purpose of this paper is to propose a comprehensive framework for assessing insurance liability risk margins and to provide practical advice on how to implement it. The key sources of uncertainty are examined and a combination of quantitative and qualitative approaches to their measurement explored.

1.2. Current approaches to assessing risk margins

In preparation for a presentation to the 2006 Reserving Seminar of the Institute of Actuaries of Australia (IAAust), the Taskforce canvassed a number of actuaries and APRA to gain a better understanding of the range of approaches used in Australia to assess risk margins. This information was supplemented with feedback from the *2006 General Insurance Claims Reserving and Risk Margins Survey*, the results of which were presented at the same seminar.

Although there appear to be a wide range of approaches used by Australian actuaries in the assessment of risk margins it is fair to say that most of the differences relate to the analysis and investigations conducted to parameterise a generally adopted risk margin calculation methodology, rather than the calculation methodology itself. The calculation methodology can be generalised as follows:

- *Coefficients of variation (CoVs)* are determined for individual valuation portfolios or groupings of portfolios, where these groupings include insurance classes made up of relatively homogeneous risks.
- A *correlation matrix* is populated with assumed *correlation coefficients* reflecting the expected correlations between valuation portfolios or groupings of portfolios.
- CoVs and correlation matrices are determined separately for outstanding claim liabilities and premium liabilities and further assumptions made about the correlation between these two components of the insurance liabilities.
- A statistical distribution is selected and combined with the adopted CoVs and correlation coefficients to determine the aggregate risk margin at a particular probability of adequacy.

The approaches used to determine CoVs vary significantly. The least sophisticated approaches involve deriving CoVs using either or both of two papers, *Research and Data Analysis Relevant to the Development of Standards and Guidelines on Liability Valuation for General Insurance* by Bateup and Reed (the Tillinghast paper) and *APRA Risk Margin Analysis* by Collings and White (the Trowbridge paper), both prepared at the end of 2001 (collectively these papers are referred to as the 2001 papers). These approaches often ignore the individual characteristics of the valuation portfolio for which risk margins are being assessed, deferring instead to the characteristics of the portfolios analysed by the authors of the two papers.

More sophisticated approaches include some form of quantitative analysis (stochastic or otherwise) supplemented by a qualitative assessment of the sources of uncertainty not captured by quantitative techniques. One such approach is discussed in the paper, *A Framework for Estimating Uncertainty in Insurance Claims Cost* by O'Dowd, Smith and Hardy, prepared for the IAAust's XVth General Insurance Seminar which was held in October 2005 (the PwC paper).

Anyone who has read the PwC paper will appreciate the similarities between the framework proposed in that paper to the framework discussed in this paper. **The Taskforce is collectively of the view that the PwC paper has significant merit and the concepts advocated by the authors of that paper have played a prominent role in the development of the framework discussed in this paper. We would encourage readers of this paper to read the PwC paper to ensure a more complete understanding of some of the concepts discussed.**

The most common approach to populating the correlation matrix with correlation coefficients is via the deployment of actuarial judgement. Usually the key risks that are considered to cause valuation portfolios to be correlated are considered in turn and the correlation between classes categorised as high, medium or low with each category having associated correlation coefficient values. The techniques deployed in the assessment of correlations range from those that are quite basic and heavily influenced by the benchmark correlation matrices discussed in the 2001 papers to those that take a more methodical approach to analysing the contribution to correlation from each key risk.

It is more the exception than the norm to include a quantitative analysis of past experience in the assessment of correlation effects. The main reason for this is that most quantitative techniques require a significant amount of data, time and cost to produce results that are sufficiently credible and intuitively justifiable. It is more common to see such techniques deployed when assessing more extreme probabilities of adequacy, i.e. well in excess of 90%, rather than probabilities of adequacy around the 75% level.

Generally, the most common distribution adopted to determine the aggregate risk margin at a particular probability of adequacy is the LogNormal distribution. The Normal distribution is also used by some actuaries, particularly at lower probabilities of adequacy where it can generate a risk margin that is higher than a heavier tailed distribution, such as the LogNormal distribution. It is uncommon for actuaries to test the adopted distribution against past experience or, taking a step further, derive a distributional form that explains the shape of the distribution of future claim cost outcomes based on past experience and/or future expectations.

The general risk margins approach adopted by most actuaries is often referred to as a *bolt-on* approach in that separate analyses are conducted to estimate the central estimate of insurance liabilities and the risk margins. The term bolt-on is also generally used to refer to any approach that does not involve the development of a single unified distribution of the entire distribution of possible future claim cost outcomes.

Judgement pervades both the central estimate assessment process and the risk margin assessment process. Also, well fitting models are those that adequately reflect past sources of uncertainty only. For these reasons, it is impossible to develop a purely quantitative model, fitted to the past data, that accurately represents the range of possible future claim cost outcomes. Rather, an approach that advocates internal consistency between the assessment of the central estimate and the sources of future uncertainty around that central estimate is important. The framework discussed in this paper is one such approach. This transparent framework combines quantitative and qualitative analysis, both of which are conducted giving full consideration to the central estimate assessment.

1.3. Practical framework for assessing risk margins

A number of key stakeholders, including Appointed Actuaries, APRA and auditors, have expressed some concern that the wide range of approaches adopted in practice to assess risk margins might lead to significant inconsistencies in the final outcomes, whether those be for regulatory or financial reporting purposes. Actuaries working in this area have also asked for guidance to help them when they are faced with analysing uncertainty. Finally, APRA have indicated that they would like to see more documentary justification of the risk margins adopted by some insurance companies.

With all of this in mind, we have prepared this paper to provide a comprehensive framework for assessing insurance liability risk margins and to provide practical advice on how to use this framework. There are a number of parts to our framework including the provision of guidance and further information on the tools, both quantitative and qualitative, that an actuary may deploy when analysing the uncertainty associated with insurance liabilities. We have included or referred to practical examples of how to deploy parts of the framework.

The proposed framework recognises that quantitative analysis of historical data cannot alone capture adequately all possible sources of future uncertainty. There will always be a need for judgement to be applied and in many situations such considerations will dominate the risk margin assessment. The application of judgement, however, is arguably the most difficult aspect of any attempt to estimate future uncertainty and assess appropriate risk margins. Our paper examines the key judgmental aspects and introduces a structured approach to combining these qualitative considerations with the results of any available quantitative analysis.

In preparing this paper the Taskforce has mainly considered, as a surrounding context, the current risk margin environment in Australia, in particular the percentile, or quantile, approach to determining margins for uncertainty. Having said this, we are aware that international developments, including proposed changes to International Financial Reporting Standards, are likely to overtake us in the not too distant future. We are of the view that the main aspects of our proposed framework can be readily adopted, altered or enhanced to complement analysis of uncertainty in the evolving wider international context.

The framework discussed in this paper can also be considered in the broader context of quantifying the uncertainty associated with reserve risk and underwriting risk for stochastic capital modelling (often referred to as *Dynamic Financial Analysis* or *Internal Capital Modelling*) purposes. In fact, when parameterising these components of a DFA model, one should draw on any analysis conducted for risk margin purposes and expand the framework to encapsulate those aspects of the parameterisation not captured by an analysis conducted specifically for risk margin purposes.

It is not proposed that this risk margin framework will have the prescriptive nature of a professional standard. Nevertheless, it is hoped that the structure and educational benefits it provides will encourage all actuaries to critically examine their current risk margin methodologies and to take from the framework those insights that are helpful to them in their particular situation. Inevitably, each actuary estimating risk margins will need to make their own judgements and this will be driven by their own knowledge and experience. The proposed framework does not attempt to usurp that process. Ultimately this framework is about enabling the profession and stakeholders to feel more confident in the quality and overall consistency of risk margins advice in future.

This is not a paper on stochastic reserving. Nor is it intended to provide all of the answers. Rather, its aim is to equip actuaries to ask the right questions and then proceed to answer these in a methodical and rigorous manner.

1.4. Structure of this paper

In **Section 2**, we present a framework which takes a methodical and rigorous approach to examining each of the key sources of uncertainty and provides a practical and user-friendly platform to help actuaries determine appropriate and justifiable risk margins for their insurance liability valuation portfolios.

Sections 3 and 4 discuss the assessment of independent risk and systemic risk, respectively, providing more practical guidance and considerations for the assessment of these sources of risk with a view to determining risk margins.

The framework is summarised in Table 1. The sections of the paper that address each step are also shown.

Table 1: Summary of risk margin analysis framework

Step	Framework component	Description	Section of paper
1	Portfolio preparation	Determine valuation portfolios, claim groups and techniques to deploy for each claim group	Section 2.3
2	Independent risk analysis	Conduct quantitative analysis, conduct benchmarking where appropriate, conduct retrospective analysis for stable periods	Sections 2.4 and 3
3	Internal systemic risk analysis	Apply balanced scorecard approach to objectively score central estimate valuation methodologies. Conduct analysis to determine appropriate CoVs to map to scores.	Sections 2.5 and 4
4	External systemic risk analysis	Identify, categorise and quantify potential future external sources of systemic risk	Sections 2.5 and 4
5	Analysis of correlation effects	Select correlation coefficients beween valuation classes and between outstanding claim and premium liabilities for internal systemic risk and for each external systemic risk category.	Sections 2.5
6	Consolidation of analysis	Consolidate CoVs and correlation coefficients. Independence assumed between three sources of uncertainty.	Section 2.6
7	Additional analysis	Conduct sensitivity testing, scenario testing, internal and external benchmarking and hindsight analysis.	Section 2.7
8	Documentation	Document the analysis and judgement relating to each step of the framework	Section 2.8
9	Review	Conduct annual reviews of key assumptions in the context of emerging experience. Full deployment of the framework at least every three years, including active interactions with business unit management.	Section 2.8

2. The proposed framework

2.1. Introduction to framework

The proposed framework provides a practical and robust platform that requires a combination of quantitative and qualitative techniques to be deployed to examine the uncertainty associated with assessing insurance liabilities with a view to determining risk margins.

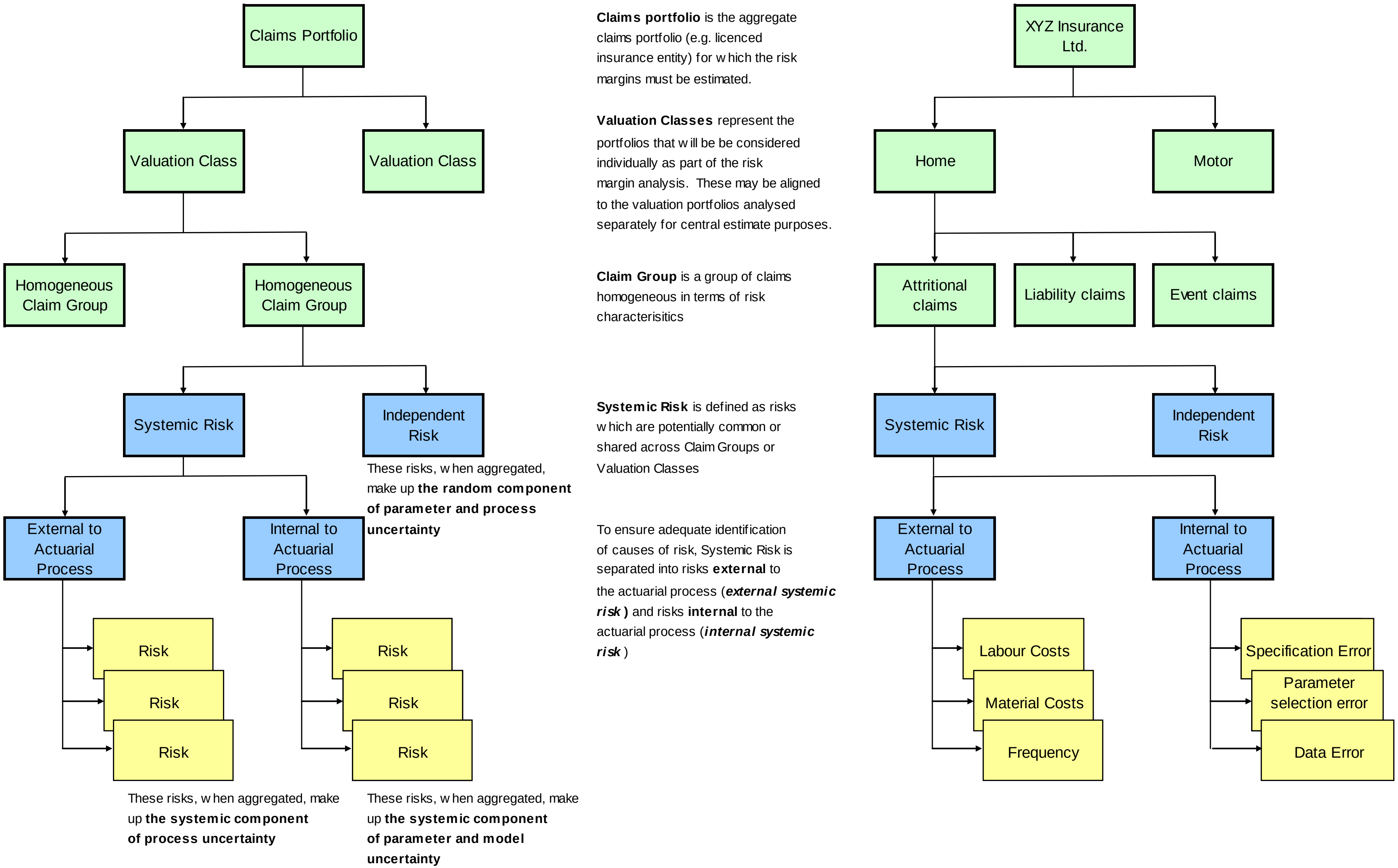
Quantitative techniques alone are insufficient to enable a complete assessment of the various sources of uncertainty. These techniques must be supplemented by qualitative analysis to ensure that all sources of uncertainty are captured. It is common practice for Australian actuaries to adjust the results obtained using quantitative techniques to allow for their known weaknesses. However, this is not always done in a rigorous manner, nor is there much consistency across the profession.

The framework is designed to introduce more rigour and consistency to the risk margin assessment process by encouraging actuaries to examine their own portfolios using a step-by-step process that requires them to ask a number of questions in the context of these portfolios. This will enable judgemental aspects of the process to be better reasoned, justified and documented and ultimately provide more structure in the application and combination of both quantitative and qualitative processes.

It is not expected that all of the techniques discussed in this paper will be used in practice for all valuation portfolios. Rather, if an actuary proceeds through the step-by-step process using techniques suited to their own portfolios, understanding the strengths and weaknesses of these techniques and asking the right questions along the way, they can only be more comfortable that the risk margins adopted are appropriate.

The framework revolves around quantifying the contribution to uncertainty from each of the main sources of uncertainty and is graphically represented in Figure 1 below.

Figure 1: Framework for determining insurance liability risk margins



2.2. Sources of uncertainty

The sources of uncertainty are the cornerstones of the framework. The framework itself has been designed to ensure alignment between the analysis and the techniques deployed with the key sources of uncertainty, ensuring a complete measurement of uncertainty.

At the highest level, the sources of uncertainty can be categorised as belonging to either the *systemic risk* source or the *independent risk* source.

Systemic risk represents those risks that are potentially common across valuation classes or claim groups. Systemic risks arise from two sources:

- Risks internal to the insurance liability valuation process, collectively referred to in this paper as *internal systemic risk*. This source of uncertainty encapsulates the extent to which the adopted actuarial valuation approach is an imperfect representation of a complex real life process. Model structure and adequacy, model parameterisation and data accuracy are all aspects of internal systemic risk. This source of uncertainty is alternatively referred to as *model specification risk*.
- Risks external to the actuarial modelling process, collectively referred to in this paper as *external systemic risk*. Even if the valuation model is an appropriate representation of reality, as it exists today, future systemic trends in claim cost outcomes that are external to the modelling process may result in actual experience differing from that expected based on the current environment and trends.

Independent risk represents those risks arising due to the randomness inherent in the insurance process. Independent risk also arises from two sources:

- The random component of *parameter risk*, representing the extent to which the randomness associated with the insurance process compromises the ability to select appropriate parameters in the valuation models.
- The random component of *process risk* being the pure effect of the randomness associated with the insurance process. Even if the valuation model was perfectly calibrated to reflect expected future outcomes, the volatility associated with the insurance process is likely to result in differences from the perfect expected outcomes.

In the detailed discussion of the framework below, quantitative and/or qualitative techniques are considered and aligned to the assessment and measurement of the internal and external sources of systemic risk and independent risk, the latter incorporating both parameter and process risk.

The nature of traditional quantitative modelling techniques, e.g. bootstrapping and stochastic chain ladder, are such that they are best suited to analysing sources of independent risk and past episodes of external systemic risk. However, they are inadequate alone to capture internal systemic risk or external systemic risk, to the extent that this latter differs from the past. For both systemic risk sources, traditional quantitative modelling techniques must be supplemented by other analysis, both quantitative and qualitative.

2.3. Preparing the claims portfolio for analysis

Before commencing any analysis one must prepare the *claims portfolio* for analysis. The claims portfolio would normally represent the aggregate insurance entity or aggregation of insurance entities for which the risk margin analysis is being conducted.

The claims portfolio should be split into appropriate *valuation classes*. A number of factors will impact how the valuation classes are selected.

An important consideration is whether the valuation portfolio split adopted to determine central estimates of insurance liabilities, or outstanding claim liabilities and premium liabilities where the split is different, should be adopted for risk margin analysis purposes. This would be preferable as it allows the risk margin analysis to be conducted in the context of the central estimate analysis and quantitative and qualitative analysis to be aligned with the key valuation drivers observed as part of the central estimate valuation. One of the attractions of the framework is that each of the sources of uncertainty being analysed can be aligned with the central estimate analysis and appropriate decisions around volatility made in the context of that analysis.

It may not be possible or particularly insightful, however, to conduct quantitative analysis at the same granular level as used for central estimate valuation purposes. The central estimate valuation portfolios may be too small for credible analysis or the valuation portfolio allocation may be at a more granular level than makes practical sense. For example, a large insurer may split its motor and home portfolios by state, product and claim type, resulting in a large number of individual central estimate valuation portfolios. The task of conducting quantitative analysis at the same granular level may be significant, costly and, considering the level of qualitative analysis that will be deployed as part of the assessment, unlikely to materially improve the final outcome. In such cases, quantitative analysis may be conducted on aggregated valuation classes and the results then allocated down, in an appropriate manner, to the valuation classes that are considered appropriate for the deployment of the framework.

In the end, the choice of valuation classes for risk margins analysis purposes will come down to a balance between the practical benefits gained from a higher level portfolio allocation and the potential additional benefit and insights gained from a more granular allocation. When making this decision consideration should be given to the need to retain as much consistency as possible between the central estimate methodology and basis and the risk margin analysis.

Once the claims portfolio has been allocated into risk margin valuation classes, consideration should be given to whether any valuation classes would benefit from a further allocation. For certain portfolios, it will be apparent that different groups of claims are materially more or less uncertain than others and should be treated separately for risk margin analysis purposes. Within each of these *claim groups* there is an element of homogeneity but between claim groups behaviour is expected to be different.

A good example of a valuation class that would normally require further segregation is a home portfolio. These portfolios are normally materially exposed to claims arising from natural peril events. The patterns of development for event claims often differ materially from those for non-event claims. Separate analysis of event and non-event claims will usually provide valuable insights into the past contribution to uncertainty from each of these claim sources with a view to making appropriate assumptions regarding future uncertainty. Also, home liability claims typically behave quite differently from other home claims and should be considered for separate analysis.

Again, a pragmatic view should be taken when considering whether groups of claims are homogeneous, a view that balances the benefits against the practicalities and cost.

For certain valuation portfolios, e.g. those with little historical data, it may not be possible to deploy all components of the framework. However, we do consider it important to consider each component in the context of each valuation portfolio as this will ensure that appropriate questions are asked as part of the analysis.

2.4. Analysing independent risk sources

Many approaches used in practice by actuaries to analyse uncertainty and assess risk margins have an element of quantitative analysis conducted using stochastic (or other) modelling techniques. Often, but not always, adjustments are made to the results from this modelling, reflecting an appreciation that it has not fully encapsulated all sources of uncertainty.

There are a number of reasons why stochastic modelling techniques do not enable a complete analysis of all sources of uncertainty:

- A good stochastic model will fit the past data well and, in doing so, *fit away* most past systemic episodes of risk external to the valuation process, leaving behind largely random sources of uncertainty. Some techniques, e.g Generalised Linear Modelling (GLM), offer more flexibility in fitting to the past experience than others, e.g. Mack method.
- Where it has not been possible to fit away all past systemic episodes of risk or where no attempt has been made to do so, the outcome of the analysis may be substantially affected by these episodes. Consideration then needs to be given to whether past episodes of systemic risk are reasonably representative of what one can expect in the future. For some portfolios this will be a very significant assumption, based solely on judgemental considerations.
- Even where one is comfortable that a model adequately reflects the volatility expected in the future from both independent and systemic sources external to the actuarial valuation process, the model is highly unlikely to incorporate uncertainty arising from sources internal to the actuarial valuation process, i.e. internal systemic risk.

The framework proposes the use of one or more stochastic modelling techniques to analyse independent sources of risk and to inform on past episodes of systemic risk external to the actuarial valuation process. There are a number of approaches that may be used to analyse independent sources of risk, including:

- Mack method;
- Bootstrapping;
- Stochastic Chain Ladder;
- Generalised Linear Modelling (GLM) techniques; and
- Bayesian techniques.

Although these techniques can be used for both outstanding claim liabilities and premium liabilities, it is possible and practically helpful to analyse independent risk as it pertains to premium liabilities using techniques specifically designed for this purpose.

The analysis of independent risk is an art in itself and actuaries will only become comfortable in this area with practical experience of working through the main issues on their own valuation portfolios. A range of stochastic techniques may be used and decisions made on the strengths and weaknesses of each approach in the context of the past experience. It may be possible to refine the modelling to focus on certain past periods with limited past episodes of systemic risk, thus largely isolating past independent risk and examining the extent to which it has impacted past volatility.

Finally, we do consider it useful to supplement any analysis of independent risk for a particular valuation portfolio with internal and external benchmarking. Benchmarking is discussed in section 2.7. The main source of external benchmarking in this regard would be the 2001 Tillinghast paper which identified the independent risk component in its overall uncertainty benchmarks. For some portfolios, benchmarking may be the only way to obtain some view of the contribution from independent risk once all other avenues have been exhausted.

2.5. Analysing systemic risk sources

The framework proposes separate analysis of internal systemic risk and external systemic risk. Qualitative approaches are proposed for this purpose. Two approaches are discussed in Section 4 of the paper, one designed to analyse internal systemic risk and the other designed to analyse external systemic risk. Introductions to these approaches are given in this sub-section. Both techniques have been designed to allow judgement to be deployed in a robust, transparent and consistent manner, giving due consideration to each of the key contributors to the two sources of systemic risk.

Internal systemic risk

Internal systemic risk refers to the uncertainty arising from the actuarial valuation models used being an imperfect representation of the insurance process as it pertains to insurance liabilities. Valuation models are designed to predict future claim cost outcomes based largely on an examination of the key predictors of claim cost, and trends in these predictors, as these have been observed in the past claim experience.

When assessing the uncertainty associated with the insurance liabilities it is important to subject the valuation methodology to objective scrutiny to assess the extent to which the quality of the insurance liability estimate may be compromised by inadequacies in the valuation process. The need to be objective as part of this process is important. Human nature is such that it is easy to become overly defensive of the modelling approach adopted for central estimate purposes. Objective comparisons and scoring of the adopted valuation methodology against best practice, irrespective of whether such best practice is possible in the context of the portfolio being analysed, is crucial to forming an appropriate view of the contribution of internal systemic risk to uncertainty.

We consider there to be three main sources of internal systemic risk. These are:

- *Specification error* **-** the error that can arise from an inability to build a model that is fully representative of the underlying insurance process. The process is likely to be too complicated to be replicated in any actuarial valuation model. Also, the information available may be such that the underlying process cannot be fully understood and the model structure is simplified as a consequence.
- *Parameter selection error* **-** the error that can arise because the model is unable to adequately measure all predictors of claim cost outcomes or trends in these predictors. Again the insurance process is such that there can be a large number of claim cost drivers that would be difficult to fully capture in an actuarial valuation model.
- *Data error* **-** the error that can arise due to poor data or unavailability of data required to conduct a credible valuation. Data error also relates to inadequate knowledge of the portfolio being analysed, including pricing, underwriting and claims management processes and strategies.

One approach to analysing internal systemic risk is discussed in detail in section 4 of the paper. This involves developing a balanced scorecard to objectively assess the model specification against a set of criteria designed to rank aspects of the modelling from worst to best practice. For each of the sources of internal systemic risk, risk indicators are developed and then scored against the adopted criteria. The scores are then aggregated for each valuation class and mapped to a quantitative measure (CoV) of the variation arising from internal systemic risk.

There are a number of subjective decisions that are required to be made as part of this process. These include the risk indicators, the measurement and scoring criteria, the importance (or weight) afforded to each risk indicator and the CoVs that map to each score from the balanced scorecard. Quantitative techniques may be used to inform aspects of these decisions.

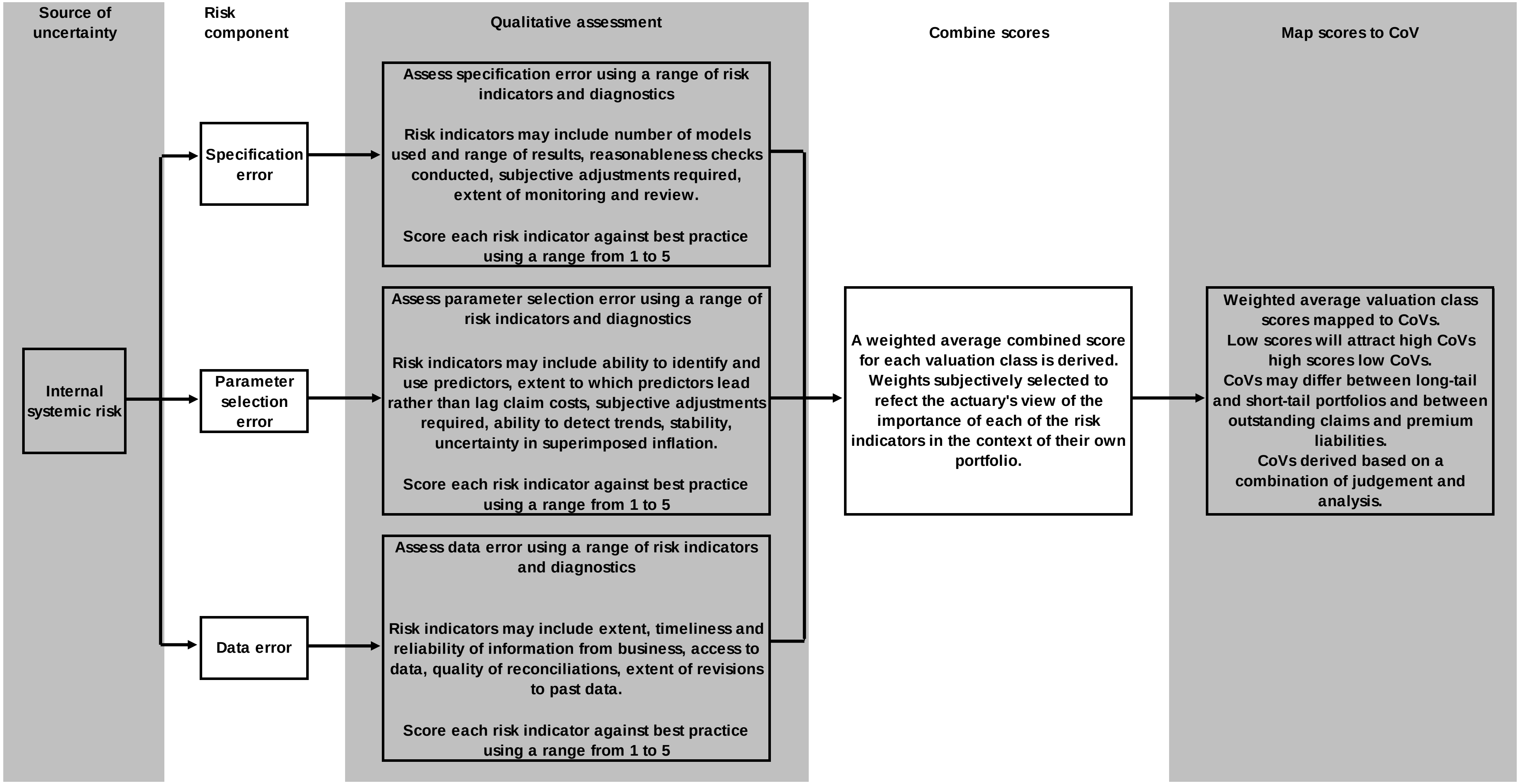
Development and deployment of a balanced scorecard approach to measuring internal systemic risk is a blend of art and science. Actuaries unfamiliar with the approach will need time to develop the skills required:

- to draw out all of the risk indicators;
- objectively score them against best practice; and
- map them to a CoV in the context of their own valuation classes.

Section 4 of the paper provides some thoughts and tools that may be used as part of such an exercise. However, it is fully expected that new techniques will emerge as experience develops and the writers of this paper welcome and encourage future contributions to the development of actuarial thinking in this area.

The analysis of internal systemic risk is summarised in Figure 2 below.

Figure 2: Internal systemic risk – systemic risk internal to the actuarial valuation process



External Systemic Risk

All of the standard quantitative modelling techniques analyse the volatility inherent in the past claim experience. As such, they can only be used to inform on the uncertainty arising from past episodes of external systemic risk. To use these techniques in isolation would require an assumption that the contribution to volatility from future external systemic risk is expected to be similar to that experienced in the past. It is quite possible, and for some valuation classes likely, that future external systemic risk will exhibit significantly different characteristics from actual past episodes.

It is, therefore, important to identify each of the main potential sources of external systemic risk and, for each of these sources, quantify their impact on the overall volatility of the insurance liabilities. The main external systemic risks for any valuation class can be categorised as belonging to a number of *risk categories*. These include:

- *Economic and social risks* – normal inflation and other social and environmental trends
- *Legislative, political risks and claim inflation risks* – relates to known or unknown changes to legislative or political environment within which each valuation portfolio currently operates and shifts or trends in the level of claim settlements (this risk category encapsulates most systemic trends normally referred to as superimposed inflation)
- *Claim management process change risk* – changes to the processes relating to claim reporting, payment, finalisation or estimation
- *Expense risk* – the uncertainty associated with the cost of managing the run off of the insurance liabilities or the cost of maintaining the unexpired risk until the date of loss
- *Event risk* – the uncertainty associated with claim costs arising from events, either natural peril events or man-made events
- *Latent claim risk* – the uncertainty associated with claims that may arise from a particular source, a source that is currently not considered to be covered
- *Recovery risk* – the uncertainty associated with recoveries, either reinsurance or non-reinsurance

Each of these risk categories will normally have been considered as part of the central estimate valuation of outstanding claim or premium liabilities. There is, therefore, a strong case for conducting the analysis of external systemic risk in conjunction with the central estimate valuation, thereby ensuring that both parts of the valuation take a consistent and complete view of all systemic risk categories.

A critical step in any valuation process is the interaction between the valuation actuary and business unit management. This is required to ensure that the valuation actuary has an appropriate level of understanding of all aspects of the insurance process, particularly as this relates to the valuation of insurance liabilities. These interactions will normally incorporate discussions about all aspects of the portfolio management process, including underwriting and risk selection, pricing, claims management, expense management, emerging portfolio trends and the environment within which the portfolio operates. It would be of great benefit to the valuation process, and not particularly onerous, to extend discussions to consider the main potential external systemic risks that may impact the portfolio. This information can then be used to inform both the central estimate valuation and in the identification and quantification of risks associated with each external systemic risk category.

For most valuation classes, the risk identification and categorisation process will identify a small number of systemic risks and categories that account for the majority of the uncertainty. For property classes, for example, *event risk* is likely to dominate the volatility of the premium liabilities whereas for long-tail portfolios *legislative, political and claims inflation risks* are likely to be the key contributors to the volatility for both outstanding claim and premium liabilities.

When analysing external systemic risk it is useful to rank each of the risk categories in descending order in terms of expected impact on insurance liability uncertainty. This ranking can then be used to guide the effort to be expended on quantifying the risks associated with each risk category. More time and effort would be spent on quantifying the uncertainty associated with material risk categories.

Section 4 of the paper discusses the assessment of external systemic risk in more detail and includes some examples of potential sources of systemic risk within each risk category.

Correlation effects

At this point in the deployment of the framework, an actuary will have derived CoVs for independent risk, internal systemic risk and for each source of external systemic risk in each systemic risk category. The next step requires making allowance for the fact that each of these sources of risk is not fully correlated either within valuation classes or between valuation classes.

At this stage, it is worth commenting that we do not consider or discuss any quantitative methods to assessing correlation effects as part of this paper. The main reasons for this are as follows:

- Available techniques tend to be technically complex and often require a substantial amount of data. The time and effort required to learn, implement and appropriately adjust these techniques may outweigh the benefits gained.
- These techniques will yield correlations that are heavily influenced by the correlations, if any, experienced in past data. Correlations associated with external systemic risk sources may differ materially from correlations associated with past episodes of systemic risk.
- Also, it is difficult, if not impossible, to separate the past correlation effects between independent risk and systemic risk or to identify the pure effect of each past systemic risk.
- Internal systemic risk cannot be modelled using standard correlation modelling techniques.
- Even if modelling of correlation effects were practical, they are unlikely to yield results that could be aligned to the outcomes of the framework discussed above in relation to independent risk, internal systemic risk and external systemic risk.

Having said this, it is not our intention to entirely rule out quantitative analysis of past correlation effects. Such analysis may provide useful insights that can help in the assessment of potential future correlation effects.

The framework can be readily extended to incorporate an appropriate allowance for correlation effects. This extension follows the spirit of the framework discussed so far and requires that correlation effects be considered in the context of each source of uncertainty and/or risk category. Again, reliance is placed on an actuary's own judgement but the actuary is encouraged to deploy their judgement in a robust and transparent manner in the context of each of the risks affecting their valuation classes.

Correlation effects can be considered in the context of each source of uncertainty. The key considerations are discussed below.

- Independent risk – as suggested by the name, this source of uncertainty can be assumed to be uncorrelated with any other source of uncertainty, either within a particular valuation class or between valuation classes.
- Internal systemic risk – this source of uncertainty can be assumed to be uncorrelated with independent risk, as discussed above, and with each potential external systemic source of risk, either within a particular valuation class or between valuation classes. Internal systemic risk contributes to correlation effects through correlation of this source of uncertainty between valuation classes or between outstanding claim and premium liabilities.
    - The *same actuary effect* and the use of template or valuation models across different valuation classes are key considerations for correlation effects between valuation classes.
    - Linkages between the premium liability methodology and outcomes from the outstanding claim valuation are key considerations for correlation effects between outstanding claim and premium liabilities.
- External systemic risk – it is reasonable to assume that the contribution to uncertainty from each risk category is uncorrelated with independent risk, internal systemic risk and with the contribution to uncertainty from each other risk category, either within a particular valuation class or between valuation classes. Correlation effects will arise from correlations between classes or between outstanding claim and premium liabilities from risks categorised as belonging to similar risk categories, e.g. claims inflation risk across long-tail portfolios or event risk across property and motor portfolios.

It is possible that external systemic risk categories may be partially correlated either within or between valuation classes. If this is the case, the correlated risk categories may be aggregated into broader categories that are not correlated with other risk categories.

For practical purpose, the correlation relationship between any two sources of uncertainty or risk categories can be considered to belong to one of a finite number of assumed correlation bands. For example, five correlation bands may be defined as *nil*, *low*, *medium*, *high* and *full* correlation. For quantification purposes one might allocate correlation coefficients of 25%, 50% and 75%, respectively, to the low, medium and high correlation bands. Having any more than five categories is likely to result in spurious accuracy attaching to what is already a largely subjective process.

The PwC paper describes a useful way of considering and assessing correlation effects. A root dummy variable, which can be considered to be the root source of correlations within a risk category, is created. Dummy variables may also be set up for groupings of valuation classes that belong to the same class of business, e.g. separate valuations may be conducted by state within a worker's compensation class of business. A hierarchical structure can then be constructed for each systemic risk category containing correlations between the following components:

- premium liabilities and outstanding claim liabilities for a particular valuation class;
- outstanding claim liabilities for individual valuation classes and the relevant class of business dummy variables; and
- class of business dummy variables and root dummy variables.

The implied correlations, both within valuation classes or classes of business and between valuation classes, can then be assessed.

2.6. Consolidation of analysis into risk margin calculation

Once an actuary has progressed through the analysis discussed above they will have the following assumptions that need to be consolidated and converted into a risk margin for the whole claims portfolio:

- CoVs in respect of independent risk for each valuation portfolio, separately for outstanding claim and premium liabilities
- CoVs in respect of internal systemic risk for each valuation portfolio, separately for outstanding claim and premium liabilities
- CoVs in respect of each potential external systemic risk category, separately for outstanding claim and premium liabilities
- Correlation coefficients between each source of uncertainty, risk category, valuation portfolio and outstanding claim/premium liability combination.

For practical purposes, we propose that a simple linear correlation dependency structure be adopted to allow for the various correlation effects. Correlation matrices are created for each of the three sources of uncertainty described in section 2.2 above. As discussed above, independent risk, internal systemic risk and external systemic risk are all assumed to be uncorrelated. As such, the contribution from each source of uncertainty to the total CoV, after correlation effects, can be calculated individually and then combined.

We consider a simple linear correlation dependency structure to be reasonable for the assessment of risk margins associated with probabilities of adequacy of up to at least 90%. Where one is faced with requirements for extreme probabilities of adequacy, e.g. for portfolios in run off or when parameterising reserve risk for DFA modelling purposes, it is recommended that other dependency structures be considered.

An example of the consolidation and risk margin calculation for an example insurer, Insurer ABC, which underwrites three classes of business, Motor, Home and CTP is shown in Figure 3 below.

Figure 3: Claims portfolio CoV and risk margin calculation for Insurer ABC

A: Proportion of insurance liabilities

	Proportion of insurance liabilities (weights)	
Class	Outstanding claims	Premium liabilities
Motor	5%	25%
Home	5%	25%
CTP	30%	10%
Total	40%	60%

B: Independent risk

	Independent risk		
Class	Outstanding claims CoV	Premium liabilities CoV	Insurance liabilities CoV
Motor	7.0%	5.0%	1.3%
Home	6.0%	5.0%	4.3%
CTP	6.0%	15.0%	5.9%
Total	4.6%	3.9%	3.0%

C: Internal systemic risk

	Internal systemic risk		
Class	Outstanding claims CoV	Premium liabilities CoV	Insurance liabilities CoV
Motor	5.5%	5.0%	4.9%
Home	5.5%	5.0%	4.9%
CTP	9.5%	8.0%	8.7%
Total	7.6%	4.2%	4.9%

Internal systemic risk correlation matrix

	Motor OSC	Motor PL	Home OSC	Home PL	CTP OSC	CTP PL
Motor OSC	100%	75%	50%	50%	25%	25%
Motor PL	75%	100%	50%	50%	25%	25%
Home OSC	50%	50%	100%	75%	25%	25%
Home PL	50%	50%	75%	100%	25%	25%
CTP OSC	25%	25%	25%	25%	100%	75%
CTP PL	25%	25%	25%	25%	75%	100%

D: External systemic risk

External systemic risk - coefficients of variation by risk category

	social, etc, risk	Legislative, political and claims inflation risk	Claim process risk	Expense risk	Event risk	Latent claim risk	Recovery risk	All risk categories
Motor OSC	1.0%	0.5%	2.0%	1.0%	1.0%	0.0%	3.0%	4.0%
Motor PL	2.0%	0.5%	2.0%	2.0%	3.0%	0.0%	5.0%	6.8%
Home OSC	1.0%	1.0%	2.0%	1.0%	2.0%	0.5%	0.5%	3.4%
Home PL	2.0%	1.0%	2.0%	2.0%	15.0%	0.5%	1.0%	15.5%
CTP OSC	3.0%	10.0%	4.0%	2.0%	0.0%	0.5%	1.0%	11.4%
CTP PL	4.0%	12.0%	4.0%	3.0%	1.0%	0.5%	2.0%	13.8%

External systemic risk - risk category correlations

Risk category	***Correlations adopted***
Economic, social and environmental risk	Nil between CTP and other, 25% PL/25% OSC between motor and home, 50% between OSC and PL within classes
Legislative, political and claims inflation risk	Nil between CTP and other, 25% PL/25% OSC between motor and home, 50% between OSC and PL within classes
Claim management process risk	25% between classes, 50% between OSC and PL within classes
Expense risk	25% between classes, 50% between OSC and PL within classes
Event risk	Nil between CTP and other, 50% PL/25% OSC between motor and home, 50% between OSC and PL within classes
Latent claim risk	Nil between classes, 50% between OSC and PL within classes
Recovery risk	Nil between classes, 50% between OSC and PL within classes

	External systemic risk		
Class	Outstanding claims CoV	Premium liabilities CoV	Insurance liabilities CoV
Motor	4.0%	6.8%	6.0%
Home	3.4%	15.5%	13.1%
CTP	11.4%	13.8%	10.7%
Total	8.6%	8.0%	6.5%

E: Consolidated CoVs

	All sources of uncertainty		
Class	Outstanding claims CoV	Premium liabilities CoV	Insurance liabilities CoV
Motor	9.8%	9.8%	7.9%
Home	8.8%	17.0%	14.6%
CTP	16.0%	21.9%	15.0%
Total	12.4%	9.9%	8.7%

F: Risk margins

Required probability of adequacy **75%**

	Risk margins - Normal distribution			Risk margins - LogNormal distribution		
Class	Outstanding claims	Premium liabilities	Insurance liabilities	Outstanding claims CoV	Premium liabilities CoV	Insurance liabilities CoV
Motor	6.6%	6.6%	5.3%	6.3%	6.3%	5.1%
Home	5.9%	11.5%	9.9%	5.7%	10.5%	9.2%
CTP	10.8%	14.8%	10.1%	9.9%	13.0%	9.4%
Total	8.4%	6.7%	5.8%	7.9%	6.3%	5.6%

The following comments are made to help in the interpretation of the example in Figure 3.

- The CoVs and correlation coefficients used and risk margins derived are indicative only. The emphasis is on demonstrating how consolidation could work in practice, rather than proposing appropriate risk margins or underlying assumptions.
- Part A gives the percentage breakdown of the total net central estimate of insurance liabilities by valuation portfolio and between outstanding claim and premium liabilities. There is no need to use actual dollar amounts in the calculation. The percentage breakdown (or weights) will suffice. For simplicity, for this example all homogeneous claim groups have been combined within the valuation classes.
- Part B shows the CoVs adopted in respect of independent risk for outstanding claim and premium liabilities following a combination of quantitative modelling and benchmarking. The insurance liability CoVs by valuation portfolio and the insurance liability, outstanding claim liability and premium liability CoVs for all valuation portfolios combined have been derived assuming independence (or nil correlation) between valuation portfolios and between outstanding claims and premium liabilities.
- Part C shows the CoVs and correlation coefficients (in correlation matrix form) adopted for outstanding claim and premium liabilities in respect of internal systemic risk. These CoVs and correlation coefficients have been derived following a qualitative analysis of internal systemic risk using a balanced scorecard approach. The insurance liability CoVs by valuation portfolio and the insurance liability, outstanding claim liability and premium liability CoVs for all valuation portfolios combined have been derived using the assumed correlations between valuation portfolios and between outstanding claim and premium liabilities. When creating any correlation matrix it is important to include a check that the matrix is positive definite.
- The first table in Part D shows the CoVs adopted in respect of each external systemic risk category. The second table summarises the adopted correlation coefficients in respect of external systemic risk. The implementation of these correlations is conducted using seven correlation matrices, one for each external systemic risk category. Each of these matrices is 6x6, similar to the correlation matrix shown in Part C for internal systemic risk. With an assumption of independence between risk categories there is no need to create a larger 42x42 matrix with a row and column representing each risk category, valuation portfolio and outstanding claim/premium liability combination. The CoVs and correlation coefficients shown in these two tables have been derived following a qualitative analysis of potential external systemic sources of risk. The third table in Part D shows the aggregate CoVs in respect of external systemic risk, derived for each valuation portfolio and for all valuation portfolios combined in respect of outstanding claim liabilities, premium liabilities and insurance liabilities.
- Part E consolidates the CoVs from each of the three sources of uncertainty, derived in Parts B to D. The key assumption underlying the derivation of consolidated CoVs is that there is independence between each of the sources of uncertainty.
- Part F converts the consolidated CoVs into risk margins assuming a required probability of adequacy of 75%. Two statistical distributions have been adopted as representative of the underlying distribution of insurance liabilities: the Normal distribution and the LogNormal distribution. At lower probabilities of adequacy, including 75%, the Normal distribution delivers a higher risk margin,

irrespective of the consolidated CoV. At higher probabilities of adequacy, including 90%, the LogNormal distribution can give a higher result, where the consolidated CoV is not too high. For particularly high CoVs, the LogNormal distribution can generate risk margins that appear unreasonable. For example, for a 75% probability of adequacy the risk margin percentage does not increase much above 25% and actually reduces as the CoV increases above 75%. Another way of looking at this is that LogNormal risk margins can reduce quite significantly as a percentage of the CoV as the latter increases whereas Normal risk margins remain unchanged as a percentage of the CoV.

- Both distributions are used in practice by actuaries with the LogNormal distribution more common for higher probabilities of adequacy and the Normal distribution, for the reasons discussed above, often given consideration at the 75% probability of adequacy. The right-tailed nature of the distribution of insurance liabilities perhaps lends itself more to a right-skewed distribution such as LogNormal. However, it does have its practical issues at lower probabilities of adequacy as discussed above. Considering the level of judgement required in the application of the framework, spending a substantial amount of time deliberating over the form of the distribution is unlikely to be of much value. An actuary should adopt a distribution that is appropriate in the context of their own claims portfolio, including the consolidated CoV assessed and probability of adequacy required. One might not be so comfortable to adopt a LogNormal or Normal distribution without further justification if the purpose of the analysis is to derive risk margins with very high probabilities of adequacy (i.e. 99.5% for portfolios in run off) or when parameterising reserve risk in a DFA modelling context.
- A spreadsheet tool has been created to do the calculation required for the consolidation shown in Figure 3. This tool has been provided as an attachment to this paper to help readers understand the key formulae underpinning the consolidation. Obviously, this tool may also be adapted for use in the deployment of the framework discussed in this paper.

2.7. Additional analysis

There are a number of areas of additional analysis that may be conducted to give an actuary further comfort regarding the outcomes from the deployment of the framework described above. These include sensitivity analysis, scenario testing, benchmarking and hindsight analysis, each of which is discussed below.

Sensitivity testing

The framework requires a substantial amount of actuarial judgement in its application. Judgement is required in all aspects of the analysis, irrespective of whether quantitative or qualitative methods have been used to assess the volatility associated with a particular source of uncertainty.

Valuable insights into the sensitivity of the final outcomes to key assumptions can be gained by varying each of the key assumptions. It is recommended that, as part of the analysis, the CoVs and correlation coefficients adopted for independent risk, internal systemic risk and each external systemic risk category be flexed and the impact on the valuation class and claims portfolio risk margins examined.

Following such an analysis, one might review certain key assumptions, particularly those that have a substantial impact on the final outcome, with a view to gaining additional comfort that the adopted assumptions are reasonable and justifiable.

As a demonstration of sensitivity testing in practice changes have been made to certain key assumptions adopted for the example in Figure 3.

- If the independent risk CoVs by valuation portfolio for outstanding claim and premium liabilities are reduced by 50%, the risk margin for the whole claims portfolio (based on the LogNormal distribution) reduces from 5.6% to 5.4%. Alternatively, doubling these CoVs increases the risk margin to 6.5%.
- If the internal systemic risk CoVs by valuation portfolio for outstanding claim and premium liabilities are increased by 50%, the risk margin for the whole claims portfolio increases from 5.6% to 6.6%. Alternatively, increasing the correlation coefficients to give full correlation across all combinations increases the risk margin to 6.3%.
- If the CoVs for the legislative, political and claims inflation systemic risk category for CTP (outstanding claims and premium liabilities) are reduced by 50%, the risk margin for the whole claims portfolio reduces from 5.6% to 5.2%. Doubling the CoV for the event systemic risk category for Home premium liabilities increases the risk margin to 7.0%. Finally, assuming full correlation, within all valuation classes and systemic risk categories, between outstanding claim and premium liabilities increases the risk margin to 5.8%.

Scenario testing

It is often insightful to tie the risk margin outcomes back to a set of valuation outcomes by strengthening some of the key assumptions adopted for central estimate purposes to align the outstanding claim liabilities and premium liabilities with the provisions assessed including risk margins. Various different assumption scenarios may be tested and valuation outcomes, including projected ultimate claim frequencies, average claim sizes, loss ratios, etc, compared for each scenario against the central estimate basis.

These (risk margin inclusive) valuation outcomes can be considered in the context of the emerging experience and what is known about the portfolio. Also, the basis changes required to deliver these outcomes can be considered in the context of the emerging experience.

Internal benchmarking

As part of the CoV selection process, the proposed CoVs should be subjected to a range of internal checks. For each source of uncertainty individually the adopted CoVs should be compared between valuation classes, particularly similar valuation classes, for outstanding claim liabilities, premium liabilities and insurance liabilities. Comparisons should also be made between outstanding claim and premium liability CoVs within classes.

For independent risk, there are two main dimensions that should be considered in the context of internal benchmarking: portfolio size and length of claim run off. The *law of large numbers* implies that the larger the portfolio, the lower the volatility arising from random effects. Also, the longer a portfolio takes to run off, the more time there is for random effects to have an impact. These considerations have a number of implications for independent risk CoV selection, including:

- Outstanding claim liability CoVs for short-tail portfolios are likely to be lower than for similar sized long-tail portfolio and substantially lower than much smaller long-tailed portfolios.
- Premium liability CoVs for long-tail portfolios would normally be higher than outstanding claim liability CoVs for the same portfolios. This is due more to the law of large numbers than any material differences in the length of the run off. The extent of the difference will depend on the size of the premium liability and outstanding claim liability with the difference being more for small portfolios which will have higher independent risk components than for large portfolios which will have smaller independent risk components.
- Premium liability CoVs for short-tail portfolios would normally be lower than outstanding claim liability CoVs for the same portfolios, assuming the same independent risk profile between outstanding claim and premium liabilities. This is due mainly to the law of large numbers. The independent risk profiles may not, however, be similar. Event risk, where material, is likely to mean that the independent risk profile of premium liabilities and outstanding claim liabilities are different. This is likely to offset the benefit that premium liabilities gain from their greater size and in any event make benchmarking problematic.

For internal systemic risk, the CoVs can be compared in the context of each valuation class. If template models are used for similar portfolios, particularly classes with homogeneous claim groups, then one would expect CoVs to be similar between classes. Also, the underlying process and the key drivers of this process are likely to be more complicated in long-tail portfolios than most short-tail portfolios. If similar valuation methodologies are applied for both short- and long-tail classes then one would expect higher internal systemic risk CoVs for the long-tail portfolios.

The main sources of external systemic risk are likely to be much more significant for long-tail portfolios with the exception of event risk for property and, to a lesser extent, motor classes and liability risk for home classes.

External benchmarking

External benchmarking refers to the use of the Tillinghast and Trowbridge 2001 papers or APRA’s November 2008 General Insurance Risk Margins Industry Report to benchmark CoVs and/or risk margins derived as part of a risk margins analysis.

APRA have indicated that a large number of actuaries rely, to varying extents, on the analysis presented in the 2001 papers in the selection of their own risk margin assumptions. This reliance ranges from those actuaries who conduct thorough analyses on their own portfolios and then benchmark the adopted risk margins with those derived from the 2001 papers to those actuaries that derive risk margins solely from the 2001 papers with little or no consideration of the reasonableness of this approach in the context of their own portfolios. The latter approach was certainly not one of the original intentions of the authors of the 2001 papers. The former approach is more consistent with the expectations of the authors.

It is not our intention to dismiss external benchmarking out of hand. Rather, we consider that this form of benchmarking has some merit when combined with a thorough analysis of a particular claims portfolio. Benchmarking will be of some benefit where there is little information available for analysis purposes, particularly for the analysis of independent risk. More generally, the use of benchmarking should be as a sanity check rather than as the entire basis of the risk margin assessment. In any deployment of benchmarking, the differences between the benchmark portfolio(s) and the claims portfolio being analysed must be considered and factored into the analysis.

The use of the Tillinghast paper in the assessment of independent risk is discussed in section 2.4 above. Before using the Tillinghast paper, however, an actuary needs to be aware of the following issues:

- The assumptions required to derive the independent component of the CoV were derived based on an analysis conducted during 2001. The independent CoVs depend on the size of the outstanding claim or premium liabilities. Inflation between 2001 and the effective date of the current valuation should be backed out of the outstanding claim and premium liabilities before calculating the independent CoV. If this is not done then the independent CoV will be understated.
- The premium liability risk margin should be calculated by applying a multiple to the outstanding claim risk margin for an outstanding claim liability **that is the same size as the premium liability**, not for the actual outstanding claim liability, irrespective of whether this is lower or higher than the premium liability.

Hindsight analysis

Hindsight analysis involves comparing past estimates of outstanding claim liabilities and premium liabilities against the latest view of the equivalent liabilities. Movements can be analysed and converted to a coefficient of variation reflective of the actual volatility observed in the past. This volatility contains a combination of past instances of independent risk, internal systemic risk and external systemic risk. Care needs to be taken in the interpretation of any hindsight analysis as the models may have changed (improved) since previous valuations were conducted. Also, future external sources of systemic risk may differ materially from past such episodes of systemic risk.

Hindsight analysis is particularly useful for short-tail valuations where there is little serial correlation between consecutive valuations. Hindsight analysis is somewhat less valuable for long-tail portfolios where there is usually significant serial correlation between consecutive valuations.

The reader is referred to the 2005 paper *An Empirical Approach to Insurance Liability Prediction Error With Application to APRA Risk Margin Determination* by Andrew Houltram for a thorough discussion of the benefits and practicalities associated with hindsight estimation.

Another form of hindsight analysis, which we will refer to as *mechanical hindsight analysis*, is one that takes a mechanical approach to estimating the outstanding claims and premium liabilities, systematically removing the most recent claims experience. An example of such an approach is as follows:

- Apply a chain ladder method on a triangulation of cumulative claim payments based on a triangulation of data at the valuation date.

- The adopted payment development factors should be calculated using an objective approach, e.g. the average of the actual experience over the last three years.
- The outstanding claim payments derived using all data to the valuation date is referred to as the 'current' estimate.
- Remove a diagonal of payment data one at a time and apply the same method objectively to derive outstanding claim payments at past valuation dates.
- Compare each of the past estimates of outstanding claim payments with the current estimate, for the equivalent accident periods and ensuring that relevant payments made between valuation dates are added to the current estimate of outstanding claim payments.
- The method can be extended to incorporate a mechanical projection of premium liabilities at each valuation date. Premium liability volatility and past levels of correlation between outstanding claim and premium liabilities can be examined.

Mechanical hindsight analysis may be used to analyse:

- independent risk, by focusing the analysis on periods where there was a degree of stability in the experience with few or no systemic trends;
- internal systemic risk, by applying this technique using a range of actuarial methods (preferably those used for central estimate valuation purpose) and observing the differences in volatility outcomes; and
- all past sources of uncertainty, by applying the approach across all past periods.

The latter is a mechanical variant of the hindsight analysis described in the first three paragraphs of this sub-section.

2.8. Documentation and regularity

Documentation

APRA have indicated that a wide range of approaches have been taken by actuaries in the documentation of risk margins analysis. Documentation ranges from that which provides a thorough discussion of approach and justification for the assumptions underpinning the adopted risk margins to that which provides very little commentary or justification.

Documentation of actuarial judgement is not necessarily an easy task. However, we believe that the framework offers actuaries an opportunity to document their analysis and key judgemental decisions in a complete and robust manner, aligned to the key steps in the framework.

Regularity and review

A full application of each step of the framework is a substantial and comprehensive undertaking. We do not consider that the framework need be applied in its entirety each time an actuary conducts a central estimate valuation of insurance liabilities.

We consider a full application of the framework at less regular intervals to be reasonable and appropriate. At the very least, however, a full application should be applied every three years. These extensive reviews should incorporate all of the steps of the framework discussed above and summarised in Table 1. They will also involve significant interaction with business unit management.

At more regular intervals, aligned to the times when central estimate valuations of insurance liabilities are conducted, a less comprehensive review of the key assumptions adopted as part of the previous full application will suffice. The key assumptions should be examined in the context of:

- any emerging trends;
- emerging systemic risks; and
- changes to valuation methodologies.

Changes to key assumptions would only be considered where there is reasonable justification for doing so, i.e. where the previous assumptions are no longer deemed appropriate. Another way of thinking of these regular reviews are as monitoring exercises where key assumptions derived from the previous full framework application are monitored against emerging experience and developing knowledge and adjusted where justified.

If new portfolios emerge in the period between full applications of the framework, one should consider applying the key steps within the framework to those portfolios.

The successful deployment of this framework will require significant interaction with business unit management. The process may benefit from a feedback and communication loop, enabling the business to provide their views on the outcomes of the analysis. This will reduce the possibility that lots of assumptions, which all make sense individually, contribute to an overall outcome that does not make sense. This communication loop may incorporate the demonstration of scenarios that would give rise to the outcome assessed at the selected probability of adequacy.

3. Independent risk assessment

Independent risk reflects the contribution to the uncertainty associated with the actual claim cost outcome from random effects. This source of risk has two components: the random component of parameter risk and the random component of process risk. It is not normally particularly enlightening or beneficial to split independent risk between these two components. Having said this, some quantitative modelling techniques do allow the split to be assessed as part of their normal application.

There are a number of approaches that may be used to analyse independent sources of risk, including::

- Mack method;
- Bootstrapping;
- Stochastic Chain Ladder;
- Generalised Linear Modelling (GLM) techniques; and
- Bayesian techniques

The bibliography includes references to a number of papers that describe these techniques.

The techniques vary in their capacity to enable actuaries to identify past levels of independent risk. In the application of most of these techniques, one is attempting to fit a model to past systemic episodes and trends and to analyse the residual volatility once these episodes and trends have been fitted away. The better the model fit is the more likely that the residual volatility observed reflects random effects alone.

An actuary faced with the task of assessing independent risk will need to decide upon which techniques to use for each of their valuation classes. This decision should consider the extent to which the independent risk for a particular valuation class is material to the overall claim portfolio risk margin, the contribution to uncertainty from internal systemic risk and external systemic risk and the cost and effort associated with applying the techniques. Where the cost and effort outweighs the potential benefit then a simpler approach, perhaps incorporating benchmarking, may be considered.

For some valuation portfolios, the data available may be too limited or volatile to enable a credible split between past episodes of systemic risk and past independent risk. In these cases, actuaries may consider using a model that does not attempt to fit away the past systemic risk and supplement this analysis with additional allowances for external systemic risk, to the extent that this is considered to differ from past systemic risk, and internal systemic risk, which cannot be modelled using standard quantitative modelling techniques.

Independent risk assessment for outstanding claim liabilities

Any of the techniques mentioned above can be used in the assessment of past independent risk for outstanding claim purposes. Some of the techniques offer more flexibility in terms of fitting to past systemic episodes and trends.

Consideration should be given to aligning the methodology adopted to analyse uncertainty with that used for central estimate purposes. For example, if the PPCI method plays an important role in the central estimate assessment and bootstrapping is the preferred approach to analysing uncertainty then the PPCI method should be bootstrapped. This will ensure that past volatility is examined and conclusions drawn in an environment that is internally consistent.

GLM techniques can be used to model individual claims or aggregate claims. These techniques are used for reserving purposes to identify the key factors that have contributed to past claim cost outcomes. Combined with a range of useful statistical diagnostics, these techniques are well placed to support the analysis of independent risk.

Bootstrapping techniques offer less flexibility than GLM techniques but can be adapted to help in the assessment of random effects. For example, if past periods that have been largely unaffected by systemic episodes can be identified then the bootstrap residuals can be calculated for these stable periods and used as part of the bootstrapping process. Plots of residuals by accident period, development period and experience period can be used to identify periods that have been affected by past systemic episodes.

Independent risk assessment for premium liabilities

The bootstrapping, GLM and Bayesian approaches may also be used for the purpose of analysing the volatility in past claim experience for the purpose of assessing the independent risk component for premium liabilities.

However, it is possible to use simpler techniques to analyse the past volatility of key components of the premium liabilities. Consider a valuation class where the central estimate of the claim cost component of the premium liabilities is assessed by combining a projected claim frequency and average claim size. The adopted claim frequency and average claim size has been selected following an analysis of output from the outstanding claim valuation supplemented by portfolio level pricing analysis.

For some valuation classes, it can be a relatively straightforward exercise to remove the impact of past systemic episodes (including seasonality) from observed claim frequencies and determine the claim frequency CoV in respect of past residual volatility. Similarly, past average claim sizes can be adjusted to remove past inflation, including both standard and superimposed, and other past systemic episodes (again including seasonality) and a CoV in respect of past residual volatility derived.

Where a loss ratio approach to projecting premium liabilities is used, allowance should be made for systemic shifts in past premium levels as well as claim costs.

Often large claims are extracted for separate analysis. Again, observations can be made as to the aspects of past experience that represent systemic episodes and those that are purely random.

The process of identifying and isolating past systemic episodes can only be enhanced if an actuary has a strong understanding of the possible systemic sources of risk for a particular portfolio. The role that product and claim management can play in improving this understanding should not be underestimated. This is discussed further in section 4.

4. Systemic risk assessment

4.1. Internal systemic risk

Internal systemic risk refers to the uncertainty arising from the actuarial valuation models used being an imperfect representation of the insurance process as it pertains to insurance liabilities.

As discussed in section 2.5, we consider there to be three main sources of internal systemic risk. These are:

- *Specification error* - the error that can arise from an inability to build a model that is fully representative of the underlying insurance process.
- *Parameter selection error* - the error that can arise because the model is unable to measure all predictors of claim cost outcomes or trends in these predictors.
- *Data error* - the error that can arise due to poor data, unavailability of data and/or inadequate knowledge of the portfolio being analysed.

When an actuary conducts an assessment of outstanding claim or premium liabilities, there are a wide range and variety of approaches and methodologies that are available. The merits of each approach will be considered in the context of the valuation classes being assessed. The characteristics of each class and the level of information available, including granularity of data, will all play a role in the decision around which approach to use.

Although care will normally be taken to ensure that the approach adopted is appropriate for the valuation class being assessed, models are likely to represent a simplified view of the insurance process. Models also range in their capacity to identify underlying trends in the claims experience. Standard triangulations methods will normally analyse predictors (e.g. claim payments, reports, finalisations, case estimates) that have been aggregated to a reasonably high level or lag rather than lead the underlying drivers of the insurance process.

In light of this, any analysis of uncertainty would be incomplete without an objective assessment of the adequacy of the modelling infrastructure and its ability to reflect and predict the underlying insurance process. In this section of our paper, we propose one approach, involving the development of a balanced scorecard, which may be used as part of such an assessment.

One other point worth making before we walk through the balanced scorecard approach in detail is that the assessment of internal systemic risk must be conducted in the context of the actual approach used to assess the central estimate of outstanding claim and premium liabilities. The strengths and weakness associated with that approach will be considered and scored with a view to determining an appropriate allowance in risk margins for internal systemic risk. Consistency between the central estimate and risk margin assessments are one outcome of a robust assessment of internal systemic risk.

The balanced scorecard was discussed in section 2.5 and presented diagrammatically in Figure 2. In summary the approach involves:

- For each of the specification, parameter and data risk components, conduct a qualitative assessment of the modelling infrastructure, considering a range of risk indicators and scoring these indicators on a scale of 1 to 5 (where 5 represents best practice).

- Apply weights to each risk indicator, reflecting its relative importance to the overall modelling infrastructure, and calculate a weighted average score representing an objective view of the quality of the modelling infrastructure for each valuation class.
- Calibrate the weighted average score derived to a CoV in respect of internal systemic risk. The development of appropriate CoVs will likely involve a substantial amount of judgement, perhaps supplemented by quantitative analysis.

In a paper entitled *Asbestos Liabilities & the New Risk Margins Framework*, prepared by Brett Riley and Bruce Watson, the authors describe an alternative approach to assessing the level of internal systemic risk. This approach specifies High and Low scenarios that ‘represent the end points of what might be considered a reasonable range of central estimates based on alternative interpretations of all available information’. The approach advocated by Messrs Riley and Watson certainly has merit and represents a reasonable alternative to the balanced scorecard approach described in this paper. It also has the appeal of being simpler and, therefore, more practical to apply.

Scoring the modelling infrastructure

We would encourage actuaries to develop a balanced scorecard approach that is suited to the characteristics of the valuation classes within their own claims portfolio including risk indicators that are most relevant in the context of these classes. Having said this, we feel that it is useful if we outline potential risk indicators that actuaries may wish to consider and develop for the purpose of their own analysis. Table 2 includes potential risk indicators and some suggested minimum requirements for a high score for each of these indicators. The characteristics that represent a poor score should be readily apparent.

Table 2: Internal systemic risk - Potential risk indicators

Risk component	Potential risk indicators	Requirements for high score
Specification error	Number of independent models used	Many different modelling approaches considered - each approach should add value by considering different dimensions of claims experience
	Extent to which models separately analyse different claim/payment types	Relevant homogeneous claim or payment types modelled separately
	Range of results produced by models	Low variations between different models in terms of past performance - take care that comparisons are appropriate (e.g. PCE vs PPCI for old accident periods for short-tail classes may not be appropriate)
	Checks made on reasonableness of results	Significant reasonableness checks conducted, including reconciliation of movement in liabilities, diagnostic checks on valuation outcomes, acceptance of results by business, expert peer review, benchmarking against industry
	Confidence in assessment of model 'goodness of fit'	Actual vs Extected close, few difficulties in selecting parameters, relevant sensitivities yield small variances in results
	Number and importance of subjective adjustments to factors	Few subjective adjustments, relevant subjective factor sensitivites yield low variances and adjustments regularly monitored and reviewed
	Extent of monitoring and review of model and assumption performance	Model and assumption performance monitored continuously and reviewed regularly
	Ability to detect trends in key claim cost indicators	Models have performed well in detecting trends in the past
	Sophistication and performance of superimposed inflation analysis	Detailed analysis of past sources of superimposed inflation and robust quantification of each past source
	Level of expense analysis to support CHE assumptions	Detailed expense analysis, including how expenses are incurred over the lifetime of claims relating to each claim type
	Ability to model using more granular data, e.g. unit record data	Unit record data is available and used to further analyse and better understand key predictors and trends in these predictors
Parameter selection error	Best predictors have been identified, whether or not they are used	Best predictors have been analysed and identified, including internal and external variables that show strong correlaton with claims experience
	Best predictors are stable over time or change due to process changes	Predictors stable over time, stabilise quickly and respond well to process changes
	Value of predictors used	Predictors are close to best predictors, lead (rather than lag) claim cost outcomes, modelled rather than subjectively allowed for and unimpaired by past systemic events
Data error	Knowledge of past processes affecting predictors	Good and credible knowledge of past processes, including changes to processes
	Extent, timeliness, consistency and reliability of information from business	Regular, complete and pro-active two-way communication between valuation actuary and claims staff/portfolio managers who understand key valuation predictors and how changes may impact or invalidate these
	Data subject of appropriate reconciliations and quality control	Reconciliations against other sources are conducted for all data sources and types, checks are conducted throughout data processing steps, reconciliations against previous valuation conducted, data and differences well understood
	Processes for obtaining and processing data are robust and replicable	No past instances of poor data understanding, no or low potential for miscoding of claim type
	Frequency and severity of past mis-estimation due to revision of data	No past instances of data revision
	Extent of current data issues and possible impact on predictors	No known current data issues

Each of the risk indicators should be considered in the context of both the outstanding claim and premium liabilities. Additional indicators may be considered for premium liabilities, for example whether the outstanding claim liabilities are used as an input to the premium liability assessment or whether credible portfolio level pricing analysis is used as an input to the premium liability assessment.

For certain short-tail portfolios, some risk indicators may not be as relevant for premium liability purposes. A large variance in the outstanding claim liabilities, which might only affect the most recent accident periods and have a relatively small impact on the projected ultimate claim frequency or average claim size, may not be material in the context of a premium liability assessment.

Table 3 shows the risk indicator scores which underpin the internal systemic risk CoVs adopted for Insurer ABC in the example in Figure 3 in section 2.6, with a particular focus on outstanding claim liabilities.

Table 3: Internal systemic risk – example balanced scorecard

Risk component	Potential risk indicators	Motor score OSC	Motor weight	Home score OSC	Home weight	CTP score OSC	CTP weight
Specification error	Number of independent models used	4	7	4	7	3	2
	Extent to which models separately analyse different claim/payment types	3	3	4.5	5	2	7
	Range of results produced by models	4	5	4	4	2	2
	Checks made on reasonableness of results	5	5	5	5	4	5
	Confidence in assessment of model 'goodness of fit'	4	5	4	5	2	7
	Number and importance of subjective adjustments to factors	5	3	4	3	3	5
	Extent of monitoring and review of model and assumption performance	4	5	4	5	5	8
	Ability to detect trends in key claim cost indicators	4	4	3	4	3	6
	Sophistication and performance of superimposed inflation analysis		0		0	4	10
	Level of expense analysis to support CHE assumptions	4	4	4	4	2	2
	Ability to model using more granular data, e.g. unit record data	2	2	2	2	5	2
Parameter selection error	Best predictors have been identified, whether or not they are used	4	3	4	5	3	7
	Best predictors are stable over time or change due to process changes	5	5	4	5	2	6
	Value of predictors used	4	5	4	5	3	5
Data error	Knowledge of past processes affecting predictors	4	8	4	8	4	8
	Extent, timeliness, consistency and reliability of information from business	4	5	4	5	4	5
	Data subject of appropriate reconciliations and quality control	4	7	4	7	4	8
	Processes for obtaining and processing data are robust and replicable	5	3	5	3	5	3
	Frequency and severity of past mis-estimation due to revision of data	5	3	3	3	5	5
	Extent of current data issues and possible impact on predictors	4	3	5	3	5	3
Total weighted average score - outstanding claims (OSC)		4.1		4.0		3.5	
Total weighted average score - premium liabilities		4.5		4.5		4.0	

The scores and weights shown in Table 3 are for illustration only and should be taken as a demonstration of concept than as a set of benchmarks that actuaries can use for such portfolios in practice.

The weights allocated to each of the risk indicators are a measure of the importance of that risk indicator, relative to the other risk indicators, in terms of its contribution to overall internal systemic risk. The weights and hence relativities between risk indicators should reflect the particular valuation infrastructure adopted for each valuation class including the relative importance of each risk indicator in the context of that valuation class.

Premium liabilities scored better than outstanding claims in this example due to the extensive use in their assessment of outcomes from the valuation of outstanding claims **and** independent and credible portfolio level pricing analyses conducted recently.

Calibrating scores to CoVs

Once a score representing an objective and qualitative view of the efficacy of the modelling infrastructure has been derived, one needs to determine a CoV that is an appropriate representation of the contribution to outstanding claim and premium liability uncertainty from internal systemic risk. This step is likely to require a significant amount of subjective judgement, supplemented by quantitative analysis.

We suggest that individual actuaries develop a CoV scale which represents their view of the uncertainty associated with internal systemic risk for the full range of possible balanced scorecard outcomes, ranging from worst practice to best practice (or 'perfect') modelling approaches. A large degree of judgement will be required to derive a reasonable range in the context of a particular claims portfolio. The analysis conducted to score the modelling infrastructure together with past model performance should provide invaluable insights into the potential variability associated with a particular modelling approach.

If more than one methodology has been deployed in the past then a hindsight analysis of the actual past performance of each method can be used to assess the relative performance of each method and the extent to which multiple models can improve the performance of the whole modelling infrastructure.

Mechanical hindsight analysis (see section 2.7) may also be used to help in the assessment of internal systemic risk. For example, a mechanical hindsight analysis can be conducted using one method with all claim or payment types aggregated. A further retrospective analysis can be conducted using multiple methods with claim or payment types separated into individual homogeneous groups. The relative difference in performance of the two modelling infrastructures over time may give some insights into the additional uncertainty associated with poor modelling approaches compared to fair or good modelling approaches.

Based on our experience, we would suggest that the minimum CoV associated with a 'perfect' model is unlikely to be much less than 5%. Even a 'perfect' model will not be able to completely replicate the true underlying insurance process or identify every possible predictor of claim cost outcomes.

If you consider a single, aggregated model with limited data or information available to populate the model, significant subjective assumptions required and few identified predictors, CoVs of 20% or above in respect of internal systemic risk are readily justifiable. For such models, it is not infeasible that internal systemic risk could be the main contributor to overall uncertainty.

Table 4 gives CoV scales used in the assessment of risk margins for Insurer ABC as part of the example in Figure 3.

Table 4: Internal systemic risk – example CoV scale

Score from balanced scorecard assessment	Motor CoV	Home CoV	CTP CoV
1.0 to 1.5	17.5%	17.5%	25.0%
1.5 to 2.0	13.0%	13.0%	20.5%
2.0 to 2.5	10.5%	10.5%	17.0%
2.5 to 3.0	8.5%	8.5%	14.0%
3.0 to 3.5	7.0%	7.0%	11.5%
3.5 to 4.0	6.0%	6.0%	9.5%
4.0 to 4.5	5.5%	5.5%	8.0%
4.5 to 5.0	5.0%	5.0%	7.0%

The CoV scale shown in Table 4 is an example only. Actuaries should select CoV scales that are appropriate in the context of their own valuation classes and the modelling infrastructure adopted for each of those valuation classes. Any hindsight analysis deployed to support the selection of appropriate CoVs should be designed to align with the actual valuation methods adopted for the valuation classes being analysed.

Further comments on the CoV scale as presented in Table 4 are:

- The scale is not linear reflecting our view that the marginal improvement in outcomes between fair and good modelling infrastructures is less than the marginal improvement between poor and fair modelling infrastructures.
- The CoVs for CTP, a long-tail portfolio, are higher than those for Motor and Home, both short-tail portfolios. For long-tail portfolios, it is generally more difficult to develop a modelling approach that is representative of the underlying insurance process. Also, key predictors are often less stable for long-tail portfolios and past episodes of systemic risk more likely to impair the ability to fit a good model.
- The scale has been used for both outstanding claim and premium liability purposes. A reasonable 'a priori' assumption is that similar scales can be used for both. Arguments can be made for premium liabilities to have higher or lower CoVs than those applying to outstanding claim liabilities, particularly for poor modelling approaches. For example, the assessment of premium liabilities may include additional uncertainty associated with the estimation of exposure or premium relating to unclosed or contractually bound future business. If this is the case then a loading on top of the outstanding claim liability CoVs may be justifiable. On the other hand, for certain stable short-tail classes, the difference between a simple loss ratio approach and a more thorough frequency/severity approach may not be material in terms of performance in the assessment of premium liabilities but the difference between a single aggregate model and multiple disaggregated models could be material in terms of performance in the assessment of outstanding claim liabilities.

4.2. External systemic risk

External systemic risk refers to the uncertainty arising from non-random risks external to the actuarial modelling process. This uncertainty encapsulates systemic episodes that have not yet occurred but may emerge in the future and those that are emerging in the recent experience but where there is some uncertainty as to how they will develop in future. The risk associated with the actuarial modelling infrastructure potentially being unable to identify emerging risks will be picked up as part of a robust internal systemic risk assessment.

Certain stochastic quantitative approaches may be used to gain insights into past and emerging sources of external systemic risk. These insights, together with those gained from the central estimate analysis, will provide useful intelligence on the type of risks that can emerge in each valuation portfolio, at least the ones that have emerged in the past. However, one cannot readily assume that past experience is a reasonable reflection of the future. A more rigorous approach should consider each of the possible future sources of external systemic risk, using a number of sources of information.

Communication with business experts

Typically actuaries will hold discussions with portfolio and claim management as part of the valuation process. These discussions normally provide valuable insights into emerging trends and possible future sources of external systemic risk. However, the focus is normally on gaining an appropriate level of portfolio understanding to enable an informed assessment of the central estimate of outstanding claim and premium liabilities. Although the information gathered will play a role in the assessment of risk margins, this tends to be more an afterthought than a key focus of discussions.

Discussions can be readily tailored to topics of relevance for both central estimate and risk margin purposes and ensure an appropriate level of focus on both aspects of the valuation process. Business management should be given time to prepare for these meetings to ensure that the valuation actuary gains the maximum possible benefit from them.

From a risk margins perspective, the focus of these meetings should be on the identification of key potential sources of systemic risk, including those that have begun to emerge and those that may emerge in future. Discussions should consider all aspects of the portfolio management process, including underwriting and risk selection, pricing, claims management, expense management, emerging portfolio trends and the environment within which the portfolio operates. Once the key sources of external systemic risk have been identified, they can be categorised for analysis purposes. As well as identifying key risks, the quantification of risk should be another key consideration for business management interactions.

Selection of assumptions

The selection of CoVs for each risk category will involve a combination of quantitative analysis and qualitative judgement. Some risk categories will be more open to quantitative analysis than others. For those categories where such analysis is more difficult, sensitivity analysis, perhaps in conjunction with business management, may shed some light on the range of possible outcomes.

In assessing CoVs in respect of each risk category, it is also important to consider the shape of the entire distribution, to the extent possible. Some risks will demonstrate characteristics that are reflective of a highly skewed distribution and, as such, may not have a material bearing on a 75^{th} percentile risk margin but may be more relevant for higher probabilities of adequacy. An example of such a risk is latent risk where the probability of such risk emerging is very low and certainly lower than 25%. Certain sources of superimposed inflation may also be considered to belong to this category.

In focusing efforts, consideration may be given to ranking individual risks for each valuation class in order of importance, separately for outstanding claim and premium liabilities. For a number of valuation classes it is quite likely that such an exercise will identify a small number of key risks and allow efforts to be focused accordingly. This might also provide justification for excluding certain risk categories that are deemed to be immaterial in terms of their contribution to the overall CoV. A scoring system, developed in conjunction with business experts, may be introduced as a convenient mechanism for ranking individual risks and checking that the contributions from individual risks to the overall CoV for external systemic risk are reasonable.

Each risk category will represent the amalgamation of a number of identified potential sources of risk. In some cases, these individual risks will be correlated and allowance will need to be made for this when combining the risks to determine a CoV for the risk category as a whole. A simple approach, similar to that discussed in section 2.5, may be used to allow for these intra-risk category correlation effects.

A key consideration when determining risk categories for a particular valuation class is whether there is any correlation between categories. The consolidation of the analysis of external systemic risk is substantially simplified if one can assume that each of the risk categories is independent. Certain risk categories may have to be combined to ensure that this assumption is valid.

In the balance of this section, we explore each of the risk categories discussed in section 2.5 with a view to providing some insights into the types of risk that may be included in each risk category and the analysis that may be conducted to estimate appropriate CoVs for each category.

Economic and social risks

This risk category incorporates a number of potential sources of external systemic risk. These sources include, but are not limited to, levels of standard inflation (AWE and CPI), general economic conditions (unemployment rates, GDP growth, interest rates, asset returns), fuel prices, driving patterns, etc.

Some of these risks can have a material impact on both outstanding claim and premium liabilities. Others are material only for premium liabilities. For example, economic conditions can have a material impact on outstanding claim and premium liabilities for professional lines and builder’s warranty valuation classes. Uncertainty around driving conditions, on the other hand is less relevant for motor outstanding claims than it is for motor premium liabilities.

Uncertainty around AWE and/or CPI will impact all valuation classes. Due to the longer term settlement for long tail classes, AWE uncertainty is somewhat more material for these classes than for short tail valuation classes. Analysis of past levels of AWE and CPI can shed some light on past systemic sources of volatility. Economic commentators often provide insights into the potential sources of volatility.

Any analysis of past levels of inflation should consider the extent to which past volatility is random and the extent to which it has been impacted by systemic events. For the purpose of analysis of systemic sources of risk, we are only interested in the latter. This applies to the analysis of past experience in respect of any systemic event in any risk category.

Potential systemic shifts in claim frequency for short tail valuation portfolios should be included in this risk category.

Legislative, political and claims inflation risk

These risks have been combined, for convenience, into one risk category since they are often correlated. For example, the risks associated with the legislative and political environment are often correlated to the drivers of non-standard claims inflation for long tail valuation classes.

This risk category is likely to be much more material for long tail valuation classes than for short tail classes. For long tail classes, in particular, a number of potentially material risks may be identified and allocated to this risk category. Some of these risks will be correlated and, as such, quantification should make allowance for this correlation.

The analysis conducted to quantify CoVs for this risk category can also be used to justify superimposed inflation assumptions for central estimate valuation purposes. After all, for long tail valuation classes, the risks in this category are normally aggregated and referred to as superimposed inflation for insurance liability valuation purposes. For each risk, one is aiming to form a view of the range of possible impacts on claim cost outcomes. The average of this range, combined across all risks, provides an estimate of superimposed inflation.

Individual actuaries will identify the key risks in this category in the context of their own claims portfolio. As a general guide, for long tail classes, this category would be expected to include some of the following sub-groups of risk:

- Impact of recent legislative amendments, including possibility of erosion of intent of amendments through assessment and threshold erosion, changes in court interpretation, etc.
- Potential for future legislative amendments with retrospective impacts.
- Precedent setting in courts, including impact of judicial decisions perhaps leading to new heads of damage.
- Changes to medical technology costs
- Changes to legal costs
- Systemic shifts in large claim frequency or severity

Typically, actuaries will have access to various forms of analysis relating to the potential impact of a specific series of legislative amendments. This information may include both external and internal analyses, the latter possibly tailored to the specifics of a particular portfolio. When supplemented by discussions with product and claim management, a sound understanding of the range of possible outcomes can be obtained, including the likelihood and potential severity of a particular outcome occurring.

For short tail classes, this risk category includes the risk that claim inflation will increase at a level different from that adopted for central estimate purposes, in addition to that arising from standard inflation (see above) or claim management process risk (see below). Claim cost reduction initiatives would normally be allocated to this category and information is sometimes available as to the range of possible outcomes from such initiatives.

Claim management process change risk

Changes to the claim management process can impact all valuation classes. Typically, however, such changes will have a more material impact on some valuation classes than others. The key here is to work closely with claim managers to gain a sound understanding of the claim management philosophy and the process that underpins that philosophy. Current or future potential changes to process should be identified as part of such discussions.

Analysis of past experience will help identify past systemic episodes that may have been impacted by the claim management process. Discussions with management may help isolate the process changes that contributed to those systemic episodes. Reporting patterns, payment patterns, finalisation and reopening rates and case estimation processes should all be considered as part of these discussions.

Sensitivity testing of key valuation assumptions, which can be useful in the assessment of CoVs for this risk category, is relatively straightforward using traditional triangulation techniques. If such analysis is conducted, sensitivities considered should be aligned with the potential sources of uncertainty identified following discussions with claim management.

Claim management process risk is likely to be more relevant for outstanding claim liabilities than for premium liabilities. For outstanding claim liabilities, particularly for short tail valuation classes, this risk can be material since it impacts the pattern of emergence of credible claim estimates. For premium liabilities, we are more interested in the extent to which changes to claim management processes can impact the magnitude of the claim cost. The impact on claim emergence is normally of secondary importance.

Expense risk

One would generally expect this to be a small contributor to total external systemic risk.

Ideally, one would spend time with product and claim management to understand the key drivers of policy maintenance and claim handling expenses. Armed with a good understanding of these drivers, a valuation actuary can identify the key sources of possible variation relative to the central estimate assumptions. Sensitivity testing around the key drivers, preferably conducted in association with informed business and process experts, and analysis of past expense levels with a view to identifying past systemic effects can be combined to help form a reasonable view as to the range of possible claim cost outcomes. Such an analysis could be conducted in conjunction with any expense analysis conducted for central estimate expense assumptions.

Event claims can have a material impact on the level of claim handling expenses. The larger an event, the smaller the fixed component of the event management cost will be as a percentage of the claim cost. In light of this, the analysis may benefit from including claim handling expenses in respect of event claims with the analysis of event risk itself.

Event risk

Event risk relates to single events which give rise to a large number of claims. This risk is likely to be material for property and, to a lesser extent, motor valuation classes but will be insignificant for most other valuation classes. Event risk also arises in medical malpractice and builders' warranty portfolios where a large number and/or cost of claims can arise from one source, i.e. a single doctor or a single builder.

The approach to assessing event risk will differ materially between outstanding claim and premium liabilities. For outstanding claim liabilities, the approach will be defined by the extent to which there are material outstanding events. If there are, then these should be analysed separately. Discussions with event claim management should be held to understand their expectations as to claim cost outcomes and to identify any specific issues that may influence outcomes. The range of development patterns for previous events may also influence the view on uncertainty.

There is often a wealth of information available to help in the quantification of event risk for premium liabilities, including:

- Past experience in respect of event claims. When analysing past experience, it is important to allow for changes in portfolio size, geographical spread, inflation, policy terms and conditions, reinsurance arrangements, etc. where these are considered to be material. It is not particularly difficult, where sufficient credible past experience is available, to build a relatively simple statistical model with key frequency and severity assumptions based on appropriately adjusted past experience. In fact, modelling of this nature may have been conducted by pricing actuaries or as part of a reinsurance placement and can be adapted for event risk analysis.
- Output from proprietary catastrophe modelling. A number of such models are used in practice, including those developed by RMS, EQE, AIR and Risk Frontiers. Insurers will normally have access to these models through their reinsurance intermediaries who are well placed to provide advice on the range of possible outcomes based on modelled events.
- Reinsurance intermediaries typically also have available models in respect of natural perils, and some man-made perils, that can be used to model perils not covered by proprietary catastrophe models. These, together with the proprietary models, will normally be used by intermediaries in support of an insurer's catastrophe reinsurance program renewal and can be readily extended to provide advice on the uncertainty associated with event risk.

Latent claim risk

Latent claim risk is negligible for most valuation classes. For some, primarily workers compensation and liability classes, the risk can be considered to be material. However, this is one of the most difficult risks to quantify. The probability of these events is low but the impact should they occur could be substantial

Purely in the context of setting risk margins it is unlikely that analysis of latent claims risk warrants a substantial commitment of resources given that it is such a low probability event. However if such risk exposure is significant enough to be a formal component of the central estimate or if the object of the exercise is modelling extreme risks for capital adequacy purposes (using a DFA approach) then a thorough examination of this risk driver is certainly warranted.

.

This risk is the one most likely to be quantified using a large degree of judgement. Discussions with underwriters may help shed some light on some potential sources and give a feel for their likelihood and potential impact. Also, casualty reinsurance underwriters often have a more informed understanding of the potential sources of latent risk claims from their dealings with a number of direct insurers globally. Using all of the information collected, scenarios may be developed to reflect a possible range of scenarios from which reasonable CoVs can be derived.

Recovery risk

This risk category encapsulates systemic uncertainty in relation to reinsurance and non-reinsurance recoveries. This category is likely to be relatively insignificant for most portfolios. One possible exception is motor valuation classes where third party recoveries are often a material consideration.

The focus here should be on systemic events that may lead to different recovery outcomes from those adopted for central estimate purposes.

An analysis of past non-reinsurance recovery rates and patterns will inform on past systemic events. Combined with discussions with claim management around current trends in recovery management and any current or planned future initiatives that may impact recovery levels, one can readily form a view as to the range of possible systemic outcomes.

Reinsurance recoverability is another potential source of external systemic risk that should be considered within this category. The extent to which this is material will depend on the reinsurance arrangements themselves. A material shift in reinsurance market conditions may significantly alter the ability to recover from reinsurers. For example, one or more catastrophic events (on a global scale) or a downturn in asset returns, or a combination of both, may substantially reduce the ability to recover from reinsurers. The probability of such events occurring and materially impacting recoveries is low but the severity, should they happen, could be high. Discussions with reinsurance management are often enlightening and can help in the identification of possible scenarios, the likelihood of them occurring and the quantitative impact should they occur.

Bibliography

Bateup R and Reed I, 2001, *Research and Data Analysis Relevant to the Development of Standards and Guidelines on Liability Valuation for General Insurance*, Proceedings of the 13th General Insurance Seminar, Institute of Actuaries of Australia.

Collings S and White G, 2001, *APRA Risk Margin Analysis*¸ Proceedings of the 13th General Insurance Seminar.

Houltram A, 2003, *Reserving Judgement – Considerations Relevant to Insurance Liability Assessment Under GPS210*, Proceedings of the 14th General Insurance Seminar.

Houltram A, 2005, *An Empirical Approach to Insurance Liability Predicition Error Assessment With Application to APRA Risk Margin Determination*, Proceedings of the 15th General Insurance Seminar.

O'Dowd C, Smith A and Hardy P, 2005, *A Framework for Estimating Uncertainty in Insurance Claims Cost*, Proceedings of the 15th General Insurance Seminar.

Riley B and Watson B, 2009, *Asbestos Liabilities & the New Risk Margins Framework*, Presented to the Institute of Actuaries of Australia 2009 Accident Compensation Seminar.

ROC Working Party, 2008, *Reserving Uncertainty*, Paper for GIRO 2008

Stochastic Reserving Papers

England P.D. and Verrall R.J., 1999, *Analytic and Bootstrap Estimatesof Prediciton Errors in Claims Reservng*, Insurance: Mathematics and Economics 25

England P.D. and Verrall R.J., 2002, *Stochastic Claims Reserving in General Insurance*, British Actuarial Journal 8, Part 3

England P.D. and Verrall R.J., 2006, *Predicitve Distributions of Outstanding Liabilities in General Insurance*, Annals of Actuarial Science 1, Part 2

Kirschner G.S., Kerley C and Isaacs B, *Two Approaches to Calculating Correlated Reserve Indications Across Multiple Lines of Business*

Mack T, 1993, *Measuring the Variability of Chain Ladder Reserve Estimates*

Taylor G and McGuire G, 2005, *Synchronous Bootstrapping of Seemingly Unrelated Regressions*

Verrall R.J., 2004, *A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving*, North American Actuarial Journal, Volume 8, Number 3

CAS MONOGRAPH SERIES
NUMBER 1

STOCHASTIC LOSS RESERVING USING BAYESIAN MCMC MODELS

Glenn Meyers, FCAS, MAAA, CERA, Ph.D.

CASUALTY ACTUARIAL SOCIETY

The emergence of Bayesian Markov Chain Monte-Carlo (MCMC) models has provided actuaries with an unprecedented flexibility in stochastic model development. Another recent development has been the posting of a database on the CAS website that consists of hundreds of loss development triangles with outcomes. This monograph begins by first testing the performance of the Mack model on incurred data, and the Bootstrap Overdispersed Poisson model on paid data. It then will identify features of some Bayesian MCMC models that improve the performance over the above models. The features examined include (1) recognizing correlation between accident years; (2) introducing a skewed distribution defined over the entire real line to deal with negative incremental paid data; (3) allowing for a payment year trend on paid data; and (4) allowing for a change in the claim settlement rate. While the specific conclusions of this monograph pertain only to the data in the CAS Loss Reserve Database, the breadth of this study suggests that the currently popular models might similarly understate the range of outcomes for other loss triangles. This monograph then suggests features of models that actuaries might consider implementing in their stochastic loss reserve models to improve their estimates of the expected range of outcomes.

Glenn Meyers, FCAS, MAAA, CERA, and Ph.D., retired from ISO in 2011 after a 37-year career as an actuary. He holds a B.S. in Mathematics and Physics from Alma College, an M.A. in Mathematics from Oakland University, and a Ph.D. in Mathematics from SUNY at Albany. A frequent speaker at Casualty Actuarial Society (CAS) meetings, he has served, and continues to serve, the CAS and the International Actuarial Association on various research and education committees. He has also served on the CAS Board of Directors. Over the years, he has published articles in the *Proceedings* of the Casualty Actuarial Society, *Variance,* and the *Actuarial Review.* His research and other service contributions have been recognized by the CAS on numerous occasions. He has received the Woodward-Fondiller Prize on three separate occasions, the Dorweiller Prize twice, the DFA Prize, the Reserves Prize, the Matthew Rodermund Service Award, and the Michelbacher Significant Achievement Award. In retirement he still devotes a good portion of his time in pursuit of his passion for actuarial research.

STOCHASTIC LOSS RESERVING USING BAYESIAN MCMC MODELS

Glenn Meyers, FCAS, MAAA, CERA, Ph.D.

Casualty Actuarial Society
4350 North Fairfax Drive, Suite 250
Arlington, Virginia 22203
www.casact.org
(703) 276-3100

Stochastic Loss Reserving Using Bayesian MCMC Models
By Glenn Meyers

Copyright 2015 by the Casualty Actuarial Society.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. For information on obtaining permission for use of material in this work, please submit a written request to the Casualty Actuarial Society.

Library of Congress Cataloging-in-Publication Data
Meyers, Glenn
Stochastic Loss Reserving Using Bayesian MCMC Models / Glenn Meyers
ISBN 978-0-9624762-7-3 (print edition)
ISBN 978-0-9624762-6-6 (electronic edition)
1. Actuarial science. 2. Loss reserving. 3. Insurance—Mathematical models.
I. Meyers, Glenn.

Copyright 2015, Casualty Actuarial Society

Contents

2015 CAS Monograph Editorial Board

C. K. Stan Khury, Editor in Chief
Emmanuel Bardis
Craig Davis
Richard Fein
Jesse Groman
Ali Ishaq
Leslie Marlo
Sholom Feldblum, consultant
Glenn Meyers, consultant
Katya Prell, consultant

Foreword

This is the inaugural volume of the new CAS Monograph Series. A CAS monograph is an authoritative, peer reviewed, in-depth work on an important topic broadly within property and casualty actuarial practice.

In this monograph Glenn Meyers introduces a novel way of testing the predictive power of two loss reserving methodologies. He first demonstrates that the method commonly used for incurred losses tends to understate the range of possible outcomes. For paid losses, both methods tend to overstate the range of expected outcomes. Then he proceeds to apply Bayesian Markov Chain Monte-Carlo models (Bayesian MCMC) to improve the predictive power by recognizing three different elements implicit in the data histories. He is careful to note that the results are based on the histories contained in the CAS Database (of loss development triangles), which prevents one from making broad unqualified statements about the conclusions drawn in this work.

This monograph lays a solid foundation for future development and research in the area of testing the predictive power of loss reserving methods generally and in the use of Bayesian MCMC models to improve confidence in the selection of appropriate loss reserving methods. Glenn Meyers manages to show the way for raising the performance standard of what constitutes a reliable loss reserving methodology in any given situation.

C. K. "Stan" Khury
Chairperson
Monograph Editorial Board

1. Introduction

The recent attempts to apply enterprise risk management principles to insurance has placed a high degree of importance on quantifying the uncertainty in the various necessary estimates with stochastic models. For general insurers, the most important liability is the reserve for unpaid losses. Over the years a number of stochastic models have been developed to address this problem. Two of the more prominent nonproprietary models are those of Mack (1993, 1994) and England and Verrall (2002).

While these, and other, models provide predictive distributions[1] of the outcomes, very little work has been done to retrospectively test, or validate,[2] the performance of these models in an organized fashion on a large number of insurers. Recently with the permission of the National Association of Insurance Commissioners (NAIC), Peng Shi and I, in Meyers and Shi (2011), were able to assemble a database consisting of a large number of Schedule P triangles for six lines of insurance. These triangles came from insurer NAIC Annual Statements reported in 1997. Using subsequent annual statements we "completed the triangle" so that we could examine the outcomes and validate, the predictive distribution for any proposed model.

Sections 3 and 4 attempt to validate the models of Mack (1993, 1994) and England and Verrall (2002). As it turns out, these models do not accurately predict the distribution of outcomes for the data included in the subject database. Explanation for these results include the following.

- The insurance loss environment is too dynamic to be captured in a single stochastic loss reserve model. I.e., there could be different "black swan" events that invalidate any attempt to model loss reserves.[3]
- There could be other models that better fit the existing data.
- The data used to calibrate the model is missing crucial information needed to make a reliable prediction. Examples of such changes could include changes in the way the underlying business is conducted, such as changes in claim processes or changes in the direct/ceded/assumed reinsurance composition of the claim values in triangles.

---

[1] In this monograph, the term "predictive distribution" will mean the distribution of a random variable, X, given observed data x. By this definition the range of outcomes, X, could be quite wide. This, in contrast to the common usage of the term "predict," connotes an ability to foresee the future and, in the context of the subject matter of this monograph, implies a fairly narrow range of expected outcomes.

[2] An explanation of "validate" will be given in Section 3.

[3] The term "black swan," as popularized by Taleb [2007], has come to be an oft-used term representing a rare high-impact event.

Possible ways to rule out the first item above are to (1) find a better model; and/or (2) find better data. This monograph examines a number of different models and data sources that are available in Schedule P. The data in Schedule P includes net paid losses, net incurred losses, and net premiums.

A characteristic of loss reserve models is that they are complex in the sense that they have a relatively large number of parameters. A major difficulty in quantifying the uncertainty in the parameters of a complex model has been that it takes a fair amount of effort to derive a formula for the predictive distribution of outcomes. See Mack's (1993, 1994) papers and Bardis, Majidi and Murphy's (2012) paper for examples of analytic solutions. Taking advantage of the ever-increasing computer speed, England and Verrall (2002) pass the work on to computers using a bootstrapping methodology with the overdispersed Poisson distribution (ODP). Not too long ago, the Bayesian models[4] were not practical for models of any complexity. But with the relatively recent introduction of Bayesian Markov Chain Monte Carlo (MCMC) models, complex Bayesian stochastic loss reserve models are now practical in the current computing environment.

Although Markov chains have long been studied by probability theorists, it took a while for their application to Bayesian statistics to be recognized. Starting in the 1930s, physicists began using statistical sampling from Markov chains to solve some of the more complex problems in nuclear physics. The names associated with these efforts include Enrico Fermi, John von Neumann, Stanislaw Ulam and Nicolas Metropolis. This led to the Metropolis algorithm for generating Markov chains. Later on, W. Keith Hastings (1970) recognized the importance of Markov chains for mainstream statistics and published a generalization of the Metropolis algorithm. That paper was largely ignored by statisticians at the time as they were not accustomed to using simulations for statistical inference. Gelfand and Smith (1990) provided the "aha" moment for Bayesian statisticians. They pulled together a relevant set of existing ideas at a time when access to fast computing was becoming widely available. In the words of McGrayne (2011, Part V): "Almost instantaneously MCMC and Gibbs sampling changed statisticians' entire method of attacking problems. In the words of Thomas Kuhn, it was a paradigm shift. MCMC solved real problems, used computer algorithms instead of theorems, and led statisticians and scientists into a world where 'exact' meant 'simulated' and repetitive computer simulations replaced mathematical equations. It was a quantum shift in statistics" (p. 225).

As was the case for the other social sciences, Bayesian MCMC should eventually have a profound effect on actuarial science. And in fact, its effect has already begun. Scollnik (2001) introduced actuaries to Bayesian MCMC models. De Alba (2002) along with Ntzoufras and Dellaportas (2002) quickly followed by applying these models to the loss reserving problem. Verrall (2007) applied them to the chain ladder model. In the time since these papers were written, the algorithms implementing

---

[4] By a "Bayesian model" I mean a model with its parameters having a prior distribution specified by the user. By "Bayesian estimation" I mean the process of predicting the distribution of a "statistic of interest" from the posterior distribution of a Bayesian model.

Bayesian MCMC models have gotten more efficient, and the associated software has gotten more user friendly.

Here is the situation we now face. First, we are able to construct a wide variety of proposed models and predict their distribution of outcomes with the Bayesian MCMC methodology. Second, we are able to validate a proposed stochastic loss reserve model using a large number of insurers on the CAS Loss Reserve Database. If the insurance loss environment is not dominated by a series of unique "black swan" events, it should be possible to systematically search for models and data that successfully validate. This monograph describes the results I have obtained to date in my pursuit of this goal.

While I believe I have made significant progress in identifying models that do successfully validate on the data I selected from the CAS Loss Reserve Database, it should be stressed that more work needs to be done to confirm or reject these results for different data taken from different time periods.

The intended audience for this monograph consists of general insurance actuaries who are familiar with the Mack (1993, 1994) and the England and Verrall (2002) models. While I hope that most sections will be readable by a "generalist" actuary, those desiring a deeper understanding should work with the companion scripts to this monograph.[5]

The computer scripts used to implement these models is written in the R programming language. To implement the MCMC calculations the R script contains another script that is written in JAGS. Like R, JAGS is an open source programming language one can download for free. For readers who are not familiar with R and JAGS, here are some links to help the reader get started.

- http://opensourcesoftware.casact.org/start This link goes to the home page of the CAS Open Source Software Committee. This page gives several other links that help one start using R and JAGS.
- http://r-project.org The home page of the R-Project.
- http://mcmc-jags.sourceforge.net/ A link to download JAGS.
- http://www.rstudio.com/ A currently popular editor for R and JAGS script.

[5] Scripts are available at www.casact.org/pubs/monographs/meyers/Monograph_Tables_and_Scripts.xlsx

2. The CAS Loss Reserve Database

In order to validate a model, one need not only the data used to build the model, but also the data with outcomes that the model was built to predict. Schedule P of the NAIC Annual Statement contains insurer-level run-off triangles of aggregated losses by line of insurance. Triangles for both paid and incurred losses (net of reinsurance) are reported in Schedule P.[6] To get the outcomes, one must look at subsequent Annual Statements.

To illustrate the calculations in this monograph, I selected incurred and paid loss triangles from a single insurer in the database, whose data are in Tables 1, 2 and 3. The data in the loss triangles above the diagonal lines are available in the 1997 Annual Statement. These data are used to build the models discussed below. The outcome data below the diagonal lines were extracted, by row, from the Annual Statements listed in the "Source" column. These data are used to validate the models.

The database, along with a complete description of how it was constructed and how the insurers were selected, is available on the CAS website at http://www.casact.org/research/index.cfm?fa=loss{us}reserves{us}data.

This monograph will fit various loss reserve models, and test the predictive distributions, to a set of 200 insurer loss triangles taken from four Schedule P (50 from each of Commercial Auto, Personal Auto, Workers Compensation and Other Liability) lines of insurance. An underlying assumption of these models is that there have not been any substantial changes in the insurer's operation. In our real world, insurers are always tinkering with their operations. Schedule P provides two hints of possible insurer operational changes:

- Changes in the net premium from year-to-year
- Changes in the ratio of net to direct premium from year to year

The criteria for selecting the 200 insurer loss triangles rests mainly on controlling for changes in the above two items. Appendix A gives the group codes for the selected insurers by line of insurance and gives a detailed description of the selection algorithm.

[6] Paid losses are reported in Part 3 of Schedule P. Incurred losses are the losses reported in Part 2 minus those reported in Part 4 of Schedule P.

Table 1. Illustrative Insurer Net Written Premium

AY	1	2	3	4	5	6	7	8	9	10
Premium	5812	4908	5454	5165	5214	5230	4992	5466	5226	4962

Table 2. Illustrative Insurer Incurred Losses Net of Reinsurance

AY/Lag	1	2	3	4	5	6	7	8	9	10	Source
1988	1722	3830	3603	3835	3873	3895	3918	3918	3917	3917	1997
1989	1581	2192	2528	2533	2528	2530	2534	2541	2538	2532	1998
1990	1834	3009	3488	4000	4105	4087	4112	4170	4271	4279	1999
1991	2305	3473	3713	4018	4295	4334	4343	4340	4342	4341	2000
1992	1832	2625	3086	3493	3521	3563	3542	3541	3541	3587	2001
1993	2289	3160	3154	3204	3190	3206	3351	3289	3267	3268	2002
1994	2881	4254	4841	5176	5551	5689	5683	5688	5684	5684	2003
1995	2489	2956	3382	3755	4148	4123	4126	4127	4128	4128	2004
1996	2541	3307	3789	3973	4031	4157	4143	4142	4144	4144	2005
1997	2203	2934	3608	3977	4040	4121	4147	4155	4183	4181	2006

Table 3. Illustrative Insurer Paid Losses Net of Reinsurance

AY/Lag	1	2	3	4	5	6	7	8	9	10	Source
1988	952	1529	2813	3647	3724	3832	3899	3907	3911	3912	1997
1989	849	1564	2202	2432	2468	2487	2513	2526	2531	2527	1998
1990	983	2211	2830	3832	4039	4065	4102	4155	4268	4274	1999
1991	1657	2685	3169	3600	3900	4320	4332	4338	4341	4341	2000
1992	932	1940	2626	3332	3368	3491	3531	3540	3540	3583	2001
1993	1162	2402	2799	2996	3034	3042	3230	3238	3241	3268	2002
1994	1478	2980	3945	4714	5462	5680	5682	5683	5684	5684	2003
1995	1240	2080	2607	3080	3678	2004	4117	4125	4128	4128	1997
1996	1326	2412	3367	3843	3965	4127	4133	4141	4142	4144	2005
1997	1413	2683	3173	3674	3805	4005	4020	4095	4132	4139	2006

3. Validating the Mack Model

Probably the two most popular nonproprietary stochastic loss reserve models are the Mack (1993, 1994) chain-ladder model and the England and Verrall (2002) bootstrap ODP model. This section describes an attempt to validate the Mack model on the incurred loss data from several insurers that are included in the CAS database. Validating the bootstrap ODP model will be addressed in the following section.

Let's begin with the classic chain-ladder model. Let $C_{w,d}$ denote the accumulated loss amount, either incurred or paid, for accident year, w, and development lag, d, for $1 \leq w \leq K$ and $1 \leq d \leq K$. $C_{w,d}$ is known for the "triangle" of data specified by $w + d \leq K + 1$. The goal of this model is to estimate the loss amounts in the last column of data, $C_{w,K}$ for $w = 2, \ldots, K$. To use the chain-ladder model, one first calculates the age to age factors given by

$$f_d = \frac{\sum_{w=1}^{K-d} C_{w,d+1}}{\sum_{w=1}^{K-d} C_{w,d}} \qquad \text{for } d = 1, \ldots, K-1.$$

The chain-ladder estimate of $C_{w,K}$ is the product of the latest reported loss, $C_{w,K+1-w}$, and the subsequent age-to-age factors $f_{K+1-w} \cdot \ldots \cdot f_{K-1}$. Putting this together, we have

$$C_{w,K} = C_{w,K+1-w} \cdot f_{K+1-w} \cdot \ldots \cdot f_{K-1}$$

Taylor (1986, p. 40) discusses the origin of the chain-ladder model and concludes that "It appears that it probably originated in the accounting literature, and was subsequently absorbed in to, or rediscovered in, the actuarial." He goes on to say that "Of course, one must bear in mind that both the chain-ladder model and estimation method are fairly obvious and might have been derived several times in past literature." Taylor believes that the rather whimsical name of the model was first used by Professor R. E. Beard as he championed the method in the early 1970s while working as a consultant to the U.K. Department of Trade.

Mack (1993, 1994) turns the deterministic chain ladder model into a stochastic model by first treating $\tilde{C}_{w,d}$ as a random variable that represents the accumulated loss amount in the (w, d) cell. He then makes three assumptions.[7]

---

[7] Depending on the context, various quantities, such as $C_{w,d}$, will represent observations, estimates or random variables. In situations where it might not be clear, let's adopt the convention that for a quantity X, $\tilde{X}$ will indicate that X is being treated as a random, or simulated, variable, $\hat{X}$ will denote an estimate of X, and a bare X will be treated as a fixed observation or parameter.

1. $\mathrm{E}\left[\tilde{C}_{w,d+1} \middle| C_{w,1}, \ldots, C_{w,d}\right] = C_{w,d} \cdot f_d$
2. For any given d, the random variables $\tilde{C}_{v,d}$ and $\tilde{C}_{w,d}$ are independent for $v \neq w$.
3. $\mathrm{Var}\left[\tilde{C}_{w,d+1} \middle| C_{w,1}, \ldots, C_{w,d}\right] = C_{w,d} \cdot \alpha_d^2$

The Mack estimate for $\mathrm{E}\left[\tilde{C}_{w,K}\right]$ for $w = 2, \ldots, K$ is given by

$$\hat{C}_{w,K} = C_{w,K+1-w} \cdot \hat{f}_{K+1-w} \cdot \cdots \cdot \hat{f}_{K-1}$$

where

$$\hat{f}_d = \frac{\sum_{w=1}^{K-d} C_{w,d+1}}{\sum_{w=1}^{K-d} C_{w,d}}$$

Given his assumptions above, Mack then derives expressions for the standard deviations $\mathrm{SD}\left[\tilde{C}_{w,K}\right]$ and $\mathrm{SD}\left[\sum_{w=2}^{K} \tilde{C}_{w,K}\right]$. Table 4 applies Mack's expressions to the illustrative insured data in Table 2 using the R "ChainLadder" package.

In addition to the loss statistics calculated by the Mack expressions, Table 4 contains the outcomes $\{C_{w,10}\}$ from Table 2. Following Mack's suggestion, I calculated the percentile of $\sum_{w=1}^{10} C_{w,10}$ assuming a lognormal distribution with matching the mean and the standard deviation.

Taken by itself, an outcome falling in the 86th percentile gives us little information, as that percentile is not unusually high. If the percentile was, say, above the 99.5th percentile, suspicion might be warranted. My intent here is to test the general applicability of the Mack model on incurred loss triangles. To do this, I selected 200 incurred loss

Table 4. Mack Model Output for the Illustrative Insurer Incurred Losses

w	$\hat{C}_{w,10}$	SD	CV	$C_{w,10}$	Percentile
1	3917	0	0.000	3917	
2	2538	0	0.000	2532	
3	4167	3	0.001	4279	
4	4367	37	0.009	4341	
5	3597	34	0.010	3587	
6	3236	40	0.012	3268	
7	5358	146	0.027	5684	
8	3765	225	0.060	4128	
9	4013	412	0.103	4144	
10	3955	878	0.222	4181	
Total	38914	1057	0.027	40061	86.03

triangles, 50 each from four different lines of insurance, and calculated the percentile of the $\sum_{w=1}^{10} C_{w,10}$ outcome for each triangle. My criteria for "general applicability of the model" is that these percentiles should be uniformly distributed. And for a sufficiently large sample, uniformity is testable! Klugman, Panjer, and Willmot (2012, Section 16.3) describe a variety of tests that can be applied in this case.

Probably the most visual test for uniformity is a plot of a histogram. If the percentiles are uniformly distributed, we should expect the height of the bars to be equal. Unless the sample size is very large, this will rarely be the case because of random fluctuations. A visual test of uniformity that allows one to test for statistical significance is the *p–p* plot combined with the Kolmogorov–Smirnov (K–S) test. Here is how it works. Suppose one has a sample of n predicted percentiles ranging from 0 to 100 and sort them into increasing order. The expected value of these percentiles is given by $\{e_i\} = 100 \cdot \{1/(n+1), 2/(n+1), \ldots, n/(n+1)\}$. One then plots the expected percentiles on the horizontal axis against the sorted predicted percentiles on the vertical axis. If the predicted percentiles are uniformly distributed, we expect this plot to lie along a 45° line. According to the K–S test as described by Klugman, Panjer, and Willmot (2012, p. 331), one can reject the hypothesis that a set of percentiles $\{p_i\}$ is uniform at the 5% level if $D \equiv \max|p_i - f_i|$ is greater than its critical value, $136/\sqrt{n}$ where $\{f_i\} = 100 \cdot \{1/n, 2/n, \ldots, n/n\}$. This is represented visually on a *p–p* plot by drawing lines at a distance $136/\sqrt{n}$ above and below the 45° line.[8] We reject the hypothesis of uniformity if the *p–p* plot lies outside the band defined by those lines. For the purposes of this monograph, a model will be deemed "validated" if it passes the K–S test at the 5% level.

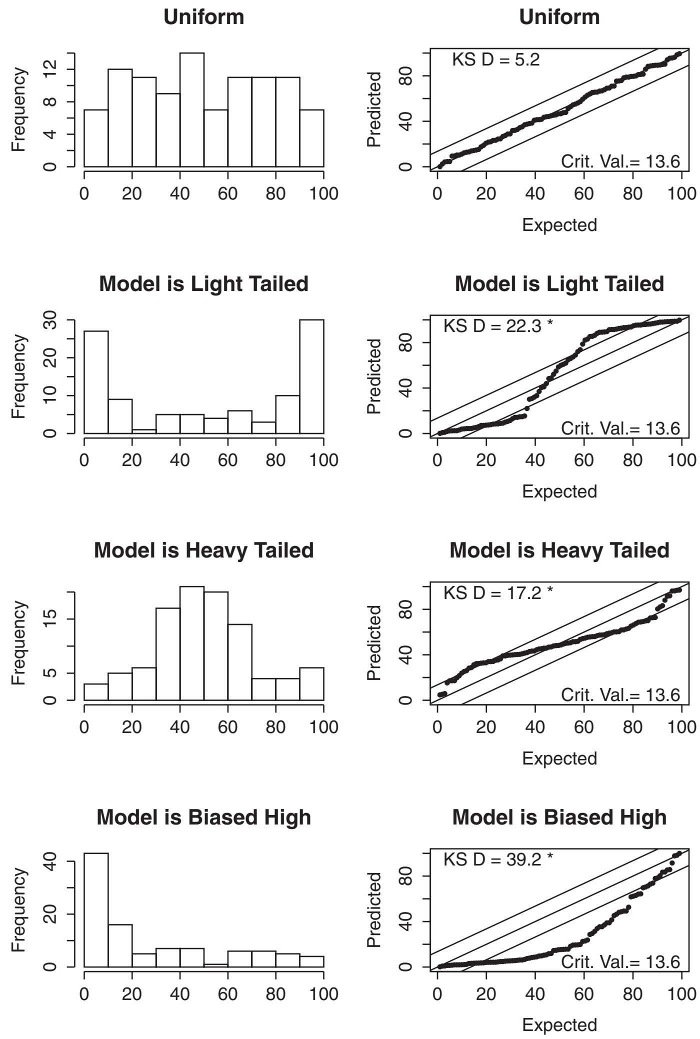
Klugman, Panjer, and Willmot (2012, p. 332) also discusses a second test of uniformity that is applicable in this situation. The Anderson–Darling (A–D) test is similar to the Kolmogorov–Smirnov test, but it is more sensitive to the fit in the extreme values (near the 0th and the 100th percentile) of the distribution. I applied the A–D test along with the K–S test on the models described in this monograph with the result that almost all A–D tests failed. If in the future someone develops a more refined model, we can raise the bar to the more stringent A–D test. Until that happens, I think the K–S test is the best tool to differentiate between models.

Figure 1 shows both histograms and *p–p* plots for simulated data with n =100. The plots labeled "Uniform" illustrate the expected result. The K–S D statistic accompanies each *p–p* plot. The "*" indicates that the D statistic is above its critical value.

Figure 1 also shows *p–p* plots for various departures from uniformity. For example, if the predicted distribution is too light in the tails, there are more than expected high and low percentiles in the predicted outcomes and we see a *p–p* plot that looks like a slanted "S" curve. If the predicted distribution is too heavy in the tails, there are more than expected middle percentiles in the predicted outcomes and we see a *p–p* plot that looks like a slanted backward "S" curve. If the model predicts results that are in general too high, predicted outcomes in the low percentiles will be more frequent.

[8] This is an approximation as $f_i \approx e_i$.

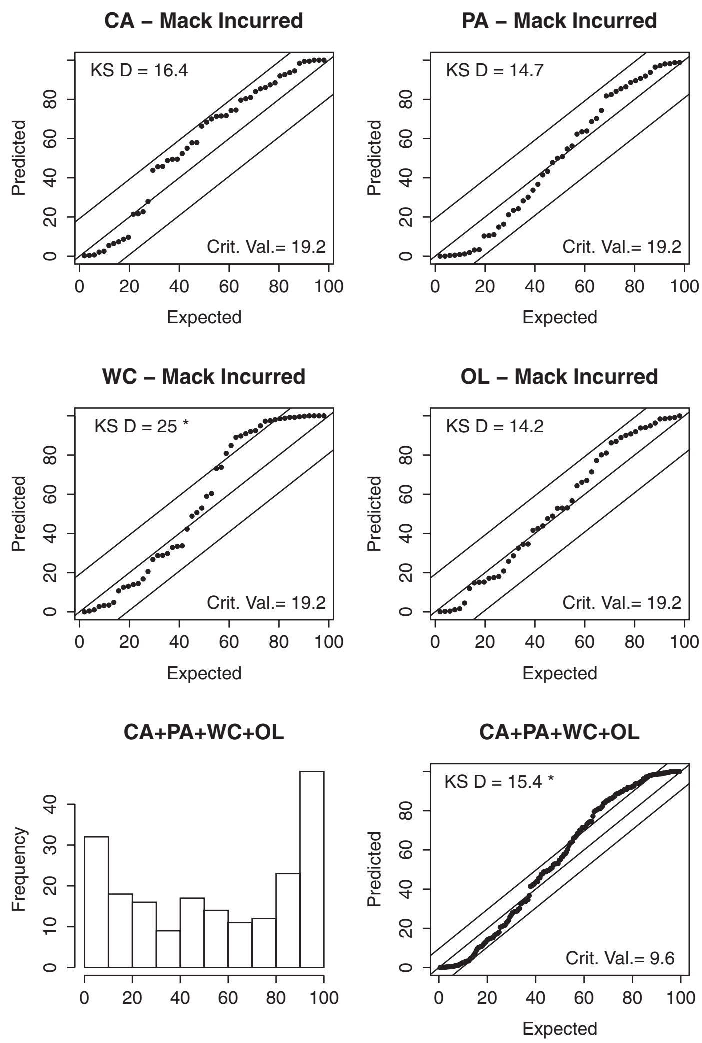
Figure 1. *p–p* Plots Test for Uniformity



To validate the Mack model, I repeated the calculations for the 200 selected incurred loss reserve triangles.

Figure 2 shows the *p–p* plots for the Mack model. The plots were first done separately for the outcome percentiles in each line of insurance. Although the plots fall inside the K–S band for three of the four lines, the plots for all four of the lines resemble the slanted "S" curve that is characteristic of a light tailed predicted distribution. When we combine the outcome percentiles of all four lines, the *p–p* plot lies outside the K–S band and we conclude that the distribution predicted by the Mack model is too light in the tails for these data. In all the validation plots below the K–S critical values are 19.2 and 9.6 for the individual lines and all lines combined respectively.

Figure 2. *p–p* Plots for the Mack Model on Incurred Loss Triangles



4. Validating the Bootstrap ODP Model

This section does an analysis similar to that done in the last section for the bootstrap ODP model as described by England and Verrall (2002) and implemented by the R "ChainLadder" package. This model was designed to work with incremental losses, $I_{w,d}$, rather than the cumulative losses $C_{w,d}$, where $I_{w,1} = C_{w,1}$ and $I_{w,d} = C_{w,d} - C_{w,d-1}$ for $d > 1$.

A key assumption made by this model is that the incremental losses are described by the overdispersed Poisson distribution with

$$E\left[\tilde{I}_{w,d}\right] = \alpha_w \cdot \beta_d \quad \text{and} \quad Var\left[\tilde{I}_{w,d}\right] = \phi \cdot \alpha_w \cdot \beta_d$$

The parameters of the model can be estimated by a standard generalized linear model (GLM) package.[9] They then use a bootstrap resampling procedure to quantify the volatility of the estimate.

England and Verrall point out that the using the ODP model on incremental losses almost all but requires one to use paid, rather than incurred, losses since the overdispersed Poisson model is defined only for nonnegative losses. Incurred losses include estimates by claims adjusters that can (and frequently do) get adjusted downward. Negative incremental paid losses occasionally occur because of salvage and subrogation, but a feature of the GLM estimation procedure allows for negative incremental losses as long as all column sums of the loss triangle remain positive.

Table 5 gives the estimates of the mean, the standard deviation for both the ODP (with 10,000 bootstrap simulations) and Mack models on the data in Table 3. The predicted percentiles of the 10,000 outcomes are also given for each model.

The validation p–p plots, similar to those done in the previous section, for both the ODP and the Mack models on paid data, are in Figures 2 and 3. The results for both models are quite similar. Neither model validates on the paid triangles. A comparison of the p–p plots in Figures 3 and 4 with the illustrative plots in Figure 1 suggests that the expected loss estimates of both models tend to be too high for these data.

Let's now consider the results of this and the prior section. These sections show that two popular models do not validate on outcomes of the 200 Schedule P triangles drawn from the CAS Loss Reserve Database. These models do not validate in different ways when we examine paid and incurred triangles. For incurred triangles, the distribution

[9] England and Verrall (2002) use a log link function in their GLM. They also note that the GLM for the ODP maximizes the quasi-likelihood, allowing the model to work with continuous (non-integer) losses.

Table 5. ODP and Mack Model Output for the Illustrative Insurer Paid Losses

	ODP			Mack			Outcome
w	$\hat{C}_{w,10}$	SD	CV	$\hat{C}_{w,10}$	SD	CV	$C_{w,10}$
1	3912	0	0	3912	0	0.0000	3912
2	2532	21	0.0083	2532	0	0.0000	2527
3	4163	51	0.0123	4162	3	0.0007	4274
4	4369	85	0.0195	4370	28	0.0064	4341
5	3554	96	0.027	3555	35	0.0098	3583
6	3211	148	0.0461	3213	157	0.0489	3268
7	5161	240	0.0465	5167	251	0.0486	5684
8	3437	332	0.0966	3442	385	0.1119	4128
9	4220	572	0.1355	4210	750	0.1781	4144
10	4635	1048	0.2261	4616	957	0.2073	4139
Total	39193	1389	0.0354	39177	1442	0.0368	40000
Percentile		73.91			72.02		

Figure 3. *p–p* Plots for the Bootstrap ODP Model on Paid Loss Triangles

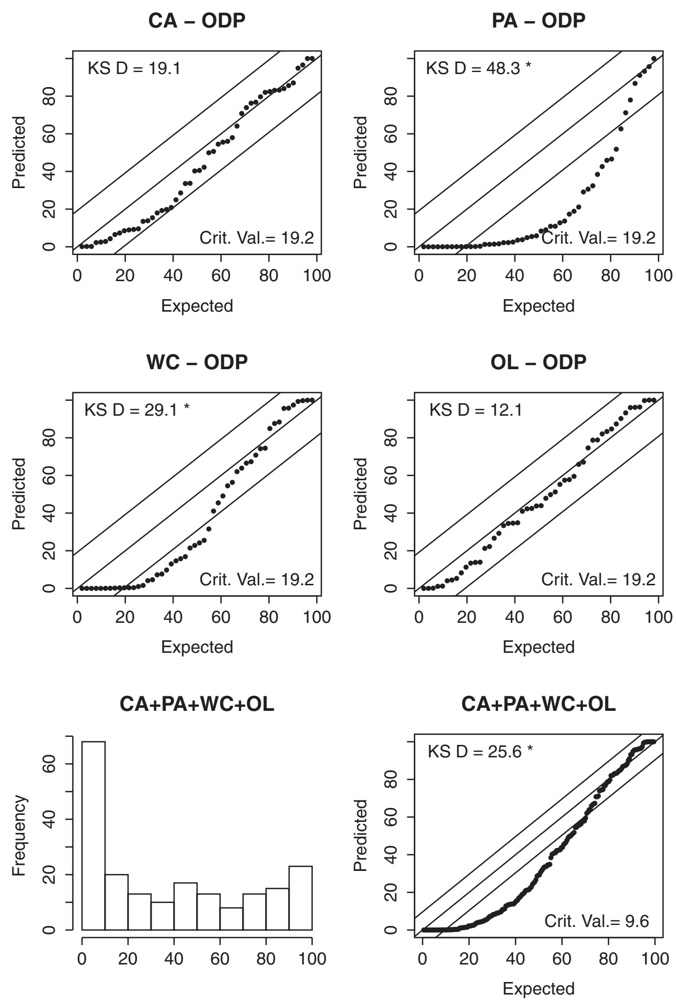
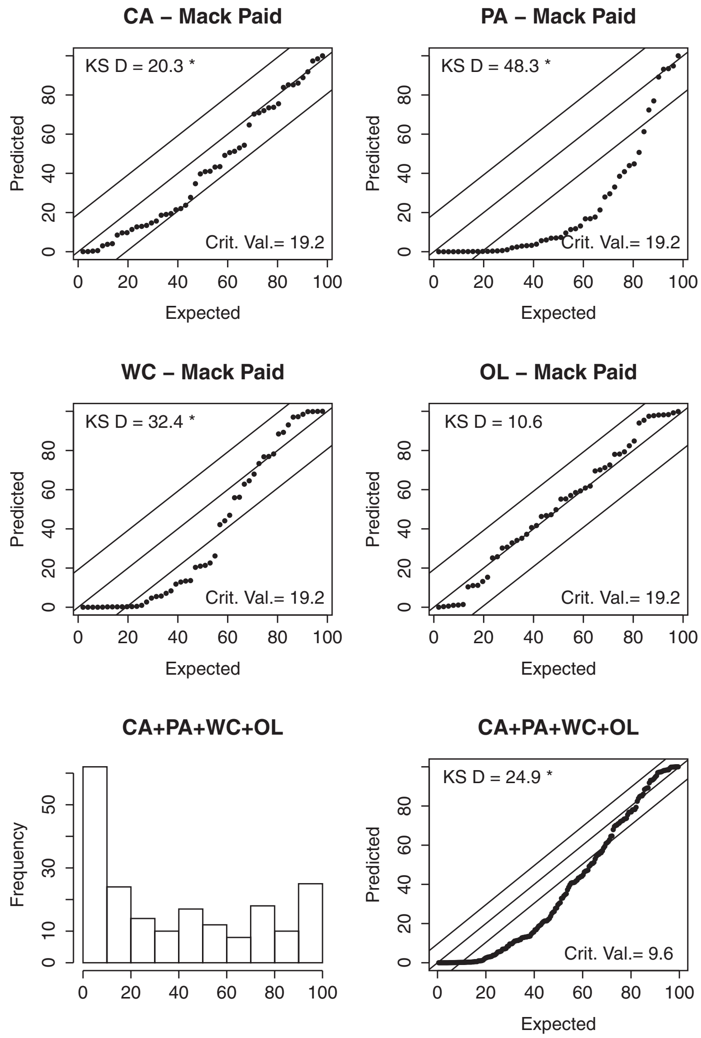


Figure 4. *p–p* Plots for the Mack Model on Paid Loss Triangles



predicted by the Mack model has a light tail. For paid triangles, the distributions predicted by both the Mack and the bootstrap ODP models tend to produce expected loss estimates that are too high. There are two plausible explanations for these observations:

1. The insurance loss environment has experienced changes that are not observable at the current time.
2. There are other models that can be validated.

To disprove the first explanation, one can develop models that do validate. Failing to develop a model that validates may give credence to, but does not necessarily confirm, that the first explanation is true. This monograph now turns to describing some efforts to find models that do validate.

5. Bayesian Models for Incurred Loss Data

I will begin this section on Bayesian MCMC models by quoting the advice of Verrall (2007). "For the readers for whom this is the first time they have encountered MCMC methods, it is suggested that they simply accept that they are a neat way to get the posterior distributions for Bayesian models and continue reading the paper. If they like the ideas and would like to find out more . . ." they should read the introduction in Appendix B. Keep in mind that the state of the art (e.g., faster multi-core personal computers, more efficient algorithms and more user-friendly software) is still rapidly advancing. Appendix C explains what I did with the current state of the art, as I perceived it, at the time I was writing this monograph.

Now let's get to the loss reserve models. As pointed out in Section 3, the Mack model did not validate on the insurers listed in Appendix A using the loss data that are in the CAS Loss Reserve Database. This section presents two Bayesian MCMC models that were proposed in an attempt to find a model that does validate on these data.

The way the Mack model did not validate, i.e., it underestimated the variability of the ultimate loss estimates, suggested a direction to go in order to fix it. Here are two ways to improve the recognition of the inherent variability of the predictive distribution.

1. The Mack model multiplies the age-to-age factors by the last observed loss, $C_{w,11-w}$. One can think of the $C_{w,11-w}$s as fixed level parameters. A model that treats the level of the accident year as random will predict more risk.
2. The Mack model assumes that the loss amounts for different accident years are independent. A model that allows for correlation between accident years could increase the standard deviation of $\sum_{w=1}^{10} \tilde{C}_{w,10}$.

I propose two different models to address the underestimation of the variability of the ultimate loss. The first model replaces the fixed level parameters, given by the last observed accident year, in the Mack model with random level parameters. As we shall see, this model improves the estimation of the variability, but does not go far enough. The second, and more complicated model, considers correlation between the accident years.

The Leveled Chain Ladder (LCL) Model

Let:

1. $\mu_{w,d} = \alpha_w + \beta_d$.
2. $\tilde{C}_{w,d}$ has a lognormal distribution with log mean $\mu_{w,d}$ and log standard deviation σ_d subject to the constraint that $\sigma_1 > \sigma_2 > \cdots > \sigma_{10}$.

To prevent overdetermining the model, set $\beta_{10} = 0$. The parameters $\{\alpha_w\}$, $\{\sigma_d\}$ and the remaining $\{\beta_d\}$ are assigned relatively wide prior distributions as follows:

1. Each $\alpha_w \sim \text{normal}\left(\log(Premium_w) + logelr, \sqrt{10}\right)$ where the parameter $logelr \sim$ uniform(−1, 0.5).[10]
2. Each $\beta_d \sim$ uniform(−5, 5) for $d < 10$.
3. Each $\sigma_d = \sum_{i=d}^{10} a_i$ where $a_i \sim$ uniform(0, 1).

The hierarchical structure of the priors in (3) above assures that $\sigma_1 > \sigma_2 > \ldots > \sigma_{10}$. The rationale behind this structure is that as d increases, there are fewer claims that are open and subject to random outcomes.

The next model adds a between-year correlation feature.[11]

The Correlated Chain-Ladder (CCL) Model

Let:

1. Each $\alpha_w \sim \text{normal}\left(\log(Premium_w) + logelr, \sqrt{10}\right)$ where the parameter $logelr \sim$ uniform(−1, 0.5).
2. $\mu_{1,d} = \alpha_1 + \beta_d$.
3. $\mu_{w,d} = \alpha_w + \beta_d + \rho \cdot \left(\log(C_{w-1,d}) - \mu_{w-1,d}\right)$ for $w > 1$.
4. $\tilde{C}_{w,d}$ has a lognormal distribution with log mean $\mu_{w,d}$ and log standard deviation σ_d subject to the constraint that $\sigma_1 > \sigma_2 > \ldots > \sigma_{10}$.

Note that the CCL model reduces to the LCL model when $\rho = 0$.

If the parameters $\{\alpha_w\}$, $\{\beta_d\}$, and ρ are given, the parameter ρ is equal to the coefficient of correlation between $\log\left(\tilde{C}_{w-1,d}\right)$ and $\log\left(\tilde{C}_{w,d}\right)$. To see this we first note that unconditionally:

$$\begin{aligned}
E\left(\log\left(\tilde{C}_{w,d}\right)\right) &= \mu_{w,d} \\
&= \alpha_w + \beta_d + \rho \cdot \left(\log\left(\tilde{C}_{w-1,d}\right) - \mu_{w-1,d}\right) \\
&= \alpha_w + \beta_d
\end{aligned}$$

Given $C_{w-1,d}$ we have that:

$$\begin{aligned}
E\left(\left(\log\left(\tilde{C}_{w,d}\right) - (\alpha_w + \beta_d)\right) \cdot \left(\log\left(C_{w-1,d}\right) - \mu_{w-1,d}\right)\right) \\
= \left(\mu_{w,d} - (\alpha_w + \beta_d)\right) \cdot \left(\log\left(C_{w-1,d}\right) - \mu_{w-1,d}\right) \\
= \rho \cdot \left(\log\left(C_{w-1,d}\right) - \mu_{w-1,d}\right)^2
\end{aligned}$$

---

[10] The JAGS expression for a normal distribution uses what it calls a "precision" parameter equal to the reciprocal of the variance. The standard deviation, $\sqrt{10}$, corresponds to the rather low precision of 0.1.

[11] Some of the models I tried before getting to this one are described in my working paper Meyers (2012). Note that what I call the LCL model in that paper is different from the LCL model above.

Then the coefficient of correlation between $\tilde{C}_{w,d}$ and $\tilde{C}_{w-1,d}$ is given by:

$$E_{C_{w-1,d}}\left(E_{C_{w,d}}\left(\frac{\log(\tilde{C}_{w,d})-(\alpha_w+\beta_d)}{\sigma_d}\cdot\frac{\log(C_{w-1,d})-\mu_{w-1,d}}{\sigma_d}\middle|C_{w-1,d}\right)\right)$$

$$=E_{C_{w-1,d}}\left(\frac{\rho\cdot\left(\log(\tilde{C}_{w-1,d})-\tilde{\mu}_{w-1,d}\right)^2}{\sigma_d^2}\right)=\rho$$

To prevent overdetermining the model, set $\beta_{10} = 0$. The parameters $\{\alpha_w\}, \{\sigma_d\}, \rho$ and the remaining $\{\beta_d\}$ are assigned relatively wide prior distributions as follows:

1. Each $\alpha_w \sim \text{normal}\left(\log(Premium_w) + logelr, \sqrt{10}\right)$ where the parameter *logelr* ~ uniform(−1, 0.5).[12]
2. Each $\beta_d \sim \text{uniform}(-5, 5)$ for $d < 10$.
3. $\rho \sim \text{uniform}(-1, 1)$—The full permissible range for ρ.
4. Each $\sigma_d = \sum_{i=d}^{K} a_i$ where $a_i \sim \text{uniform}(0,1)$.

I deliberately chose rather diffuse[13] prior distributions since I had no direct knowledge of the claims environment other than the data that are reported in Schedule P. While preparing annual statements, actuaries with more direct knowledge of the claims environment normally attempt to reflect this knowledge in their unpaid loss estimates. Bornhuetter and Ferguson (1972) describe a very popular method where one can reflect knowledge of an insurer's expected loss ratio in their estimates. With minor modifications of the JAGS script, one can reflect this knowledge by specifying more restrictive priors for $\{\alpha_w\}$ parameters and the *logelr* parameter.

The predictive distribution of outcomes is a mixed distribution where the mixing is specified by the posterior distribution of parameters. Here is what the script for the CCL model does.

The predictive distribution for $\sum_{w=1}^{10} C_{w,10}$ is generated by a simulation. For each parameter set $\{\alpha_w\}, \{\beta_d\}, \{\sigma_d\}$ and $\{\rho\}$, start with the given $C_{1,10}$ and calculate the mean, $\mu_{2,10}$. Then simulate $\tilde{C}_{2,10}$ from a lognormal distribution with log mean, $\mu_{2,10}$, and log standard deviation, σ_{10}. Similarly, use the result of this simulation to simulate $\tilde{C}_{2,10}, \ldots, \tilde{C}_{10,10}$. Then form the sum $C_{1,10} + \sum_{w=2}^{10} \tilde{C}_{w,10}$. The script generates 10,000 simulations that make up a sample from the predictive distribution from which one can calculate various statistics such as the mean, standard deviation and the percentile of the outcome. Here is a more detailed explanation of this process.

Given the group code for an insurer in the CAS Loss Reserve Database, the R script for the CCL Model performs the following steps:

1. Reads in the data triangle $\{C_{w,d}\}$ for the insurer identified by the group code.
2. Runs the JAGS script and gets a sample of 10,000 parameter sets, $\{\alpha_w\}, \{\beta_d\}, \{\sigma_d\}$ and ρ from the posterior distribution of the CCL model.

[12] The JAGS expression for a normal distribution uses what it calls a "precision" parameter equal to the reciprocal of the variance. The standard deviation, corresponds to the rather low precision of 0.1.

[13] One might also use a "noninformative" prior distribution. Noninformative prior distributions are usually attached to a specific mathematical objective. See, for example, Section 3.3 of Berger (1985).

3. Simulates 10,000 copies, one for each parameter set in (2) above, of $\left\{\tilde{C}_{w,10}\right\}_{d=2}^{10}$. The simulation proceeds as follows.
   - Set $\mu_{1,10} = \alpha_1 + \beta_{10}$. Recall that $C_{1,10}$ is given in the original data.
   - Set $\tilde{\mu}_{2,10} = \alpha_2 + \beta_{10} + \rho \cdot \left(\log\left(C_{1,10}\right) - \mu_{1,10}\right)$. Simulate $\tilde{C}_{2,10}$ from a lognormal distribution with log mean $\mu_{2,10}$ and log standard deviation σ_{10}.
   - Set $\tilde{\mu}_{3,10} = \alpha_3 + \beta_{10} + \rho \cdot \left(\log\left(\tilde{C}_{2,10}\right) - \mu_{2,10}\right)$. Simulate $\tilde{C}_{3,10}$ from a lognormal distribution with log mean $\mu_{2,10}$ and log standard deviation σ_{10}.
   - . . .
   - Set $\tilde{\mu}_{10,10} = \alpha_{10} + \beta_{10} + \rho \cdot \left(\log\left(\tilde{C}_{9,10}\right) - \tilde{\mu}_{9,10}\right)$. Simulate $\tilde{C}_{10,10}$ from a lognormal distribution with log mean $\mu_{10,10}$ and log standard deviation σ_{10}.
4. For each w, calculate summary statistics $\hat{C}_{w,10} = \text{mean}\left(\tilde{C}_{w,10}\right)$ and SD = standard deviation $\left(\tilde{C}_{w,10}\right)$. Calculate similar statistics for the total $C_{1,10} + \sum_{w=2}^{10} \tilde{C}_{w,10}$.
5. Calculate the percentile of the outcome by counting how many of the 10,000 instances of $\sum_{w=2}^{10} \tilde{C}_{w,10}$ are $\leq$ the actual outcomes $\sum_{w=2}^{10} C_{w,10}$.

Table 6 gives the results from the first five MCMC samples produced by the script for the CCL model applied to the losses for the illustrative insurer in Table 2. The top 31 rows of that table were generated in Step 2 of the simulation above. The remaining rows were generated in Step 3.

Table 6. Illustrative MCMC Simulations

	MCMC Sample Number				
	1	2	3	4	5
α_1	8.2763	8.2452	8.2390	8.2591	8.2295
α_2	7.8226	7.7812	7.8008	7.8048	7.7810
α_3	8.2625	8.3200	8.2929	8.2883	8.2642
α_4	8.3409	8.3286	8.3539	8.3622	8.3159
α_5	8.2326	8.1166	8.1093	8.1855	8.1523
α_6	8.1673	8.0307	8.0491	8.1727	8.0470
α_7	8.6403	8.4776	8.4113	8.5815	8.4871
α_8	8.2177	8.2488	8.2708	8.0752	8.1763
α_9	8.3174	8.2007	8.2589	8.3744	8.2653
α_{10}	7.4101	8.0036	8.7584	8.4241	8.8420
β_1	–0.5125	–0.5180	–0.6504	–0.4947	–0.7384
β_2	–0.2756	–0.1014	–0.1231	–0.2138	–0.0844
β_3	–0.1271	–0.0313	–0.0622	–0.0758	–0.0498
β_4	–0.1013	–0.0090	0.0165	0.0439	0.0479
β_5	0.0518	–0.0109	0.0060	0.0034	0.0610
β_6	0.0180	0.0885	0.0139	0.0175	0.0709
β_7	0.0105	0.0583	0.0205	0.0427	0.0362

(continued on next page)

Table 6. Illustrative MCMC Simulations *(continued)*

	MCMC Sample Number				
	1	2	3	4	5
β_8	0.0400	−0.0090	0.0612	0.0444	0.0338
β_9	0.0005	0.0287	0.0419	0.0116	0.0333
β_{10}	0.0000	0.0000	0.0000	0.0000	0.0000
σ_1	0.3152	0.2954	0.3164	0.1895	0.2791
σ_2	0.2428	0.1982	0.2440	0.1858	0.1711
σ_3	0.1607	0.1632	0.2078	0.1419	0.1089
σ_4	0.1245	0.1133	0.0920	0.0842	0.0800
σ_5	0.0871	0.0830	0.0694	0.0747	0.0794
σ_6	0.0733	0.0649	0.0626	0.0508	0.0463
σ_7	0.0324	0.0281	0.0294	0.0368	0.0352
σ_8	0.0279	0.0247	0.0172	0.0270	0.0330
σ_9	0.0171	0.0239	0.0130	0.0267	0.0329
σ_{10}	0.0170	0.0237	0.0105	0.0241	0.0244
ρ	0.1828	0.4659	0.4817	0.1901	0.2155
$\mu_{1,10}$	8.2763	8.2452	8.2390	8.2591	8.2295
$C_{1,10}$	3917	3917	3917	3917	3917
$\tilde{\mu}_{2,10}$	7.8221	7.7942	7.8172	7.8074	7.7904
$\tilde{C}_{2,10}$	2520	2468	2480	2432	2453
$\tilde{\mu}_{3,10}$	8.2643	8.3278	8.2924	8.2862	8.2674
$\tilde{C}_{3,10}$	3893	4190	3939	4090	3802
$\tilde{\mu}_{4,10}$	8.3414	8.3345	8.3474	8.3679	8.3107
$\tilde{C}_{4,10}$	4229	4212	4233	4346	4075
$\tilde{\mu}_{5,10}$	8.2341	8.1219	8.1109	8.1873	8.1527
$\tilde{C}_{5,10}$	3761	3285	3269	3597	3676
$\tilde{\mu}_{6,10}$	8.1670	8.0192	8.0400	8.1728	8.0593
$\tilde{C}_{6,10}$	3450	3127	3120	3552	3196
$\tilde{\mu}_{7,10}$	8.6365	8.4910	8.4140	8.5819	8.4893
$\tilde{C}_{7,10}$	5488	4719	4441	5299	4765
$\tilde{\mu}_{8,10}$	8.2129	8.2340	8.2634	8.0739	8.1720
$\tilde{C}_{8,10}$	3652	3847	3933	3295	3708
$\tilde{\mu}_{9,10}$	8.3156	8.2106	8.2655	8.3794	8.2752
$\tilde{C}_{9,10}$	4112	3538	3949	4426	3914
$\tilde{\mu}_{10,10}$	7.4112	7.9853	8.7659	8.4271	8.8414
$\tilde{C}_{10,10}$	1613	3001	6511	4507	6763

Table 7. CCL, LCL, and Mack Models on Illustrative Insurer Incurred Data

	CCL			LCL			Mack			Outcome
w	$C_{w,10}$	SD	CV	$C_{w,10}$	SD	CV	$C_{w,10}$	SD	CV	$C_{w,10}$
1	3917	0	0.000	3917	0	0.000	3,917	0	0.000	3,917
2	2545	57	0.022	2544	59	0.023	2,538	0	0.000	2,532
3	4110	113	0.028	4110	106	0.026	4,167	3	0.001	4,279
4	4314	130	0.030	4307	122	0.028	4,367	37	0.009	4,341
5	3549	123	0.035	3545	115	0.032	3,597	34	0.010	3,587
6	3319	146	0.044	3317	132	0.040	3,236	40	0.012	3,268
7	5277	292	0.055	5315	265	0.050	5,358	146	0.027	5,684
8	3796	331	0.087	3775	301	0.080	3,765	225	0.060	4,128
9	4180	622	0.149	4203	561	0.134	4,013	412	0.103	4,144
10	4155	1471	0.354	4084	1157	0.283	3,955	878	0.222	4,181
Total	39161	1901	0.049	39116	1551	0.040	38,914	1,057	0.027	40,061
Percentile		73.72			76.38			86.03		

Table 7 gives the estimates of the mean and standard deviation, by accident year and in total, for the LCL, the CCL, and the Mack Models for the illustrative insurer. The predicted percentiles of the 40,061 outcome are also given for each model. Note that the standard deviations of the predicted outcomes were significantly higher for the CCL and the LCL models than they were for the Mack Model. This is generally the case, as can be seen in Figure 5. This figure plots the standard deviations (in the log scale) of the CCL and LCL models against those of the Mack Model for the 200 loss triangles listed in Appendix A. The higher standard deviations of the CCL model over the LCL model can be attributed to the generally positive correlation parameters that are shown in Figure 6 for the illustrative insurer. Generally this is the case for other insurers as can be seen in Figure 7.

The validation *p–p* plots for the LCL and CCL models run on the selected 200 triangles are given in Figures 8 and 9. For the LCL model:

- The *p–p* plots combined lines of insurance lie within the Kolmogorov–Smrinov bounds for Commercial Auto, Personal Auto and Workers Comp.
- All four lines have the slanted S pattern that characterizes models that are too thin in the tails. This pattern is reinforced in the combined plot, and the resulting plot does not lie within the Kolmogorov–Smirnov bounds. But the combined plot is an improvement over the corresponding Mack *p–p* plot.

For the CCL Model:

- The *p–p* plots for all four lines lie within the Kolmogorov–Smirnov bounds, but just barely so for the Other Liability line.
- While the combined *p–p* plot lies within the Kolmogorov–Smirnov bounds, the slanted S pattern indicates a mildly thin tail predicted by the model.

Figure 5. Compare Standard Deviation for CCL and LCL with Mack

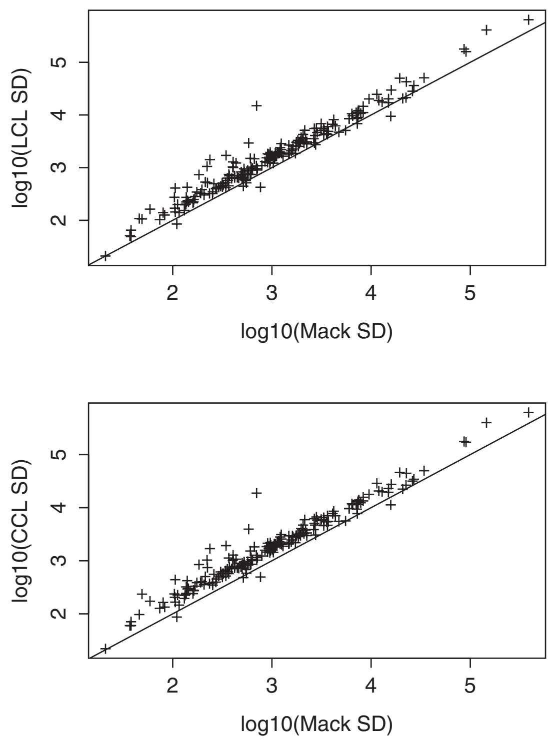


Figure 6. Posterior Distribution of ρ for the Illustrative Insurer

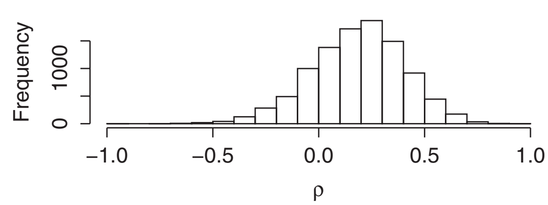


Figure 7. Posterior Mean of ρ for the 200 Incurred Loss Triangles

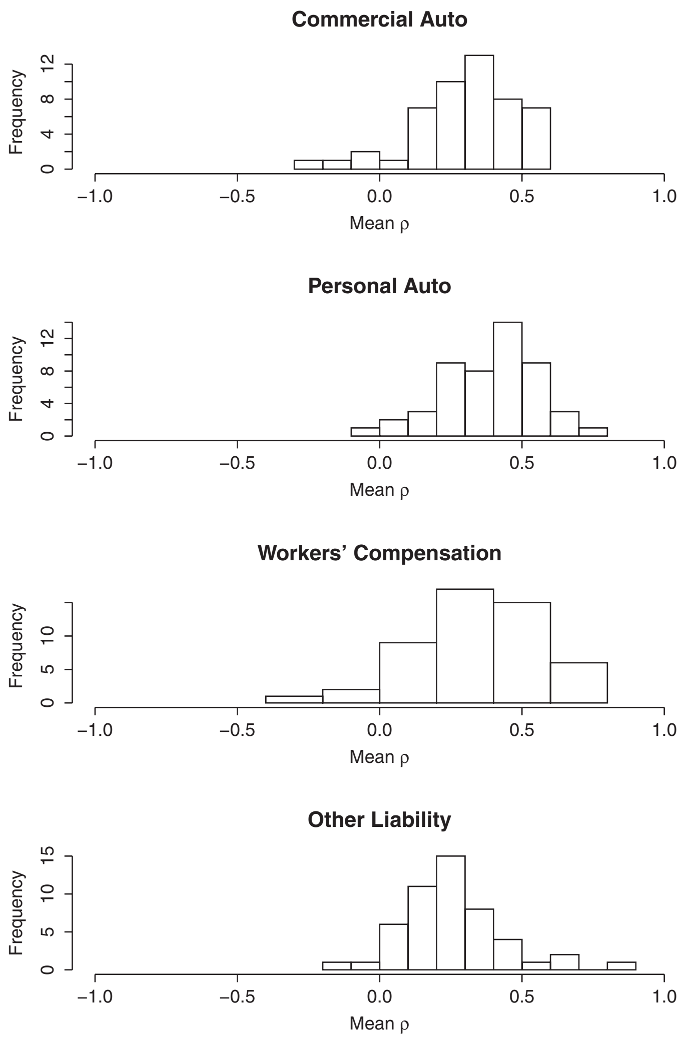


Figure 8. *p–p* Plots for the LCL Model on the Incurred Loss Triangles

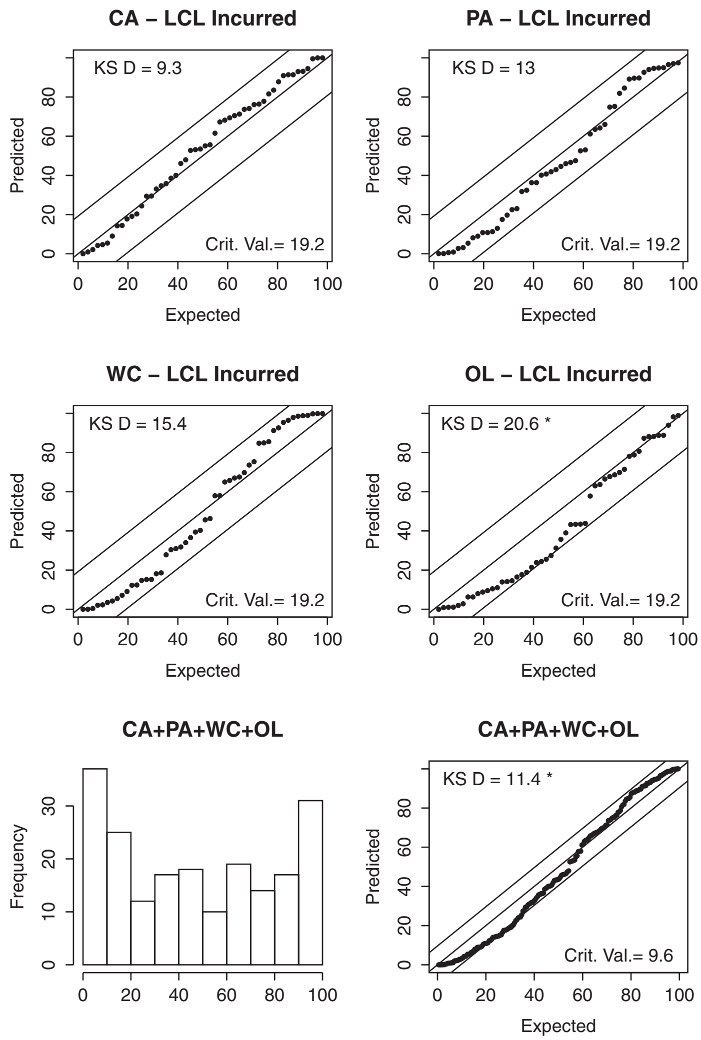
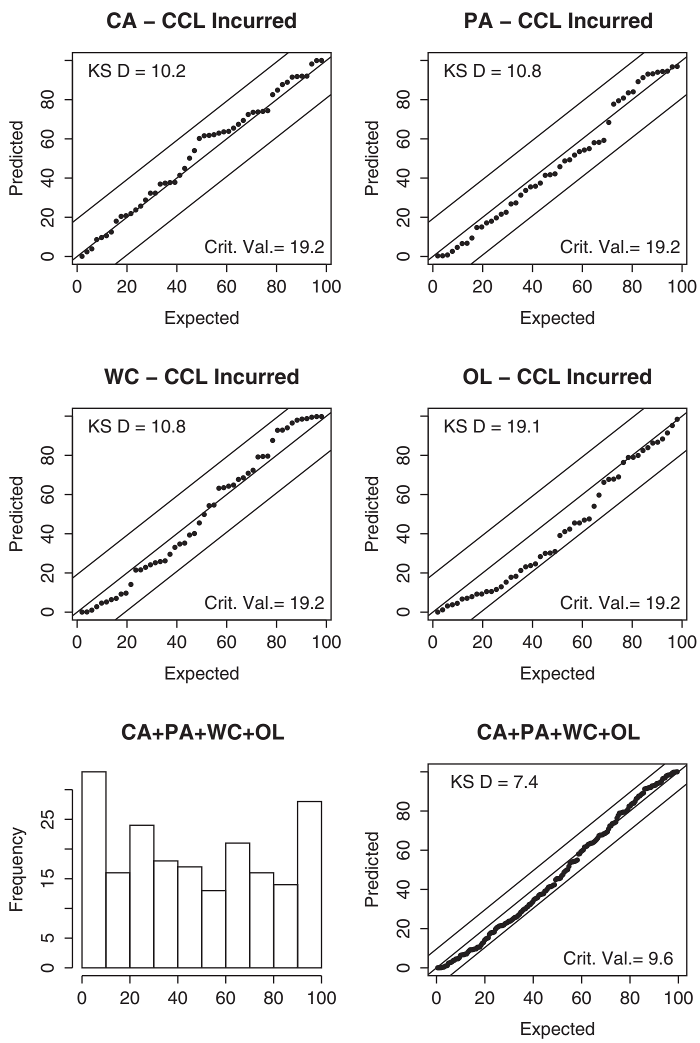


Figure 9. *p–p* Plots for the CCL Model on the Incurred Loss Triangles



6. Bayesian Models for Paid Loss Data

Given the improved validation of the CCL model on incurred loss data, it seems appropriate to try it out on paid loss data. Table 8 shows the CCL and ODP estimates. As should be expected given the results in Section 5, the standard deviation of the outcomes produced by the CCL model are noticeably higher than those produced by the ODP model.

The validation *p–p* plots for the CCL model applied to paid data are in Figure 10. When comparing this plot with the validation *p–p* plots for the ODP model (Figure 3) and the Mack model (Figure 4), we see that all three models show tend to produced estimates that are too high for these loss triangles.

Given the improved validation of the CCL model with incurred loss data, it is tempting to conclude that the incurred loss data contains crucial information that is not present in the paid loss data. However, there is also the possibility that a model other than the ODP or the CCL may be appropriate. A feature of such a model might be that it has a trend along the payment year $(= w + d - 1)$. Models with a payment year trend have been proposed in the writings of Ben Zehnwirth over the years. See, for example, Barnett and Zehnwirth (2000). The inclusion of a payment year trend in a model has two important consequences.

1. The model should be based on incremental paid loss amounts rather than cumulative paid loss amounts. Cumulative losses include settled claims which do not change with time.
2. Incremental paid loss amounts tend to be skewed to the right and are occasionally negative. We need a loss distribution that allows for these features.

One distribution that has these properties is the skew normal distribution. This distribution is starting to be applied in actuarial settings. See, for example, Pigeon, Antonio and Denuit (2013) Here is a description of this distribution taken from Frühwirth-Schnatter and Pyne (2010). This distribution has three parameters.

1. μ—the location parameter.
2. ω—the scale parameter, with $\omega > 0$.
3. δ—the shape parameter, with $\delta \in (-1, 1)$.[14]

[14] The reference calls the shape parameter α and then define $\delta = \alpha / \sqrt{1+\alpha^2}$. The parameter designation, α, was already taken in this monograph.

Table 8. CCL and ODP Models on Illustrative Insurer Paid Data

	CCL			ODP			Outcome
w	$\hat{C}_{w,10}$	SD	CV	$\hat{C}_{w,10}$	SD	CV	$C_{w,10}$
1	3912	0	0	3912	0	0.0000	3912
2	2568	114	0.0444	2532	21	0.0083	2527
3	4157	199	0.0479	4163	51	0.0123	4274
4	4330	234	0.0540	4369	85	0.0195	4341
5	3574	212	0.0593	3554	96	0.0270	3583
6	3417	259	0.0758	3211	148	0.0461	3268
7	5235	465	0.0888	5161	240	0.0465	5684
8	3664	463	0.1264	3437	332	0.0966	4128
9	4444	870	0.1958	4220	572	0.1355	4144
10	5036	1961	0.3894	4635	1048	0.2261	4139
Total	40337	2692	0.0667	39193	1389	0.0354	40000
Percentile		49.18			73.91		

The skew normal distribution is defined as the sum of two random variables

$$X \sim \mu + \omega \cdot \delta \cdot Z + \omega \cdot \sqrt{1-\delta^2} \cdot \varepsilon$$

where $Z \sim$ truncated $\text{normal}_{[0,\infty)}(0,1)$ and $\varepsilon \sim \text{normal}(0,1)$. This distribution can also be expressed as a mixed truncated normal-normal distribution by setting

$$X \sim \text{normal}\left(\mu + \omega \cdot \delta \cdot Z, \omega \cdot \sqrt{1-\delta^2}\right).$$

In looking at either expression for the skew normal distribution one can see that when $\delta = 0$, the skew normal becomes a normal distribution. As δ approaches one, the distribution becomes more skewed and becomes a truncated normal distribution when $\delta = 1$. Figure 11 plots[15] the the density functions for $\mu = 0$, $\omega = 15$ and δ close to one.[16]

It should be apparent that the coefficient of skewness can never exceed the coefficient of skewness of the truncated normal distribution, which is equal to 0.995. As it turns out, this constraint is important. I have fit models with the skew normal distribution that otherwise are similar to what will be described below and found that for most triangles, δ is very close to its theoretical limit. This suggests that a distribution with a higher coefficient of skewness is needed.

[15] Using the R "sn" package.

[16] The parameters in Figures 11 and 12 are representative of what one could expect in the later settlement lags where negative incremental losses frequently occur.

Figure 10. *p–p* Plots for the CCL Model on Paid Loss Triangles

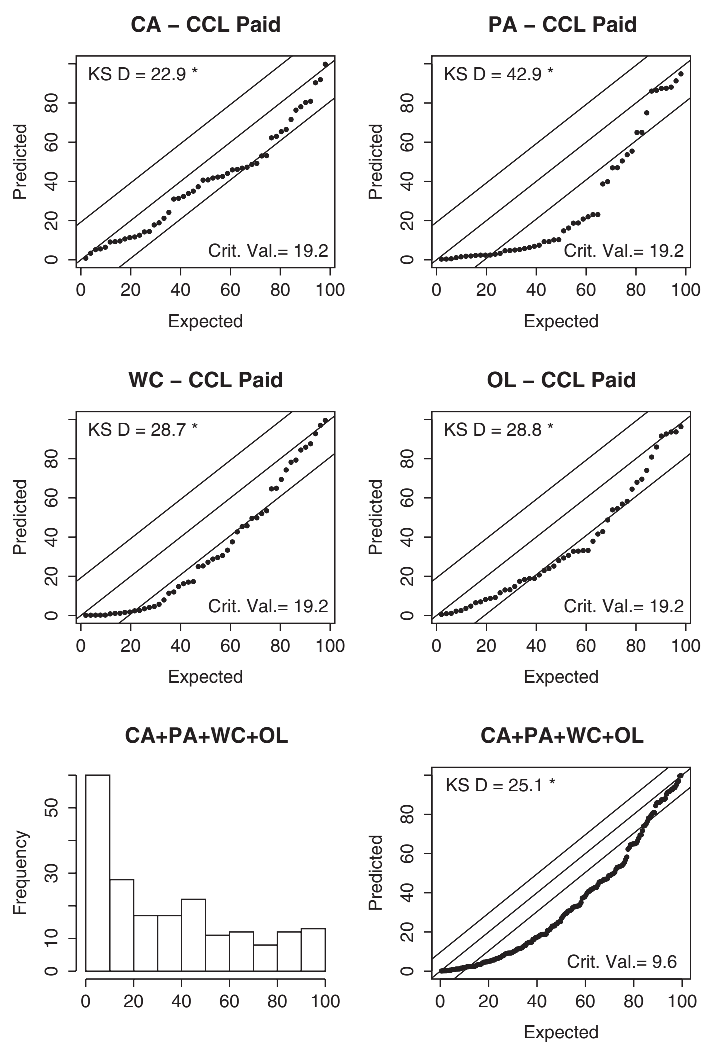


Figure 11. The Skew Normal Distribution

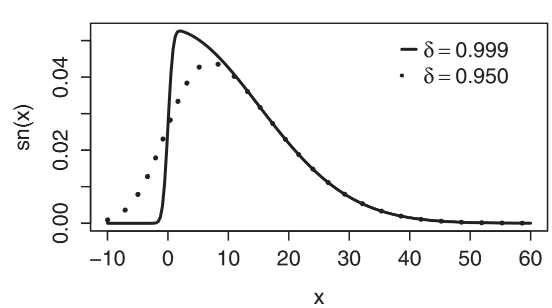
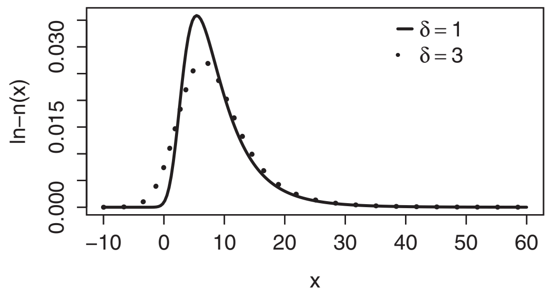


Figure 12. The Mixed Lognormal-Normal Distribution



The formulation of the skew normal distribution described by Frühwirth-Schnatter and Pyne (2010) suggests an alternative. Simply replace the truncated normal distribution with another skewed distribution, such as the lognormal distribution. Here is one way to do that. Define

$$X \sim \text{normal}(Z, \delta), \text{ where } Z \sim \text{lognormal}(\mu, \sigma).$$

Let's call this distribution the mixed lognormal-normal (ln-n) distribution with parameters given by δ, μ and σ. Figure 12 plots the density functions for $\mu = 2$, $\sigma = 0.6$, and two different values of δ.

Now that we have a loss distribution with the desired features of skewness and a domain that includes negative numbers, let's describe a model for incremental paid losses with a calendar-year trend.

The Correlated Incremental Trend (CIT) Model

Let:

1. $\mu_{w,d} = \alpha_w + \beta_d + \tau \cdot (w + d - 1)$.
2. $Z_{w,d} \sim \text{lognormal}(\mu_{w,d}, \sigma_d)$ subject to the constraint that $\sigma_1 < \sigma_2 < \ldots < \sigma_{10}$.
3. $\tilde{I}_{1,d} \sim \text{normal}(Z_{1,d}, \delta)$.
4. $\tilde{I}_{w,d} \sim \text{normal}\left(Z_{w,d} + \rho \cdot \left(\tilde{I}_{w-1,d} - Z_{w-1,d}\right) \cdot e^{\tau}, \delta\right)$ for $w > 1$.

When comparing the CIT model with the CCL model (as it might be applied to incremental losses) there are some differences to note.

- The CCL model was applied to cumulative losses. One should expect σ_d to decrease as d increases as a greater proportion of claims are settled. In the CIT model, one should expect that the smaller less volatile claims to be settled earlier. Consequently, σ_d should increase as d increases.
- In the CCL model, the autocorrelation feature was applied to the logarithm of the cumulative losses. Since there is the possibility of negative incremental losses, it was necessary to apply the autocorrelation feature in Step 4 above after leaving the "log" space. The hierarchical feature of the mixed lognormal-normal distribution

provides the opportunity to do this. For a given set of parameters, ρ is the coefficient of correlation between $\tilde{I}_{w-1,d}$ and $\tilde{I}_{w,d}$.

- The trend factor, τ, is applied additively in the "log" space in Step 1 above. As the autocorrelation feature in Step 4 above is applied outside of the "log" space, it is necessary to trend the prior payment year's difference by multiplying that difference by e^{τ}.

To prevent overdetermining the model, set $\beta_{10} = 0$. The parameters $\{\alpha_w\},\{\sigma_d\},\rho$, and the remaining $\{\beta_d\}$ are assigned prior distributions as follows:

1. Each $\alpha_w \sim \text{normal}\left(log\left(Premium_w\right) + logelr, \sqrt{10}\right)$ where $logelr \sim \text{uniform}(-5,1)$.
2. Each $\beta_d \sim \text{uniform}(0, 10)$ for $d = 1$ to 4 and $\beta_d \sim \text{uniform}(0, \beta_{d-1})$ for $d > 4$. This assures that β_d decreases for $d > 4$.
3. $\rho \sim \text{uniform}(-1, 1)$—The full permissible range for ρ.
4. $\tau \sim \text{normal}(0, 0.0316)$—corresponding to a precision parameter used by JAGS of 1000.
5. $\sigma_1^2 \sim \text{uniform}(0,0.5), \sigma_d^2 \sim \text{uniform}\left(\sigma_{d-1}^2, \sigma_{d-1}^2 + 0.1\right)$.
6. $\delta \sim \text{uniform}(0, \text{Average Premium})$

There are two deviations from the selection of diffuse prior distributions that are in the CCL model.

- I first tried a wider prior for τ. In examining the MCMC output I noticed that quite often, the value of τ was less than −0.1, which I took to be unreasonably low. This low value was usually compensated for by offsetting high values for the α and/or β parameters. This could have a noticeable effect on the final result, so I decided to restrict the volatility of τ to what I considered to be a reasonable range of payment year changes.
- In examining the MCMC output, I noticed that, occasionally, high values of σ_d would occur. This led to unreasonably high simulated losses in the output, so I decided to limit how fast σ_d could increase with *d.*

The predictive distributions of the sum, $\sum_{d=1}^{10} \tilde{I}_{w,d}$ for each *w,* and the overall sum, $\sum_{w=1}^{10}\sum_{d=1}^{10} \tilde{I}_{w,d}$ are simulated 10,000 times with a Bayesian MCMC model. The details are very similar to those described in Section 5 and will not be given here.

By setting the prior distribution of ρ equal to zero, we eliminate the between accident year correlation. Following the naming convention of the last section, let's call this model the Leveled Incremental Trend (LIT) model.

Table 9 shows the estimates of for the illustrative insurer with the CIT and the LIT model on paid data.

Before producing these distributions, I had no particular expectation of how ρ would be distributed for paid data. However, I did expect τ to be predominantly negative since the *p–p* plots in Figures 3, 4 and 10 indicted that the all the other models predicted results that were too high.

Let's first examine the effects of between-year correlation in the CIT model. Figure 13 gives the posterior distributions for ρ for the illustrative insurer. Figure 14 gives the histograms of the posterior means ρ for each insurer by line of business.

Table 9. CIT and LIT Models on Illustrative Insurer Paid Data

	CIT			LIT			Outcome
w	$\hat{C}_{w,10}$	SD	CV	$\hat{C}_{w,10}$	SD	CV	$C_{w,10}$
2	2539	9	0.0035	2538	9	0.0035	2527
3	4183	21	0.0050	4185	20	0.0048	4274
4	4395	40	0.0091	4393	32	0.0073	4341
5	3553	42	0.0118	3566	32	0.0090	3583
6	3063	101	0.0330	3151	40	0.0127	3268
7	5062	123	0.0243	5065	111	0.0219	5684
8	3512	514	0.1464	3355	234	0.0697	4128
9	4025	707	0.1757	4138	594	0.1435	4144
10	4698	1482	0.3155	4703	1489	0.3166	4139
Total	38942	1803	0.0463	39006	1723	0.0442	40000
Percentile		79.04			79.69		

As seen in Figure 14, the posterior means of ρ for the paid data were not as overwhelmingly positive as we saw in the incurred data shown in Figure 7. Figure 15 shows a small but noticeable difference between the standard deviations of the CIT and LIT models.

My efforts to rein in the correlation between the $\{\alpha_w\}$, the $\{\beta_d\}$, and the τ parameters were, at best, only partially successful, as Figure 16 indicates. The analogous plot for the LIT model is very similar. With the given data, it is hard for the CIT and the LIT models to sort out the effects of the level plus the development and the trend.

As seen in Figure 17, the posterior means of τ were predominantly negative. But as pointed out above, a negative might be offset by higher $\{\alpha_w\}$s and $\{\beta_d\}$s. Figure 18 shows only a handful of triangles where there was a noticeable decrease in the final expected loss estimates. And most of those differences appeared in the Other Liability line of business.

Figures 19 and 20 show the validation *p–p* plots for the CIT and the LIT models. As do the Mack, ODP and CCL models on paid data indicate, the predictive distributions for the CIT and LIT models tend to overstate the estimates of the expected loss.

Figure 13. Posterior Distribution of ρ and for the Illustrative Insurer

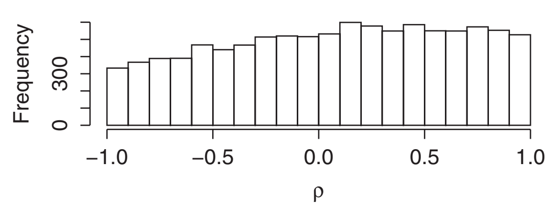


Figure 14. Posterior Mean of ρ by Line and Insurer for Paid Loss Data

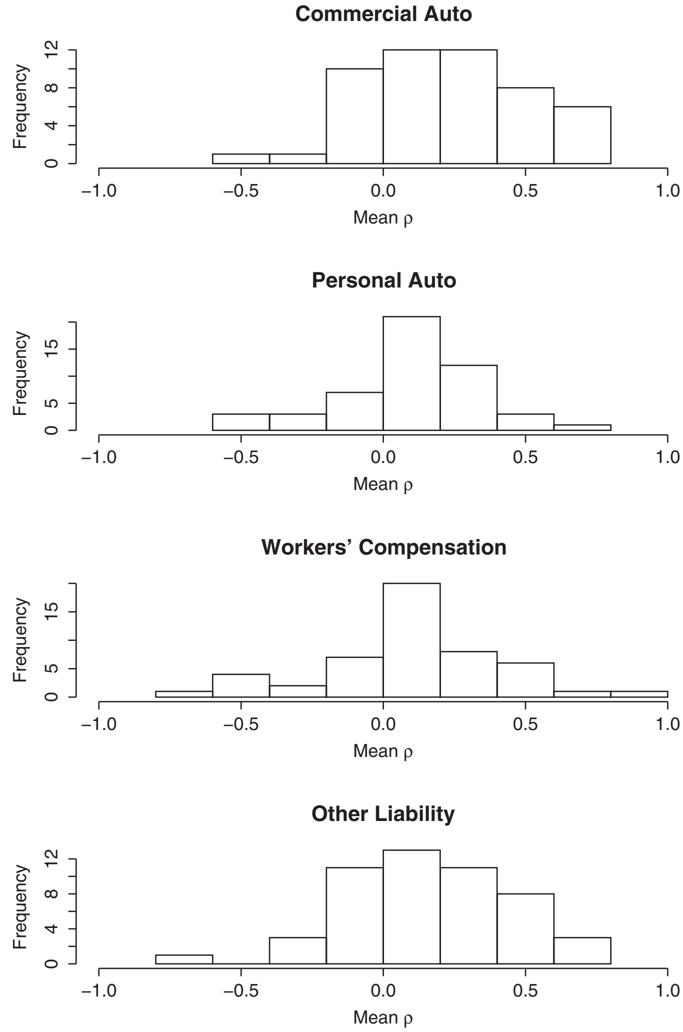


Figure 15. Compare the Standard Deviations of CIT and LIT Models for Paid Loss Data

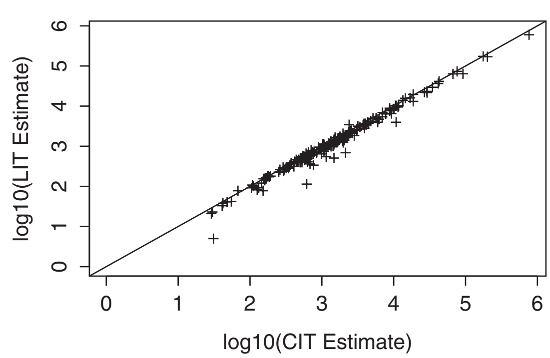


Figure 16. Correlations Between Parameters in the CIT Model for the Illustrative Insurer

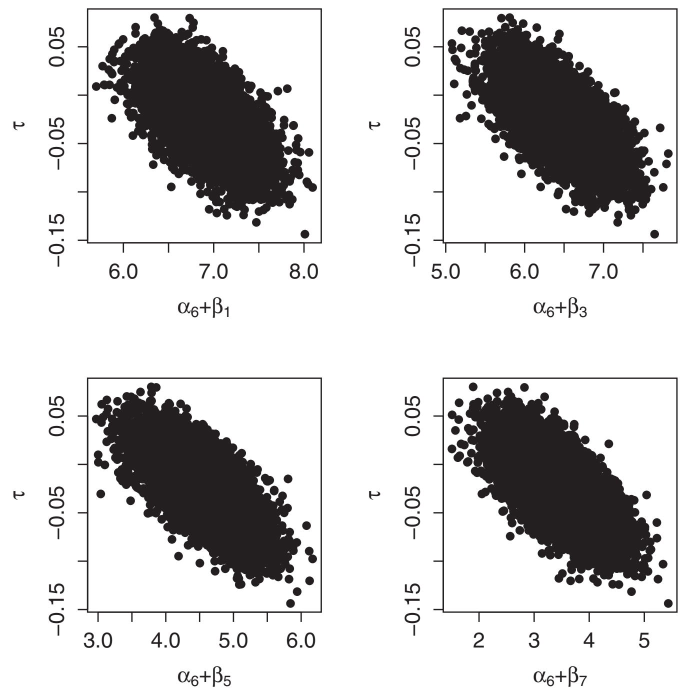
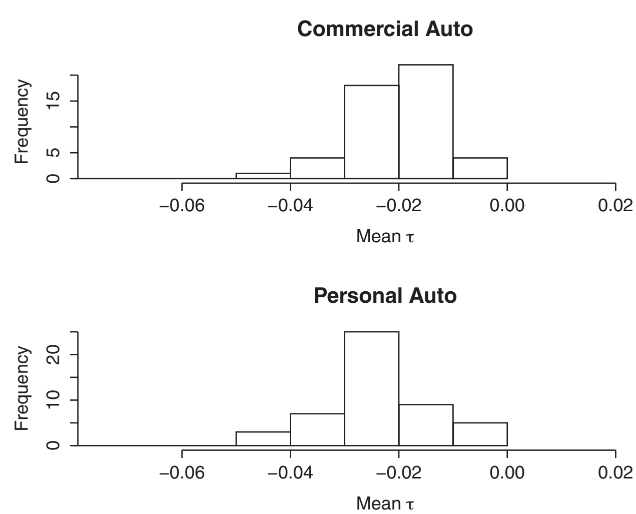


Figure 17. Posterior Mean of τ by Line and Insurer for Paid Loss Data



(continued on next page)

Figure 17. Posterior Mean of τ by Line and Insurer for Paid Loss Data *(continued)*

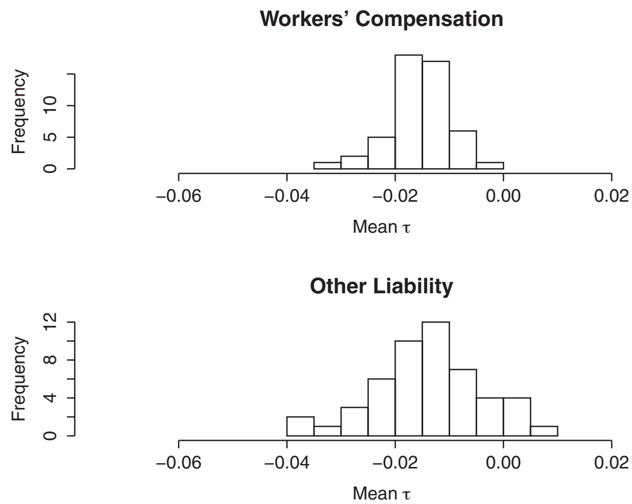


Figure 18. Comparing Estimates for the CCL and the CIT Models for Paid Data

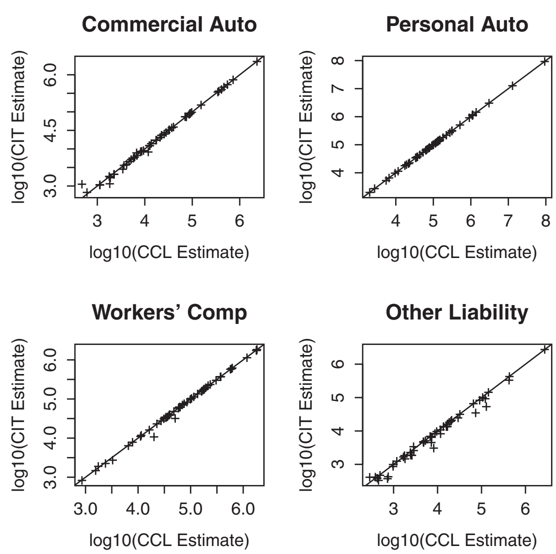


Figure 19. *p–p* Plots for the CIT Model

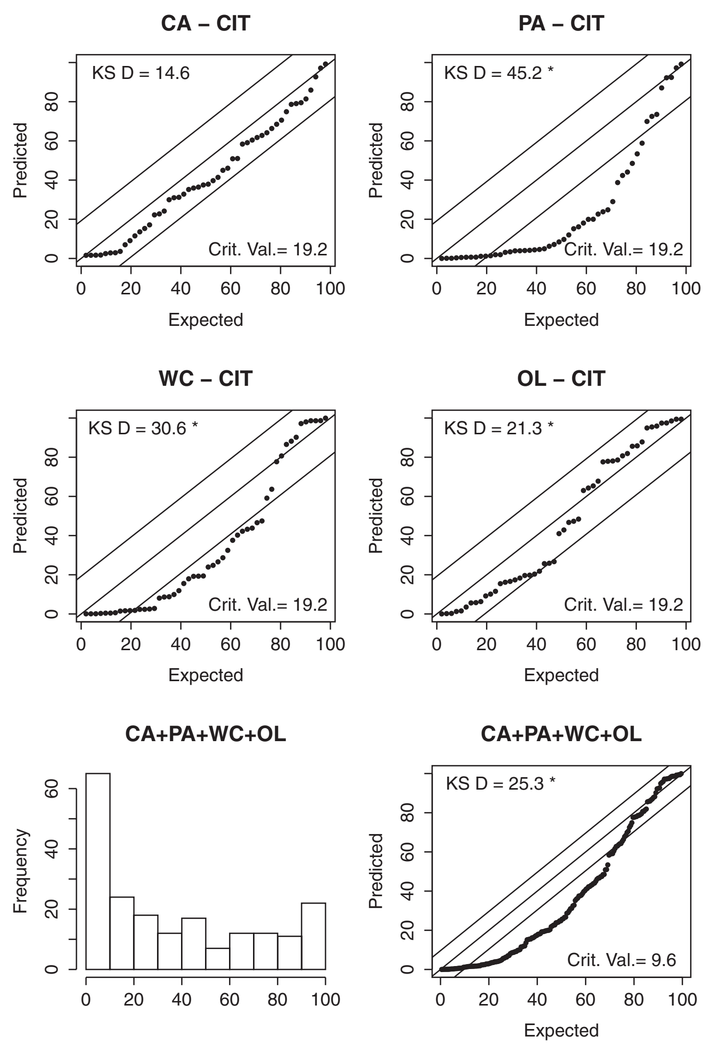
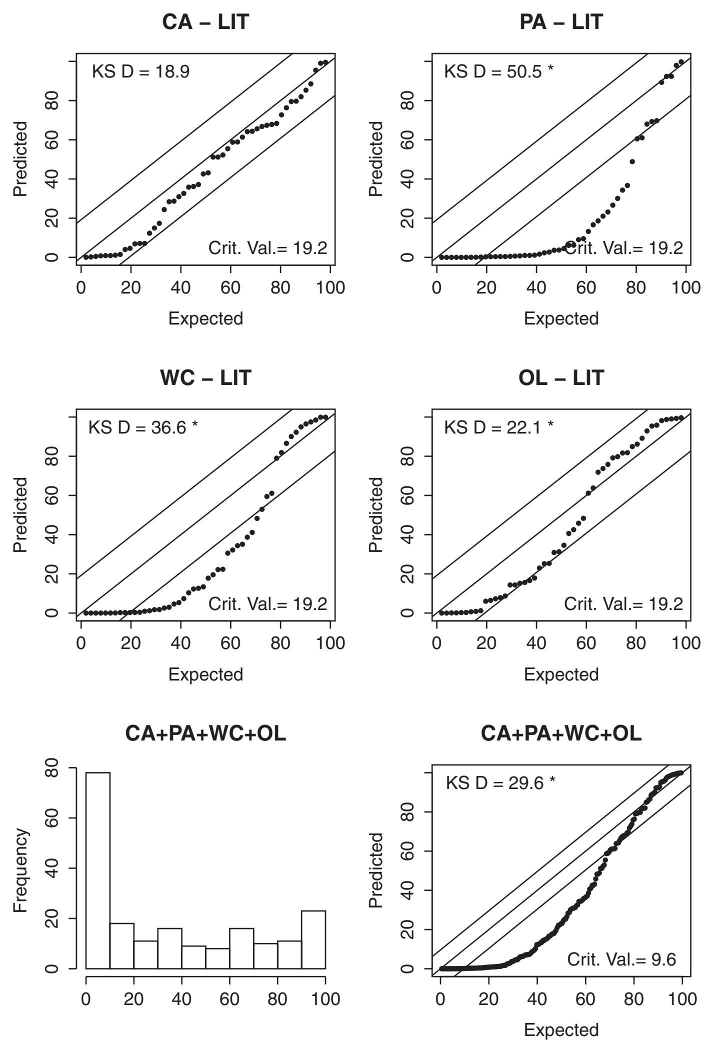


Figure 20. *p–p* Plots for the LIT Model



So, in spite of a serious attempt to improve on the results produced by the earlier models on paid data, the CIT and LIT models did not achieve the desired improvement. This result tends to support the idea that is generally accepted, that the incurred data reflects real information that is not in the paid data.

A reviewer of this monograph checked with some colleagues and found that claims are "reported and settled faster today due to technology," and suggested that the CIT model might not fully reflect this change. A model that addresses the possibility of a speedup of claim settlement is the following.

The Changing Settlement Rate (CSR) Model

Let:

1. Each $\alpha_w \sim \text{normal}\left(\log(Premium_w) + logelr, \sqrt{10}\right)$ where the parameter $logelr \sim$ uniform$(-1,0.5)$.
2. $\beta_d \sim$ uniform$(-5,5)$ for $d = 1, \ldots, 9$, $\beta_{10} = 0$.

3. $\mu_{w,d} = \alpha_w + \beta_d \cdot (1-\gamma)^{(w-1)}$ $\gamma \sim$ normal(0, 0.025).
4. Each $\sigma_d = \sum_{i=d}^{10} a_i$ where $a_i \sim$ uniform(0, 1).
5. $\tilde{C}_{w,d}$ has a lognormal distribution with log mean $\mu_{w,d}$ and log standard deviation σ_d subject to the constraint that $\sigma_1 > \sigma_2 > \ldots > \sigma_{10}$.

Since $\beta_{10} = 0$ and cumulative paid losses generally increase with the development year, d, β_d for $d < 10$ is usually negative. Then for each $d < 10$, a positive value of γ will cause $\beta_d \cdot (1-\gamma)^{(w-1)}$ to increase with w and thus indicate a speedup in claim settlement. Similarly, a negative value of γ will indicate a slowdown in claim settlement.

Table 10 shows the results for the CSR model on the illustrative insurer.

Figure 21 shows that the posterior distribution of γ is predominantly positive. This confirms the reviewer's contention that the claim settlement rate is, in general, increasing.

The validation *p–p* plots in Figure 22 shows that for three of the four lines of insurance, the CSR model corrects the bias found in the earlier models. This model also correctly predicts the spread of the predicted percentile of the outcomes for those lines. While the CSR model still exhibits bias for the personal auto line of business, the bias is significantly smaller than that of the other models.

It appears that the incurred loss data recognized the speedup in claim settlements.

Table 10. CIT and CSR Models on Illustrative Insurer Paid Data

	CIT			CSR			Outcome
w	$\hat{C}_{w,10}$	SD	CV	$\hat{C}_{w,10}$	SD	CV	$C_{w,10}$
1	3912	0	0	3912	0	0	3912
2	2539	9	0.0035	2559	103	0.0403	2527
3	4183	21	0.0050	4135	173	0.0418	4274
4	4395	40	0.0091	4285	198	0.0462	4341
5	3553	42	0.0118	3513	180	0.0512	3583
6	3063	101	0.0330	3317	216	0.0651	3268
7	5062	123	0.0243	4967	404	0.0813	5684
8	3512	514	0.1464	3314	402	0.1213	4128
9	4025	707	0.1757	3750	734	0.1957	4144
10	4698	1482	0.3155	3753	1363	0.3632	4139
Total	38942	1803	0.0463	37506	2247	0.0599	40000
Percentile		79.04			87.62		

Figure 21. Posterior Mean of γ by Line and Insurer for Paid Loss Data

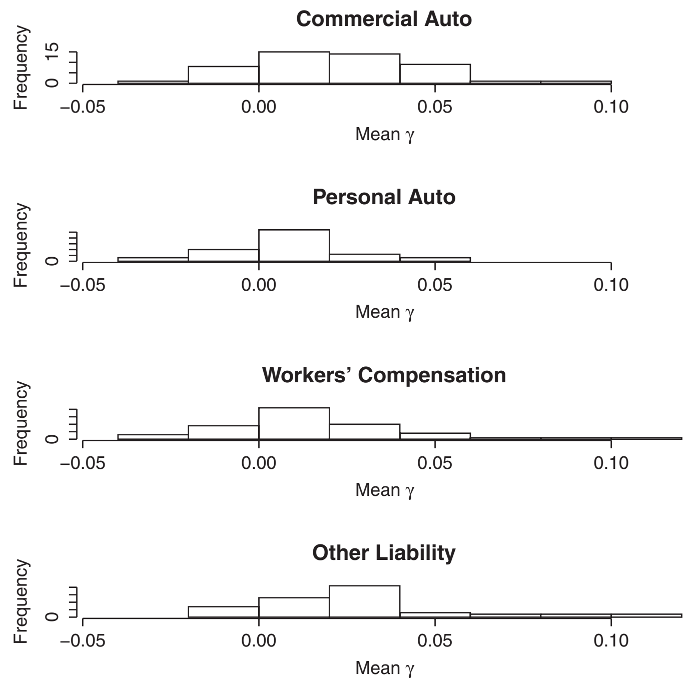
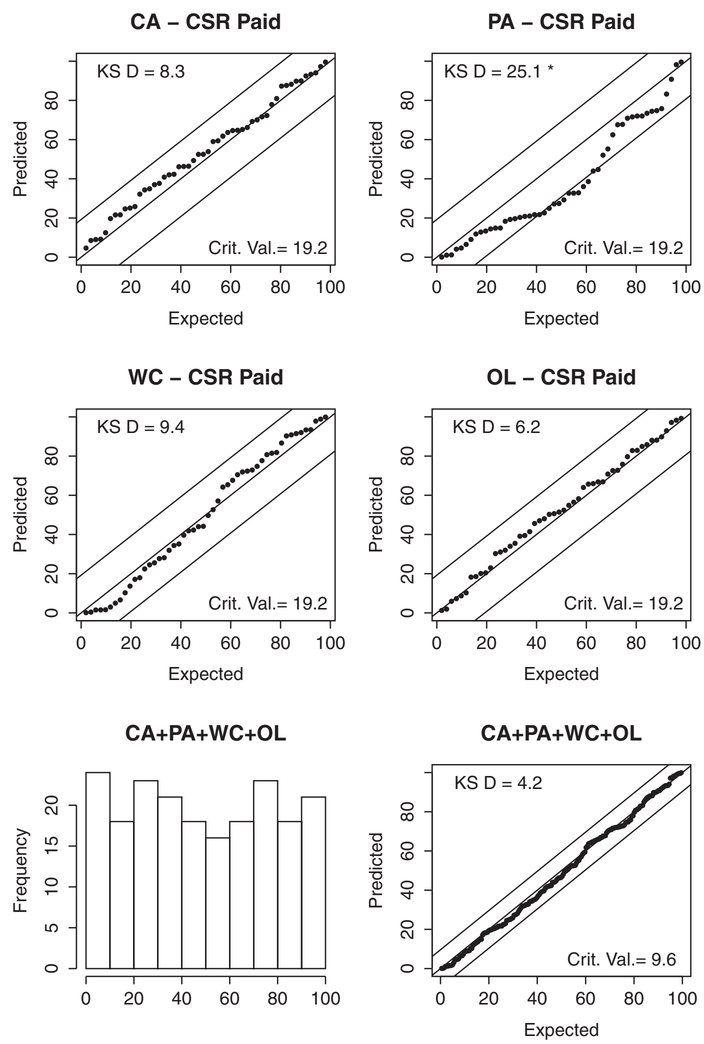


Figure 22. *p–p* Plots for the CSR Model



7. Process Risk, Parameter Risk and Model Risk

Let us now address a topic that frequently comes up in stochastic modeling discussions – process risk, parameter risk and model risk. One way to describe process and parameter risk is to consider the relationship for a random variable X conditioned on a parameter θ.

$$Var[X] = E_\theta[Var[X|\theta]] + Var_\theta[E[X|\theta]].$$

Let's call the left side of the above equation the "Total Risk." Let's call the first term of the right side the "Process Risk" as it represents the average variance of the outcomes from the expected result. Finally, let's call the second term the "Parameter Risk" as it represents the variance due to the many possible parameters in the posterior distribution. Another often-used term that overlaps with parameter risk is the "range of reasonable estimates."

For the CCL model, the parameter θ is represented by the vector

$$(\alpha_1, \ldots, \alpha_{10}, \beta_1, \ldots, \beta_9, \sigma_1, \ldots, \sigma_{10}, \rho).$$

The MCMC sample simulates 10,000 parameters denoted by θ_i. We then have the illustrative insurer:

$$\text{Total Risk} = Var\left[\sum_{w=1}^{10} \tilde{C}_{w,10}\right] = 1901^2.$$

The random variables $\mu_{w,10}$ are derived from the posterior distribution of the α_w. One can then use the formula for the mean of a lognormal distribution to calculate:

$$\text{Parameter Risk} = Var_\theta\left[E\left[\sum_{w=1}^{10} \tilde{C}_{w,10}|\theta\right]\right] = Var\left[\sum_{w=1}^{10} e^{\mu_{w,10}+\frac{\sigma_{10}^2}{2}}\right] = 1893^2.$$

For this example, the parameter risk is very close to the total risk, and hence there is minimal process risk. I have repeated this calculation on several (including some very large) insurers and I obtained the same result that process risk is minimal.

Model risk is the risk that one did not select the right model. If the possible models fall into the class of "known unknowns" one can view model risk as parameter risk. It is possible to formulate a model as a weighted average of the candidate models, with the weights as parameters. If the posterior distribution of the weights assigned to each model has significant variability, this is an indication of model risk. Viewed in this light, model risk is a special case of parameter risk.

As a thought experiment, one can consider what happens if we were to run this model on a very large dataset. The parameter risk will shrink towards zero and any remaining risk, such as model risk, will be interpreted as process risk.

This thought experiment is of largely academic interest since all aggregated loss triangles one finds in practice are small datasets. But it does serve to illustrate some of the theoretical difficulties that occur when one tries to work with the parameter/process/model classifications of risk. My own preference is to focus on total risk, as that it is the only risk that we can test by looking at actual outcomes.

8. Summary and Conclusions

The central thrust of this monograph is twofold.

- It implements the idea of large-scale retrospective testing of stochastic loss reserve models on real data. The goal is not to comment on the reserves of individual insurers. Instead the goal is to test the predictive accuracy of specific models.
- As shortcomings in existing models are identified, it demonstrates that Bayesian MCMC models can be developed to overcome some of these shortcomings.

The principle behind the retrospective testing is that a specific model is built with data that we customarily observe. The model is used to predict a distribution of outcomes that we will observe in the future. When we do observe outcomes for a large number of predictions, we expect the percentiles of the outcomes to be uniformly distributed. If they are not uniformly distributed, we look for a better model. We may or may not find one.

The data used in this study comes from the CAS Loss Reserve Database. It consists of hundreds of paid and incurred loss triangles that Peng Shi and I obtained from a proprietary database maintained by the NAIC. We are grateful that the NAIC allowed us to make these data available to the public. The data I used to build the models came from the 1997 NAIC Annual Statements. The outcomes came from subsequent statements.

Here is a high-level summary of the results obtained with these data.

- For incurred data, the variability predicted by Mack model is understated. One of its key assumptions is that the losses from different accident years are independent. This monograph proposes the correlated chain ladder (CCL) model as an alternative. This model allows for a particular form of dependency between accident years. It finds that the CCL model predicts the distribution of outcome correctly within a specified confidence level.
- For paid data, the bootstrap ODP model, the Mack model and the CCL model tend to give estimates of the expected ultimate loss that are high. This suggests that there is a change in the loss environment that is not being captured in these models. This monograph proposes three models, the Leveled Incremental Trend (LIT), the Correlated Incremental Trend (CIT) model, and the Changing Settlement Rate (CSR) as alternatives. The first two models allow for payment year trends. While the introduction of a payment year trend seems plausible given the bias identified in the earlier models, the performance of the LIT and CIT models are similar to the earlier models in the validation *p*–*p* plots. The CSR model corrects the bias identified in the previous models for three of the four lines of insurance, and has significantly less bias on the fourth line of insurance.

- Note that for the "Other Liability" line of insurance, the Mack and ODP models validate better than any of the new models proposed in this monograph. While it might be a small sample problem, the sample is not all that small. This suggests that more study is needed. Note that these results are for a specific annual statement year—1997. Studies such as this should be repeated on other annual statement years to see if the above conclusions still hold.

In preparing this monograph I have made every effort to adhere to the "open source" philosophy. The data is publicly available. The software is publicly available for free. The R and JAGS scripts used in creating these models are to be made publicly available. I have purposely restricted my methods to widely used software (R, JAGS and RStudio) in order to make it easy for others to duplicate and improve on these results.

In building the Bayesian models I used prior distributions that were as diffuse as I could make them. The restrictions I did make (for example, the restriction that $\sigma_1 > \sigma_2 > \ldots > \sigma_{10}$ in the CCL model) reflected my experience over several years of general model building. I did not have intimate knowledge of each insurer's business operations. Those with knowledge of an insurer's business operation should be able to incorporate this knowledge to obtain better results. As all probabilities are conditional, the Bayesian methodology allows for one to incorporate additional information by adjusting the prior distributions. I made every effort to code the models transparently so that such adjustments are easy to make.

The models proposed in this monograph are offered as demonstrated improvements over current models. I expect to see further improvements over time. The Bayesian MCMC methodology offers a flexible framework with which one can make these improvements.

References

Bardis, Emmanuel Theodore, Ali Majidi, and Daniel M. Murphy. 2012. "A Family of Chain-Ladder Factor Models for Selected Link Ratios." *Variance* 6:143–160.

Barnett, Glen, and Ben Zehnwirth. 2000. "Best Estimates for Reserves." *Proceedings of the Casualty Actuarial Society* 87:245–321.

Berger, James O. 1985. *Statistical Decision Theory and Bayesian Analysis.* 2nd ed. New York: Springer-Verlag.

Bornhuetter, Ronald L., and Ronald E. Ferguson. 1972. "The Actuary and IBNR." *Proceedings of the Casualty Actuarial Society* 59:181–195.

Brooks, Steve, Andrew Gelman, Galin L. Jones, and Xiao-Li Meng, eds. 2011. *Handbook of Markov Chain Monte Carlo.* London: Chapman and Hall/CRC.

De Alba, Enrique. 2002. "Bayesian Estimation of Outstanding Claims Reserves." *North American Actuarial Journal* 6(4):1–20.

England, P. D., and R. J. Verrall. 2002. "Stochastic Claims Reserving in General Insurance." Meeting paper. Presented to the Institute and Faculty of Actuaries, 28 January. http://www.actuaries.org.uk/research-and-resources/documents/stochastic-claims-reserving-general-insurance.

Gelfand, A. E., and A. F. M. Smith. 1990. "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association* 85(410):398–409.

Frühwirth-Schnatter, Sylvia, and Saumyadipta Pyne. 2010. "Bayesian Inference for" Mixtures of Univariate and Multivariate Skew-Normal and Skew-t Distributions." *Biostat* 11(2):317–336.

Hartman, Brian. 2014. "Bayesian Computational Methods." In *Predictive Modeling Applications in Actuarial Science,* vol. 1, *Predictive Modeling Techniques,* eds. Edward W. Frees, Richard A. Derrig and Glenn Meyers, 315–333. New York: Cambridge University Press.

Hastings, W. K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57(1):97–109.

Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences.* Hoboken, N. J.: Wiley.

Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. 2012. *Loss Models, From Data to Decisions.* 4th ed., Hoboken, N.J.: Wiley.

Mack, Thomas. 1993. "Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates." *ASTIN Bulletin* 23(2):213–225.

Mack, Thomas. 1994. "Measuring the Variability of Chain Ladder Reserve Estimates." Casualty Actuarial Society *Forum* (Spring):101–182.

McGrayne, Sharon Bertsch. 2011. *The Theory That Would Not Die.* New Haven, Conn.: Yale University Press.

Metropolis, N., and S. Ulam. "The Monte Carlo Method." *Journal of the American Statistical Association* 44(247):335–341.

Meyers, Glenn G. 2009. "Pure Premium Regression with the Tweedie Model." *The Actuarial Review* 36(2):12.

Meyers, Glenn G., and Peng Shi. 2011. "The Retrospective Testing of Stochastic Loss Reserve Models." Casualty Actuarial Society *E-Forum* (Summer):1–37.

Meyers, Glenn G. 2012. "The Leveled Chain Ladder Model for Stochastic Loss Reserving." Casualty Actuarial Society *E-Forum* (Summer):1–33.

Ntzoufras, I., and P. Dellaportas. 2002. "Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty." *North American Actuarial Journal* 6(1): 113–128.

Pigeon, Mathiew, Katrien Antonio, and Michel Denuit. 2013. "Individual Loss Reserving with the Multivariate Skew Normal Framework." *ASTIN Bulletin* 43(3): 399–428.

Scollnik, D. P. M. 2001. "Actuarial Modeling with MCMC and BUGS." *North American Actuarial Journal* 5(2):96–124.

Taleb, N. N. 2010. *The Black Swan: The Impact of the Highly Improbable.* 2nd ed. New York: Random House.

Taylor, G. C. 1986. *Claims Reserving In Non-Life Insurance.* New York: Elsevier.

Verrall, Richard. 2007. "Obtaining Predictive Distributions for Reserves Which Incorporate Expert Opinion." *Variance* 1:53–80.

Appendix A. The Data Selection Process

When selecting the loss triangles to use in this monograph my overriding consideration was that the process should be mechanical and well defined. There are two potential mistakes one can make in selecting the insurers to analyze.

- If one were to take all the insurers in the database, or randomly select the insurers, there could be some insurers who made significant changes in their business operations that could violate the assumptions underlying the models.
- If one is too selective, one runs the risk of selecting only those data that best fit a chosen model. For example, let's suppose that I wanted the CCL model to fit the incurred data even better than it does. As an extreme case, noting that CCL model still appears to be a bit light in the tails, I could have replaced some of the insurers that have outcomes in the tail with other insurers that have outcomes in the middle.

While I did not have inside information on any changes in the business operations, Schedule P provides some hints in their reporting of both net and direct earned premium by accident year. Both of these data elements are in the CAS Loss Reserve Database.

- If an insurer makes significant changes in its volume of business over the 10-year period covered by Schedule P, a change in business operation could be inferred.
- If an insurer makes significant changes in its net to direct premium ratio over the 10-year period, a change in its reinsurance strategy could be inferred.

To carry out an analysis of this sort, I needed a large number of insurers. After looking at the quality and consistency of the data available in the CAS Loss Reserve Database, I decided to use 50 insurers in each of four major lines of insurance—Commercial Auto, Personal Auto, Workers Compensation, and Other Liability. Early on I concluded that there were an insufficient number of insurers in the Products Liability and the Medical Malpractices lines to obtain an adequately sized selection.

To implement these considerations, I calculated the coefficients of variation for the net earned premiums and the net to direct premium ratios over the ten available years. By trial and error, I then set up limits on these coefficients (CV) of variation that obtained the desired number of insurers. This procedure should have eliminated some of the insurers that changed their business operations.

After some provisional testing, I eliminated insurer group 38997 from the Personal Auto and Workers Comp lines, and insurer groups 16373, 44598 and 14885 from the

Other Liability line because the R "ChainLadder" package produced "NA" results for the Mack calculation of the standard deviation. I also eliminated insurer group 14451 from the Other Liability line because the MCMC algorithm took very long to converge for paid losses. After eliminating these insurer groups I adjusted the CV limits to give 50 insurers for each line. The final CV limits are given in Table 11. The final list of the selected insurer groups are in Table 12.

Table 11. CV Limits for Insurer Triangles

	Commercial Auto	Personal Auto	Workers' Comp	Other Liability
CV(Premium)	<0.399	<0.450	<0.950	<0.390
CV(Net/Direct)	<0.125	<0.125	<0.200	<0.125

Table 12. Group Codes for Selected Insurers

Commercial Auto		Personal Auto		Workers' Comp		Other Liability	
353	13420	353	13641	86	13528	620	14370
388	13439	388	13889	337	14176	669	14915
620	13641	620	14044	353	14320	671	15113
833	13889	692	14176	388	14508	683	15148
1066	14044	715	14257	671	14974	715	15210
1090	14176	1066	14311	715	15148	833	15571
1538	14257	1090	14443	1252	15199	1538	17043
1767	14320	1538	15199	1538	15334	1767	17450
2135	14974	1767	15393	1767	18309	2003	17493
2208	15199	2003	15660	2135	18538	2135	18163
2623	18163	2143	16373	2712	18767	2143	18686
2712	18767	3240	16799	3034	18791	2208	24830
3240	19020	4839	18163	3240	21172	3240	26797
3492	21270	5185	18791	5185	23108	5185	27065
4839	25275	6947	23574	6408	26433	5320	28550
5185	27022	7080	25275	6807	27529	6459	30449
6408	27065	8427	27022	7080	30589	6947	30651
6459	29440	8559	27065	8559	32875	7625	32301
6947	31550	10022	27499	8672	33499	10657	33049
7080	32301	13420	27766	9466	34576	13501	36315
8427	34606	13439	29440	10385	35408	13919	38733
10022	35483	13501	31550	10699	37370	13994	41068
10308	37036	13528	34509	11347	38687	14044	41580
11037	38733	13587	34592	11703	38733	14176	42439
11118	44598	13595	34606	13439	41300	14257	43354

Appendix B. Introduction to Bayesian MCMC Models

Since the recognition of Markov Chain Monte Carlo as a powerful tool for doing Bayesian analyses in 1990, there have been many efforts to create software to aid in these analyses. Progress in making the available software faster and more user friendly is still being made. In spite of this progress, I believe that it is necessary for an actuary to have a picture of what is happening inside the black box. The purpose of this appendix is to provide a brief description of what is inside the black box.

A Markov chain is a random process where the transition to the next state depends only on its current state, and not on prior states. Formally, a Markov chain, X_t, for $t = 1, 2, \ldots$ is a sequence of vectors satisfying the property that

$$\Pr(X_{t+1} = x | X_1 = x_1, X_2 = x_2, \ldots, X_t = x_t) = \Pr(X_{t+1} = x | X_t = x_t).$$

The properties of Markov chains have been well studied by scholars. Those interested in these studies can start with Chapter 4 of Jackman (2009). What actuaries need to know about Markov chains in Bayesian MCMC analyses can be summarized as follows.

- There is a certain class of Markov chains, generally called "ergodic," for which the vectors, $\{X_t\}$, approaches a limiting distribution. That is to say that as T increases, the distribution of $\{X_t\}$ for all $t > T$ approaches a unique limiting distribution.
- The Markov chains used in Bayesian MCMC analyses, such as the Metropolis Hastings algorithm, are members of this class.
- Let x be a vector of observations and let y be a vector of parameters in a model. In Bayesian MCMC analyses, the Markov chain is defined in terms of the prior distribution, $p(y)$, and the conditional distribution, $f(x|y)$. The limiting distribution is the posterior distribution, $f(y|x)$. That is to say, if we let the chain run long enough, the chain will randomly visit all states with a frequency that is proportional to their posterior probabilities.

The operative phrase in the above is "long enough." In practice we want to: (1) develop an algorithm for obtaining a chain that is "long enough" as quickly as possible; and (2) develop criteria for being "long enough."

Here is how Bayesian MCMC analyses work in practice.

1. The user specifies the prior distribution, $p(y)$, and the conditional distribution, $f(x|y)$.

2. The user selects a starting vector, x_1, and then, using a computer simulation, runs the Markov chain through a sufficiently large number, t_1, of iterations. This first phase of the simulation is called the "adaptive" phase, where the algorithm is automatically modified to increase its efficiency.
3. The user then runs an additional t_2 iterations. This phase is called the "burn-in" phase. t_2 is selected to be high enough so that a sample taken from subsequent t_3 periods represents the posterior distribution.
4. The user then runs an additional t_3 iterations and then takes a sample, $\{x_t\}$, from the $(t_2 + 1)^{th}$ step to the $(t_2 + t_3)^{th}$ step to represent the posterior distribution $f(y|x)$.
5. From the sample, one then constructs various "statistics of interest" that are relevant to the problem addressed by the analysis.

The most common algorithms for generating Bayesian Markov chains are variants of the Metropolis-Hastings algorithm.

Given a prior distribution, $p(y)$, and a conditional distribution, $f(x|y)$, the Metropolis-Hastings algorithm introduces a third distribution, $J(y_t|y_{t-1})$, called the "proposal" or "jumping" distribution. Given a parameter vector, y_{t-1}, the algorithm generates a Markov chain by the following steps.

1. Select a candidate value, y^*, at random from the proposal distribution, $J(y_t|y_{t-1})$.
2. Compute the ratio

$$R \equiv R_1 \times R_2 = \frac{f(x|y^*) \cdot p(y^*)}{f(x|y_{t-1}) \cdot p(y_{t-1})} \times \frac{J(y_{t-1}|y^*)}{J(y^*|y_{t-1})}.$$

3. Select U at random from a uniform(0,1) distribution.
4. If $U < R$ then set $y_t = y^*$. Otherwise set $y_t = y_{t-1}$.

The first part of the ratio, R_1, represents the ratio of the posterior probability of the proposal, y^*, to the posterior probability of y_{t-1}. The higher the value of R_1, the more likely will be accepted into the chain. Regardless of how the proposal density distribution is chosen, the distribution of y_t can be regarded as a sample from the posterior distribution, after a suitable burn-in period.

To see the issues that can arise when implementing the Metropolis-Hastings algorithm, let us examine the following made-up example.

Sample Claim Data

484	1407	2262	5015	6500
603	1565	2654	5354	6747
631	1894	2672	5464	9143
1189	2140	4019	5598	12782
1229	2244	4318	6060	18349

We want to model the losses using a lognormal distribution with unknown parameter μ and known parameter $\sigma = 1$.

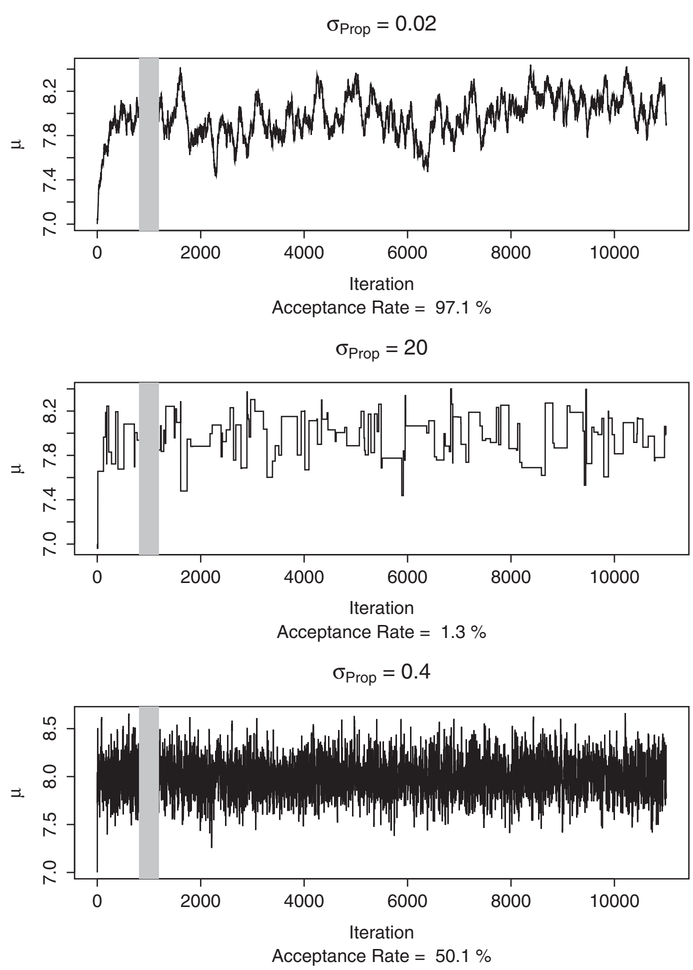
The prior distribution of μ is a normal distribution with mean 8 and standard deviation 1. For the proposal distribution of $(\mu_t|\mu_{t-1})$, I chose a normal distribution with mean μ_{t-1} and standard deviation σ_{Prop}. The starting value, μ_1, was set equal to 7.00. For this example, there is no adaptive phase and the burn-in phase was 1,000 iterations.

To illustrate the effect of the choice of the proposal distribution, I ran the Metropolis-Hastings algorithm using the normal proposal distributions with $\sigma_{Prop} = 0.02$ (low volatility), $\sigma_{Prop} = 20$ (high volatility) and $\sigma_{Prop} = 0.4$ (volatility just about right). Figure 23 shows plots of the value of μ_t as the chain progresses for each choice of σ_{Prop}. These plots are generally called trace plots in the MCMC literature.

Note that while the starting value $\mu_1 = 7$ was outside of the high density region of the posterior distribution of μ, as t increases μ_t moves rather quickly into the high density region for $\sigma_{Prop} = 20$ and $\sigma_{Prop} = 0.4$. It takes a bit longer for $\sigma_{Prop} = 0.02$, as the differences between μ^* and μ_{t-1} tend to be small.

If $\sigma_{Prop} = 0.02$, μ^* will be close to μ_{t-1} and the ratio in Step 2 of the Metropolis–Hastings algorithm will be relatively high and thus μ_t will be close to μ_{t-1}. In the first

Figure 23. Trace Plot 1: Metropolis—Hastings Example



trace plot of Figure 23 we see a high degree of autocorrelation between successive iterations. If $\sigma_{Prop} = 20$, μ^* could be quite far from μ_{t-1} and the ratio in Step 2 could be relatively low and thus μ_t will equal μ_{t-1}. In the second trace plot of Figure 23 we still see a high degree of autocorrelation. If $\sigma_{Prop} = 0.4$, μ^* can be far enough away from μ_{t-1} to reduce the autocorrelation, and close enough to avoid rejection and the setting of $\mu_t = \mu_{t-1}$. Getting a good value for σ_{Prop} is balancing act. The third trace plot in Figure 23 shows a relatively low degree of autocorrelation and suggests that μ_t for $t = 1001, \ldots, 11000$ is a representative sample from the posterior distribution.

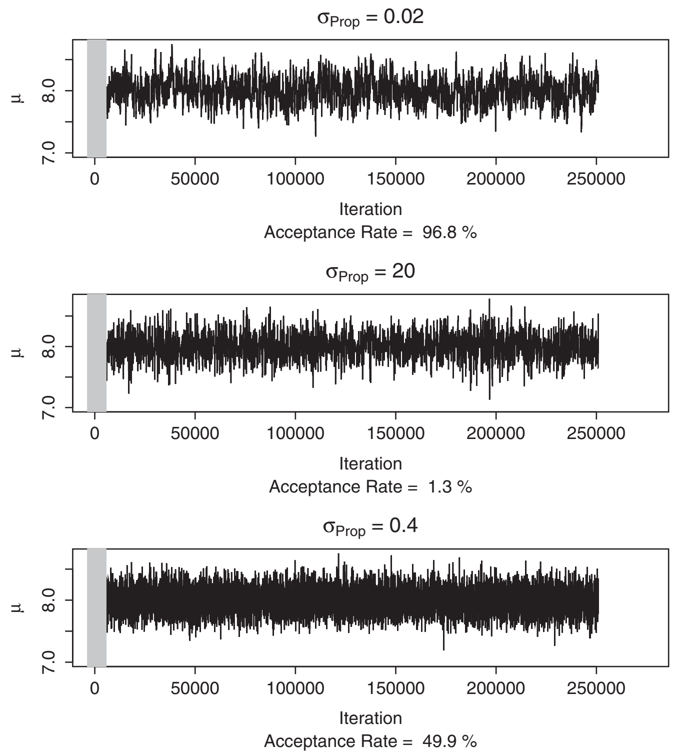
For a single parameter model, like the one in this example, it is relatively easy to scale the proposal distribution by trial and error to minimize autocorrelation. For models with many parameters, like the ones in the next section, such manual scaling is not practical. This problem has been studied extensively and here is a short description of the current state of the art.

A good statistic to look at when trying to minimize autocorrelation in the Metropolis-Hastings algorithm is the acceptance rate of y^* into the Markov chain. I have scanned a number of sources, e.g., Chapter 5 in Jackman (2009), or Chapter 4 of Brooks et al. (2011), that suggest that an acceptance rate of about 50% is near optimal for a one parameter model. The optimal acceptance rate decreases to about 25% as we increase the number of parameters in our model. Also, the researchers have developed methods to automatically adjust the proposal density function in the Metropolis-Hastings algorithm. Chapter 4 of Brooks et al. (2011) provides a recent description of the state of the art. We shall see below that all this has been mechanized in JAGS. The phase of generating the Markov chain where the proposal density function is optimized is called the "adaptive" phase.

As models become more complex, adaptive MCMC may not be good enough to eliminate the autocorrelation. While the theory on Markov chain convergence still holds, there is no guarantee on how fast it will converge. So if one observes significant autocorrelation after the best scaling effort, the next best practice is to increase t_3 until there are a sufficient number of ups and downs in the trace plot and then take a sample of the $t_1 + t_2 + 1$ to $t_1 + t_2 + t_3$ iterations. This process is known as "thinning." Figure 24 shows what happens when we increase t_3 to 250,000 and record every 25th observation.

Before leaving this example, let us examine how one might turn the posterior distribution of μ into something of interest to actuaries. One reason actuaries fit a lognormal distribution to a set of claims is that they want to determine the cost of an excess layer. Given the parameters μ and σ of a lognormal distribution, there are formulas in Appendix A of Klugman, Panjer, and Willmot (2012) that give the cost of an excess layer of loss. The functions that calculate these formulas are included in the R "actuar" package. As the posterior distribution of μ reflects the parameter risk in our model, it is also possible to reflect the parameter risk in the expected cost of a layer by calculating the expected cost of the layer for each μ in the simulated posterior distribution. Also, it is possible to simulate an actual outcome of a loss, X, in a layer given each μ in the posterior distribution. The distribution of X calculated in this way

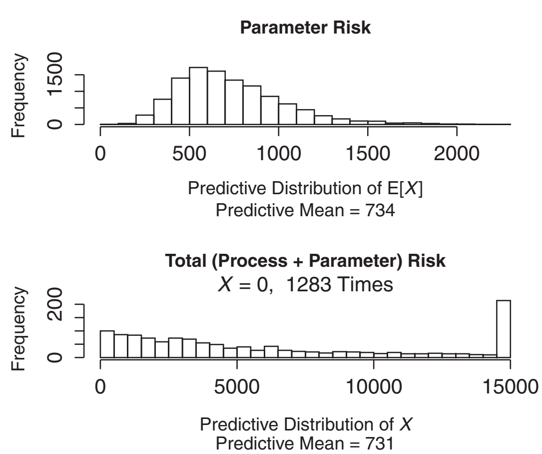
Figure 24. Trace Plot 2: Metropolis—Hastings Example with Thinning



reflects both the parameter risk and the process risk in the model. Figure 25 shows the predictive distribution of the expected cost of the layer between 10,000 and 25,000, $E[X]$, and the predicted outcome of losses X in that layer.

As statisticians and practitioners became aware of the potential for Bayesian MCMC modeling in solving real-world problems, a general software initiative to implement Bayesian MCMC analyses, called the BUGS project, began. BUGS is an

Figure 25. Predictive Distributions of the Cost of a Layer between 10,000 and 25,000



acronym for **B**ayesian inference **U**sing **G**ibbs **S**ampling.[17] The project began in 1989 in the MRC Biostatistics Unit, Cambridge, and led initially to the 'Classic' BUGS program, and then onto the WinBUGS software developed jointly with the Imperial College School of Medicine at St Mary's, London. The project's web site is at http://www.mrc-bsu.cam.ac.uk/bugs/. The various software packages associated with the BUGS project have captured many of the good techniques involved in Bayesian MCMC modeling.

On the advice of some colleagues I chose to use the JAGS (**J**ust **A**nother **G**ibbs **S**ampler) package. It has the additional feature that it runs on a variety of platforms (Window, Mac, Linux and several varieties of Unix). Like R, it can be downloaded for free.

I use JAGS with R. My typical MCMC program begins by reading in the data, calling the JAGS script using the R package "runjags." I then fetch the sample of the posterior back into the R program where I calculate various "statistics of interest."

While I realize that JAGS is doing something more sophisticated, I find it helpful to "think" of JAGS as using a simple version of the Metropolis–Hasting algorithm similar to that illustrated in the example above. Once a model is specified, there are three stages in running a JAGS program:

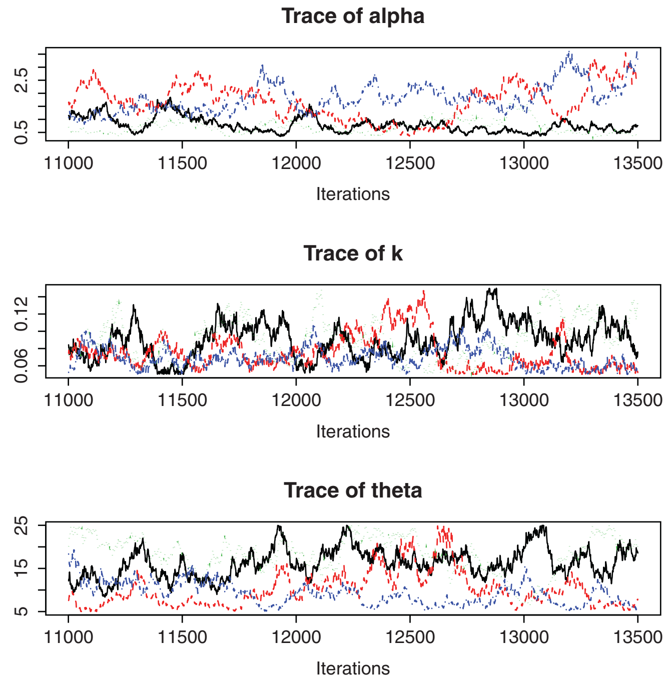
1. The adaptive stage where JAGS modifies the proposal distribution in the Metropolis-Hastings algorithm. JAGS will issue a warning if it thinks that you haven't allowed enough iterations for adapting. Let's denote the number of iterations for scaling by t_1.
2. The burn-in stage runs until we have reached the limiting posterior distribution. JAGS has diagnostics (described below) that indicate convergence. The burn-in stage runs from iterations $t_1 + 1$ to $t_1 + t_2$.
3. The sampling stage that produces the sample of the posterior distribution. The sampling stage runs from iterations $t_1 + t_2 + 1$ to $t_1 + t_2 + t_3$.

JAGS has a number of convergence diagnostics that are best illustrated with an example. We are given the total losses from a set of thirty insurance policies in the following table.

Exposure	Loss	Exposure	Loss	Exposure	Loss
51	23	226	273	368	410
66	138	231	275	374	482
119	53	254	259	377	500
125	88	255	200	381	424
131	80	258	123	392	242
152	136	268	275	444	431
196	165	279	327	449	337
197	136	295	509	478	399
225	328	340	457	484	458
225	347	364	317	495	553

[17] Gibbs sampling is an MCMC algorithm that is a special case of the Metropolis Hastings algorithm. This is demonstrated in Chapter 1 of Brooks *et al.* (2011).

Figure 26. Trace Plots Without Thinning—CRM Example



Our task is to use these data to estimate the expected cost of losses in excess of 1000 for an insurance policy with an exposure of 800. Note that in our data, there is no insurance policy with an exposure as high as 800, and no loss over 1000.

Let's use the collective risk model with a Poisson distribution for the claim count, and a distribution for the claim severity. Here is the description of the model using the notation in Klugman, Panjer, and Willmot (2012).[18]

1. $\lambda = k \cdot \text{Exposure}$
2. $n \sim \text{Poisson}(\lambda)$
3. $Loss \sim \Gamma(n \cdot \alpha, \theta)$
4. $k \sim \text{Uniform}(0.05, 0.15)$
5. $\alpha \sim \text{Uniform}(0.1, 10)$
6. $\theta \sim \text{Uniform}(5,200)$

In JAGS, the script looks pretty much like the model description above after a change in notation for the distribution parameters. Let's first consider convergence diagnostics. First of all, with JAGS one can run multiple independent chains. I first ran this model with 1,000 iterations for the adaptive stage, 10,000 iterations for the burn-in stage and then 2,500 iterations for the sampling stage. JAGS then produces trace plots for all four chains, colored differently, superimposed on each other. A visual indication of convergence is that all the chains bounce around in the same general area. Figure 26 shows the trace plots produced by JAGS for the three parameters in this example.

[18] This particular version of the collective risk model is called a Tweedie distribution. See Meyers (2009).

As we can see from the trace plots, the chains are very distinct, so we should conclude that the chains have not converged.

A second diagnostic provided by JAGS is the Gelman-Rubin statistic for each parameter. Here is a heuristic description of the statistic.[19] First estimate the within chain variability, W, and the between chain variability, B. Gelman and Rubin then recommend that one use the statistic

$$\sqrt{\hat{R}} = \sqrt{\frac{\widehat{W} + \hat{B}}{\widehat{W}}}.$$

The $\sqrt{\hat{R}}$ is called the "potential scale reduction (or 'shrink') factor." or PSRF. This statistic will approach one as the number of iterations increases, since the between chain variability will approach zero. What we need to know is how long the chains have to be before we can stop and get a representative sample of the posterior distribution. Chapter 6 of Brooks et al. (2011) recommends that we accept convergence if the PSRF is 1.1 or below for all parameters. The default for the "runjags" package is 1.05, which is what I used in for the models in this monograph. The PSRFs for this JAGS run were 1.87, 1.21 and 1.92 for the parameters α, k and θ, respectively.

Continuing the example, I reran the JAGs model with same parameters but thinned the chains to take every 25th iteration. The results are in Figure 27. The PSRFs for this JAGS run were 1.03, 1.02 and 1.01 for the parameters α, k and θ respectively. So we can accept that the run has converged.

JAGS then sent 10,000 parameter sets $\{\alpha_t, k_t, \theta_t\}$ back to the R script. R then simulated losses to the insurance policy as follows.

For t = 1 to 10,000.

1. Set $\lambda = k_t \cdot 800$.
2. Select n_t at random from a Poisson distribution with mean λ.
3. Select $Loss_t$ at random from a $\Gamma(n_t \cdot \alpha_t, \theta_t)$ distribution.[20]

Figure 28 shows a histogram of the ground up losses from the above simulation and the expected cost of the layer in excess of 1,000.

The examples in this appendix illustrate the ideas behind Bayesian MCMC models, those being the adaptive phase, the burn-in phase, the sampling phase, and convergence testing. Understanding these concepts should enable one to start running these kinds of models. When running these models one should keep in mind that the state of the art is still evolving, so one should periodically check the current literature and software developments on Bayesian MCMC modeling for recent developments.

---

[19] See Jackman (2009, Section 6.2) or Hartman (2014) for a more detailed description of this statistic.

[20] If each X_i has a $\Gamma(\alpha, \theta)$ distribution, then $X_1 + \cdots + X_n$ has a $\Gamma(n \cdot \alpha, \theta)$ distribution.

Figure 27. Trace Plots With Thinning—CRM Example

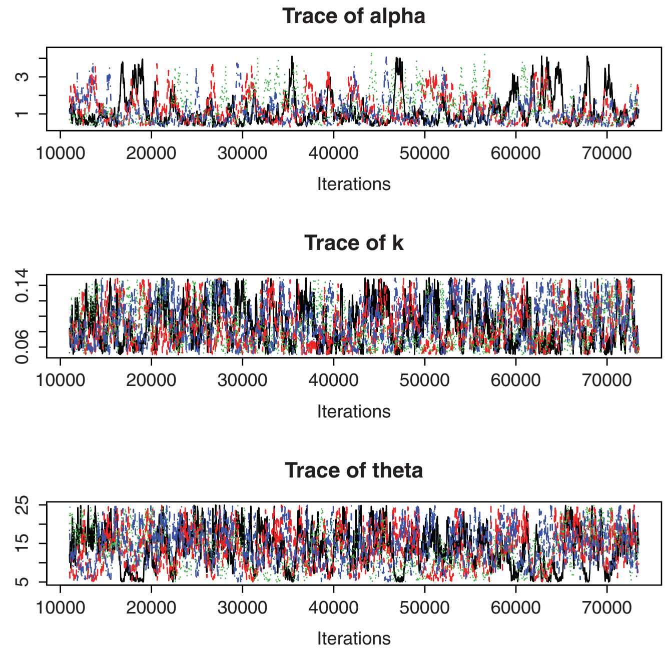
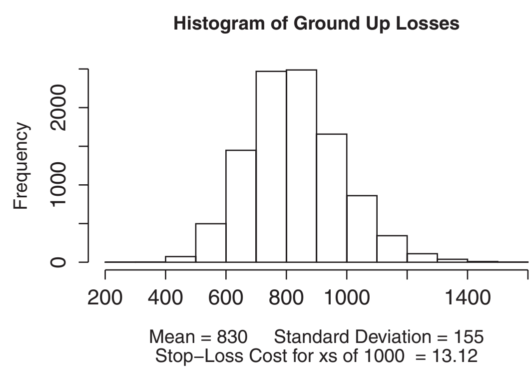


Figure 28. Output—CRM Example



Appendix C. Bayesian MCMC Model Implementation

The state of the art and the software for Bayesian MCMC modeling is still evolving. Since there may be upgrades by the time the reader sees this monograph, I think that it is important for me to describe the computing environment in which I ran the models in this monograph.

My computer was a Macbook Pro with a quad core processor. On this computer I used R version 3.0.2 and JAGS version 3.3, implementing JAGS with the "runjags" package. The main consideration in selecting the "runjags" package was that made it easy to run the four chains in parallel with my quad core computer. Running the chains in parallel made a significant improvement in the run time.

For the LCL, CCL, and CSR models I used 1,000 iterations for the adaptive phase, and 10,000 iterations for the burn-in phase. I ran the model inside a loop, with the sampling phase initially set at 10,000 iterations with a thinning parameter equal to four. If the maximum PSRF for the parameters I monitored was greater than 1.05, I doubled the number of iterations in the sampling phases and the thinning parameter and ran the simulation again—continuing until the target PSRF target was achieved.

For most of the LCL and CCL models on incurred data, the initial run achieved the PSRF target. The highest thinning parameter was 32. Convergence was somewhat slower for the CCL and CSR models on the paid data. There was one triangle that required a thinning parameter equal to 512.

For the CIT and LIT models on the paid data, I increased the burn-in to 50,000 iterations. Convergence was noticeably slower. Far fewer triangles met the PSRF target with a thinning parameter set equal to four.

The R/JAGS scripts for all models are in a spreadsheet that will be distributed with this monograph. For each model, I put these scripts inside a loop that ran all 200 triangles while I was otherwise occupied. Summary statistics for all 200 triangles are also included in the spreadsheet and because I fixed the random number seed, the scripts are able to reproduce the summary statistics for any of the triangles.

ABOUT THE SERIES:

CAS monographs are authoritative, peer-reviewed, in-depth works focusing on important topics within property and casualty actuarial practice. For more information on the CAS Monograph Series, visit the CAS website at www.casact.org.

Expertise. Insight.
Solutions.

www.casact.org

(Revised January 2, 2013)[1]

Claims Development by Layer: The Relationship between Claims Development Patterns, Trend and Claim Size Models

Rajesh Sahasrabuddhe, FCAS, MAAA

The purpose of Charles Cook's 1970 paper *Trend and Loss Development Factors* was to address the "overlap fallacy." That is, the focus of that paper was to demonstrate that trend and claims development were mutually exclusive adjustments. While this is certainly true, it should also be understood that there is a relationship between limited claims development patterns and trend factors. The "connector" between claims development patterns and trend is the claim size model. This relationship is critical to analyzing "real word" data which is rarely available on a ground-up, unlimited basis and where the implicit assumption of trend in a single direction may not be appropriate.

This paper presents a demonstration of that relationship and also provides an approach to adjust development patterns for a particular claim size layer in order to calculate a development pattern for any other layer. As importantly, the approach discussed is designed to produce models that are internally consistent with respect to development patterns, trend factors and size of loss models (increased / decreased limit factors).

Keywords development patterns, excess layer

1. INTRODUCTION

The purpose of this paper is to demonstrate the relationship between claims development, trend and claim size factors. Those relationships are then explored in order to provide a practical approach for adjusting a development pattern appropriate for any claim layer to produce a development pattern for any other layer. The approach also allows for adjustments related to cost level assumptions implicit in development patterns and ensures that assumptions related to claim size models, claims development and trend are internally consistent.

The procedure may be applied to either paid claims or reported claims. Additionally, although we use "claims" in the discussion, the procedure may also be applied to claims and allocated claim adjustment expenses (or only allocated claim adjustment expenses) assuming that all parameters and assumptions are defined consistently.

[1] A previous revision dated November 25, 2012 corrected minor typographical errors in Equations 2.3 and 3.6, and the cross reference for the calculation of item D1 in Examples 1 and 2.
This January 2, 2013 revision includes exhibits that were inadvertently excluded from the November 25, 2013 version. Those exhibits include a minor correction to Example 3.

1.1 Research Context

The current approach for estimating excess layer development is based on Emanuel Pinto and Daniel Gogol's paper, "*An Analysis of Excess Loss Development.*" The focus of that paper is the fitting of observed development factors as a function of retentions. The observed factors were developed using an analysis of a large industry database. Pinto/Gogol then present an approach for calculating excess layer development in Section 5 and this approach is explored further in George M. Levine's review. However, this approach requires that the actuary first calculate excess layer development using their fitting approach.

Many actuaries would not have access to such industry data and as such the Pinto/Gogol approach would not be practical. In addition to this issue, the methodology does not use the inherent relationship of claims size models, trend and claims development patterns.

1.2 Scope and Objective

This paper includes comments related to assumptions implicit in the determination of development patterns, trend and claim size distributions in practice. However, the development of these actuarial models and their parameters is beyond the scope of this paper. The objective of this paper is to provide a methodology to calculate development factors by layer once the actuary has already determined his/her assumptions with respect to a "base" development pattern, trend and claim size models.

1.3 Outline

The paper presents a discussion of a robust approach and then provides an example that incorporates simplifying assumptions that are common in actuarial practice. The remainder of the paper proceeds as follows. Section 2 will provide notation and define important algebraic definitions of model factors. Section 3 provides the discussion of the inter-relationship between claims development, trend and claim size models. Section 4 will provide implementation examples to the oft-studied *Mack* triangle and a simpler approach that may be sufficient for many analyses.

2. BACKGROUND

We begin by examining the implicit and explicit assumptions of claims development, trend and claim size models.

The discussion will assume that we are analyzing an $n \times n$ claims triangle. We generalize our discussion to allow for data that is truncated from below at d and censored from above at p. This is typical of data subject to deductibles and policy limits. Of course, if $d = 0$ and $p = \infty$, then the claims data is provided on a ground-up, unlimited (GUU) basis. The notation used in this paper is as follows:

$C_{i,j}^{\boldsymbol{L}}$ = Cumulative claims in the layer L, for exposure period i as of the end of development interval j

$C_{i,\infty}^{\boldsymbol{L}}$ = Ultimate claims in the layer L, for exposure period i ($j = \infty$)

$L(d, p)$ = Claims layer truncated from below at d and censored from above at p where $0 \leq d < p \leq \infty$

Though it will be obvious that this is not a necessary assumption, in order to simplify notation, we will assume claims layer $\boldsymbol{L}$ is consistent throughout the data triangle. Claims data is typically organized as presented in Table 1.

TABLE 1
CUMULATIVE CLAIMS DATA

Exposure Period (i)	Development Interval (j): 1	2	3	…	n
1	$C_{1,1}^{\boldsymbol{L}}$	$C_{1,2}^{\boldsymbol{L}}$	$C_{1,3}^{\boldsymbol{L}}$	…	$C_{1,n}^{\boldsymbol{L}}$
2	$C_{2,1}^{\boldsymbol{L}}$	$C_{2,2}^{\boldsymbol{L}}$	$C_{2,3}^{\boldsymbol{L}}$	…	
3	$C_{3,1}^{\boldsymbol{L}}$	$C_{3,2}^{\boldsymbol{L}}$	$C_{3,3}^{\boldsymbol{L}}$	…	
…	…				
n	$C_{n,1}^{\boldsymbol{L}}$				

Below we first discuss trend, claims size models and development patterns separately and then discuss their relationships.

2.1 Trend Factors

Trend rates typically refer to the annual change in cost level for a particular claims layer. In practice, trend rates often do not vary between accident periods. In addition, trend that acts in the development period or calendar period direction is often not considered. Finally, the consideration of the varying effects of trend applicable to different claims layer is often nonexistent.

Rather than using annual rates of change, we will use cost level indices, $\boldsymbol{T}$. Cost level indices are determined so as to apply to cumulative claims for accident year i as of development maturity j. The indices are an accumulation of the incremental changes relative to a "base cost level." Any accident year and maturity combination can be considered the "base." In practice, the base cost level will typically be defined as the cost level associated with ultimate claims for the oldest exposure period.

Our trend is explicitly defined to apply to the ground-up, unlimited claims layer. This is consistent with approaches in practice where the trend assumption is based on external cost information such as the Consumer Price Index. If trend is estimated from claims data that is subject to policy limits or deductibles then we will first need to adjust the data to a ground-up, unlimited basis using the claim size model.

Our model allows for trend that acts in multiple directions. We use the following notation for cost level indices.

$T_{i,j}$ = Trend indices for cumulative GUU claims for exposure period *i* at the end of development interval *j*

TABLE 2
COST LEVEL INDICES

Exposure Period (*i*)		Development Interval (*j*) 1	2	3	…	*n*
	1	$T_{1,1}$	$T_{1,2}$	$T_{1,3}$	…	$T_{1,n}$
	2	$T_{2,1}$	$T_{2,2}$	$T_{2,3}$	…	
	3	$T_{3,1}$	$T_{3,2}$	$T_{3,3}$	…	
	…	…				
	n	$T_{n,1}$				

2.2 Claim Size Model

The claim size model describes the distribution of claim sizes. Though we do not restrict claim size models with respect to complexity, for practicality we require the following:

- that claims size model parameters can be adjusted for the impact of inflation (includes most common claim size models such as the lognormal and exponential)
- that limited expected values and unlimited means (first moments) can be calculated with reasonable effort.

2.2.1 Limit Adjustment Factors

The limit adjustment factors, $S(a,b)$, represents the ratio of expectations of claims between layer L_a and L_b.

$$S_{i,\infty}(\boldsymbol{L}_a, \boldsymbol{L}_b) = \{LEV(p_a; \boldsymbol{\Phi}_{i,\infty}) - LEV(d_a; \boldsymbol{\Phi}_{i,\infty})\} \;/\; \{LEV(p_b; \boldsymbol{\Phi}_{i,\infty}) - LEV(d_b; \boldsymbol{\Phi}_{i,\infty})\} \tag{2.1}$$

$$S_{i,j}(\boldsymbol{L}_a, \boldsymbol{L}_b) = \{LEV(p_a; \boldsymbol{\Phi}_{i,j}) - LEV(d_a; \boldsymbol{\Phi}_{i,j})\} \;/\; \{LEV(p_b; \boldsymbol{\Phi}_{i,j}) - LEV(d_b; \boldsymbol{\Phi}_{i,j})\} \tag{2.2}$$

$$S_{i,j}(L_a, L_b) = \mathrm{E}[C_{i,j}^{\boldsymbol{L}_a} / C_{i,j}^{\boldsymbol{L}_b}] \tag{2.3}$$

where LEV is the characteristic limited expected value function for the claim size model and $\boldsymbol{\Phi}$ represents the "name" (e.g. lognormal, Pareto, exponential) and parameters of the claim size model. We also acknowledge that the parameters of the claims size model, $\boldsymbol{\Phi}$, will vary by exposure period i and development interval j as a result of differences in cost level.

In later sections, we will use the notation $LEV(\boldsymbol{L}; \boldsymbol{\Phi})$ to refer to the limited expected value for the layer $L(d, p)$. This is calculated as follows:

$$LEV(\boldsymbol{L}; \boldsymbol{\Phi}) = LEV(p; \boldsymbol{\Phi}) - LEV(d; \boldsymbol{\Phi}) \tag{2.4}$$

2.2.2 Gross-up Factors

In the special case where $p_a=\infty$ and $d_a=0$, $S(a,b)$ simplifies to a factor to gross-up claims to a GUU basis. We can then use the characteristic first moment (mean) function, M, in the numerator rather than the limited expected value function.

$$G_{i,.}(b) = M(\boldsymbol{\Phi}_{i,\infty}) \;/\; \{LEV(p_b; \boldsymbol{\Phi}_{i,\infty}) - LEV(d_b; \boldsymbol{\Phi}_{i,\infty})\} \tag{2.4}$$

$$G_{i,j}(b) = M(\boldsymbol{\Phi}_{i,j}) \;/\; \{LEV(p_b; \boldsymbol{\Phi}_{i,j}) - LEV(d_b; \boldsymbol{\Phi}_{i,j})\} \tag{2.5}$$

2.3 Claims Development

Claims development factors, F, represent the expected ratios of ultimate claims to claims at maturities prior to ultimate. That is:

$$F_{i,j}^{\boldsymbol{L}} = \mathrm{E}[\, C_{i,\infty}^{\boldsymbol{L}} \;/\; C_{i,j}^{\boldsymbol{L}} \,] \tag{2.6}$$

3. RESULTS AND DISCUSSION

We can now explore the relationships between claims development, trend, and claim size models. The discussion assumes that we have been provided with unlimited claims trend factors and that we have developed the cost level indices as presented in Table 2.

3.1 Claim Size and Trend

As per the requirements of Section 2.2, for our selected claim size model, we can calculate model parameters for prior or future exposure periods using the trend indices.

$$\boldsymbol{\Phi}_{i,j} \sim f\left(\boldsymbol{\Phi}_{n,j}, T_{i,j}, T_{n,j}\right) \tag{3.1}$$

3.2 Claim Development Patterns, Claim Size and Trend

In practice, claims development patterns are estimated from unadjusted data and are applied to claims for all exposure periods. We should acknowledge that this is not appropriate unless (i) claims data are provided on a GUU basis and (ii) trend acts only in the accident year direction. Since this is oftentimes not the case, we address these issues by adjusting the triangle of claims data prior to analysis. Specifically, we adjust observed claim amounts for differences in cost level and limit using the limited expected value function.

3.2.1 Development of Basic Limit Claims Development Pattern, Exposure Year *n* Cost Level

We first select a Basic Limit, $\boldsymbol{B}$, which is the threshold at which we believe the data is sufficiently credible for the purpose of estimating claims development patterns. Recall from Table 1 that $\boldsymbol{L}$ represents the layer for which data is available. We then adjust each observation of cumulative claims as follows[2]:

$$E\left[\acute{C}_{i,j}^{\boldsymbol{B}} | C_{i,j}^{\boldsymbol{L}}\right] = C_{i,j}^{\boldsymbol{L}} \times LEV(\boldsymbol{B}; \boldsymbol{\Phi}_{n,j}) / LEV(\boldsymbol{L}; \boldsymbol{\Phi}_{i,j}) \tag{3.2}$$

We note that there is no restriction that $\boldsymbol{B} \neq \boldsymbol{L}$. We should recognize that if $\boldsymbol{B} = \boldsymbol{L}$, then we are simply adjusting the data for differences due to the impact of trend in the layer. (Note the difference between the first subscript of $\boldsymbol{\Phi}$ in the numerator and denominator of Equation 3.2).

We then analyze this adjusted data, $\acute{C}_{i,j}^{\boldsymbol{B}}$, in order to estimate development patterns at a common (basic) limit and an exposure period i=n cost level. This pattern is denoted $F_{n,j}^{\boldsymbol{B}}$ and we have the following relationship:

$$F_{n,j}^{\boldsymbol{B}} = E\left[\acute{C}_{n,\infty}^{\boldsymbol{B}} / \acute{C}_{n,j}^{\boldsymbol{B}}\right] \tag{3.3}$$

As you review the following sections, keep in mind that this basic limit development pattern at exposure year *n* cost level will now be used to calculate basic limit development for any other layer and exposure period (cost level).

---

[2] We presume that a triangle at the basic limit is not readily available.

3.2.2 Calculation of Claims Development Pattern for Any Layer and Cost Level

Equation 3.2 also provides an important general relationship applicable to any layer $\boldsymbol{X}$ if we have data for layer $\boldsymbol{L}$.

$$E\left[C_{i,j}^{\boldsymbol{X}} | C_{i,j}^{\boldsymbol{L}}\right] = C_{i,j}^{\boldsymbol{L}} \times LEV(\boldsymbol{X}; \boldsymbol{\Phi}_{i,j})/LEV(\boldsymbol{L}; \boldsymbol{\Phi}_{i,j}) \tag{3.4}$$

$$C_{i,j}^{\boldsymbol{L}} \times S_{i,j}(\boldsymbol{X}, \boldsymbol{L}) \tag{3.5}$$

Using this general relationship, we can calculate basic limit development factors for any exposure period for any layer $\boldsymbol{X}$ from the development factor for $\boldsymbol{B}$ at exposure year n cost levels:

$$F_{i,j}^{\boldsymbol{X}} = E\left[\frac{C_{i,\infty}^{\boldsymbol{X}}}{C_{i,j}^{\boldsymbol{X}}}\right] = E\left[\frac{C_{n,\infty}^{B}}{C_{n,j}^{B}} \times \frac{LEV(\boldsymbol{X}; \boldsymbol{\Phi}_{i,\infty})/LEV(\boldsymbol{B}; \boldsymbol{\Phi}_{n,\infty})}{LEV(\boldsymbol{X}; \boldsymbol{\Phi}_{i,j})/LEV(\boldsymbol{B}; \boldsymbol{\Phi}_{n,j})}\right] \tag{3.6}$$

$$F_{i,j}^{\boldsymbol{X}} = F_{n,j}^{\boldsymbol{B}} \times \frac{LEV(\boldsymbol{X}; \boldsymbol{\Phi}_{i,\infty})/LEV(\boldsymbol{B}; \boldsymbol{\Phi}_{n,\infty})}{LEV(\boldsymbol{X}; \boldsymbol{\Phi}_{i,j})/LEV(\boldsymbol{B}; \boldsymbol{\Phi}_{n,j})} \tag{3.7}$$

$$F_{i,j}^{\boldsymbol{X}} = F_{n,j}^{\boldsymbol{B}} \times \frac{S_{i,\infty}(\boldsymbol{X}, \boldsymbol{B})}{S_{i,j}(\boldsymbol{X}, \boldsymbol{B})} \tag{3.8}$$

However, as we demonstrated in Equation 3.1, $\boldsymbol{\Phi}_{i,j}$ is a function of trend indices and $\boldsymbol{\Phi}_{n,j}$. So, substituting Equation 3.1 into Equation 3.7, we have:

$$\boxed{F_{i,j}^{\boldsymbol{X}} = F_{n,j}^{\boldsymbol{B}} \times \frac{LEV(\boldsymbol{X}; T_{i,\infty}, T_{n,\infty}, \boldsymbol{\Phi}_{n,\infty})/LEV(\boldsymbol{B}; \boldsymbol{\Phi}_{n,\infty})}{LEV(\boldsymbol{X}; T_{i,j}, T_{n,j}, \boldsymbol{\Phi}_{n,j})/LEV(\boldsymbol{B}; \boldsymbol{\Phi}_{n,j})}} \tag{3.9}$$

Equations 3.8 and 3.9 are the primary findings of this research: **Development factors at different cost levels and different layers are related to each other based on claim size models and trend.**

3.3 Other Practical Uses

Oftentimes, we are simply provided with a development pattern. Although we are typically aware of the limits associated with the triangle and/or pattern, it is not stated at any particular cost level.

In Equation 3.9, we demonstrated that, for limited claims data, development patterns will vary with cost level. However, this relationship is often ignored usually because it is presumed immaterial. For convenience, we will simply assert that the cost level is that of the latest exposure period.

We also typically have a claim size model at ultimate (e.g. increased limit factors), but size models by age are usually not available. Let us also assume that we are only concerned with estimating development factors applicable to claims at the latest valuation date.

We can use a variation of Equation 3.6 to develop claims development patterns:

$$F_{i,j}^{\boldsymbol{X}} = F_{n,j}^{B} \times \frac{LEV(\boldsymbol{X}; \boldsymbol{\Phi}_{i,\infty})/LEV(\boldsymbol{B}; \boldsymbol{\Phi}_{n,\infty})}{R_j\ (X, B)} \tag{3.10}$$

The primary difference between Equations 3.8 and Equation 3.6 is that rather than using claim size models by age in the denominator, we use a quantity, $R_j\ (X, B)$, that is simpler to estimate approximately.

$R_j\ (X, B)$ is the ratio between limited expected values for layer $\boldsymbol{X}$ and $\boldsymbol{B}$ at the end of development interval *j*. $R_j\ (X, B)$ is only evaluated along a single diagonal since we typically have at least one diagonal (usually the current diagonal) where we can observe ratios of claims at various limits. It should be noted that *R* carries only one subscript, that for maturity. In using this latter approach, we assume that differences in cost level are immaterial to the calculation of **ratios** of claims by layer[3].

For the moment, we will ignore the possibility of negative development and assume that $R_j\ (X, B)<1$. The latter assumption indicates that we are trying to develop an estimate for a pattern at a lower layer given a pattern at a higher layer. We should recognize that $\boldsymbol{R}$ will have the following properties:

i. $\boldsymbol{R}_a > \boldsymbol{R}_b$ for a<b - At early maturities, there will be less development in the excess layer than at later maturities.

ii. $\boldsymbol{R}_a \geq U$, where $\boldsymbol{U} = \lim_{a\to\infty} \boldsymbol{R}_a$ - We should recognize that *U* can be calculated as the product of $\boldsymbol{R}$ and the ratio of ultimate claim development factors layer $\boldsymbol{X}$ and $\boldsymbol{B}$. Until we reach ultimate, the reported ratio will always be greater than ultimate ratio. This is because the there is more development associated with the denominator of $\boldsymbol{R}$ (claims in layer $\boldsymbol{B}$, the higher limit) than the numerator of $\boldsymbol{R}$ (claims in layer $\boldsymbol{X}$, the lower limit) and at ultimate $\boldsymbol{R} = \boldsymbol{U}$.

iii. If our base development pattern is provided on an unlimited basis (i.e. $\boldsymbol{B}$=GUU), then the maximum value for $\boldsymbol{R}$ may be calculated as *U*Claims Development Factor*. The derivation of this maximum is presented in Appendix A.

It should be recognized that these conditions will be violated if there is negative development or if we assume that an excess layer might develop more quickly than a working layer. These conditions are not necessary for application of this approach. However, it is useful to review the results under the typical considerations described above to provide a more intuitive understanding of the dynamics of the calculation.

[3] Note that we are not asserting that they are immaterial with respect to absolute limited expected values.

In the third example presented in Section 4, we use a simpler approach to calculating $\boldsymbol{R}$[4] which is then used to calculate development factors for a layer other than the layer associated with the development pattern provided.

3.4 Issues

Relative to common development method projections, the procedure described above requires additional assumptions and calculations. The use of certain assumptions and calculations would not appear to be overly onerous:

1. The procedure requires that the actuary select a basic limit. However, actuaries either explicitly or implicitly select a basic limit in applying the development method. That is, whenever a development triangle is analyzed there is an implicit assumption that the limit associated with that triangle is sufficiently credible to produce development factors.

2. The procedure requires the use of a(n ultimate) claim size model in order to implement a development method analysis. This may or may not result in an additional burden on the actuary. Oftentimes, claim size information (such as increased limit factors) or a claim size model is already available to the actuary. If not, we would submit that knowledge of the distribution of claim sizes is important in understanding the dynamics of claims development.

   We should also recognize that we use the claim size model only to calculate relative limited expected values near the deductible, basic limit, policy limit and limit underlying the development data. Deductibles generally would not be an issue for the types of exposures for which the actuary would be willing to invest the effort required of this approach. As such, what is important is that our claim size model produces reasonable ratios of limited expectations to unlimited means at higher values. It is less important that the absolute limited expected values are accurate and therefore a simpler size of loss model may be sufficient though we need to recognize its shortcomings and not use that model out of context.

3. The procedure requires that the data triangle be adjusted to a basic limit and common cost level. As demonstrated in Examples 1 & 2 of Section 4, given claim size and trend information, the calculation and application of adjustment factors would not seem to create a significant additional burden.

---

[4] Simpler than calculating claim size models by age.

There are however two sets of assumptions that could be perceived as resulting in a significant additional burden.

1. Claim size models at maturities prior to ultimate are generally not available. In addition, these models would have limited application outside of this context. However, understanding changes in claims size models over time would be a significant benefit for actuaries to understand excess layer development.

   With an insurance company database or even a self-insured risk of sufficient size, we believe that an algorithm could be reasonably programmed to calculate these claim size models.

   Although a robust claim size model is required for full implementation of this approach (Examples 1 & 2), it should be recognized that only the ratio of expected values is required to adjust development patterns from one layer to another. This is a significantly reduced burden as will be demonstrated in Example 3 in the next section.

2. The procedure requires the calculation of a triangle of trend indices in order to implement a development method analysis. We would expect that a trend assumption exists in the analysis. The trend indices specify the cost level associated with cumulative claim observations. This becomes somewhat difficult to conceptualize in two respects:

   a. Trend typically acts on incremental activity.

   b. The impact of trend on reported incurred claims and, more specifically, the timing of the effect of trend on case basis reserves, is difficult to ascertain.

   These difficulties are not an issue if we assume that development only acts in the exposure period direction. Even if we have trend also acting across calendar periods, we would submit that this will require the actuary to confront the assumption with respect to the direction(s) in which trend acts or (more importantly) does not act. In addition documenting this assumption produces greater transparency and better informs the consumer of actuarial information.

4. EXAMPLES

We now present three examples that implement the concepts described in Section 3. The first two examples are based on the oft-studied claims triangle included in the *Distribution-Free Calculation* *of the Standard Error of Chain Ladder Reserve Estimates* by Thomas Mack. Example 1 and Example 2 are identical except that in Example 1, the Basic Limit is well above the working claims layer; in Example 2, the Basic Limit is within the working layer. The third example presents the approach discussed in Section 3.3 where we adjust a development pattern provided to us to determine patterns for other layers.

4.1 Example 1 & 2

For Examples 1 & 2, we provide the following additional (contrived) information about the Mack triangle. This information is intended to be typical of that which might apply to actual data:

- We have selected a basic limit of $500 thousand
- The policy limit is $2 million
- The data in the triangle is for the ground-up layer to $1 million
- Trend acts at a rate of 2% each exposure period; but there was a one-time increase to 5% between exposure period 6 and 7.
- Trend acts at a rate of 1% each calendar period; but there was a one-time decrease of 5% between calendar period 2 and 3.

The calculations in the examples are presented as follows:

- In Section A, we present the claims data and relevant information. Both exposure periods and development intervals are annual. However, since this is not a strict requirement of our approach, we have retained the more generic labels: "Exposure Period" and "Development Interval."
- In Section B, we present the calculation of trend indices.
- In Section C, we present the claim size model. Section C1 provides the claim size model at Exposure Period 10 cost level. We use an exponential model for simplicity of presentation; however any model that meets the requirements of Section 2.2 could be used.

  In Section C2, we present the calculation of adjusted exponential parameters based on the Exposure Year 10 parameters and trend indices.

  In Sections C4 through C6, we present the calculation of limited expected values using the characteristic function of the exponential model.
- In Section D1, we present the adjusted cumulative claims triangle. This triangle adjusts all historical observations to the basic limit at Exposure Period 10 cost levels. The

adjustments are based on ratios of limited expected values. In Sections D2 and D3, we calculate the incremental and cumulative development patterns.

- In Section E, we apply Equation 3.7 to calculate development factors for various layers at appropriate exposure year cost levels. In Section E7, we present the differences between factors calculated through examination of the (unadjusted) triangle in Section A1 and the factors resulting from our approach.

  Factors for certain excess layers are presented as "very large." This occurs since the expectation of claim in the layer at early maturities is very small.

  We note that the differences presented in Section of E7 of Example 1 are quite small. The differences will grow with the expectation of claims in the layer between the basic limit and layer under review. This is demonstrated in Example 2, where the resulting differences are quite a bit greater. We should also recognize that layers that are excess layers for an insurer (or self-insured) become working layers for reinsurers (excess insurers).

  It will also grow in situations where trend and/or development act over longer periods or at higher rates.

4.2 Example 3

The third example presents the approach described in Section 3.3. This approach is intended to provide a simpler application of the theory in Section 3. As presented in Example 1, if the basic limit is sufficiently high and trend is contained, the impact of data adjustments is minimal.

The calculations in Example 3 are reasonably self-explanatory. However, readers should note the following:

- At ultimate, all claims development factors equal unity and the ratio at age (col. 9) equals the ratio at ultimate (col. 8).
- The x axis is labeled "maturity," not exposure period. The observed pattern should be viewed as one observation of a random process at a particular maturity and not viewed as the ratio applicable to an exposure period.
- We use an algorithm to select ratios by age. At the earliest maturity, we know that the ratio should be "high." That is because claims emergence in excess layers is still "low."

  Our selected ratios are calculated as follows:

Selected Ratio = Ultimate Ratio + (1-Ultimate Ratio) * Decay Factor

This approach recognizes that we want to "keep" a portion of the distance between the ultimate ratio and the maximum ratio (unity). This portion is determined through the use of a decay model where we keep most of the difference at the earliest maturity and none at ultimate.

In practice, assuming we are analyzing development patterns at limits at or above the working layer, the ratios will be close to unity and the amount of error that could possibly be created by this approach is minimal.

5. CONCLUSION

In this paper we have demonstrated that there is a relationship between claim development patterns by layer and that that relationship is a function of trend and claim size models. This relationship can be used to calculate development patterns for a claims layer from a development pattern for any other claims layer.

These relationships also demonstrate that limited development factors are a function of not only maturity but also cost level. Therefore, the same pattern of limited factors should not always be applied to all exposure periods under review.

With short development patterns, low trend rates and limits above the working layer, the adjustment is small and often immaterial. Not all exposures exhibit these characteristics and for these exposures, the adjustments may be meaningful. For exposures where the adjustment may not be meaningful, we provided an alternative simpler approach to adjust development patterns.

Acknowledgment

The author acknowledges Katy Siu and Jason Shook, for their reviews of this paper. Any remaining errors in the paper are solely the responsibility of the author.

Appendix A: Calculation of Maximum Ratios of Basic Limit to Unlimited Claims

The maximum ratio is represented by the limiting case where all development in the unlimited layer occurs above the basic limit. The maximum ratio is calculated as follows:

Notation:

R = Ultimate ratio of basic limit to unlimited claims
A = Ratio of basic limit to unlimited claims prior to ultimate
D = Unlimited claim development factor

Claims

	Prior to Ultimate	At Ultimate
Limited to Basic Limit	B_a	B_r
Excess of Basic Limit	X_a	X_r
Unlimited	C_a	C_r

Identities:

I1: $B_a = B_r$ (All development in excess layer; basic limit layer at ultimate)
I2: $R = B_r / C_r$
I3: $C_r = C_a * D$

Then under maximum conditions:

$A_{\max} = B_a / C_a$
$A_{\max} = B_a / (C_r/D)$ « per I3 »
$A_{\max} = D * B_a / C_r$
$A_{\max} = D * B_r / C_r$ « per I1 »
$A_{\max} = D * R$ « per I2 »

REFERENCES

C Cook, "Trend and Loss Development Factors," *PCAS* **1970**, Vol. LVII, 1-14.
Gogol, Daniel F, and Emanuel Pinto, "An Analysis of Excess Loss Development," *PCAS* **1987**, Vol. LXXIV, 227-255.
G Levine, "An Analysis of Excess Loss Development (Discussion)," *PCAS* **1987**, Vol. LXXIV, 256-271.
T Mack, "Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates," *ASTIN Bulletin* **1993**, Vol. 23:2, 213-225.

Biography of the Author

Rajesh Sahasrabuddhe is a Principal with Oliver Wyman Actuarial Consulting in Philadelphia. He is responsible for providing actuarial consulting services to a variety of clients. He graduated *summa cum laude* with a BS degree in Mathematics-Actuarial Science from the University of Connecticut. He is a Fellow of the CAS and a Member of the American Academy of Actuaries. He participates on the CAS Syllabus Committee in a leadership role.

Current contact information is available to members of the CAS through its website (www.casact.org).

Claims Development by Layer

Example 1

A. Data and Information

1 Cumulative Development Triangle (C_{ij})

Development Interval (j)

Exposure Period (i)	1	2	3	4	5	6	7	8	9	10
1	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
2	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	
3	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315		
4	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268			
5	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311				
6	396,132	1,333,217	2,180,715	2,985,752	3,691,712					
7	440,832	1,288,463	2,419,861	3,483,130						
8	359,840	1,421,128	2,864,498							
9	376,686	1,363,294								
10	344,014									

2 Limit of Data in Triangle 1,000,000

3 Selected Basic Limit 500,000

4 Policy Limit 2,000,000

B. Trend Indices

1 Exposure Period Trend Index [2% EP Trend; 5% between EP 6 and 7]

Development Interval (j)

Exposure Period (i)	1	2	3	4	5	6	7	8	9	10
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.020	1.020	1.020	1.020	1.020	1.020	1.020	1.020	1.020	1.020
3	1.040	1.040	1.040	1.040	1.040	1.040	1.040	1.040	1.040	1.040
4	1.061	1.061	1.061	1.061	1.061	1.061	1.061	1.061	1.061	1.061
5	1.082	1.082	1.082	1.082	1.082	1.082	1.082	1.082	1.082	1.082
6	1.104	1.104	1.104	1.104	1.104	1.104	1.104	1.104	1.104	1.104
7	1.159	1.159	1.159	1.159	1.159	1.159	1.159	1.159	1.159	1.159
8	1.182	1.182	1.182	1.182	1.182	1.182	1.182	1.182	1.182	1.182
9	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206
10	1.230	1.230	1.230	1.230	1.230	1.230	1.230	1.230	1.230	1.230

2 Calendar Period Trend Index [1% Calendar Period Trend; -5% between CP 2 and 3]

Development Interval (j)

Exposure Period (i)	1	2	3	4	5	6	7	8	9	10
1	1.000	1.010	0.960	0.969	0.979	0.989	0.998	1.008	1.019	1.029
2	1.010	0.960	0.969	0.979	0.989	0.998	1.008	1.019	1.029	1.039
3	0.960	0.969	0.979	0.989	0.998	1.008	1.019	1.029	1.039	1.049
4	0.969	0.979	0.989	0.998	1.008	1.019	1.029	1.039	1.049	1.060
5	0.979	0.989	0.998	1.008	1.019	1.029	1.039	1.049	1.060	1.070
6	0.989	0.998	1.008	1.019	1.029	1.039	1.049	1.060	1.070	1.081
7	0.998	1.008	1.019	1.029	1.039	1.049	1.060	1.070	1.081	1.092
8	1.008	1.019	1.029	1.039	1.049	1.060	1.070	1.081	1.092	1.103
9	1.019	1.029	1.039	1.049	1.060	1.070	1.081	1.092	1.103	1.114
10	1.029	1.039	1.049	1.060	1.070	1.081	1.092	1.103	1.114	1.125

3 Combined Trend Index [B1 * B2]

Development Interval (j)

Exposure Period (i)	1	2	3	4	5	6	7	8	9	10
1	1.000	1.010	0.960	0.969	0.979	0.989	0.998	1.008	1.019	1.029
2	1.030	0.979	0.988	0.998	1.008	1.018	1.029	1.039	1.049	1.060
3	0.998	1.008	1.018	1.029	1.039	1.049	1.060	1.070	1.081	1.092
4	1.028	1.039	1.049	1.060	1.070	1.081	1.092	1.103	1.114	1.125
5	1.059	1.070	1.081	1.092	1.102	1.114	1.125	1.136	1.147	1.159
6	1.091	1.102	1.113	1.125	1.136	1.147	1.159	1.170	1.182	1.194
7	1.157	1.169	1.181	1.193	1.204	1.217	1.229	1.241	1.253	1.266
8	1.192	1.204	1.216	1.229	1.241	1.253	1.266	1.278	1.291	1.304
9	1.228	1.241	1.253	1.266	1.278	1.291	1.304	1.317	1.330	1.344
10	1.266	1.278	1.291	1.304	1.317	1.330	1.343	1.357	1.370	1.384

Claims Development by Layer

Example 1

C. Claim Size Model (Apply to Cumulative Claims)

1 Claims Size Model Parameters at Exposure Year 10 Cost Level [via claim size modeling]

	Development Interval (*j*)									
Exponential (θ)	1	2	3	4	5	6	7	8	9	10
i =10	28,138	84,242	133,998	182,460	204,649	228,245	252,830	265,063	275,707	280,000

2 Claims Size Model Parameters [C1 * $B3_{i,j}$ / $B3_{10,j}$]

Exposure Period (*i*)	Development Interval (*j*)									
Exponential (θ)	1	2	3	4	5	6	7	8	9	10
1	22,233	66,564	99,590	135,608	152,099	169,636	187,908	197,000	204,911	208,101
2	22,905	64,501	102,598	139,703	156,693	174,759	193,583	202,949	211,099	214,386
3	22,195	66,449	105,696	143,922	161,425	180,037	199,429	209,078	217,475	220,861
4	22,865	68,455	108,888	148,268	166,300	185,474	205,452	215,392	224,042	227,531
5	23,555	70,523	112,177	152,746	171,322	191,075	211,657	221,897	230,808	234,402
6	24,267	72,653	115,564	157,359	176,496	196,846	218,049	228,598	237,779	241,481
7	25,735	77,048	122,556	166,879	187,174	208,755	231,241	242,429	252,164	256,090
8	26,512	79,375	126,257	171,919	192,827	215,059	238,224	249,750	259,780	263,824
9	27,313	81,772	130,070	177,111	198,650	221,554	245,418	257,292	267,625	271,792
10	28,138	84,242	133,998	182,460	204,649	228,245	252,830	265,063	275,707	280,000

3 Unlimited Means

Exposure Period (*i*)	Development Interval (*j*)									
	1	2	3	4	5	6	7	8	9	10
1	22,233	66,564	99,590	135,608	152,099	169,636	187,908	197,000	204,911	208,101
2	22,905	64,501	102,598	139,703	156,693	174,759	193,583	202,949	211,099	214,386
3	22,195	66,449	105,696	143,922	161,425	180,037	199,429	209,078	217,475	220,861
4	22,865	68,455	108,888	148,268	166,300	185,474	205,452	215,392	224,042	227,531
5	23,555	70,523	112,177	152,746	171,322	191,075	211,657	221,897	230,808	234,402
6	24,267	72,653	115,564	157,359	176,496	196,846	218,049	228,598	237,779	241,481
7	25,735	77,048	122,556	166,879	187,174	208,755	231,241	242,429	252,164	256,090
8	26,512	79,375	126,257	171,919	192,827	215,059	238,224	249,750	259,780	263,824
9	27,313	81,772	130,070	177,111	198,650	221,554	245,418	257,292	267,625	271,792
10	28,138	84,242	133,998	182,460	204,649	228,245	252,830	265,063	275,707	280,000

4 Limited Expected Values at Policy Limits

Exposure Period (*i*)	Development Interval (*j*)									
	1	2	3	4	5	6	7	8	9	10
1	22,233	66,564	99,590	135,607	152,099	169,635	187,904	196,992	204,899	208,087
2	22,905	64,501	102,598	139,703	156,692	174,757	193,577	202,938	211,083	214,367
3	22,195	66,449	105,696	143,922	161,424	180,034	199,420	209,063	217,453	220,835
4	22,865	68,455	108,888	148,268	166,299	185,470	205,440	215,372	224,013	227,496
5	23,555	70,523	112,177	152,746	171,321	191,070	211,640	221,870	230,769	234,356
6	24,267	72,653	115,564	157,358	176,494	196,838	218,026	228,562	237,726	241,420
7	25,735	77,048	122,556	166,878	187,170	208,740	231,200	242,365	252,074	255,987
8	26,512	79,375	126,257	171,917	192,821	215,039	238,170	249,667	259,662	263,690
9	27,313	81,772	130,070	177,109	198,642	221,527	245,348	257,184	267,473	271,619
10	28,138	84,242	133,998	182,456	204,638	228,209	252,737	264,922	275,512	279,779

5 Limited Expected Values at Limits of Data Triangle

Exposure Period (*i*)	Development Interval (*j*)									
	1	2	3	4	5	6	7	8	9	10
1	22,233	66,564	99,586	135,522	151,887	169,169	186,990	195,770	203,355	206,398
2	22,905	64,501	102,592	139,594	156,428	174,187	192,478	201,479	209,249	212,366
3	22,195	66,449	105,688	143,784	161,096	179,340	198,105	207,328	215,285	218,474
4	22,865	68,455	108,877	148,094	165,893	184,629	203,871	213,318	221,461	224,723
5	23,555	70,523	112,161	152,527	170,822	190,056	209,778	219,448	227,777	231,112
6	24,267	72,652	115,544	157,085	175,885	195,621	215,826	225,719	234,233	237,640
7	25,735	77,048	122,521	166,462	186,279	207,020	228,179	238,510	247,385	250,932
8	26,512	79,375	126,211	171,407	191,748	213,003	234,644	245,194	254,248	257,866
9	27,313	81,772	130,010	176,486	197,356	219,126	241,247	252,014	261,246	264,932
10	28,138	84,241	133,921	181,699	203,105	225,390	247,987	258,969	268,375	272,128

6 Limited Expected Values at Basic Limit

Exposure Period (*i*)	Development Interval (*j*)									
	1	2	3	4	5	6	7	8	9	10
1	22,233	66,528	98,933	132,211	146,418	160,735	174,776	181,433	187,052	189,273
2	22,905	64,473	101,813	135,805	150,248	164,762	178,957	185,674	191,337	193,574
3	22,195	66,413	104,764	139,462	154,134	168,836	183,176	189,948	195,652	197,903
4	22,865	68,409	107,785	143,181	158,075	172,956	187,431	194,253	199,993	202,256
5	23,555	70,464	110,876	146,960	162,068	177,119	191,718	198,586	204,358	206,632
6	24,267	72,578	114,037	150,798	166,111	181,322	196,035	202,944	208,743	211,026
7	25,735	76,931	120,483	158,539	174,229	189,725	204,633	211,606	217,447	219,744
8	26,512	79,229	123,851	162,538	178,404	194,029	209,019	216,018	221,873	224,175
9	27,313	81,591	127,286	166,587	182,619	198,360	213,422	220,441	226,307	228,611
10	28,138	84,019	130,788	170,682	186,869	202,716	217,839	224,873	230,745	233,050

Claims Development by Layer

Example 1

D. Calculation of Development Factors at Basic Limit

1 Cumulative Triangle Exposure Year 10 Cost Levels and Basic Limit ($C_{i,j}$) [$A1_{i,j}$ * $C6_{10,j}$ / $C5_{i,j}$]

Exposure Period (i)	Development Interval (j) 1	2	3	4	5	6	7	8	9	10
1	452,881	1,419,731	2,279,040	2,793,772	3,377,947	3,978,376	4,038,185	4,142,394	4,349,865	4,405,265
2	432,566	1,610,197	2,766,439	4,100,116	4,538,378	4,794,873	5,260,266	5,484,617	5,887,561	
3	368,296	1,634,013	2,745,402	3,840,401	4,623,708	4,671,636	5,090,015	5,324,759		
4	382,236	1,741,436	2,636,785	4,330,559	4,539,482	4,811,270	4,902,613			
5	529,368	1,353,815	2,481,777	3,242,747	3,722,312	4,131,335				
6	459,320	1,541,795	2,468,413	3,244,186	3,922,258					
7	481,990	1,405,037	2,583,135	3,571,425						
8	381,903	1,504,277	2,968,365							
9	388,062	1,400,758								
10	344,014									

2 Exposure Year 10 Incremental Basic Limit Development Factors [per D1; Volume Weighted Averages]

	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8	8 to 9	9 to 10
i=10	3.511	1.714	1.399	1.147	1.076	1.057	1.039	1.063	1.013

3 Exposure Year 10 Cumulative Development Factors [per D2]

	1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
i=10	12.291	3.501	2.042	1.460	1.273	1.183	1.119	1.077	1.013	1.000

E. Calculation of Development Factors by Layer

1 Basic Limit [D_{3j} * ($C6_{i,10}$/$C6_{10,10}$) / ($C6_{i,j}$/$C6_{10,j}$)]

Exposure Period (i)	Development Interval (j) 1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
1	12.633	3.590	2.193	1.531	1.319	1.211	1.133	1.084	1.015	1.000
2	12.541	3.789	2.179	1.524	1.315	1.209	1.132	1.083	1.014	
3	13.232	3.761	2.165	1.517	1.310	1.206	1.130	1.083		
4	13.126	3.731	2.151	1.510	1.306	1.203	1.129			
5	13.017	3.701	2.136	1.503	1.301	1.200				
6	12.904	3.669	2.121	1.496	1.296					
7	12.671	3.605	2.090	1.482						
8	12.547	3.571	2.074							
9	12.421	3.536								
10	12.291									

2 Basic Limit to Policy Limit [D_{3j} * (($C4_{i,10}$-$C6_{i,10}$) / $C6_{10,10}$) / (($C4_{i,j}$-$C6_{i,j}$/$C6_{10,j}$)]

Exposure Period (i)	Development Interval (j) 1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
1	very large	652.420	32.802	5.924	3.380	2.175	1.499	1.257	1.057	1.000
2	very large	946.242	30.374	5.704	3.293	2.140	1.488	1.252	1.056	
3	very large	807.075	28.187	5.498	3.210	2.107	1.477	1.247		
4	very large	691.561	26.215	5.305	3.132	2.075	1.466			
5	very large	595.239	24.431	5.124	3.058	2.044				
6	very large	514.560	22.814	4.954	2.987					
7	very large	390.710	20.042	4.647						
8	very large	341.869	18.820							
9	very large	300.278								
10	very large									

3 Policy Limit to Unlimited [D_{3j} * (($C3_{i,10}$-$C4_{i,10}$) / $C6_{10,10}$) / (($C3_{i,j}$-$C4_{i,j}$) / $C6_{10,j}$)]

Exposure Period (i)	Development Interval (j) 1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
1	very large	very large	84,538.278	279.503	48.056	11.155	3.254	1.887	1.183	1.000
2	very large	very large	62,192.336	240.423	43.321	10.464	3.157	1.857	1.178	
3	very large	very large	46,166.664	207.723	39.172	9.835	3.066	1.829		
4	very large	very large	34,571.138	180.241	35.524	9.261	2.979			
5	very large	very large	26,108.458	157.047	32.309	8.735				
6	very large	very large	19,880.311	137.391	29.466					
7	very large	very large	11,880.695	106.724						
8	very large	very large	9,257.797							
9	very large	very large								
10	very large									

Claims Development by Layer

Example 1

4 Limit of Data in Triangle [D_{3j} * ($C5_{i,10}$ / $C6_{10,10}$) / ($C5_{i,j}$/$C6_{10,j}$)]

Development Interval (j)

Exposure Period (i)	1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
1	13.776	3.913	2.375	1.629	1.387	1.255	1.155	1.096	1.018	1.000
2	13.759	4.155	2.372	1.627	1.385	1.254	1.154	1.095	1.018	
3	14.607	4.149	2.369	1.625	1.384	1.253	1.154	1.095		
4	14.584	4.143	2.366	1.623	1.382	1.252	1.153			
5	14.559	4.136	2.362	1.620	1.381	1.251				
6	14.532	4.128	2.357	1.618	1.379					
7	14.469	4.110	2.347	1.612						
8	14.433	4.100	2.342							
9	14.394	4.089								
10	14.352									

5 Unadjusted Incremental Development Factors at Limits of Data Triangle[per A1; Volume Weighted Averages]

1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8	8 to 9	9 to 10
3.490	1.747	1.457	1.174	1.104	1.086	1.054	1.077	1.018

6 Unadjusted Cumulative Development Factors [per E5]

1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
14.445	4.139	2.369	1.625	1.384	1.254	1.155	1.096	1.018	1.000

Differences [E6 / E4, last diagonal -1]

7	+0.7%	+1.2%	+1.2%	+0.8%	+0.4%	+0.3%	+0.1%	+0.1%	+0.0%	+0.0%

Claims Development by Layer

Example 2

A. Data and Information

1 Cumulative Development Triangle (C_{ij})

Development Interval (j), by Exposure Period (i)

Exposure Period (i)	1	2	3	4	5	6	7	8	9	10
1	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
2	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	
3	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315		
4	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268			
5	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311				
6	396,132	1,333,217	2,180,715	2,985,752	3,691,712					
7	440,832	1,288,463	2,419,861	3,483,130						
8	359,840	1,421,128	2,864,498							
9	376,686	1,363,294								
10	344,014									

2 Limit of Data in Triangle 1,000,000

3 Selected Basic Limit 500,000

4 Policy Limit 2,000,000

B. Trend Indices

1 Exposure Period Trend Index [2% EP Trend; 5% between EP 6 and 7]

Development Interval (j), by Exposure Period (i)

Exposure Period (i)	1	2	3	4	5	6	7	8	9	10
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.020	1.020	1.020	1.020	1.020	1.020	1.020	1.020	1.020	1.020
3	1.040	1.040	1.040	1.040	1.040	1.040	1.040	1.040	1.040	1.040
4	1.061	1.061	1.061	1.061	1.061	1.061	1.061	1.061	1.061	1.061
5	1.082	1.082	1.082	1.082	1.082	1.082	1.082	1.082	1.082	1.082
6	1.104	1.104	1.104	1.104	1.104	1.104	1.104	1.104	1.104	1.104
7	1.159	1.159	1.159	1.159	1.159	1.159	1.159	1.159	1.159	1.159
8	1.182	1.182	1.182	1.182	1.182	1.182	1.182	1.182	1.182	1.182
9	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206	1.206
10	1.230	1.230	1.230	1.230	1.230	1.230	1.230	1.230	1.230	1.230

2 Calendar Period Trend Index [1% Calendar Period Trend; -5% between CP 2 and 3]

Development Interval (j), by Exposure Period (i)

Exposure Period (i)	1	2	3	4	5	6	7	8	9	10
1	1.000	1.010	0.960	0.969	0.979	0.989	0.998	1.008	1.019	1.029
2	1.010	0.960	0.969	0.979	0.989	0.998	1.008	1.019	1.029	1.039
3	0.960	0.969	0.979	0.989	0.998	1.008	1.019	1.029	1.039	1.049
4	0.969	0.979	0.989	0.998	1.008	1.019	1.029	1.039	1.049	1.060
5	0.979	0.989	0.998	1.008	1.019	1.029	1.039	1.049	1.060	1.070
6	0.989	0.998	1.008	1.019	1.029	1.039	1.049	1.060	1.070	1.081
7	0.998	1.008	1.019	1.029	1.039	1.049	1.060	1.070	1.081	1.092
8	1.008	1.019	1.029	1.039	1.049	1.060	1.070	1.081	1.092	1.103
9	1.019	1.029	1.039	1.049	1.060	1.070	1.081	1.092	1.103	1.114
10	1.029	1.039	1.049	1.060	1.070	1.081	1.092	1.103	1.114	1.125

3 Combined Trend Index [B1 * B2]

Development Interval (j), by Exposure Period (i)

Exposure Period (i)	1	2	3	4	5	6	7	8	9	10
1	1.000	1.010	0.960	0.969	0.979	0.989	0.998	1.008	1.019	1.029
2	1.030	0.979	0.988	0.998	1.008	1.018	1.029	1.039	1.049	1.060
3	0.998	1.008	1.018	1.029	1.039	1.049	1.060	1.070	1.081	1.092
4	1.028	1.039	1.049	1.060	1.070	1.081	1.092	1.103	1.114	1.125
5	1.059	1.070	1.081	1.092	1.102	1.114	1.125	1.136	1.147	1.159
6	1.091	1.102	1.113	1.125	1.136	1.147	1.159	1.170	1.182	1.194
7	1.157	1.169	1.181	1.193	1.204	1.217	1.229	1.241	1.253	1.266
8	1.192	1.204	1.216	1.229	1.241	1.253	1.266	1.278	1.291	1.304
9	1.228	1.241	1.253	1.266	1.278	1.291	1.304	1.317	1.330	1.344
10	1.266	1.278	1.291	1.304	1.317	1.330	1.343	1.357	1.370	1.384

Claims Development by Layer

Example 2

C. Claim Size Model (Apply to Cumulative Claims)

1 Claims Size Model Parameters at Exposure Year 10 Cost Level [via claim size modeling]

	Development Interval (*j*)									
Exponential (θ)	1	2	3	4	5	6	7	8	9	10
i =10	60,295	168,483	267,996	348,332	409,299	456,490	505,660	530,125	551,415	565,000

2 Claims Size Model Parameters [C1 * $B3_{i,j}$ / $B3_{10,j}$]

Exposure Period (*i*)	Development Interval (*j*)									
Exponential (θ)	1	2	3	4	5	6	7	8	9	10
1	47,642	133,128	199,180	258,887	304,199	339,272	375,816	393,999	409,822	419,919
2	49,081	129,001	205,195	266,705	313,386	349,518	387,166	405,898	422,199	432,600
3	47,560	132,897	211,392	274,760	322,850	360,073	398,858	418,156	434,949	445,665
4	48,996	136,911	217,776	283,058	332,600	370,948	410,904	430,784	448,085	459,124
5	50,476	141,046	224,353	291,606	342,644	382,150	423,313	443,794	461,617	472,990
6	52,000	145,305	231,129	300,413	352,992	393,691	436,097	457,197	475,558	487,274
7	55,146	154,096	245,112	318,588	374,348	417,509	462,481	484,857	504,329	516,754
8	56,812	158,750	252,514	328,209	385,654	430,118	476,448	499,500	519,560	532,360
9	58,527	163,544	260,140	338,121	397,300	443,108	490,837	514,585	535,250	548,437
10	60,295	168,483	267,996	348,332	409,299	456,490	505,660	530,125	551,415	565,000

3 Unlimited Means

Exposure Period (*i*)	Development Interval (*j*)									
	1	2	3	4	5	6	7	8	9	10
1	47,642	133,128	199,180	258,887	304,199	339,272	375,816	393,999	409,822	419,919
2	49,081	129,001	205,195	266,705	313,386	349,518	387,166	405,898	422,199	432,600
3	47,560	132,897	211,392	274,760	322,850	360,073	398,858	418,156	434,949	445,665
4	48,996	136,911	217,776	283,058	332,600	370,948	410,904	430,784	448,085	459,124
5	50,476	141,046	224,353	291,606	342,644	382,150	423,313	443,794	461,617	472,990
6	52,000	145,305	231,129	300,413	352,992	393,691	436,097	457,197	475,558	487,274
7	55,146	154,096	245,112	318,588	374,348	417,509	462,481	484,857	504,329	516,754
8	56,812	158,750	252,514	328,209	385,654	430,118	476,448	499,500	519,560	532,360
9	58,527	163,544	260,140	338,121	397,300	443,108	490,837	514,585	535,250	548,437
10	60,295	168,483	267,996	348,332	409,299	456,490	505,660	530,125	551,415	565,000

4 Limited Expected Values at Policy Limits

Exposure Period (*i*)	Development Interval (*j*)									
	1	2	3	4	5	6	7	8	9	10
1	47,642	133,128	199,171	258,773	303,774	338,338	373,981	391,539	406,709	416,332
2	49,081	129,001	205,183	266,558	312,855	348,374	384,956	402,957	418,499	428,352
3	47,560	132,897	211,376	274,570	322,191	358,680	396,209	414,655	430,569	440,653
4	48,996	136,911	217,754	282,816	331,786	369,258	407,742	426,635	442,921	453,234
5	50,476	141,045	224,323	291,300	341,645	380,112	419,557	438,896	455,554	466,096
6	52,000	145,305	231,088	300,027	351,770	391,243	431,653	451,439	468,466	479,234
7	55,146	154,096	245,042	317,989	372,558	414,040	456,358	477,020	494,769	505,979
8	56,812	158,749	252,423	327,468	383,496	426,005	469,287	490,388	508,497	519,926
9	58,527	163,543	260,021	337,208	394,713	438,252	482,494	504,028	522,492	534,136
10	60,295	168,482	267,843	347,214	406,209	450,779	495,975	517,938	536,750	548,605

5 Limited Expected Values at Limits of Data Triangle

Exposure Period (*i*)	Development Interval (*j*)									
	1	2	3	4	5	6	7	8	9	10
1	47,642	133,056	197,865	253,447	292,836	321,470	349,552	362,866	374,105	381,110
2	49,081	128,946	203,626	260,430	300,495	329,523	357,913	371,347	382,674	389,729
3	47,560	132,826	209,527	267,544	308,268	337,673	366,352	379,896	391,303	398,402
4	48,996	136,819	215,569	274,786	316,150	345,913	374,861	388,507	399,986	407,123
5	50,476	140,928	221,752	282,155	324,136	354,239	383,436	397,173	408,715	415,886
6	52,000	145,156	228,075	289,646	332,222	362,644	392,070	405,887	417,485	424,685
7	55,146	153,862	240,967	304,782	348,458	379,450	409,265	423,212	434,894	442,134
8	56,812	158,458	247,701	312,616	356,809	388,057	418,038	432,035	443,746	450,999
9	58,527	163,183	254,572	320,556	365,237	396,720	426,844	440,882	452,614	459,875
10	60,295	168,038	261,575	328,598	373,738	405,433	435,677	449,745	461,490	468,753

6 Limited Expected Values at Basic Limit

Exposure Period (*i*)	Development Interval (*j*)									
	1	2	3	4	5	6	7	8	9	10
1	47,641	130,016	182,998	221,361	245,406	261,556	276,465	283,245	288,835	292,261
2	49,079	126,327	187,251	225,794	249,827	265,920	280,743	287,475	293,020	296,416
3	47,558	129,810	191,537	230,232	254,237	270,263	284,992	291,670	297,168	300,532
4	48,994	133,360	195,853	234,671	258,632	274,581	289,207	295,830	301,277	304,609
5	50,473	136,974	200,196	239,109	263,009	278,872	293,389	299,953	305,347	308,645
6	51,997	140,651	204,561	243,542	267,367	283,134	297,533	304,035	309,374	312,637
7	55,140	148,090	213,237	252,269	275,900	291,453	305,601	311,973	317,197	320,386
8	56,803	151,944	217,653	256,671	280,182	295,615	309,626	315,928	321,092	324,242
9	58,516	155,855	222,080	261,056	284,435	299,739	313,608	319,837	324,938	328,049
10	60,280	159,819	226,514	265,423	288,655	303,824	317,544	323,700	328,736	331,806

Claims Development by Layer

Example 2

D. Calculation of Development Factors at Basic Limit

1 Cumulative Triangle Exposure Year 10 Cost Levels and Basic Limit ($C_{i,j}$) [$A1_{i,j}$ * $C6_{10,j}$ / $C5_{i,j}$]

Exposure Period (i)	Development Interval (j) 1	2	3	4	5	6	7	8	9	10
1	452,767	1,351,032	1,986,583	2,323,084	2,706,395	3,137,755	3,148,934	3,217,037	3,368,618	3,396,735
2	432,457	1,532,103	2,413,944	3,417,616	3,649,372	3,798,743	4,123,637	4,283,518	4,586,542	
3	368,203	1,554,934	2,398,380	3,209,537	3,732,390	3,718,630	4,012,222	4,183,099		
4	382,140	1,657,379	2,306,488	3,629,412	3,679,452	3,848,804	3,886,711			
5	529,235	1,288,675	2,174,035	2,725,977	3,030,200	3,322,067				
6	459,205	1,467,893	2,165,788	2,736,050	3,207,585					
7	481,869	1,338,349	2,274,720	3,033,321						
8	381,807	1,433,334	2,619,477							
9	387,965	1,335,194								
10	343,928									

2 Exposure Year 10 Incremental Basic Limit Development Factors [per D1; Volume Weighted Averages]

	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8	8 to 9	9 to 10
i=10	3.344	1.578	1.341	1.109	1.061	1.046	1.035	1.061	1.008

3 Exposure Year 10 Cumulative Development Factors [per D2]

	1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
i=10	9.639	2.883	1.827	1.363	1.229	1.158	1.107	1.069	1.008	1.000

E. Calculation of Development Factors by Layer

1 Basic Limit [D_{3j} * ($C6_{i,10}$/$C6_{10,10}$) / ($C6_{i,j}$/$C6_{10,j}$)]

Exposure Period (i)	Development Interval (j) 1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
1	10.743	3.121	1.992	1.439	1.273	1.185	1.120	1.077	1.011	1.000
2	10.576	3.258	1.975	1.431	1.269	1.182	1.119	1.076	1.011	
3	11.066	3.215	1.957	1.423	1.264	1.179	1.117	1.075		
4	10.888	3.172	1.940	1.415	1.259	1.177	1.116			
5	10.709	3.129	1.923	1.407	1.255	1.174				
6	10.529	3.086	1.906	1.400	1.250					
7	10.175	3.004	1.874	1.385						
8	9.996	2.963	1.858							
9	9.817	2.923								
10	9.639									

2 Basic Limit to Policy Limit [D_{3j} * (($C4_{i,10}$-$C6_{i,10}$) / $C6_{10,10}$) / (($C4_{i,j}$-$C6_{i,j}$/$C6_{10,j}$)]

Exposure Period (i)	Development Interval (j) 1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
1	very large	55.346	9.569	3.616	2.273	1.714	1.348	1.195	1.052	1.000
2	very large	68.491	9.178	3.529	2.238	1.697	1.342	1.192	1.050	
3	very large	63.024	8.810	3.445	2.205	1.681	1.335	1.189		
4	very large	58.114	8.465	3.366	2.172	1.665	1.329			
5	very large	53.693	8.140	3.289	2.141	1.649				
6	very large	49.704	7.834	3.216	2.111					
7	very large	42.907	7.279	3.079						
8	very large	39.930	7.020							
9	very large	37.221								
10	very large									

3 Policy Limit to Unlimited [D_{3j} * (($C3_{i,10}$-$C4_{i,10}$) / $C6_{10,10}$) / (($C3_{i,j}$-$C4_{i,j}$) / $C6_{10,j}$)]

Exposure Period (i)	Development Interval (j) 1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
1	very large	very large	515.543	34.213	9.036	4.072	2.071	1.521	1.151	1.000
2	very large	very large	441.637	31.367	8.568	3.939	2.037	1.507	1.147	
3	very large	very large	380.045	28.831	8.138	3.814	2.005	1.494		
4	very large	very large	328.487	26.566	7.741	3.697	1.974			
5	very large	very large	285.139	24.538	7.373	3.586				
6	very large	very large	248.540	22.717	7.034					
7	very large	very large	191.734	19.638						
8	very large	very large	169.080							
9	very large	very large								
10	very large									

Claims Development by Layer

Example 2

4 Limit of Data in Triangle [D_{3j} * ($C5_{i,10}$ / $C6_{10,10}$) / ($C5_{i,j}$/$C6_{10,j}$)]

Development Interval (*j*)

Exposure Period (*i*)	1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
1	14.008	3.977	2.403	1.639	1.392	1.257	1.155	1.096	1.018	1.000
2	13.905	4.197	2.387	1.632	1.387	1.254	1.154	1.095	1.017	
3	14.669	4.165	2.372	1.623	1.382	1.251	1.152	1.094		
4	14.551	4.132	2.356	1.615	1.377	1.248	1.151			
5	14.429	4.098	2.339	1.607	1.372	1.245				
6	14.302	4.063	2.323	1.599	1.367					
7	14.040	3.990	2.289	1.582						
8	13.902	3.952	2.271							
9	13.760	3.913								
10	13.614									

5 Unadjusted Incremental Development Factors at Limits of Data Triangle[per A1; Volume Weighted Averages]

1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8	8 to 9	9 to 10
3.490	1.747	1.457	1.174	1.104	1.086	1.054	1.077	1.018

6 Unadjusted Cumulative Development Factors [per E5]

1 to ult	2 to ult	3 to ult	4 to ult	5 to ult	6 to ult	7 to ult	8 to ult	9 to ult	10 to ult
14.445	4.139	2.369	1.625	1.384	1.254	1.155	1.096	1.018	1.000

Differences [E6 / E4, last diagonal -1]

7	+6.1%	+5.8%	+4.3%	+2.8%	+1.3%	+0.7%	+0.3%	+0.1%	+0.0%	+0.0%

Paper #5-2 Example 3
1/1/2013 11:24 AM

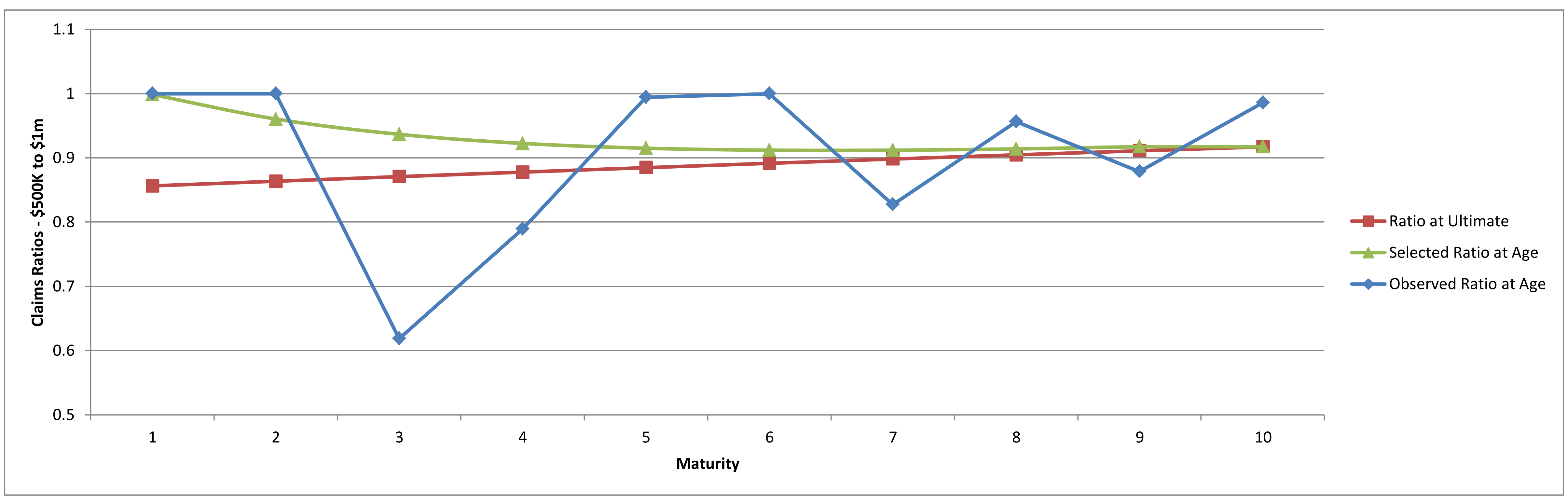
Example 1
Page 1

Claims Development by Layer

Example 3

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
								$500K to $1m				
Exposure Period (*i*)	Maturity	Claims, Limited to $1m, as of End of EP 10	Claims, Limited to Basic Limit ($500K), as of End of EP 10	Observed Ratio	Exponential Claim Size Model Parameter (θ)	Limited Expected Value at Basic Limit at Ultimate	Limited Expected Value at $1m Limit at Ultimate	Ratio at Ultimate	Selected Ratio at Age	Ultimate Claims Development Factor at $1m	Ultimate Claims Development Factor at $500K	Ultimate Claims Development Factor $500K to $1m
1	10	3,901,463	3,846,592	0.986	208,000	189,203	206,301	0.917	0.917	1.000	1.000	1.000
2	9	5,339,085	4,692,053	0.879	214,985	193,978	212,932	0.911	0.917	1.018	1.011	1.094
3	8	4,909,315	4,695,780	0.957	222,204	198,788	219,736	0.905	0.914	1.096	1.085	1.213
4	7	4,588,268	3,795,644	0.827	229,665	203,628	226,713	0.898	0.912	1.155	1.137	1.335
5	6	3,873,311	3,873,311	1.000	237,377	208,493	233,862	0.892	0.912	1.254	1.226	1.546
6	5	3,691,712	3,670,631	0.994	245,348	213,379	241,183	0.885	0.915	1.384	1.339	1.878
7	4	3,483,130	2,750,008	0.790	253,587	218,283	248,672	0.878	0.923	1.625	1.546	2.563
8	3	2,864,498	1,771,896	0.619	262,102	223,199	256,328	0.871	0.937	2.369	2.202	4.833
9	2	1,363,294	1,363,294	1.000	270,903	228,123	264,147	0.864	0.961	4.139	3.721	14.296
10	1	344,014	344,014	1.000	280,000	233,050	272,128	0.856	0.999	14.445	12.389	1,444.501

(5) = (4) / (3)
(6) Via claim size model
(7) LEV [exponential(θ);x]= q * (1 - exp (x/θ))
(8) LEV [exponential(θ);x]= q * (1 - exp (x/θ))
(9) = (7) / (8)
(10) See Section 4.2
(11) Provided
(12) = (11) * (9) / (10)
(13) = (11) * (1- (9)) / (1- (10))



CAS MONOGRAPH SERIES
NUMBER 4

USING THE ODP BOOTSTRAP MODEL: A PRACTITIONER'S GUIDE

Mark R. Shapland

CASUALTY ACTUARIAL SOCIETY

There are many papers that describe the over-dispersed Poisson (ODP) bootstrap model, but these papers are either limited to the basic calculations of the model or focus on the theoretical aspects of the model and always implicitly assume that the ODP bootstrap model is perfectly suited to the data being analyzed. In order to use the ODP bootstrap model on real data, the analyst must first test and review the assumptions of the model and may need to consider various modifications to the basic algorithm in order to put the ODP bootstrap model to practical use. This monograph starts by gathering the evolutionary changes from different papers into a complete ODP bootstrap modeling framework using a standard notation. Then it generalizes the basic model into a more flexible framework. Next it describes the adjustments or enhancements required for practical use and addresses the diagnostic testing of the model assumptions. While this monograph is focused on the ODP bootstrap model, we must recognize that it is a special subset of a larger framework of models and that there are a wide variety of other stochastic models that should also be considered. However, since no single model is perfect we also explore ways to combine or credibility weight the ODP bootstrap model results with various other models in order to arrive at a "best estimate" of the distribution, similar to how a deterministic best estimate is generally derived in practice. Finally, the monograph will also extend the model to illustrate the GLM Bootstrap and the model output to address other risk management issues and suggest areas for future research.

Keywords. Bootstrap, Over-Dispersed Poisson, Reserve Variability, Reserve Range, Distribution of Possible Outcomes, Generalized Linear Model, Best Estimate.

Availability of Excel workbooks. In lieu of technical appendices, several companion Excel workbooks are included that illustrate the calculations described in this monograph. The companion materials are summarized in the Supplementary Materials section and are available at https://www.casact.org/sites/default/files/2021-02/practitionerssuppl-shaplandmonograph04.zip. Other sources of ODP bootstrap modeling software that could be used for educational purposes would include working parties and other industry groups in North America and Europe, including but not limited to models freely available in the R statistical software package.

USING THE ODP BOOTSTRAP MODEL: A PRACTITIONER'S GUIDE

Mark R. Shapland

Casualty Actuarial Society
4350 North Fairfax Drive, Suite 250
Arlington, Virginia 22203
www.casact.org
(703) 276-3100

Using the ODP Bootstrap Model: A Practitioner's Guide
By Mark R. Shapland

Copyright 2016 by the Casualty Actuarial Society

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. For information on obtaining permission for use of material in this work, please submit a written request to the Casualty Actuarial Society.

Library of Congress Cataloguing-in-Publication Data
Shapland, Mark R.
Using the ODP Bootstrap Model: A Practitioner's Guide
ISBN 978-0-9968897-4-2 (print edition)
ISBN 978-0-9968897-5-9 (electronic edition)
1. Actuarial science. 2. Loss reserving. 3. Insurance—Mathematical models.
I. Shapland, Mark

Contents

Foreword

The concept of bootstrapping generally invokes the idea that once a process has been started, it can replicate without additional external input. Disciplines from biology and physics to business and statistics use bootstrapping to analyze numerous processes. For example, in statistics, bootstrapping involves starting with one sample and using it to derive many more subsamples drawn from the original sample. A specialized application within actuarial science involves derivation of a distribution of possible outcomes for each step in the loss development process.

Considerable literature has been developed over the past twenty-plus years regarding bootstrapping as it relates to actuarial science and the loss reserving process. In this work, Mr. Shapland collects the research from this vast literature base and frames it in one comprehensive presentation. The result is a complete over-dispersed Poisson (ODP) bootstrap model. At the same time, those who have worked with ODP bootstrapping know that these models have limitations when using real-world data. Mr. Shapland's work also proposes modifications and enhancements that allow more practical application of the ODP bootstrap model. In addition, he provides details on generalized linear models, of which the ODP bootstrap is one form.

With the knowledge that model risk is a real risk—no single model is perfect—Mr. Shapland further explores ways to combine the results of ODP bootstrapping with other types of models in an effort to determine a true "best estimate" of the distribution.

A set of illustrative Excel files, along with detailed instructions on how to use them, complements this monograph. With these files, the reader can follow through, step by step, the theory presented in monograph.

This monograph provides a one-stop shop for practical application of bootstrapping for the loss reserving process. The Monographs Editorial Board thanks the author for a valuable contribution to the casualty actuarial literature.

Leslie R. Marlo
Chairperson
Monograph Editorial Board

2016 CAS Monograph Editorial Board

Leslie R. Marlo, Editor in Chief
Emmanuel Bardis
Brendan P. Barrett
Craig C. Davis
Ali Ishaq
C. K. Stan Khury, consultant
Glenn G. Meyers, consultant
Katya Ellen Press, consultant

1. Introduction

The term "bootstrap" has a colorful history that dates back to German folk tales of the 18th century. It is aptly conveyed in the familiar cliché admonishing laggards to "pull oneself up by their own bootstraps." A physical paradox and virtual impossibility, the idea has nonetheless caught the imagination of scientists in a broad array of fields, including physics, biology and medical research, computer science, and statistics.

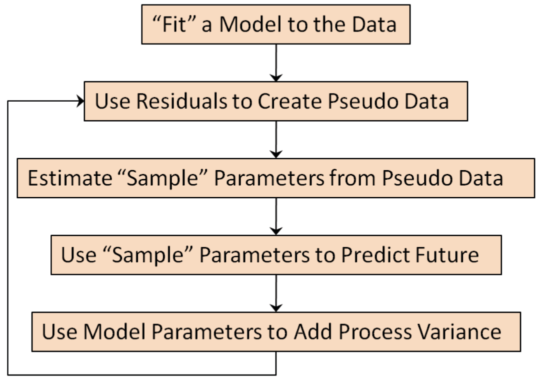
Bradley Efron (1979), Chairman of the Department of Statistics at Stanford University, is most often associated as the source of expanding bootstrapping into the realm of statistics, with his notion of taking one available sample and using it to arrive at many others through resampling.

In actuarial science, the concept of bootstrapping has become increasingly common in the process of loss reserving. The most commonly cited examples are England and Verrall (1999; 2002), Pinheiro, et al. (2003), and Kirschner, et al. (2008), who combine the bootstrap concept with a basic chain ladder model. These papers detail a form of the model where the incremental losses are modeled as over-dispersed Poisson random variables. In this monograph, it is called the over-dispersed Poisson bootstrap model, or the ODP bootstrap. The goal of the ODP bootstrap model is to generate a distribution of possible outcomes, rather than a point estimate, providing more information about the potential results.

At the present time, the vast majority of reserving actuaries in the U.S. are focused on deterministic point estimates. This is not surprising as the American Academy of Actuaries' primary standard of practice for reserving, ASOP 36, is focused on deterministic point estimates and the actuarial opinion required by regulators is also focused on deterministic estimates. However, actuaries are moving towards estimating an unpaid claim distribution, encouraged by the following factors:

- ASOP 43 defines "actuarial central estimate" in such a way that it could include either deterministic point estimates or a first moment estimate from a distribution;
- the SEC is looking for more reserving risk information in the 10-K reports filed by publicly traded companies;
- all of the major rating agencies have built or are building dynamic risk models to help with their insurance rating process and welcome the input of company actuaries regarding unpaid claim distributions;
- companies that use dynamic risk models to help their internal risk management processes need unpaid claim distributions;

Figure 1.1. Stochastic Model Diagram



- The Solvency II regime in Europe is moving many insurers towards unpaid claim distributions; and
- International Financial Accounting Standards, while still being discussed, shows actuaries that the future of insurance accounting may rely on unpaid claim distributions for booked reserves.

1.1. Objectives

One objective of this monograph is to provide more practical details on the Generalized Linear Model (GLM), of which the ODP bootstrap model[1] is a specific form. A GLM allows the user to "fit" the model to the data, as illustrated in Figure 1.1. The benefit of a GLM is that it can be specifically tailored to the statistical features found in the data under analysis. In contrast, consider algorithms that essentially force the data to be "fit" to a static method in order to predict the future as illustrated in Figure 1.2.[2]

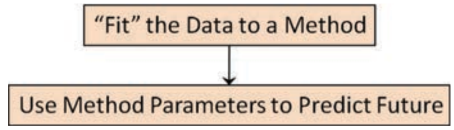
If a method does not use parameters or assumptions that fit the statistical features of the data then it may not project a reasonable point estimate. Similarly, if model assumptions and parameters do not fit the statistical features found in the data then the results of a simulation may not be a very good estimate of the distribution of possible outcomes. Thus, the modeling framework must be able to adapt to or "fit" the model to the data so this point will be elaborated on in later sections.

Another objective of this monograph is to show how the ODP bootstrap modeling framework can be used in practice, to help the wider adoption of unpaid claim distributions. Most of the papers describing stochastic models, including the ODP bootstrap model, tend to focus primarily on the theoretical aspects of the model while ignoring the data issues that commonly arise in practice. As a result the models can be quite elegantly implemented yet suffer from practical limitations such as only being useful

[1] Some authors define a model as having a defined structure and error distribution, so under this more restrictive definition bootstrapping would be considered to be a method or algorithm. However, using a less restrictive definition of a model as an algorithm that produces a distribution, bootstrapping would be defined as a model.

[2] For most deterministic reserving methods diagnostic tools can be used to test assumptions, adjust parameters and "fit" the method to the data, but not all assumptions can be adjusted and blindly applying a method is equivalent to a static method.

Figure 1.2. Static Method Diagram



for complete triangles or only for positive incremental values. Thus, while keeping as close to the theoretical foundation as possible, another objective is to illustrate how practical adjustments can be made to accommodate common data issues and allow the model to "fit" the data. As a practical matter, it is also possible that the model does not fit the data very well, or less well than other models, so the process of diagnosing the assumptions will inform the actuary's judgment when considering how much weight, if any, to give the model in relation to other models.

Another potential roadblock seems to be the notion that actuaries are still searching for the perfect model to describe "the" distribution of unpaid claims, as if imperfections in a model remove it from all consideration since it can't be "the one." This notion can also manifest itself when an actuary settles for a model that seems to work the best or is the easiest to use, or with the idea that each model must be used in its entirety or not at all. Interestingly, this notion was dispelled long ago with respect to deterministic point estimates as actuaries commonly use many different methods, which range from easy to complex, and judgmentally weight the results to arrive at their best estimate.

Model risk—the risk that the model you have chosen is not the same as the one that generates future losses—is very real and weighting or combining multiple estimates is a very practical way of addressing model risk. Thus, another objective of this monograph is to show how stochastic reserving can be similar to deterministic reserving when it comes to analyzing and using the best parts of multiple models by illustrating how the results from an ODP bootstrap model can be weighted together with other models. More importantly, the monograph hopes to illustrate the advantage of using a more complete set of risk estimation tools (which can include both stochastic models and deterministic methods) to arrive at an actuarial best estimate of the distribution of possible outcomes, rather than to focus on deterministic methods to select the "mean" and then simply "add on" a simple approximation or use only a favorite model to turn that selected mean into a distribution.

2. Notation

The papers that describe the basic ODP bootstrap model use different notation, despite sharing common steps. Rather than pick the notation in one of the papers, the notation from the CAS Working Party on Quantifying Variability in Reserve Estimates Summary Report (CAS Working Party 2005) will be used since it is intended to serve as a basis for further research.

Many models visualize loss data as a two-dimensional array, (w, d) with accident period or policy period w, and development age d (think w = "when" and d = "delay"). For this discussion, we assume that the loss information available is an "upper triangular" subset for rows $w = 1, 2, \ldots, n$ and for development ages $d = 1, 2, \ldots, n - w + 1$. The "diagonal" for which $w + d$ equals the constant, k, represents the loss information for each accident period w as of accounting period k.[3]

For purposes of including tail factors, the development beyond the observed data for periods $d = n + 1, n + 2, \ldots, u$, where u is the ultimate time period for which any claim activity occurs—i.e., u is the period in which all claims are final and paid in full, must also be considered.

The monograph uses the following notation for certain important loss statistics:

$c(w, d)$: cumulative loss from accident[4] year w as of age d.

$q(w, d)$: incremental loss for accident year w from $d - 1$ to d.

$c(w, n) = U(w)$: total loss from accident year w when claims are at ultimate values at time n,[5] or

$c(w, u) = U(w)$: total loss from accident year w when claims are at ultimate values at time u.

$R(w)$: future development after age d for accident year w, i.e., $= U(w) - c(w, d)$.

$f(d)$: factor applied to $c(w, d)$ to estimate $q(w, d + 1)$ or can be used more generally to indicate any factor relating to age d.

[3] For a more complete explanation of this two-dimensional view of the loss information, see the *Foundations of Casualty Actuarial Science* (2001), Chapter 5, particularly pages 210–226.

[4] The use of accident year is used for ease of discussion. All of the discussion and formulas that follow could also apply to underwriting year, policy year, report year, etc. Similarly, year could also be half-year, quarter or month.

[5] This would imply that claims reach their ultimate value without any tail factor. This is generalized by changing n to $n + t = u$, where t is the number of periods in the tail.

$F(d)$: factor applied to $c(w, d)$ to estimate $c(w, d+1)$ or $c(w, n)$ or can be used more generally to indicate any cumulative factor relating to age d.

$G(w)$: factor relating to accident year w—capitalized to designate ultimate loss level.

$h(k)$: factor relating to the diagonal k along which $w + d$ is constant.[6]

$e(w, d)$: a random fluctuation, or error, which occurs at the w, d cell.

$E(x)$: the expectation of the random variable x.

$Var(x)$: the variance of the random variable x.

x^*: a randomly sampled value of the variable x.

What are called factors here could also be summands, but if factors and summands are both used, some other notation for the additive terms would be needed. The notation does not distinguish paid vs. incurred, but if this is necessary, capitalized subscripts P and I could be used.

[6] Some authors define $d = 0, 1, \ldots, n-1$ which intuitively allows $k = w$ along the diagonals, but in this case the triangle size is $n \times n - 1$ which is not intuitive. With $d = 1, 2, \ldots, n$ defined as in this monograph, the triangle size $n \times n$ is intuitive, but then $k = w + 1$ along the diagonals is not as intuitive. A way to think about this which helps tie everything together is to assume the w variables are the beginning of the accident periods and the d variables are at the end of the development periods. Thus, if we are using years then cell $c(n, 1)$ represents accident year n evaluated at $12/31/n$, or essentially $1/1/n + 1$.

3. The Bootstrap Model

Although many variations of a bootstrap model framework are possible, this monograph will focus on the most common example which reproduces the basic chain ladder method—the ODP bootstrap model. Let's briefly review the assumptions of the basic chain ladder method, because these assumptions are important in understanding the distribution created by the ODP bootstrap model.

Start with a triangle array of cumulative data:

		d					
		1	**2**	**3**	. . .	**n–1**	**n**
w	**1**	c(1, 1)	c(1, 2)	c(1, 3)	. . .	c(1, n–1)	c(1, n)
	2	c(2, 1)	c(2, 2)	c(2, 3)	. . .	c(2, n–1)	
	3	c(3, 1)	c(3, 2)	c(3, 3)	. . .		
	. . .	. . .	. . .				
	n–1	c(n–1, 1)	c(n–1, 2)				
	n	c(n, 1)					

A typical deterministic analysis of this data will start with an array of development ratios or development factors:

$$F(w,d)=\frac{c(w,d)}{c(w,d-1)}. \tag{3.1}$$

Then two key assumptions are made in order to make a projection of the known elements to their respective ultimate values. First, it is assumed that each accident year has the same development factor. Equivalently, for each $w = 1, 2, \ldots, n$:

$$F(w,d)=F(d).$$

Under this first assumption, one of the more popular estimators for the development factor is the weighted average:

$$\hat{F}(d)=\frac{\sum_{w=1}^{n-d+1}c(w,d)}{\sum_{w=1}^{n-d+1}c(w,d-1)}. \tag{3.2}$$

Certainly there are other popular estimators in use, but they are beyond our scope at this stage yet most are still consistent with our first assumption that each accident year has the same factor. Projections of the ultimate values, or $\hat{c}(w, n)$ for $w = 1, 2, \ldots, n$ are then computed using:

$$\hat{c}(w,n) = c(w,d)\prod_{i=d+1}^{n}\hat{F}(i), \text{ for all } d = n - w + 1. \tag{3.3}$$

This part of the claim projection algorithm relies explicitly on the second assumption, namely that each accident year has a parameter representing its relative level. These level parameters are the current cumulative values for each accident year, or $c(w, n - w + 1)$. Of course variations on this second assumption are also common, but the point is that every model has explicit assumptions that are an integral part of understanding the quality of that model.

One variation on the second assumption is to assume that the accident years are completely homogeneous.[7] In this case we would estimate the level parameter of the accident years using:

$$\frac{\sum_{w=1}^{n-d+1} c(w,d)}{n-d+1}. \tag{3.4}$$

Complete homogeneity implies that the observations $c(1, d)$, $c(2, d)$, . . . , $c(n - d + 1, d)$ are generated by the same mechanism. Thus, the column averages from (3.4) would replace the last actual values along the diagonal to calculate an estimate assuming homogeneity of accident years.

Interestingly, the basic chain ladder algorithm treats the processes generating the observations as NOT homogeneous[8] and effectively that "pooling" of the data does not provide any increased efficiency.[9] In contrast, it could be argued that the Bornhuetter-Ferguson (1972) and Cape Cod methods are a "blend" of these two extremes as the homogeneity of the future expected result depends on the consistency of the *a priori* loss ratios and decay rate, respectively.

3.1. Origins of Bootstrapping

Possibly the earliest development of a stochastic model for the actuarial array of cumulative development data is attributed to Kremer (1982) and the earliest discussion of bootstrapping is in Ashe (1986). The basic model used by Kremer is described by England and Verrall (1999) and Zehnwirth (1989), so there will be no further elaboration here. It should be noted, however, that this model can be extended by considering alternatives which are discussed in Barnett and Zehnwirth (2000) and Zehnwirth (1994), Renshaw (1989), Christofides (1990), and Verrall (1991; 2004), among others.

---

[7] Homogeneous data can have a different meaning in mathematics, but here we are defining it to mean having consistent or the same underlying exposures.

[8] Meaning the underlying exposures are changing over time and thus the current cumulative results (observation) for each year are more appropriate for projecting an estimate.

[9] For a more complete discussion of these assumptions of the basic chain ladder model see Zehnwirth (1989).

3.2. The Over-Dispersed Poisson Model

The genesis of this model into an ODP bootstrap framework originated with Renshaw and Verrall (1994) when they proposed modeling the incremental claims $q(w, d)$ directly as the response, with the same linear predictor as Kremer (1982), but using a generalized linear model (GLM) with a log-link function and an over-dispersed Poisson (ODP) error distribution.[10] Then, England and Verrall (1999) discuss how a specific form of this model is identical to the volume weighted chain ladder model, and use bootstrapping (sampling the residuals with replacement) to estimate a distribution of point estimates[11] instead of simulating from a multivariate normal distribution for a GLM. More formally, the following formulas are used to parameterize the GLM.

$$E\left[q(w,d)\right] = m_{w,d} \text{ and } Var\left[q(w,d)\right] = \phi E\left[q(w,d)\right] = \phi m_{w,d}^{z} \tag{3.5}$$

$$\ln\left[m_{w,d}\right] = \eta_{w,d} \tag{3.6}$$

$$\eta_{w,d} = c + \alpha_w + \beta_d, \text{ where: } w = 1, 2, \ldots, n;\ d = 1, 2, \ldots, n; \text{ and } \alpha_1 = \beta_1 = 0. \tag{3.7}$$

In this case the α parameters function as adjustments to the constant, c, level parameter and the β parameters adjust for the development trends after the first development period. The power, z, is used to specify the error distribution with:

$z = 0$ for Normal,
$z = 1$ for Poisson,
$z = 2$ for Gamma, and
$z = 3$ for inverse Gaussian.

Thus, the z parameter specifies not only the mean-variance relationship, but the whole shape of the distribution, including higher moments. Alternatively, we can remove the constant, c, which will cause the α parameters to function as individual level parameters while the β parameters continue to adjust for the development trends after the first development period:

$$\eta_{w,d} = \alpha_w + \beta_d, \text{ where: } w = 1, 2, \ldots, n; \text{ and } d = 2, 3, \ldots, n. \tag{3.8}$$

Standard statistical software can be used to estimate parameters and goodness of fit measures. The parameter ϕ is a scale parameter that is estimated as part of the fitting procedure while setting the variance proportional to the mean (thus "over-dispersed" Poisson for $z = 1$)[12]. For educational purposes, the calculations to solve these equations

[10] Generalized Linear Modeling can be done with and without link functions and with a variety of error distributions. We are only describing here the particular GLM model that leads to the replication of the chain ladder results. For a more complete treatise on Generalized Linear Modeling, see McCullagh and Nelder (1989).

[11] Some authors refer to this as the standard deviation of the posterior distribution.

[12] While over-dispersed Poisson, or ODP, are commonly used terms for this model, it is certainly possible for the scale parameter to be less than one and thus "under-dispersed" Poisson would be more technically correct in that case. Alternatively, the more general term quasi-Poisson could be used. In addition, we note that the z parameter in equation 3.5, and some later formulas, could be removed for simplicity since the primary focus of this monograph is the ODP Bootstrap model, but it is included so we do not lose sight of the fact that the ODP Bootstrap model is a specialized case of a larger family of models.

for a 10×10 triangle are included in the "Bootstrap Models.xlsm" file, but here, and in the "GLM Framework.xlsm" file, the calculations are illustrated for a 3×3 triangle for ease of exposition. Consider the following incremental data triangle:

	1	2	3
1	$q(1, 1)$	$q(1, 2)$	$q(1, 3)$
2	$q(2, 1)$	$q(2, 2)$	
3	$q(3, 1)$		

In order to set up the GLM model to fit parameters to the data we need to do a log-link or transform which results in:

	1	2	3
1	$\ln[q(1, 1)]$	$\ln[q(1, 2)]$	$\ln[q(1, 3)]$
2	$\ln[q(2, 1)]$	$\ln[q(2, 2)]$	
3	$\ln[q(3, 1)]$		

The model, as described in (3.8), is then specified using a system of equations with vectors of α_w and β_d parameters as follows:

$$\begin{aligned}
\ln[q(1,1)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 0\beta_2 + 0\beta_3 \\
\ln[q(2,1)] &= 0\alpha_1 + 1\alpha_2 + 0\alpha_3 + 0\beta_2 + 0\beta_3 \\
\ln[q(3,1)] &= 0\alpha_1 + 0\alpha_2 + 1\alpha_3 + 0\beta_2 + 0\beta_3 \\
\ln[q(1,2)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 1\beta_2 + 0\beta_3 \\
\ln[q(2,2)] &= 0\alpha_1 + 1\alpha_2 + 0\alpha_3 + 1\beta_2 + 0\beta_3 \\
\ln[q(1,3)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 1\beta_2 + 1\beta_3.
\end{aligned} \tag{3.9}$$

Converting this to matrix notation we have:

$$Y = X \times A \tag{3.10}$$

Where:

$$Y = \begin{bmatrix} \ln[q(1,1)] \\ \ln[q(2,1)] \\ \ln[q(3,1)] \\ \ln[q(1,2)] \\ \ln[q(2,2)] \\ \ln[q(1,3)] \end{bmatrix}, \tag{3.11}$$

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}, \text{ and} \tag{3.12}$$

$$A = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_2 \\ \beta_3 \end{bmatrix}. \tag{3.13}$$

In this form we can use iteratively weighted least squares or maximum likelihood[13] to solve for the parameters in the A vector (3.13) that minimize the squared difference between the Y matrix (3.11) and the solution matrix (3.14):

$$\begin{bmatrix} \ln[m_{1,1}] \\ \ln[m_{2,1}] \\ \ln[m_{3,1}] \\ \ln[m_{1,2}] \\ \ln[m_{2,2}] \\ \ln[m_{1,3}] \end{bmatrix}. \tag{3.14}$$

After solving the system of equations we will have:

$$\begin{aligned} \ln[m_{1,1}] &= \eta_{1,1} = \alpha_1 \\ \ln[m_{2,1}] &= \eta_{2,1} = \alpha_2 \\ \ln[m_{3,1}] &= \eta_{3,1} = \alpha_3 \\ \ln[m_{1,2}] &= \eta_{1,2} = \alpha_1 + \beta_2 \\ \ln[m_{2,2}] &= \eta_{2,2} = \alpha_2 + \beta_2 \\ \ln[m_{1,3}] &= \eta_{1,3} = \alpha_1 + \beta_2 + \beta_3. \end{aligned} \tag{3.15}$$

This solution can then be shown as a triangle:

	1	2	3
1	$\ln[m_{1,1}]$	$\ln[m_{1,2}]$	$\ln[m_{1,3}]$
2	$\ln[m_{2,1}]$	$\ln[m_{2,2}]$	
3	$\ln[m_{3,1}]$		

[13] Other methods, such as orthogonal decomposition or Newton-Raphson, can also be used to solve for the parameters. Iteratively weighted least squares and maximum likelihood are both illustrated in the companion Excel files.

These results can then be exponentiated to produce the fitted, or expected, incremental results of the GLM model:

	1	2	3
1	$m_{1,1}$	$m_{1,2}$	$m_{1,3}$
2	$m_{2,1}$	$m_{2,2}$	
3	$m_{3,1}$		

This monograph will refer to this as the "GLM framework" and it is illustrated for a simple 3×3 triangle in the "GLM Framework.xlsm" file. While the GLM framework is used to solve these equations for the fitted results, the usefulness of this framework is that the fitted incremental values (with the Poisson error distribution assumption) will equal the incremental values that can be derived from volume-weighted average development factors, as shown in the "GLM Framework.xlsm" file.[14] That is, it can be reproduced by using the last cumulative diagonal, dividing backwards successively by each volume-weighted average development factor and subtracting to get the fitted incremental results. This monograph will refer to this method as the "simplified GLM" or "ODP Bootstrap." This has three very useful consequences.

First, the GLM portion of the algorithm can be replaced with a simpler development factor algorithm while still being based on the underlying GLM framework. Second, the use of the development factors serves as a "bridge" to the deterministic framework and allows the model to be more easily explainable to others. And, third, for the GLM algorithm the log-link process means that negative incremental values can often cause the algorithm to not have a solution, whereas using development factors will generally allow for a solution.[15]

With a model fitted to the data, the ODP bootstrap process involves sampling with replacement from the residuals. England and Verrall (1999) note that the deviance, Pearson, and Anscombe residuals could all be considered for this process, but the Pearson residuals are the most desirable since they are calculated consistently with the scale parameter. The unscaled Pearson residuals, $r_{w,d}$, and scale parameter, ϕ, are calculated as follows:

$$r_{w,d} = \frac{q(w,d) - m_{w,d}}{\sqrt{m_{w,d}^{z}}}. \tag{3.16}$$

$$\phi = \frac{\sum r_{w,d}^{2}}{N - p}. \tag{3.17}$$

[14] Using other than the Poisson assumption (i.e., $z \neq 1$), the incremental values may be close to the values from the development factors, but they will not be equal.

[15] More specifically, individual negative cell values may not be a problem (by using the negative of the log of the absolute value in 3.14). If the total of all incremental cell values in a development column is negative, then the GLM algorithm will fail. This situation will not cause a problem fitting the model as a link ratio less than one will be perfectly useful. However, this may still cause other problems, e.g., the "GLM framework" and "simplified GLM" may not be equivalent, which we will address in Section 4.

Where N = the number of observations, or incremental data cells in the triangle, which is typically equal to $n \times (n + 1) \div 2$, and p = the number of parameters, which is typically equal to $2 \times (n - 1)$.[16] Sampling with replacement from the residuals can then be used to create new sample triangles of incremental values using formula 3.18. Sampling with replacement assumes that the residuals are independent and identically distributed, but it does not require the residuals to be normally distributed. Indeed, this is often cited as an advantage of the ODP bootstrap model since whatever distributional form the residuals have will flow through to the simulation process. Some authors have referred to this as a "semi-parametric" bootstrap model since we are not parameterizing the residuals.

$$q^*(w,d) = r^* \times \sqrt{m_{w,d}^z} + m_{w,d}. \tag{3.18}$$

The sample triangle of incremental values can then be cumulated, new average development factors can be calculated for the sample and applied to calculate a point estimate for this data, resulting in a distribution of point estimates for some large number of samples. In England and Verrall (1999) this is the end of the process, but at the end of the appendix they note that you should also adjust the resulting distribution by the degrees of freedom adjustment factor (3.19) and the Scale Parameter (3.17), to effectively allow for over-dispersion of the residuals in the sampling process and add process variance to approximate a distribution of possible outcomes.

$$f^{DoF} = \sqrt{\frac{N}{N-p}}. \tag{3.19}$$

Later, in England and Verrall (2002), the authors note that the Pearson residuals (3.16) could be multiplied by the degrees of freedom adjustment factor (3.19) to include the over-dispersion in the residuals. As calculated in (3.20), these adjusted residuals are referred to as scaled Pearson residuals. They also expand the simulation process by adding process variance to the future incremental values from the point estimates. To add this process variance, they assume that each future incremental value $m_{w,d}$ is the mean and the mean times the scale parameter, $\phi m_{w,d}$, is the variance of a gamma distribution.[17] This revised model could now be described as estimating a distribution of possible outcomes, which incorporates process variance and parameter variance in the simulation of the historical and future data.[18]

[16] The number of data cells could be less than $n \times (n + 1) \div 2$ and the number of parameters could be less than $2 \times (n - 1)$. For example, if the incremental values are zeros for the last three columns in a triangle then these cells would not be included in the total for N and there will be three fewer β parameters since none are needed to fit to these zero values as the development process is completed already.

[17] The Poisson distribution could be used to remain more consistent with the underlying theory of the GLM framework, but it is considerably slower to simulate, so gamma is a close substitute that performs much faster in simulation although it can be more skewed than the Poisson. Indeed, other distributions could be used as well to better approximate the observed "skewness" of the residuals from the diagnostics.

[18] Some authors refer to this as the full predictive distribution of the cash flows.

$$r_{w,d}^{S} = \frac{q(w,d) - m_{w,d}}{\sqrt{m_{w,d}^{z}}} \times f^{DoF}. \tag{3.20}$$

However, Pinheiro et al. (2001; 2003) noted that the bias correction for the residuals using the degrees of freedom adjustment factor (3.20) does not create standardized residuals, which is an important step for making sure that the residuals all have the same variance. In order to have standardized Pearson residuals, the GLM framework requires the use of a hat matrix adjustment factor (3.23).

$$H = X\left(X^{T}WX\right)^{-1}X^{T}W. \tag{3.21}$$

First, the hat matrix (3.21) is calculated using matrix multiplication of the design matrix (3.12) and the weight matrix (3.22).

$$W = \begin{bmatrix} m_{1,1} & 0 & 0 & 0 & 0 & 0 \\ 0 & m_{2,1} & 0 & 0 & 0 & 0 \\ 0 & 0 & m_{3,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & m_{1,2} & 0 & 0 \\ 0 & 0 & 0 & 0 & m_{2,2} & 0 \\ 0 & 0 & 0 & 0 & 0 & m_{1,3} \end{bmatrix} \tag{3.22}$$

$$f_{w,d}^{H} = \sqrt{\frac{1}{1 - H_{i,i}}}. \tag{3.23}$$

The hat matrix adjustment factor (3.23) uses the diagonal of the hat matrix (3.21). In Pinheiro et al. (2003) the authors note two important points about the ODP bootstrap process as described by England and Verrall (1999; 2002). First, the sampling of the residuals should not include any zero-value residuals, which are typically in the corners of the triangle.[19] The exclusion of the zero-value residuals is accounted for in the hat matrix adjustment factor (3.23), but another common explanation is that the zero-value cells will have some variance but we just don't know what it is yet so we should sample from the remaining residuals but not the zeros. Second, the hat matrix adjustment factor (3.23) is a replacement for, and an improvement on, the degrees of freedom factor (3.19).[20]

Thus, the scaled Pearson residuals (3.20) should be replaced by the standardized Pearson residuals:

$$r_{w,d}^{H} = \frac{q(w,d) - m_{w,d}}{\sqrt{m_{w,d}^{z}}} \times f_{w,d}^{H}. \tag{3.24}$$

---

[19] Technically, the two "corner" residuals are zero because they each have a parameter that is unique to that incremental value which causes the fitted incremental value to exactly equal the actual incremental value.

[20] This second point was not addressed clearly in Pinheiro et al. (2001), but as the authors updated and clarified the monograph in Pinheiro et al. (2003) this issue was more clearly addressed.

However, the scale parameter (3.17) is still calculated as before, although the standardized Pearson residuals could be used to approximate the scale parameter as follows:

$$\phi^H = \frac{\sum \left(r_{w,d}^H \right)^2}{N}. \tag{3.25}$$

At this point we have a complete basic "ODP bootstrap" model, as it is often referred to. It is also important to note that the two key assumptions mentioned earlier, each accident year has the same development factor and each accident year has a parameter representing its relative level, are equally applicable to this model.

In order for the reader to test out the different "combinations" of this modeling process the "Bootstrap Models.xlsm" file includes options to allow these historical algorithms to be simulated. The purpose for describing this evolution of the ODP bootstrap model framework is threefold: first, to allow the interested reader to better understand the details of the algorithm and how these papers and their authors have contributed to the evolution of this model framework; second, to illustrate the value of collaborative research via different published papers and the contributions of different authors; and, third, to provide a solid foundation for continuing the evolutionary process and to discuss practical adjustments.

3.3. Variations on the ODP Model

When estimating insurance risk it is generally considered desirable to focus on the claim payment stream in order to measure the variability of the actual cash flows that directly affect the bottom line. Clearly, changes in case reserves and IBNR reserves will also impact the bottom line, but to a considerable extent the changes in IBNR are intended to counter the impact of the changes in case reserves. To some degree, then, the total reserve movements can act to mask the underlying changes due to cash flows. On the other hand, the case reserves contain valuable information about potential future payments so we should not ignore them and use only paid data.

3.3.1. Bootstrapping the Incurred Loss Triangle

The ODP bootstrap model can be used to model both paid and incurred loss data. Using incurred data incorporates case reserves, thus perhaps improving the ultimate estimates. However, the resulting distribution from using incurred data will be possible outcomes of the IBNR, not a distribution of the unpaid.[21] There are two possible approaches for modeling an unpaid loss distribution using incurred loss data: modeling incurred data and convert the ultimate values to a payment pattern, or, modeling paid and case reserves separately.

Using the first approach, a convenient way of converting the results of an incurred data model to a payment stream is to run the paid data model in parallel with the

[21] Using incurred data will also create issues in weighting the results of different models which will be discussed in Section 6.

incurred data model, and use the random payment pattern from each iteration from the paid data model to convert the ultimate values from each corresponding iteration from the incurred data to a payment pattern for each iteration (for each accident year individually). The "Bootstrap Models.xlsm" file illustrates this concept. It is worth noting, however, that this process allows the "added value" of using the case reserves to help predict the ultimate results to work its way into the calculations, thus perhaps improving the ultimate estimates, while still focusing on the payment stream for measuring risk. In effect, it allows a distribution of IBNR to become a distribution of IBNR and case reserves. This process could be made more sophisticated by correlating some part of the paid and incurred models (e.g., the residual sampling and/or process variance portions), but that is beyond the scope of this monograph.

The second approach could be accomplished by applying the ODP bootstrap to the Munich chain ladder model. This has the advantage over the first approach of not modeling the paid losses twice, and of explicitly measuring and imposing a framework around the correlation of the paid and outstanding losses. Since it is so well detailed in Liu and Verrall (2010), it will not be discussed in detail here.

3.3.2. Bootstrapping the Bornhuetter-Ferguson and Cape Cod models

Another common issue with using the ODP bootstrap model is that the distribution for the most recent accident years can produce results with more variance than you would expect when compared to earlier accident years. This is usually because more development factors are used to extrapolate the sampled values for the most recent accident years which, when coupled with random samples of incremental values, can result in more extreme fluctuations in point estimates. This is analogous to one of the weaknesses of the deterministic paid chain ladder method—a low, or high, initial observation can lead to an abnormally low, or high, projected ultimate, respectively.

To help alleviate this problem, the Bornhuetter-Ferguson (1972) or generalized Cape Cod (Struzzieri and Hussian 1998) deterministic methods can be worked into the underlying ODP bootstrap model, and the deterministic assumptions of these methods can also be converted to stochastic assumptions. For example, instead of using deterministic *a priori* loss ratios for the Bornhuetter-Ferguson model, the *a priori* loss ratios can be simulated from a distribution. Similarly, the Cape Cod algorithm can be applied to every ODP bootstrap model iteration to produce a stochastic Cape Cod projection that reflects the unique characteristics of each sample triangle.[22]

The "Bootstrap Models.xlsm" file also illustrates these Bornhuetter-Ferguson and Cape Cod ODP bootstrap models.[23]

---

[22] In addition to being consistent between paid and incurred data, to the extent there is commonality with deterministic methods the assumptions should also be consistent. For example, it would not make sense to use one set of *a priori* loss ratio assumptions for a deterministic Bornhuetter-Ferguson method and a different set of mean assumptions for a modified ODP bootstrap model.

[23] More complex implementations of these models could include modifying the underlying assumptions of the GLM framework which would result in a completely different set of residuals, but that is beyond the scope of this monograph.

3.4. The GLM Bootstrap Model

Two limitations of the chain-ladder model, and hence the ODP bootstrap of the chain-ladder model, is that it does not measure or adjust for calendar-year effects, and it includes a significant number of parameters and many would argue that it over-fits the model to the data.

Another approach is to go back to the original GLM framework. Returning to formulas (3.5) to (3.8), the GLM framework does not require a certain number of parameters so we are free to specify only as many parameters as we need to get a robust model, which can address the over-fitting issue. Indeed, it is ONLY when we specify a parameter for EVERY accident year and EVERY development year and specify a Poisson error distribution that we end up exactly replicating the volume weighted average development factors that allow us to substitute the deterministic algorithm instead of solving the GLM fit.

Thus, using the original GLM framework, which this monograph will refer to as the "GLM Bootstrap" model, we can specify a model with only a few parameters, but there are two drawbacks to doing so.[24] First, the GLM must be solved for each iteration of the bootstrap model (which may slow down the simulation process) and, second, the model is no longer directly explainable to others using development factors.[25] While the impact of these drawbacks should be considered, the potential benefits of using the GLM bootstrap can be much greater.

First, having fewer parameters will help avoid over-parameterizing the model.[26] For example, if we use only one accident year parameter then the model specified using a system of equations is as follows (which is analogous to formula 3.9):

$$
\begin{aligned}
\ln[q(1,1)] &= 1\alpha_1 + 0\beta_2 + 0\beta_3 \\
\ln[q(2,1)] &= 1\alpha_1 + 0\beta_2 + 0\beta_3 \\
\ln[q(3,1)] &= 1\alpha_1 + 0\beta_2 + 0\beta_3 \\
\ln[q(1,2)] &= 1\alpha_1 + 1\beta_2 + 0\beta_3 \\
\ln[q(2,2)] &= 1\alpha_1 + 1\beta_2 + 0\beta_3 \\
\ln[q(1,3)] &= 1\alpha_1 + 1\beta_2 + 1\beta_3
\end{aligned} \tag{3.26}
$$

In this case we will only have one accident year parameter and $n - 1$ development trend parameters, but it will only be coincidence that we would end up with the equivalent of average development factors. Interestingly, this model parameterization moves us away from one of the common basic assumptions (i.e., each accident year has its own level) and substitutes the assumption that all accident years are homogeneous.

[24] Using the GLM framework allows for many other variations in the specification of models and then bootstrapping as described in more detail in England and Verrall (1999; 2002) and others, but this monograph will focus on variations consistent with the framework underpinning the ODP bootstrap model.

[25] However, age-to-age factors could be calculated for the fitted data to compare to the actual age-to-age factors and used as an aid in explaining the model to others.

[26] Over-parameterization will be addressed more completely in Section 5.

Another example of using fewer parameters would be to only use one development year parameter (while continuing to use an accident-year parameter for each year), which would equate to the system of equations in (3.27).

$$\begin{aligned}
\ln[q(1,1)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 0\beta_2 \\
\ln[q(2,1)] &= 0\alpha_1 + 1\alpha_2 + 0\alpha_3 + 0\beta_2 \\
\ln[q(3,1)] &= 0\alpha_1 + 0\alpha_2 + 1\alpha_3 + 0\beta_2 \\
\ln[q(1,2)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 1\beta_2 \\
\ln[q(2,2)] &= 0\alpha_1 + 1\alpha_2 + 0\alpha_3 + 1\beta_2 \\
\ln[q(1,3)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 2\beta_2
\end{aligned} \tag{3.27}$$

In this example the model parameterization moves away from the other common basic assumption (i.e., each accident year has its own level, but the same development parameter is used for all periods), and again it would be pure coincidence to end up with the equivalent of average development factors.[27] It is also interesting to note that for both of these two examples there will be one additional non-zero residual that can be used in the simulations because in each case one of the incremental values no longer has a unique parameter—i.e., for (3.26) $q(3, 1)$ is no longer uniquely defined by α_3, and for (3.27) $q(1, 3)$ is no longer uniquely defined by β_3.

Of course we can take this simplification to its logical extreme and use a model with only one accident year parameter and one development year parameter, which would result in the system of equations in as shown in (3.28).

$$\begin{aligned}
\ln[q(1,1)] &= 1\alpha_1 + 0\beta_2 \\
\ln[q(2,1)] &= 1\alpha_1 + 0\beta_2 \\
\ln[q(3,1)] &= 1\alpha_1 + 0\beta_2 \\
\ln[q(1,2)] &= 1\alpha_1 + 1\beta_2 \\
\ln[q(2,2)] &= 1\alpha_1 + 1\beta_2 \\
\ln[q(1,3)] &= 1\alpha_1 + 2\beta_2
\end{aligned} \tag{3.28}$$

In this example the model parameterization moves away from both of the common basic assumptions (i.e., each accident year has its own level, and the different development parameter is used for all periods), and again it would be pure coincidence to end up with the equivalent of average development factors. In this most "basic" model it is interesting to note that both of the "zero residuals" will now be non-zero and can be used in the simulations because both corners no longer have a unique parameter.

This flexibility allows the modeler to use enough parameters to capture the statistically relevant level and trend changes in the data without forcing a specific number of parameters.[28]

[27] While we have only one parameter to describe the development period trends, if we convert these to development factors there will be a different factor for each period.

[28] How to determine which parameters are statistically relevant will be discussed in Section 5.

The second benefit, and depending on the data perhaps the most significant, is that this framework affords us the ability to add parameters for calendar-year trends. Adding diagonal, or calendar year trend, parameters to (3.8) we now have:

$$\eta_{w,d} = \alpha_w + \beta_d + \gamma_k \text{, where: } w = 1, 2, \ldots, n;\ d = 2, 3, \ldots, n; \text{ and } k = 2, 3, \ldots, n. \tag{3.29}$$

A complete system of equations for the (3.29) framework would look like the following:

$$\begin{aligned}
\ln[q(1,1)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 0\beta_2 + 0\beta_3 + 0\gamma_2 + 0\gamma_3 \\
\ln[q(2,1)] &= 0\alpha_1 + 1\alpha_2 + 0\alpha_3 + 0\beta_2 + 0\beta_3 + 1\gamma_2 + 0\gamma_3 \\
\ln[q(3,1)] &= 0\alpha_1 + 0\alpha_2 + 1\alpha_3 + 0\beta_2 + 0\beta_3 + 1\gamma_2 + 1\gamma_3 \\
\ln[q(1,2)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 1\beta_2 + 0\beta_3 + 1\gamma_2 + 0\gamma_3 \\
\ln[q(2,2)] &= 0\alpha_1 + 1\alpha_2 + 0\alpha_3 + 1\beta_2 + 0\beta_3 + 1\gamma_2 + 1\gamma_3 \\
\ln[q(1,3)] &= 1\alpha_1 + 0\alpha_2 + 0\alpha_3 + 1\beta_2 + 1\beta_3 + 1\gamma_2 + 1\gamma_3
\end{aligned} \tag{3.30}$$

However, there is no unique solution for a system with seven parameters and six equations, so some of these parameters will need to be removed. A logical starting point would be to start with a "basic" model with one accident year (level) parameter, one development trend parameter and one calendar trend parameter and then add or remove parameters as needed.[29] The system of equations for this basic model is as follows:

$$\begin{aligned}
\ln[q(1,1)] &= 1\alpha_1 + 0\beta_2 + 0\gamma_2 \\
\ln[q(2,1)] &= 1\alpha_1 + 0\beta_2 + 1\gamma_2 \\
\ln[q(3,1)] &= 1\alpha_1 + 0\beta_2 + 2\gamma_2 \\
\ln[q(1,2)] &= 1\alpha_1 + 1\beta_2 + 1\gamma_2 \\
\ln[q(2,2)] &= 1\alpha_1 + 1\beta_2 + 2\gamma_2 \\
\ln[q(1,3)] &= 1\alpha_1 + 2\beta_2 + 2\gamma_2
\end{aligned} \tag{3.31}$$

A third benefit of the GLM bootstrap model is that it can be used to model data shapes other than triangles. For example, missing incremental data for the first few diagonals would mean that the cumulative values could not be calculated and the remaining values in those first few rows would not be useful for the ODP bootstrap. However, since the GLM bootstrap uses the incremental values the entire trapezoid can be used to fit the model parameters.[30]

---

[29] A simple algorithm to add and/or remove parameters in a search for the "optimal" set of parameters is included in the "Bootstrap Models.xlsm" file, but more complex algorithms are outside the scope of this monograph. We focus on the "mechanical" aspects of searching for the "optimal" set of parameters in Section 5 in order to enhance the educational benefits.

[30] This issue will be examined in more detail in Section 4.

It should also be noted that the GLM bootstrap model allows the future expected values to be directly estimated from the parameters of the model for each sample triangle in the bootstrap simulation process. However, we must solve the GLM within each iteration for the same parameters as we originally set up for the model rather than using development factors to project future expected values (which is a way of fitting the model to each sample triangle).

The additional modeling power that this flexible GLM bootstrap model adds to the actuary's toolkit cannot be overemphasized. Not only does it allow one to move away from the two basic assumptions of a deterministic chain ladder method, it allows for the ability to match the model parameters to the statistical features you find in the data, rather than "force" the data to fit the model, often with far fewer parameters and to extrapolate those features. For example, modeling with fewer development trend parameters means that the last parameter can be assumed to continue past the end of the triangle which will give the modeler a "tail" of the incremental values beyond the end of the triangle without the need for a specific tail factor.

While the monograph continues to illustrate the GLM bootstrap with a 3 × 3 triangle, also included in the companion Excel files are a set of "GLM Bootstrap 6___.xlsm" files that illustrate the calculations for these different models using a 6 × 6 triangle. Also, the "Bootstrap Models.xlsm" file contains a "GLM bootstrap" model for a 10 × 10 triangle that can be used to specify any combination of accident year, development year, and calendar year parameters, including setting parameters to zero. The GLM bootstrap model is akin to the incremental log model described in Barnett and Zehnwirth (2000), so we will leave it to the reader to explore this flexibility by using the Excel file.

4. Practical Issues

Now that the basic ODP bootstrap model has been expanded in a variety of ways, it is important to address some of the key assumptions of the ODP model and some common data issues.

4.1. Negative Incremental Values

As noted in Section 3.2, because of the log-link used in the GLM framework the incremental values must be greater than zero in order to parameterize a model. However, a slight modification to the log-link function will help this common problem become a little less restrictive. If we use (4.1) as the log-link function, then individual negative values are only an issue if the total of all incremental values in a development column is negative, as the GLM algorithm will not be able to find a solution in that case.

$$\begin{gathered} \ln\left[q(w,d)\right] \text{ for } q(w,d)>0, \\ 0 \text{ for } q(w,d)=0, \\ -\ln\left[abs\{q(w,d)\}\right] \text{ for } q(w,d)<0. \end{gathered} \tag{4.1}$$

Using (4.1) in the GLM bootstrap will help in many situations, but it is quite common for entire development columns of incremental values to be negative, especially for incurred data. To give the GLM framework the ability to solve for a solution in this case we need to make another modification to the basic model to include a constant. Whenever a column or columns of incremental values sum to a negative value, we can find the largest negative[31] in the triangle, set ψ equal to the largest negative and adjust the log-link function by making all the incremental values positive.

$$\begin{gathered} q^{+}(w,d)=q(w,d)-\psi \\ \ln\left[q^{+}(w,d)\right] \text{ for all } q(w,d) \end{gathered} \tag{4.2}$$

Using the adjusted log-link function (4.2) we can solve the GLM using formulas (3.7), (3.8), or (3.27). Then we use (4.3) to adjust the fitted incremental values

[31] The largest negative value can either be the largest negative among the sums of development columns (in which case there may still be individual negative values in the adjusted triangle) or the largest negative incremental value in the triangle.

and the constant ψ is used to reduce each fitted incremental value by the largest negative.

$$m_{w,d} = \overset{+}{m}_{w,d} + \psi \tag{4.3}$$

The combination of formulas (4.2) and (4.3) allow the GLM bootstrap to handle all negative incremental values, which overcomes a common criticism of the ODP bootstrap. Incidentally, these formulas can also be used to allow the incremental log model described by Barnett and Zehnwirth (2000) to handle negative incremental values. As long as these formulas are used sparingly, the author believes that the resulting distribution will not be adversely affected.

When using the ODP bootstrap simulation process, the solution to negative incremental values needs to focus on the residuals and sampled incremental values since a development factor less than 1.00 will create negative incremental values in the fitted values. More specifically, we need to modify formulas (3.16) and (3.18) as follows: [32]

$$r_{w,d} = \frac{q(w,d) - m_{w,d}}{\sqrt{abs\{m^{z}_{w,d}\}}}. \tag{4.4}$$

$$q^*(w,d) = r^* \times \sqrt{abs\{m^{z}_{w,d}\}} + m_{w,d}. \tag{4.5}$$

While the fitted incremental values and residuals using the development factor simplification (ODP bootstrap) will generally not match the GLM framework solution using (4.1) or (4.2) and (4.3) they should be reasonably close. While the purists may object to these practical solutions, we must keep in mind that every model is an approximation of reality so our goal is to find reasonably close models that replicate the statistical features in the data rather than only restrict ourselves to "pure" models. After all, the assumptions of the "pure" models are themselves approximations.

4.1.1. Negative Values During Simulation

Even though we have solved problems with negative values when parameterizing a model, negative values can still affect the process variance in the simulation process. When each future incremental value (using $m_{w,d}$ as the mean and the mean times the scale parameter, $\phi m_{w,d}$, as the variance) is sampled from a gamma distribution to add process variance, the parameters of a gamma distribution must be positive. In this case we have two options for using the gamma distribution to simulate from a negative incremental value, $m_{w,d}$.

$$-Gamma\left[abs\{m_{w,d}\}, \phi abs\{m_{w,d}\}\right] \tag{4.6}$$

$$Gamma\left[abs\{m_{w,d}\}, \phi abs\{m_{w,d}\}\right] + 2m_{w,d} \tag{4.7}$$

[32] The use of other types of residuals, as noted in Section 3.2, may also help address the issue of negative incremental values, but their exposition is left to the interested reader.

Using formula (4.6) is more intuitive as we are using absolute values to simulate from a gamma distribution and then changing the sign of the result. However, since the gamma distribution is skewed to the right, the resulting distribution using (4.6) will be skewed to the left. Using formula (4.7) is a little less intuitive, but seems more logical since adding twice the mean, $m_{w,d}$, will result in a distribution with a mean of $m_{w,d}$ while keeping it skewed to the right (since $m_{w,d}$ is negative).

Negative incremental values can also cause extreme outcomes. This is most prevalent when resampled triangles are created with negative incremental losses in the first few development periods, causing one column of cumulative values to sum close to zero and the next column to sum to a much larger number and, consequentially, produce development factors that are extremely large. This can result in one or more extreme iterations in a simulation (for example, outcomes that are multiples of 1,000s of the central estimate). These extreme outcomes cannot be ignored, even if the high percentiles are not of interest, because they may significantly affect the mean of the distribution.

In these instances, you have several options. You can 1) remove these iterations from your simulation and replace them with new iterations, 2) recalibrate your model, or 3) limit incremental values to a minimum of zero (or some other minimum value).

The first option is to identify the extreme iterations and remove them from your results. Care must be taken that only truly unreasonable extreme iterations are removed, so that the resulting distribution does not understate the probability of extreme outcomes.

The second option is to recalibrate the model to fix this issue. First you must identify the source of the negative incremental losses. The most theoretically sound method to deal with negative incremental values is to consider the source of these losses. For example, it may be from the first row in your triangle, which was the first year the product was written, and therefore exhibit sparse data with negative incremental amounts. One option is to remove this row from the triangle if it is causing extreme results and does not improve the parameterization of the model. Or, if they are caused by reinsurance or salvage and subrogation, then you can model the losses gross of salvage and subrogation, model the salvage and subrogation separately, and combine the iterations assuming the values are correlated.

The third option is to constrain the model output by limiting incremental losses to a minimum of zero, where any negative incremental is replaced with a zero incremental.[33] For each of these options, keep in mind that this is a form of diagnosing a model by reviewing the simulated results and then searching for a practical solution before abandoning a model altogether. This does not mean that you should never abandon a model in favor of a practical adjustment. Indeed, the higher the frequency of the underlying issue (negative incremental values in this case) the more likely that the model does not really fit the data.

---

[33] While zero is a convenient minimum or lower bound, a small positive number could also be used, in which case any values less than the minimum are changed to the minimum.

4.2. Non-Zero Sum of Residuals

The standardized residuals that are calculated in the ODP bootstrap model are essentially error terms, and should in theory be independent and identically distributed with a mean of zero. However, the residuals are random observations of the true residual distribution, so the average of all the residuals is usually non-zero. If significantly different than zero, then the fit of the model should be questioned. If the average of the residuals is close to zero, then the question is whether they should be adjusted so that their average is zero. For example, if the average of the residuals is positive, then re-sampling from the residual pool will not only add variability to the resampled incremental losses, but may increase the resampled incremental losses such that the average of the resampled loss will be greater than the fitted loss.

It could be argued that the non-zero average of residuals is a characteristic of the data set, and therefore should not be removed. For example, standardized residuals implies a normal distribution with zero mean, but skewness in the residuals does not necessarily imply an average of zero. However, if a zero residual average is desired, then one option is the addition of a single constant to all non-zero residuals, such that the sum of the shifted residuals is zero.

4.3. Using an *N*-Year Weighted Average

It is quite common for actuaries to use weighted averages that are less than all years in their chain-ladder and related methods. Similarly, both the ODP bootstrap and the GLM bootstrap can be adjusted to only consider the data in the most recent diagonals. For the GLM framework (and the GLM bootstrap model), we can use only the most recent $L + 1$ diagonals (since an L-year average uses $L + 1$ diagonals) to parameterize the model. The shape of the data to be modeled essentially becomes a trapezoid instead of a triangle, the excluded diagonals are given zero weight in the model and we have fewer calendar year trend parameters if we are using formula (3.29). When running the GLM bootstrap simulations we will only need to sample residuals for the trapezoid that was used to parameterize the model as that is all that will be needed to estimate parameters for each iteration.

For the ODP bootstrap model, we can calculate L-year average factors instead of all-year factors and only have residuals for the most recent $L + 1$ diagonals. However, when running the ODP bootstrap simulations we would still need to create a whole resampled triangle so that we can calculate cumulative values.[34] But, for consistency, we would want to use L-year average factors for projecting the future expected values from these resampled triangles.

The calculations for the GLM bootstrap are illustrated in the companion "GLM Bootstrap 6 with 3yr avg.xlsm" file. Note that because the GLM bootstrap estimates parameters for the incremental data, the fitted values will no longer match the fitted values from the ODP bootstrap using volume-weighted average development factors.

[34] The fitted values for the "unused" diagonals would be calculated using the L-year average ratios, but the corresponding residuals for those diagonals are all excluded from the sampling process.

Depending on the data, the fitted values from the simplified GLM (ODP bootstrap) may or may not be a reasonable approximation to the GLM framework (GLM bootstrap).

Note that this discussion of using L-year average factors assumes volume weighted averages to be consistent with the GLM framework. This also assumes that all of the diagnostic tests will be adjusted to reflect the use of the last $L + 1$ diagonals, although this is beyond the scope of the monograph. Finally, other types of averages could be used (i.e., straight average, average excluding high & low, etc.) to be more consistent with what actuaries might use in a deterministic analysis, but these typically move further away from the GLM framework and are beyond the scope of this monograph.

4.4. Missing Values

Sometimes the loss triangle will have missing values. For example, values may be missing from the middle of the triangle, or a triangle may be missing the oldest diagonals, if loss data was not kept in the early years of the book of business.

If values are missing, then the following calculations will be affected:

- Loss development factors
- Fitted triangle—if the missing value lies on the most recent diagonal
- Residuals
- Degrees of freedom

There are several solutions. The missing value may be estimated using the surrounding values. Or, the loss development factors can be modified to exclude the missing values, and there will not be a corresponding residual for those missing values. Subsequently, when triangles are resampled, the simulated incremental corresponding to the missing value should still be resampled so that the cumulative values in those rows can be calculated, but they would still be excluded from the projection process (i.e., not included with the sample age-to-age factors) to reproduce the uncertainty in the original dataset.

If the missing value lies on the most recent diagonal, the fitted triangle cannot be calculated in the usual way. A solution is to estimate the value, or use the value in the second most recent diagonal to construct the fitted triangle. These are not strictly mathematically correct solutions, and judgment will be needed as to their effect on the resulting distribution. Of course for the GLM bootstrap model, the missing data only reduces the number of observations used in the model.

4.5. Outliers

There may be a few extreme or incorrect values in the original triangle dataset that could be considered outliers. These may not be representative of the variability of the dataset in the future and, if so, the modeler may want to remove their impact from the model.

There are several solutions. These values could be removed, and dealt with in the same manner as missing values. Another alternative is to identify outliers and exclude them from the average development factors (either the numerator, denominator, or both) and residual calculations, as when dealing with missing values, but re-sample the corresponding incremental when re-sampling triangles. In this case we have removed the extreme impact of the incremental cell, but we still want to include a non-extreme variability, which is different from a missing data cell since in that case the additional uncertainty of that missing data can be included by continuing to exclude that cell in the projection process.

The calculations for the GLM bootstrap are illustrated in the companion "GLM Bootstrap 6 with Outlier.xlsm" file. Again the GLM bootstrap fitted values will no longer exactly match the fitted values from the ODP Bootstrap using volume weighted average development factors, but they should normally be close.

If there are a significant number of outliers, then this could be an indication that the model is not a good fit to the data. With the GLM bootstrap, new parameters could be chosen, or the distribution of the error term can be changed (i.e., change the z parameter). Under the ODP bootstrap model, an L-year weighted average could be used, instead of an all year weighted average, which may provide a better fit to the data, or, heteroscedasticity may exist. Remember, though, that for the ODP bootstrap model there is no distribution assumption for the residuals so a significant number of residual outliers could just mean that the residuals are quite skewed. One of the nice features of the ODP bootstrap is that the skewness in the residuals will be reflected in the simulation process which will result in a skewed distribution of possible outcomes.[35] Thus, removing any outliers (i.e., giving them zero weight) should be done with caution and would most commonly be done only after understanding the underlying data.

4.6. Heteroscedasticity

As noted earlier, the ODP bootstrap model is based on the assumption that the standardized Pearson residuals are independent and identically distributed. It is this assumption that allows the model to take a residual from one development period/accident period and apply it to the fitted loss in any other development period/accident period, to produce the sampled values. In statistical terms this is referred to as homoscedasticity (the residuals have the same variance) and it is important that this assumption is validated.

A common problem is when some development periods have residuals that appear to be more variable than others—i.e., they appear to have different variances. This is referred to as heteroscedasticity. With heteroscedasticity, it is no longer possible to take a residual from one development/accident period and deem it suitable to be applied to any other development/accident period. In making this assessment, you must account for the credibility of the observed differences in variance, and also to note that there are fewer residuals as the development years become older, so comparing development years is difficult, particularly near the tail-end of the triangle.[36]

[35] Other methods of handling outliers could also be introduced, e.g., tempering residuals that are further away from the interquartile range, but the key to any approach is to understand what the residuals represent so an explicit assumption can be made and the "best" solution can be used.

[36] Section 5 will illustrate how to use residual graphs and other statistical tests to evaluate heteroscedasticity.

The existence of heteroscedasticity may suggest that the model is not a good fit for the data. Under an ODP bootstrap, there are a number of ad-hoc adjustments that can be made to address heteroscedasticity, but they may or may not improve the fit of the model to the data. They also often result in even more parameters in a model which could already be over-parameterized. In contrast, under a GLM bootstrap the flexibility of choosing the number of parameters to use, the ability to account for any calendar year trends, and the flexibility to choose the distribution of the error term mean that there are many ways within the model framework itself to improve the fit to the data. Therefore, this flexibility could remove the heteroscedasticity problem or at least reduce it.

Nevertheless, if the ODP bootstrap model is still to be used, then to adjust for heteroscedasticity in your data there are at least three options, 1) stratified sampling, 2) calculating hetero-adjustment (or variance) parameters, or 3) calculate non-constant scale parameters. Stratified sampling is accomplished by grouping those development periods with homogeneous variances and then sampling only from the residuals in each group. While this process is straightforward, some groups may only have a few residuals in them, which limits the amount of variability in the possible outcomes compared to the other two options and at least partially defeats the benefits of random sampling with replacement.

The second option is to group those development periods with homogeneous variances and calculate the standard deviation of the residuals in each of the groups. Then calculate h_i, which is the "hetero-adjustment" factor, for each group, i:

$$h_i = \frac{stdev\left(\bigcup_1^j r_{w,d}^H\right)}{stdev\left(\bigcup_i r_{w,d}^H\right)}, \text{ for each } 1 \le i \le j \tag{4.8}$$

All residuals in group i are multiplied by h_i.

$$r_{w,d}^{iH} = \frac{q(w,d) - m_{w,d}}{\sqrt{m_{w,d}}} \times f_{w,d}^H \times h^i \tag{4.9}$$

Now all groups have the same standard deviation and we can sample with replacement from among all $r_{w,d}^{iH}$. The original distribution of residuals has been altered, but this can be remedied. When the adjusted residuals are resampled, the residual is divided by the hetero-adjustment factor, h_i, that applies to the development year of the incremental loss, as shown in (4.10).

$$q^{i^*}(w,d) = \frac{r^*}{h^i} \times \sqrt{m_{w,d}} + m_{w,d}. \tag{4.10}$$

By doing this, the heteroscedastic variances we observed in the data are replicated when the sample triangles are created, but we are able to freely resample with replacement from the entire pool of heteroscedasticity adjusted residuals. Also note that these factors are new parameters so it will affect the degrees of freedom, which impacts the scale parameter (3.17) and the degrees of freedom adjustment factor (3.19).[37] Finally, the hetero-adjustment factors should also be used to adjust the variance by development period when simulating the future process variance.

The third option is to modify the formula for the scale parameter (3.17) so that we have a different scale parameter for each hetero group, as illustrated in (4.11) and (4.12).[38] In (4.12) n_i is the number of residuals in each hetero group.

$$\phi = \frac{\sum r_{w,d}^2}{N-p} = \frac{N}{N-p} \times \frac{\sum r_{w,d}^2}{N} = \frac{\sum\left(\sqrt{\frac{N}{N-p}} \times r_{w,d}\right)^2}{N} \tag{4.11}$$

$$\phi_i = \frac{\sum_{i=1}^{n_i}\left(\sqrt{\frac{N}{N-p}} \times r_{w,d}\right)^2}{n_i} \tag{4.12}$$

For this option, the different scale parameters also amount to new parameters so the degrees of freedom adjustment factor would likewise be impacted. In this case, the scale parameters adjust the future process variance, but we also need to calculate parameters to adjust the residuals as shown in (4.13). These hetero-adjustment factors, h_i, can also be used to adjust the residuals in (4.9) and used in calculating the resampled loss in (4.10), similar to the second option.

$$h_i = \frac{\sqrt{\phi}}{\sqrt{\phi_i}} \tag{4.13}$$

While the hetero-adjustment factors in (4.13) are a bit more theoretically sound, in practice the factors in (4.8) are likely to be very close so the differences are not likely to have much impact. Both of these options are illustrated in the "Bootstrap Models.xlsm" file.

Of course no matter which formula is used, care needs to be exercised as hetero groups are used toward the tail of the triangle where fewer and fewer observations stretch the credibility of the resulting factors.[39] Finally, while use of the GLM bootstrap should reduce the need for hetero factors, the same three options could also be used for that model too.

4.7. Heteroecthesious Data

The basic ODP bootstrap model requires both a symmetrical shape (e.g., annual by annual, quarterly by quarterly, etc. triangles) and homoecthesious data (i.e., similar

[37] Some authors have suggested adding a factor for each development period to insure homoscedasticity. However, this adds many more parameters to a model that can already suffer from the criticism of over-parameterization. Thus, a balance between the need for hetero parameters and parsimony is appropriate. This will be discussed in more detail in Section 5.

[38] For a more detailed development of this third option see England and Verrall (2006). In particular, see Appendix A.1 on pages 266–268.

[39] In the discussion of diagnostics in Section 5 it will be noted that the use of the AIC and BIC statistics will effectively reflect the credibility of the development periods.

exposures).[40] As discussed above, using an L-year weighted average in the ODP bootstrap model or adjusting to a trapezoid shape allow us to "relax" the requirement of a symmetrical shape. Other non-symmetrical shapes (e.g., annual × quarterly data) can also be modeled with either the ODP bootstrap or GLM bootstrap, but they will not be discussed in detail in this monograph.

Most often, the actuary will encounter heteroecthesious data (i.e., incomplete or uneven exposures) at interim evaluation dates, with the two most common data triangles being either a partial first development period or a partial last calendar period. For example, with annual data evaluated as of June 30, partial first development period data would have all development periods ending at 6, 18, 30, etc. months, while partial last calendar period data would have development periods as of 12, 24, 36, etc. months for all of the data in the triangle except the last diagonal, which would have development periods as of 6, 18, 30, etc. months. In either case, not all of the data in the triangle has full annual exposures—i.e., it is heteroecthesious data.

4.7.1. Partial First Development Period Data

For partial first development period data, the first development column has a different exposure period than the rest of the columns (e.g., in the earlier example the first column has six months of development exposure while the rest have 12). In a deterministic analysis this is not a problem as the development factors will reflect the change in exposure. For parameterizing an ODP bootstrap model, it also turns out to be a moot issue, since the Pearson residuals use the square root of the fitted value to make them all "exposure independent."

The only adjustment for this type of heteroecthesious data is the projection of future incremental values. In a deterministic analysis, the most recent accident year needs to be adjusted to remove exposures beyond the evaluation date. For example, continuing the previous example the development periods at 18 months and later are all for an entire year of exposure whereas the six month column is only for six months of exposure. Thus, the 6–18 month development factor will effectively extrapolate the first six months of exposure in the latest accident year to a full accident year's exposure. Accordingly, it is common practice to reduce the projected future payments by half to remove the exposure from June 30 to December 31.[41]

The simulation process for the ODP bootstrap model can be adjusted similarly to the way a deterministic analysis would be adjusted. After the development factors from each sample triangle are used to project the future incremental values the last accident year's values can be reduced (in the previous example by 50%) to remove the future exposure and then process variance can be simulated as before. Alternatively, the future incremental values can be reduced after the process variance step.

---

[40] To the author's knowledge, the terms *homoecthesious* and *heteroecthesious* are new. They are a combination of the Greek *homos* (or όμός) meaning the same or *hetero* (or έτερο) meaning different and the Greek *ekthesë* (or έκθεση) meaning exposure.

[41] Reduction by half is actually an approximation since we would also want to account for the differences in development between the first and second half years.

4.7.2. Partial Last Calendar Period Data

For partial last calendar period data, most of the data in the triangle has annual exposures and annual development periods, except for the last diagonal which, continuing our example, only has a 6-month development period. For a deterministic analysis, it is common to exclude the last diagonal when calculating average development factors, then interpolate those factors to project the future values. Similarly to the adjustments for partial first development period data, we can adjust the calculations and steps in the ODP bootstrap model. Instead of ignoring the last diagonal during the parameterization of the model, an alternative is to adjust or annualize the exposures in the last diagonal to make them consistent with the rest of the triangle. The fitted triangle can be calculated from this annualized triangle to obtain residuals.

During the ODP bootstrap simulation process, development factors can be calculated from the fully annualized sample triangles and interpolated. Then, the last diagonal from the sample triangle can be adjusted to de-annualize the incremental values in the last diagonal—i.e., reversing the annualization of the original last diagonal. The new cumulative values can be multiplied by the interpolated development factors to project future values. Again, the future incremental values for the last accident year must be reduced (in the previous example by 50%) to remove the future exposure.[42]

4.8. Exposure Adjustment

Another common issue in real data is exposures that have changed dramatically over the years. For example, in a line of business that has experienced rapid growth or is being run off. If the earned exposures exist for this data, then a useful option for the ODP bootstrap model is to divide all of the claim data by the exposures for each accident year—i.e., effectively using pure premium development instead of total loss development. This may improve the fit of the model to the data.

During the ODP bootstrap simulation process, all of the calculations would be done using the exposure-adjusted data and only after the process variance step has been completed would you multiply the results by the exposures by year to restate them in terms of total values again.

When adjusting the GLM bootstrap for exposure, the model is fitted to exposure adjusted losses, similar to the ODP bootstrap model using exposure. However, under the GLM, the fit to the exposure adjusted losses are also exposure-weighted. That is, exposure adjusted losses with higher exposure are assumed to have lower variance. For more details, see Anderson et al. (2007).

For the GLM bootstrap, exposure adjustment could allow fewer accident year parameter(s) to be used.

4.9. Tail Factors

One of the most common data issues is that claim development is not complete within the loss triangle and tail factors are commonly used to extrapolate beyond the end of

[42] These heteroecthesious data issues are not illustrated in the "Bootstrap Models.xlsm" file.

the data triangle. There are many common methods for calculating tail factors and a useful reference in this regard is the CAS Tail Factor Working Party Report (2013). Tail factors can be added to the ODP bootstrap algorithm and converted from deterministic to stochastic by assuming that the tail factor parameter follows a distribution. Once this is added, other considerations such as process variance, hetero-adjustment factors, etc. can all be extended to include the tail factors.

A key ingredient for all of these considerations is to verify that the simulations in the tail are reasonable. For example, the tail factor itself represents the accumulation of incremental factors (i.e., an age-to-ultimate factor) and using just a single factor may not produce appropriate incremental results so the "extrapolation" of "incremental tail factors" may be more appropriate. In the "Bootstrap Models.xlsm" file, the tail factors can be extrapolated for up to 5 years so that one possibility for how these concepts can be implemented is included in the companion files.

A rough rule of thumb for the tail factor standard deviation is 50% or less of the tail factor minus one (assuming the tail factor is greater than one). However, this should be compared to the standard deviations of the age-to-age factors leading up to the tail in both the actual data triangle and in the simulated results.

As noted at the end of Section 3.4, for the GLM bootstrap model the last development parameter can continue to apply past the end of the data triangle until the trend results in no further claim activity, thus indirectly creating a tail factor. In addition to the last development parameter, the last calendar period parameter would also extend past the end of the tail until the combination of the two trends resulted in no further claim activity.

4.10. Fitting a Distribution to ODP Bootstrap Residuals

Because the number of data points used to parameterize the ODP bootstrap model are limited (in the case of a 10×10 triangle to 55 data points or 53 residuals), it is hard to determine whether the most extreme observation is a one-in-100 or a one-in-1,000 event (or simply, in this example, a one-in-53 event). Of course, the nature of the extreme observations in the data will also affect the level of extreme simulations in the results. Judgment is involved here, but the modeler will either need to be satisfied with the level of extreme simulations in the results or modify the ODP bootstrap algorithm.

One way to overcome a lack of extreme residuals for the ODP bootstrap model would be to fit a distribution to the residuals and sample from the distribution instead of from the residuals themselves (e.g., use a normal distribution if the residuals are found to be normally distributed). This option is beyond the scope of the companion Excel files, but this could be referred to as parametric bootstrapping of the ODP bootstrap model. Note however, that as there are a wide variety of other types of models that can be bootstrapped, either with or without residuals, parametric bootstrapping can be done in other ways.

5. Diagnostics

The quality of any model depends on the quality of the underlying assumptions. When a model fails to "fit" the data, it cannot produce a good estimate of the distribution of possible outcomes.[43] However, a balance must be considered for parsimony of parameters and the goodness-of-fit. Over-parameterization may cause the model to be less predictive of future losses. On the other hand, no model will perfectly "fit" the data, so the best you can hope for with any model is that it reasonably represents the data and your understanding of the processes that impact the data. Therefore, diagnostically evaluating the assumptions underlying a model is important for evaluating whether it will produce reasonable results or not and whether it should stay in your selected group of reasonable models which could receive some weight.

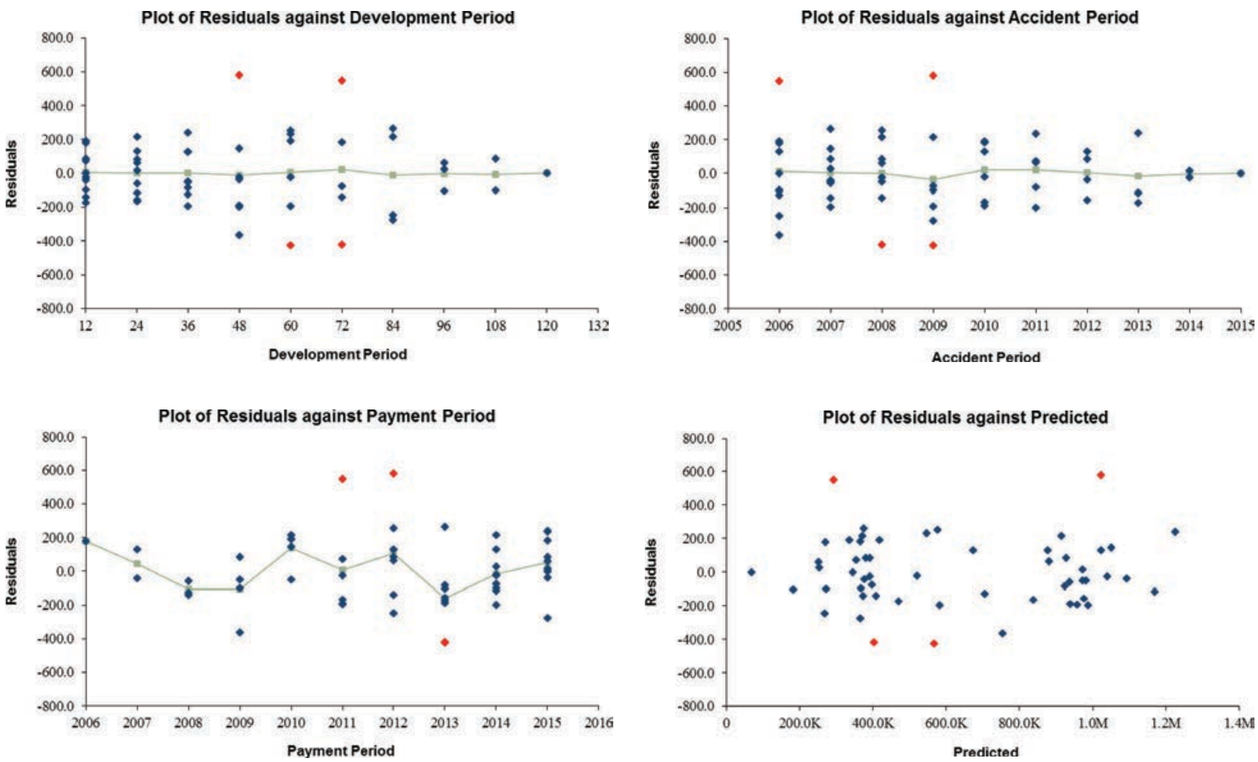
The CAS Working Party, in the third section of their report on quantifying variability in reserve estimates (2005), identified 20 criteria or diagnostic tools for gauging the quality of a stochastic model. The Working Party also noted that, in trying to determine the optimal fit of a model, or indeed an optimal model, no single diagnostic tool or group of tools can be considered definitive. Depending on the statistical features found in the data, a variety of diagnostic tools are necessary to best judge the quality of the model assumptions and to adjust the parameters of the model. This monograph will discuss some of these tools in detail as they relate to the ODP bootstrap and the GLM bootstrap models.

The key diagnostic tests are designed for three purposes: to test various assumptions in the model, to gauge the quality of the model fit to the data, and/or to help guide the adjustment of model parameters. Some tests are relative in nature, enabling results from one set of model parameters to be compared to those of another, for a specific model, allowing a modeler to improve the fit of the model. For the most part, however, the tests can't be used to compare different models. The objective, consistent with the goals of a deterministic analysis, is ***not*** to find the one best model, but rather a set of reasonable models.

Some diagnostic measures include statistical tests, providing a pass/fail determination for some aspects of the model assumptions. This can be useful even though a "fail" does not necessarily invalidate an entire model; it only points to areas where improvements can be made to the model or its parameterization. The goal is to find the sets of models

[43] While the examples are different, significant portions of Sections 5 and 6 are based on Milliman (2014) and IAA (2010).

Figure 5.1. Residual Graphs Prior to Heteroscedasticity Adjustment



and parameters that will yield the most realistic, most consistent simulations, based on statistical features found in the data.

To illustrate some of the diagnostic tests for the ODP bootstrap model we will consider data from England and Verrall (1999).[44]

5.1. Residual Graphs

The ODP bootstrap model does not require a specific type of distribution for the residuals, but they are assumed to be independent and identically distributed. Because residuals will be sampled with replacement during the simulations, this requirement is important and thus it is necessary to test this assumption. Graphing residuals is a good way to do this.

Going clock-wise, and starting from the lower-left-hand corner, the graphs in Figure 5.1 show the residuals (blue and red dots[45]) by calendar period, development period, and accident period and against the fitted incremental loss (in the lower-right-hand corner). In addition, the graphs include a trend line (in green) that highlights the averages for each period.

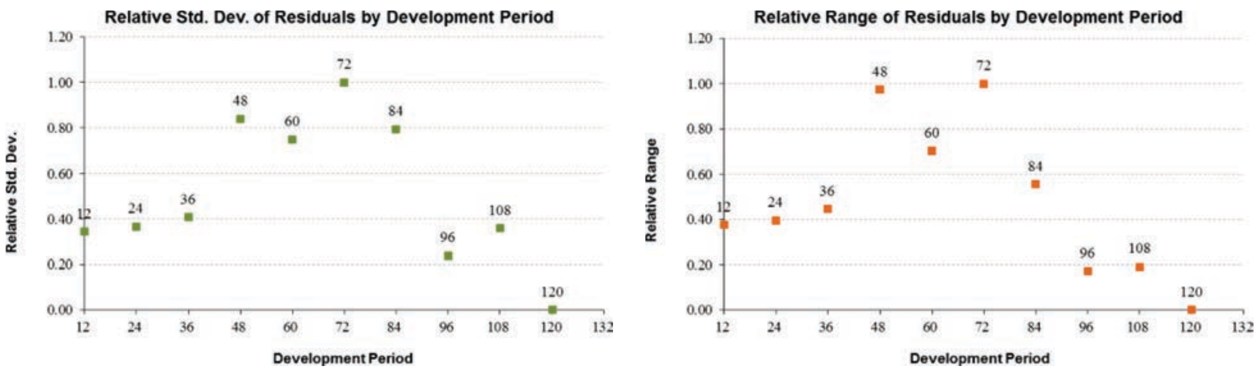
At first glance, the residuals in the graphs appear reasonably random, indicating the model is likely a good fit of the data. But a closer look may also reveal potential features in the data that may indicate ways to improve the model fit.

---

[44] The data triangle was originally used by Taylor and Ashe (1983) and has been used by other authors. This data is included in the "Bootstrap Models.xlsm" file.

[45] In the graphs that follow, the red dots are outliers as identified in Figure 5.7.

Figure 5.2. Residual Relativities Prior to Heteroscedasticity Adjustment



The graphs in Figure 5.1 do not appear to indicate issues with un-modeled trends by accident period or development period (that is, the green "average" lines appear flat at zero). That's because the ODP bootstrap specifies a parameter for every accident and development period. The development-period graph does, however, reveal a potential heteroscedasticity issue associated with the data—i.e., different variances. Note how the upper left graph appears to show a variance of the residuals in the first three periods that differs from those of the middle four or last two periods.

Adjustments for heteroscedasticity can be made with the "Bootstrap Models.xlsm" file, which enables us to recognize groups of development periods and then adjust the residuals to a common standard deviation value, as described in Section 4.6. As an aid to visualizing how to group the development periods into "hetero" groups, graphs of the standard deviation and range relativities can be developed. Figure 5.2 represents pre-adjusted relativities for the residuals shown in Figure 5.1 (i.e., prior to adjustment for factors calculated using either formulas 4.8 or 4.13 and 4.9).

The relativities illustrated in Figure 5.2 help to clarify the changing variability. However, further testing will be required to assess the optimal groups, which can be performed using the other diagnostic tests noted below.

The residual plots in Figure 5.3 originate from the same data model after adjusting for heteroscedasticity using the third option described in Section 4.6 (i.e., using formulas 4.13 and 4.9). The "hetero" groups chosen are for the first three, middle four, and last two development periods, respectively. Determining whether this adjustment has improved the model will require review of other diagnostic tests.

Comparing the residual plots in Figures 5.1 and 5.3 shows that the residuals now appear to exhibit the same standard deviation, or homoscedasticity. More consistent relativities may also be seen in a comparison of the residual relativities in Figures 5.2 and 5.4.

5.2. Normality Test

The ODP bootstrap model does not depend on the residuals being normally distributed, but even so, comparing residuals against a normal distribution remains a useful test, enabling comparison of parameter sets and gauging skewness of the residuals. This test uses both graphs and calculated test values. Figure 5.5 is based on the data used earlier, before and after the adjustment for heteroscedasticity.

Figure 5.3. Residual Graphs After Heteroscedasticity Adjustment

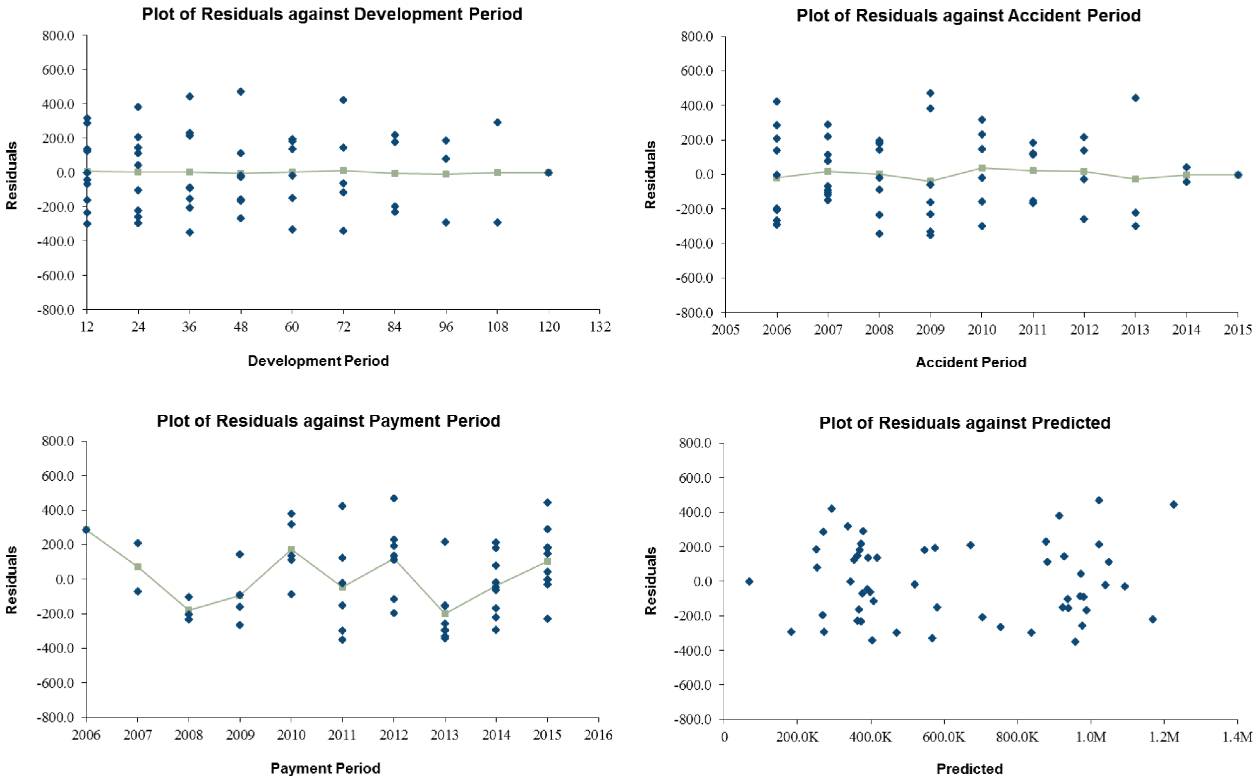


Figure 5.4. Residual Relativities After Heteroscedasticity Adjustment

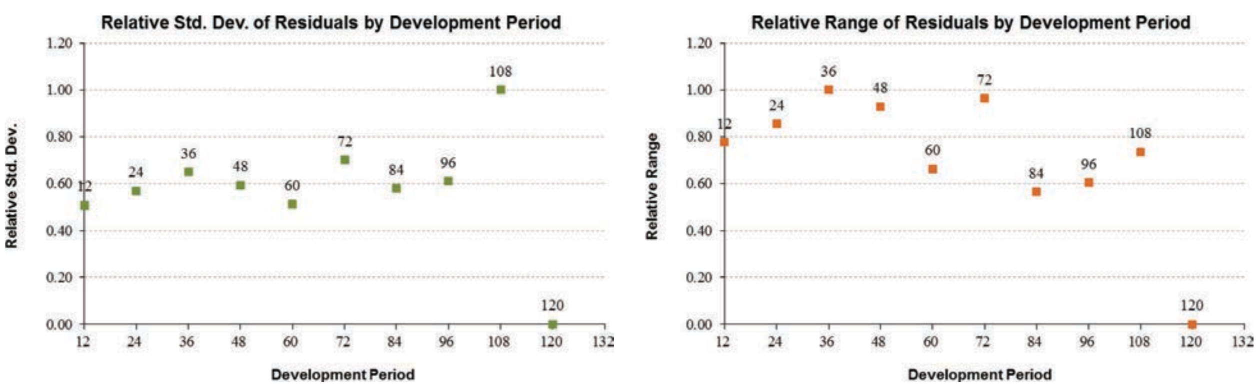
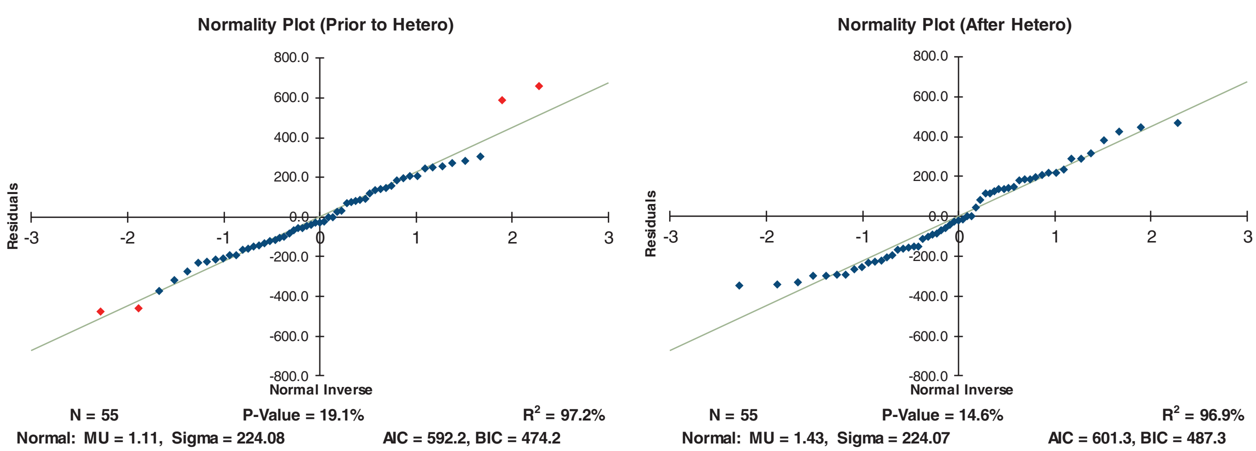


Figure 5.5. Normality Plots Prior to and After Heteroscedasticity Adjustment



Even before the heteroscedasticity adjustment, the residual plots appear close to normally distributed, with the data points tightly distributed around the diagonal line. The *P*-value, a statistical pass-fail test for normality, came in at 19.1%, which exceeds the value generally considered a "passing" score of the normality test, which is greater than 5.0%.[46] The graphs in Figure 5.5 also show *N* (the number of data points) and the R^2 test. After the hetero adjustment, the *P*-value and R^2 get slightly worse, which indicates that the heteroscedasticity adjustment has not improved the results of the diagnostic tests.

While the *P*-value and R^2 tests assess the goodness of fit of the model to the data, they do not penalize for added parameters. Adding more parameters will almost always improve the fit of the model to the data, but the goal is to have a good fit with as few parameters as possible. Two other tests, the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC), address this limitation, using the difference between each residual and its normal counterpart from the normality plot to calculate the Residual Sum Squared (RSS) and include a penalty for additional parameters, as shown in (5.1) and (5.2), respectively.[47]

$$\text{AIC} = 2 \times p + n \times \left[\ln\left(\frac{2 \times \pi \times \text{RSS}}{n} \right) + 1 \right] \tag{5.1}$$

$$\text{BIC} = n \times \ln\left(\frac{\text{RSS}}{n} \right) + p \times \ln(n) \tag{5.2}$$

A smaller value for the AIC and BIC tests indicate residuals that fit a normal distribution more closely, and this improvement in fit overcomes the penalty of adding a parameter.

In our example, with some trial and error, a better "hetero" grouping was found with the diagnostic results shown in Figure 5.6.[48] For the new "hetero" groups, all of the statistical tests improved significantly.

While it might be tempting to add a hetero group for each development column to improve normality, in general normality can be improved with far fewer groups which also helps keep the model from being over-parameterized. As an example, if we use 9 hetero groups for the Taylor and Ashe (1983) data the *P*-value is 14.3%, which is worse than no groups and only slightly better than the original 3 groups, but the AIC and BIC increase significantly.

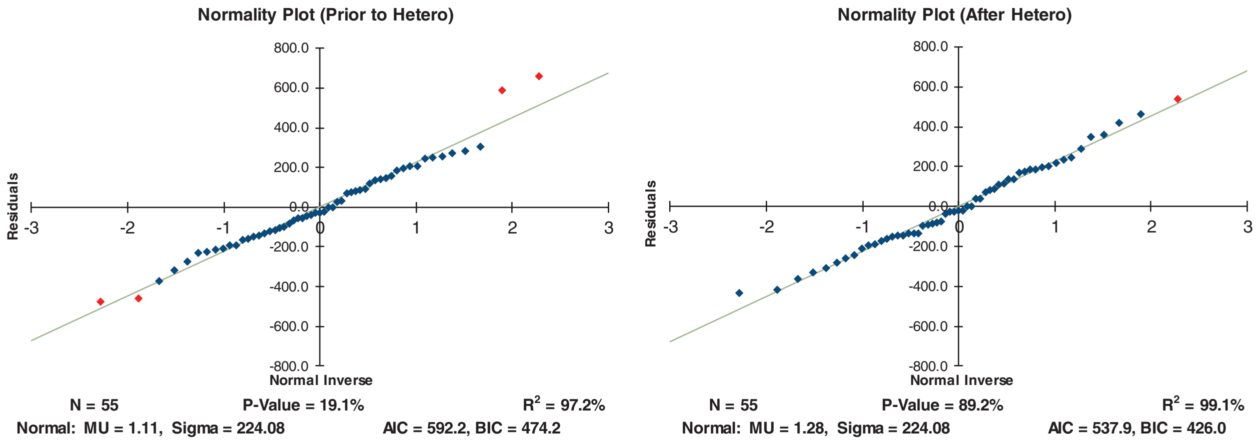
---

[46] Remember that this doesn't indicate whether the ODP bootstrap model itself passes or fails—the ODP bootstrap model doesn't require the residuals to be normally distributed. While not included in the "Bootstrap Models.xlsm" file, as discussed in Section 4.10 it could be used to determine whether to switch to a parametric bootstrap process using a normal distribution.

[47] There are different versions of the AIC and BIC formula from various authors and sources, but the general idea of each version is consistent. Other similar formulas could also be used.

[48] In the "Bootstrap Models.xlsm" file the Taylor & Ashe data was entered as both paid and incurred. The first set of "hetero" groups are illustrated for the "paid" data and the second set of "hetero" groups are illustrated for the "incurred" data. The "best" groups were found using the optimization tool shown in the "Groups" sheet.

Figure 5.6. Normality Plots Prior to and After Heteroscedasticity Adjustment



5.3. Outliers

Identifying outliers in the data provides another useful test in determining model fit. Outliers can be represented graphically in a box-whisker plot, which shows the inter-quartile range (the 25th to 75th percentiles) and the median (50th percentile) of the residuals—the so-called box. The whiskers then extend to the largest values within three times this inter-quartile range.[49] Values beyond the whiskers may generally be considered outliers and are identified individually with a point.

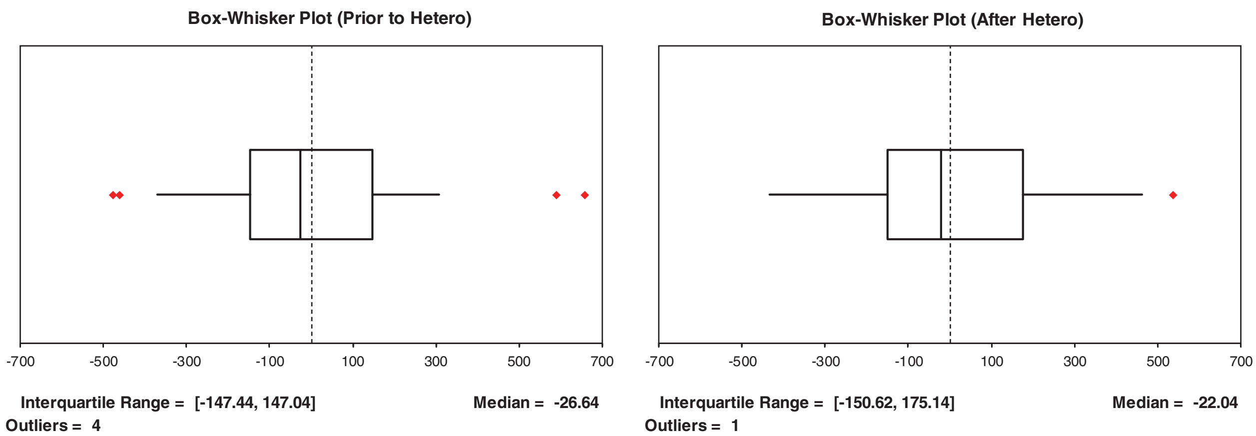
Figure 5.7 shows an example of the residuals for the second set of "hetero" groups (Figure 5.6). A pre-hetero adjustment plot returns four outliers (red dots) in the data model, corresponding to the two highest and two lowest values in the previous graphs in Figures 5.1, 5.3, 5.5, and 5.6.

Even after the hetero adjustment, the residuals still appear to contain one outlier. Now comes a very delicate and often tricky matter of actuarial judgment. If the data in those cells genuinely represent events that cannot be expected to happen again, the outlier(s) may be removed from the model (by giving it/them zero weight). But extreme caution should be taken even when the removal of outliers seems warranted. The possibility always remains that apparent outliers may actually represent realistic extreme values, which, of course, are critically important to include as part of any sound analysis.

Additionally, when residuals are not normally distributed a significant number of outliers tend to result, which may only be an artifact of the distributional shape of the residuals. In this case it is preferable to let these stand in order to enable the simulation process to replicate this shape. Finally, a significant number of residuals can also mean the underlying model is not a good fit to the data so other models should be used (see Section 4.5 for a discussion) or this model given less weight (see Section 6).

[49] Various authors and textbooks use widths for the whiskers which tend to span from 1.5 to 3 times the inter-quartile range. Changing the multiplier will therefore make the box-whisker plot more or less sensitive to outliers. It is also possible to illustrate "mild" outliers with a multiplier of 1.5 and the more "extreme" outliers with a multiplier of 3 using different colors and/or symbols in the graphs. Of course the actual multipliers can be adjusted based on personal preference.

Figure 5.7. Box-Whisker Plots Prior to and After Heteroscedasticity Adjustment



While the three diagnostic tests shown above demonstrate techniques commonly used with most types of models, they are not the only tests available.[50] Next, we'll take a look at the flexibility of the GLM bootstrap and some of the diagnostic elements of the simulation results. For a more extensive list of other tests available, see the report, CAS Working Party on Quantifying Variability in Reserve Estimates (2005).

5.4. Parameter Adjustment

As noted in Section 5.1 the relatively straight average lines in the development and accident period graphs are a reflection of having a parameter for every accident and development period. In most instances, this is also a strong indication that the model may be over-parameterized. Using the "GLM Bootstrap" model in the "Bootstrap Models.xlsm" file we can illustrate the power of removing some of the parameters.

Starting with a "basic" model which includes only one parameter for accident, development and calendar periods (i.e., only one α, β and γ parameter), and adding vertical brown bars to signify a parameter and vertical red lines to signify no parameter (i.e., parameter of zero), the residual graphs for the "GLM Bootstrap" model are shown in Figure 5.8.

The brown bars in the basic model residual graphs represent the parameters and statistics shown in Table 5.1.

Now for this "basic" model the green average lines show trends in the underlying data that are not yet captured by the model as well as a parameter for calendar year trend that is not significant. For example, the overall development period trend parameter is −11%, but the underlying data shows a positive trend for the first 2 or 3 periods followed by a stronger negative trend for the remaining development periods. Another way to see that this basic model does not yet provide a good fit to the underlying data is to compare the implied development pattern with that of the ODP bootstrap model, as shown in Figure 5.9.

---

[50] For example, see Venter (1998).

Figure 5.8. Residual Graphs for "Basic" GLM Bootstrap Model

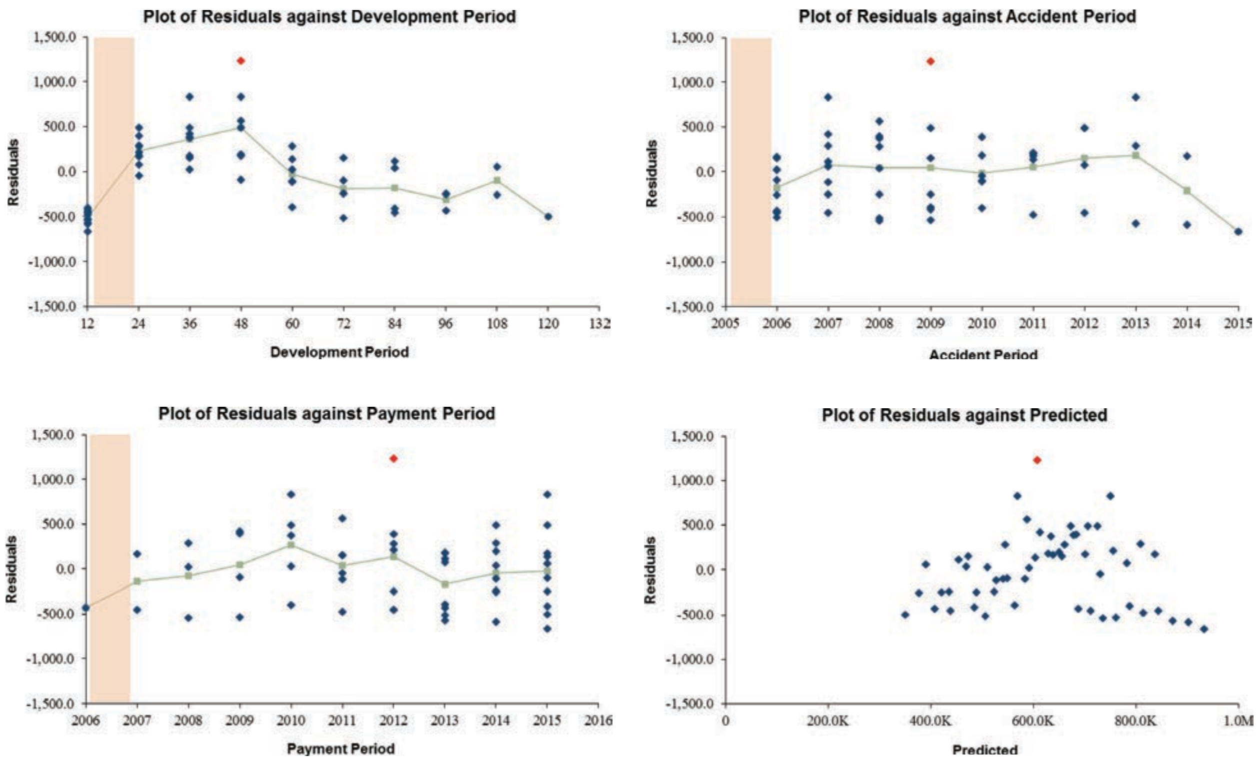
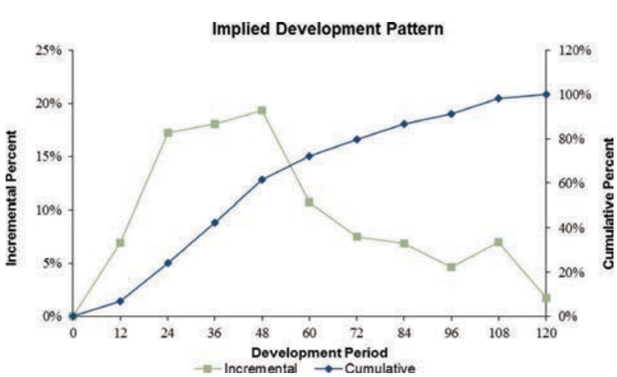


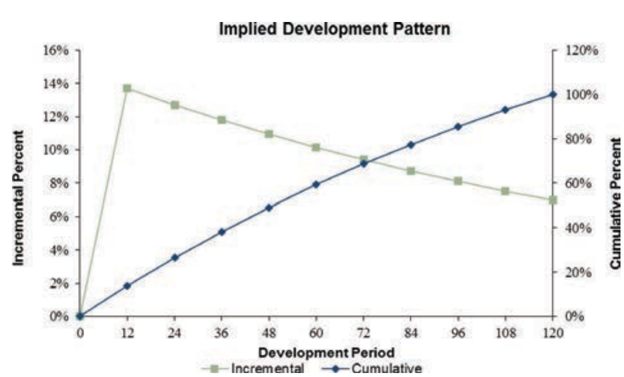
Table 5.1. Parameters and Statistics for "Basic" GLM Bootstrap Model

Parm	Value	Exp(Value)	*t*-Stat	Periods
α_1	13.44	686,938	73.92	Accident Years 2006–2015
β_1	(0.11)		(3.19)	Development Periods 12–132
γ_1	0.03		1.08	Calendar Years 2006–2015

Figure 5.9. Implied Development Patterns

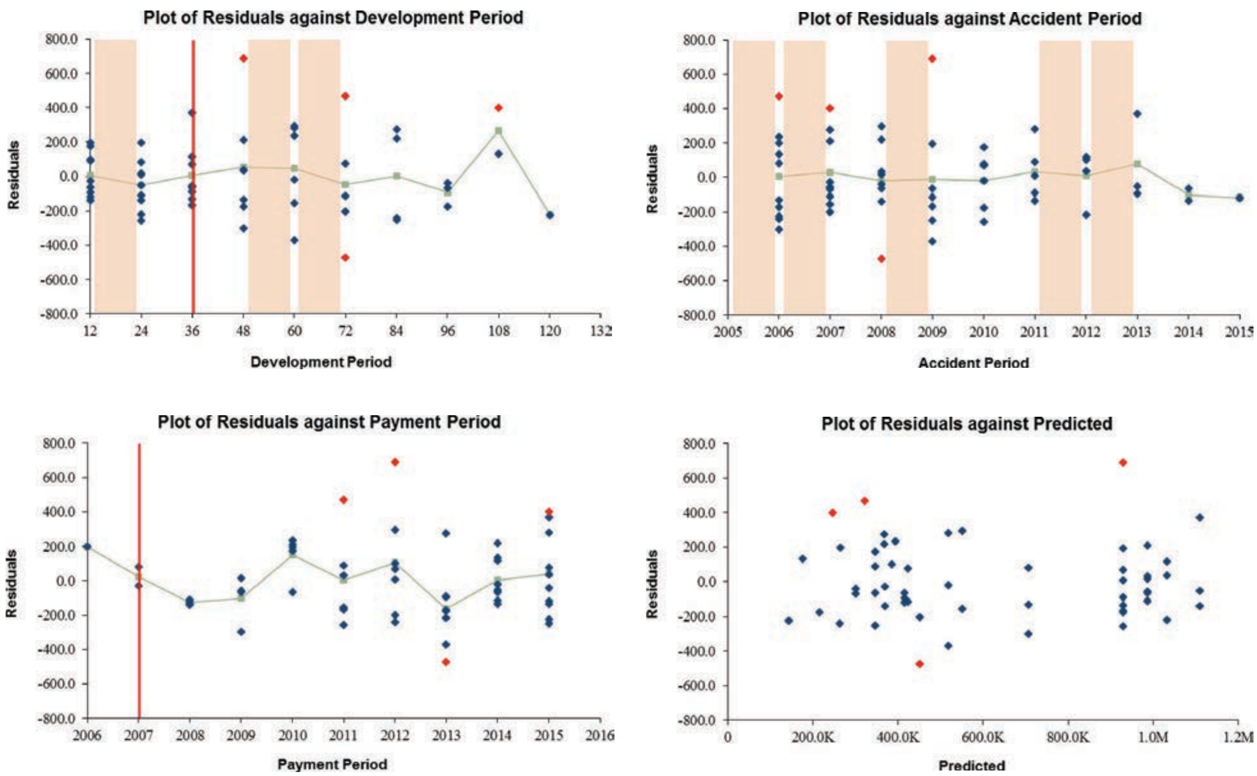


ODP Bootstrap Model



"Basic" GLM Bootstrap Model

Figure 5.10. Residual Graphs for GLM Bootstrap Model



With a little trial and error we can find a reasonably good fit to the data using only five accident, three development and no calendar parameters as shown in Figure 5.10.[51]

In addition to checking the remaining trends in the data with the green average lines, *t*-statistics for each new parameter can be checked to make sure each parameter is statistically significant.[52] The final parameters and statistics for the GLM Bootstrap model are shown in Table 5.2.

Using the "optimal" set of "hetero" groups we can also check the normality graphs and statistics in Figure 5.11 and outliers in Figure 5.12.[53] Comparing the statistics to the ODP bootstrap values shown in Figures 5.6 and 5.7, most values improved while some did not, yet the GLM Bootstrap model is far more parsimonious.

---

[51] In the "Bootstrap Models.xlsm" file the optimization tool in the "GLM" sheet can be used to help find a good fit for the parameters of the GLM bootstrap. The algorithm for this tool starts with the ODP bootstrap parameters and then removes the least significant parameters until only significant parameters remain. Then, if there are few enough Alpha and Beta parameters, the Gamma parameters are added and removed if not significant. The tool does not test to see if a parameter should be zero, so some improvements can sometimes occur by forcing parameters to equal zero (e.g., compare the parameters from Figure 5.10 to the parameters in the optimization tool). Finally, it is possible to have a better model fit (i.e., lower AIC and/or BIC) with more parameters even though some of the parameters may not be significant, so judgment is still appropriate for selection of parameters.

[52] The *t*-statistic indicates that a parameter is statistically significant if the absolute value is greater than 2.

[53] When using the GLM bootstrap, any selected outliers and hetero groups used for the ODP bootstrap should be reset and then re-evaluated as they will likely be different for the GLM bootstrap. For the "after hetero" portions of Figures 5.11 and 5.12 the optimization tool in the "Groups" sheet was used.

Table 5.2. Parameters and Statistics for GLM Bootstrap Model

Parm	Value	Exp(Value)	*t*-Stat	Periods
α_1	12.48	264,036	79.26	Accident Year 2006
α_2	12.82	368,718	2.48	Accident Years 2007–2008
α_3	12.76	347,009	2.11	Accident Years 2009–2011
α_4	12.86	385,644	2.35	Accident Year 2012
α_5	12.93	414,414	3.29	Accident Years 2013–2015
β_1	0.98		7.88	Development Periods 12–24
—	0.00			Development Periods 24–48
β_2	(0.58)		(4.88)	Development Periods 48–60
β_3	(0.20)		(3.29)	Development Periods 60–132
—	0.00			Calendar Year 2006–2015

Figure 5.11. Normality Plots for GLM Bootstrap Model

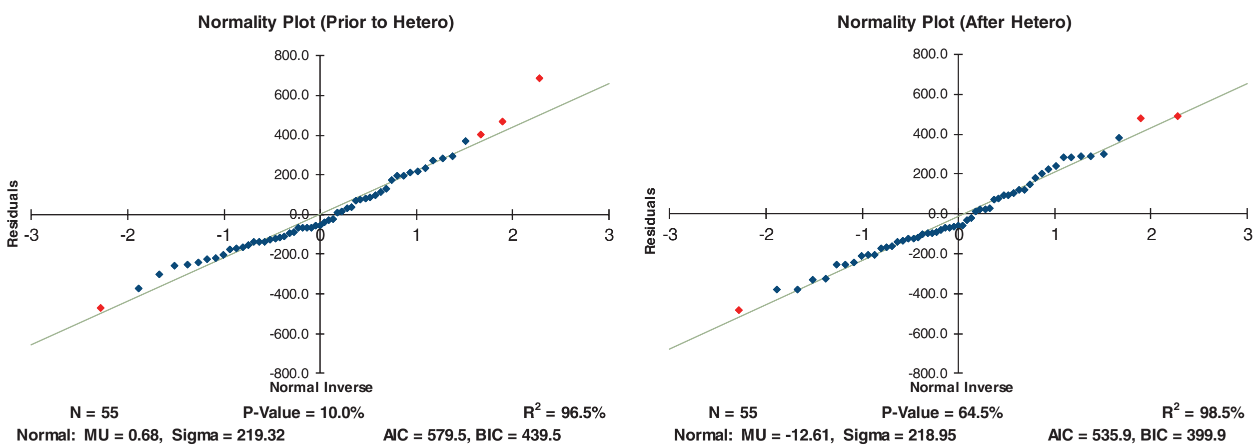


Figure 5.12. Box-Whisker Plots for "GLM Bootstrap" Model

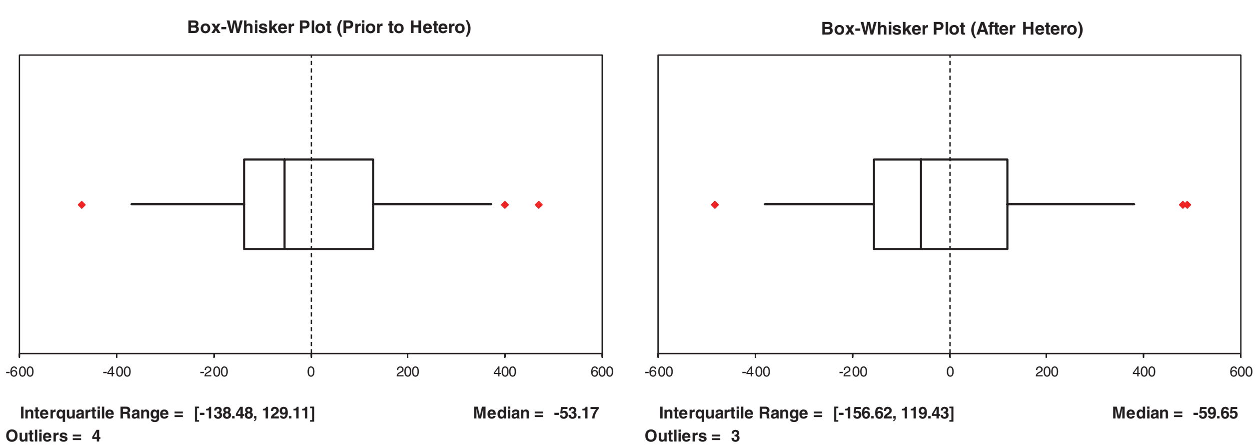
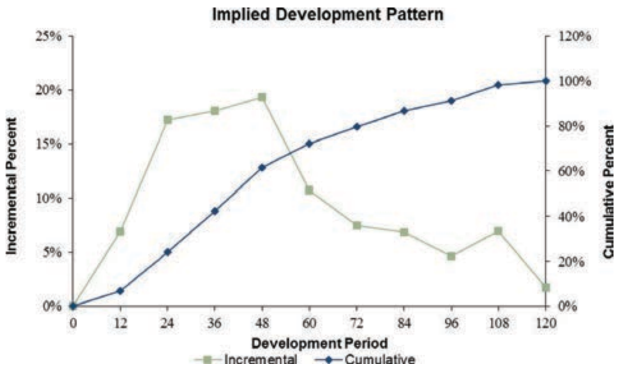
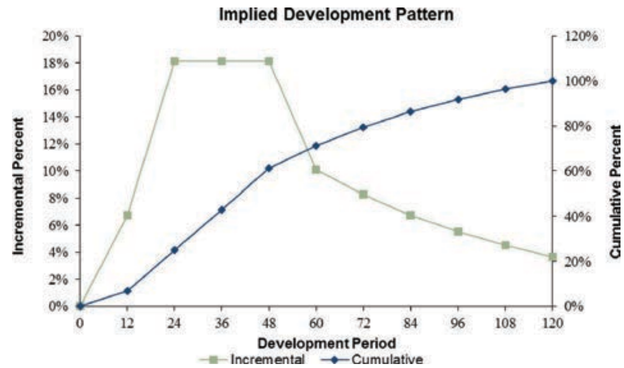


Figure 5.13. Implied Development Patterns



ODP Bootstrap Model



"Basic" GLM Bootstrap Model

As one final check on the trends in this GLM bootstrap model, we can compare a graph of the implied development patterns with the patterns from the chain ladder in the ODP bootstrap model, as shown in Figure 5.13. Because the chain ladder model used a parameter for each development period the implied development pattern can appear a bit jagged, which is why it is often "smoothed" out in practice by selecting development factors. Interestingly, the GLM bootstrap model looks quite similar, yet with much smoother trends in the development patterns. As noted earlier, the last GLM bootstrap development (and calendar trend) parameter can be assumed to extend until the projected model incremental values equal zero which could then be compared to tail factors used in the ODP bootstrap model.[54]

5.5. Model Results

Once the parameter diagnostics have been reviewed, simulations should be run for each model. These simulation results provide an additional diagnostic tool to aid in evaluation of the model, as described in Section 3 of CAS Working Party (2005). As an example, we will review the results for the Taylor and Ashe (1983) data using the ODP bootstrap model. The estimated-unpaid results shown in Figure 5.14 were simulated using 10,000 iterations with the hetero adjustments from Figure 5.6.

5.5.1. Estimated-Unpaid Results

It's recommended to start a diagnostic review of the estimated unpaid results with the standard error (standard deviation) and coefficient of variation (standard error divided by the mean), shown in Figure 5.14. Keep in mind that the standard error should increase when moving from the oldest years to the most recent years, as the standard errors (value scale) should follow the magnitude of the mean of unpaid estimates. In Figure 5.14, the standard errors conform to this pattern. At the same time, the standard error for the total of all years should be larger than any individual year.

[54] Results for the GLM bootstrap model, as illustrated in Figures 5.9 through 5.12, are shown in Appendix E, although no extrapolation was included to be consistent with the ODP bootstrap results.

Figure 5.14. Estimated Unpaid Model Results

Taylor & Ashe Data
Accident Year Unpaid
Paid Chain Ladder Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-	-	-	-	-	-	-	-
2007	94,649	96,571	102.0%	(119,298)	541,054	71,176	147,232	278,360	374,056
2008	473,619	199,302	42.1%	(25,494)	1,217,544	454,644	590,676	830,916	1,018,835
2009	714,763	250,044	35.0%	140,156	1,642,391	684,461	882,486	1,146,017	1,396,652
2010	981,305	271,726	27.7%	324,024	2,062,359	951,467	1,148,872	1,475,857	1,731,112
2011	1,414,007	364,527	25.8%	468,645	2,829,838	1,392,288	1,642,974	2,059,339	2,349,855
2012	2,173,552	489,442	22.5%	806,008	4,293,160	2,142,306	2,489,525	3,033,205	3,345,364
2013	3,969,749	768,637	19.4%	1,655,462	6,369,285	3,913,503	4,501,100	5,307,862	5,989,765
2014	4,317,349	887,688	20.6%	1,874,779	7,677,306	4,260,113	4,898,209	5,844,560	6,516,905
2015	4,703,420	2,176,343	46.3%	445,056	13,859,166	4,493,023	6,127,676	8,500,947	10,529,157
Totals	18,842,414	2,902,735	15.4%	11,312,275	29,464,222	18,594,140	20,734,478	23,885,153	26,388,103

Also, the coefficients of variation should generally decrease when moving from the oldest years to the more recent years and the coefficient of variation for all years combined should be less than for any individual year. With the exception of the 2014 and 2015 accident years, the coefficients of variation in Figure 5.14 seem to also conform, although some random fluctuations may be seen.

The main reason for the decrease in the coefficient of variation has to do with the independence in the incremental claim-payment stream. Because the oldest accident year typically has only a few incremental payments remaining, or even just one, the variability is nearly all reflected in the coefficient. For more current accident years, random variations in the future incremental payment stream may tend to offset one another, thereby reducing the variability of the total unpaid loss.[55]

While the coefficients of variation should go down, they could also start to rise again in the most recent years, as seen in Figure 5.14 for 2014 and 2015. Such reversals are from a couple of issues:

- With an increasing number of parameters used in the model, the parameter uncertainty tends to increase when moving from the oldest years to the more recent years. In the most recent years, parameter uncertainty can grow to overpower process uncertainty, which may cause the coefficient of variation to start rising again. At a minimum, increasing parameter uncertainty will slow the rate of decrease in the coefficient of variation.
- The model may be overestimating the uncertainty in recent accident years if the increase is significant. In that case, another model algorithm (e.g., Bornhuetter-Ferguson or Cape Cod) may need to be used instead of a chain-ladder model.

Keep in mind also that the standard error or coefficient of variation for the total of all accident years will be less than the sum of the standard error or coefficient of variation for the individual years. This is because the model assumes that accident years are independent.

[55] To visualize this reducing Coefficient of Variation, recall that the standard deviation for the total of several independent variables is equal to the square root of the sum of the squares.

Minimum and maximum results are the next diagnostic element in our analysis of the estimated unpaid claims in Figure 5.14, representing the smallest and largest values from all iterations of the simulation. These values will need to be reviewed in order to determine their veracity. If any of them seem implausible, the model assumptions would need to be reviewed. Their effects could materially alter the mean indication. Sometimes implausible extreme iterations are the result of negative incremental values in those "rare" iterations and the limiting incremental value options discussed in Section 4.1 can be used to constrain the model simulation process.

5.5.2. Mean, Standard Deviation and CoV of Incremental Values

The mean, standard deviation and coefficients of variation for every incremental value from the simulation process also provide useful diagnostic results, enabling us to dig deeper into potential coefficient of variation issues that may be found in the estimated unpaid results. Consider, for example, the mean, standard deviation and coefficient of variation results shown in Figures 5.15, 5.16 and 5.17, respectively.

The mean values in Figure 5.15 appear consistent throughout and support the increases in estimated unpaid by accident year that are shown in Figure 5.14. In fact, the future mean values, which lay beyond the stepped diagonal line in Figure 5.15, sum to the results in Figure 5.14. The standard deviation values in Figure 5.16 also

Figure 5.15. Mean of Incremental Values

Taylor & Ashe Data
Accident Year Incremental Values by Development Period
Paid Chain Ladder Model

Accident	Mean Values									
Year	**12**	**24**	**36**	**48**	**60**	**72**	**84**	**96**	**108**	**120+**
2006	278,309	678,678	706,559	769,219	414,449	296,763	266,301	182,021	270,614	66,922
2007	380,244	940,173	979,875	1,076,297	588,887	408,707	372,144	251,228	381,983	94,649
2008	376,488	936,096	971,651	1,038,686	584,856	405,028	367,109	256,617	379,226	94,393
2009	358,750	918,068	955,061	1,023,741	565,152	405,626	367,359	249,479	372,693	92,592
2010	328,119	837,454	881,193	941,139	514,722	373,168	332,243	226,148	339,996	82,918
2011	353,894	879,226	924,325	986,018	540,281	386,069	348,473	234,329	357,224	87,913
2012	386,915	980,382	1,016,136	1,104,983	595,138	436,918	393,002	267,350	389,062	92,083
2013	477,460	1,175,498	1,227,022	1,334,527	739,306	511,050	461,997	320,655	480,476	121,737
2014	396,237	973,510	1,023,124	1,106,316	597,274	431,428	390,159	264,060	404,885	100,103
2015	342,385	875,509	913,011	977,993	539,429	389,906	344,466	230,160	347,729	85,218

Figure 5.16. Standard Deviation of Incremental Values

Taylor & Ashe Data
Accident Year Incremental Values by Development Period
Paid Chain Ladder Model

Accident	Standard Error Values									
Year	**12**	**24**	**36**	**48**	**60**	**72**	**84**	**96**	**108**	**120+**
2006	132,756	127,296	126,502	280,755	159,020	136,284	105,608	84,429	104,410	50,555
2007	154,318	150,888	145,947	329,237	187,519	159,277	117,183	101,220	122,902	96,571
2008	151,882	147,943	153,986	332,283	193,190	160,114	121,272	101,681	167,482	98,760
2009	146,220	150,178	149,690	327,782	186,733	158,176	119,597	125,035	171,497	98,042
2010	145,531	138,894	144,262	300,660	178,639	151,920	139,437	118,041	156,163	87,924
2011	146,339	141,271	148,740	317,044	185,534	183,768	145,838	122,161	155,734	95,224
2012	153,454	152,178	153,054	338,980	242,220	199,497	163,440	139,434	168,716	97,719
2013	173,003	165,002	168,993	447,745	261,465	215,809	165,965	141,086	201,662	121,867
2014	156,130	151,172	235,610	410,825	235,604	210,647	163,372	131,985	177,616	103,507
2015	142,319	418,523	436,805	577,315	322,537	254,332	205,111	153,930	222,174	98,010

Figure 5.17. Coefficient of Variation of Incremental Values

Taylor & Ashe Data
Accident Year Incremental Values by Development Period
Paid Chain Ladder Model

Accident	Coefficient of Variation Values									
Year	12	24	36	48	60	72	84	96	108	120+
2006	47.7%	18.8%	17.9%	36.5%	38.4%	45.9%	39.7%	46.4%	38.6%	75.5%
2007	40.6%	16.0%	14.9%	30.6%	31.8%	39.0%	31.5%	40.3%	32.2%	102.0%
2008	40.3%	15.8%	15.8%	32.0%	33.0%	39.5%	33.0%	39.6%	44.2%	104.6%
2009	40.8%	16.4%	15.7%	32.0%	33.0%	39.0%	32.6%	50.1%	46.0%	105.9%
2010	44.4%	16.6%	16.4%	31.9%	34.7%	40.7%	42.0%	52.2%	45.9%	106.0%
2011	41.4%	16.1%	16.1%	32.2%	34.3%	47.6%	41.9%	52.1%	43.6%	108.3%
2012	39.7%	15.5%	15.1%	30.7%	40.7%	45.7%	41.6%	52.2%	43.4%	106.1%
2013	36.2%	14.0%	13.8%	33.6%	35.4%	42.2%	35.9%	44.0%	42.0%	100.1%
2014	39.4%	15.5%	23.0%	37.1%	39.4%	48.8%	41.9%	50.0%	43.9%	103.4%
2015	41.6%	47.8%	47.8%	59.0%	59.8%	65.2%	59.5%	66.9%	63.9%	115.0%

appear consistent, although the future periods seem to have larger standard deviations than historical periods. But the standard deviations can't be added because the standard deviations in Figure 5.14 represent those for aggregated incremental values by accident year, which are less than perfectly correlated.

The differences between the future and historical coefficients of variation in Figure 5.17 help clarify any issues with the model results. For example, notice how the differences by development period are more significant in the bottom two rows in Figure 5.17. This is consistent with the increases in the accident year 2014 and 2015 coefficients of variation noted in Figure 5.14, so they can be used to diagnose the causes noted above when compared to the same results for different models.

6. Using Multiple Models

So far we have focused only on one model. In practice, multiple stochastic models should be used in the same way that multiple methods should be used in a deterministic analysis. First the results for each model must be reviewed and finalized, after an iterative process of diagnostic testing and reviewing model output to make sure the model "fits" the data, has reasonable assumptions and produces reasonable results. Then these results can be combined by assigning a weight to the results of each model.

Two primary methods exist for combining the results for multiple models:

- **Run models with the same random variables.** For this algorithm, every model uses the exact same random variables. In the "Bootstrap Models.xlsm" file, the random values are simulated before they are used to simulate results, which means that this algorithm may be accomplished by reusing the same set of random variables for each model. At the end, the incremental values for each model, for each iteration by accident year (that have a partial weight), can be weighted together.
- **Run models with independent random variables.** For this algorithm, every model is run with its own random variables. In the "Bootstrap Models.xlsm" file the random values are simulated before they are used to simulate results, which means that this algorithm may be accomplished by simulating a new set of random variables for each model.[56] At the end, the weights are used to randomly select a model for each iteration by accident year so that the result is a weighted "mixture" of models.

Both algorithms are similar to the process of weighting the results of different deterministic methods to arrive at an actuarial best estimate. The process of weighting the results of different stochastic models produces an actuarial best estimate of a distribution. In practice it is also common to further "adjust" or "shift" the weighted results by year after considering case reserves and the calculated IBNR. This "shifting" can also be done for weighted distributions, either additively to maintain the exact shape and width of the distribution by year or multiplicatively to maintain the exact shape of the distribution but adjusting the width of the distribution.

[56] In general, in order to simulate new random values a new seed value must be selected, otherwise the same random values will be simulated. In the "Bootstrap Models.xlsm" file the seed value is incremented for each model and data type so that different seed values are being used as long as new random numbers are generated for each model and data type.

Figure 6.1. Model Weights by Accident Year

Accident	Model Weights by Accident Year								
Year	Paid CL	Incd CL	Paid BF	Incd BF	Paid CC	Incd CC	Paid GLM	Incd GLM	TOTAL
2006	50.0%	50.0%							100.0%
2007	50.0%	50.0%							100.0%
2008	50.0%	50.0%							100.0%
2009	50.0%	50.0%							100.0%
2010	50.0%	50.0%							100.0%
2011	50.0%	50.0%							100.0%
2012	50.0%	50.0%							100.0%
2013	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%			100.0%
2014	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%			100.0%
2015	16.7%	16.7%			16.7%	16.7%	16.7%	16.7%	100.0%

The second method of combining multiple models will be illustrated using combined Schedule P data for five top 50 companies.[57] Data for all Schedule P lines with 10 years of history may be found in the "Industry Data.xlsm" file, but this example will be confined to Parts A, B, and C. For each line of business ODP bootstrap models were run for paid and incurred data (labeled Chain Ladder), as well as paid and incurred data for the Bornhuetter-Ferguson and Cape Cod models described in Section 3.3 and the GLM bootstrap model described in Section 3.4.[58] For this section, only the results for Part A (Homeowners/Farmowners) will be reviewed.[59]

By comparing the results for all eight models (or fewer, depending on how many are used)[60] a qualitative assessment of the relative merits of each model may be determined. Bayesian methods can be used to determine weighting based on the quality of each model's forecasts. The weights can be determined separately for each year. The table in Figure 6.1 shows an example of weights for the Part A data.[61] The weighted results are displayed in the "Best Estimate" column of Figure 6.2. As a parallel to a deterministic analysis, the means from the eight models could be used to derive a reasonable range from the modeled results (i.e., from $4,099 to $5,650) as shown in Figure 6.3. Alternatively, if we only consider results by accident year which are given some weight when deriving the best estimate, then the "weighted range" may be a more representative view of the uncertainty of the actuarial central estimate.[62]

When selecting weights for stochastic models, the standard deviations should also be considered in addition to the means by model since the weighted best estimate should reflect the actuary's judgments about the entire distribution not just a central

[57] The five companies represent large, medium and smaller companies that have been combined to maintain anonymity. For each Part, a unique set of five companies were used.

[58] An additional benefit of converting the incurred data models to a random payment stream as discussed in Section 3.3.1 is that they can be combined with other model results.

[59] Only selected weighted results are displayed and discussed in Section 6. A more complete set of results, including results for each model, are included in Appendix A.

[60] Other models in addition to the ODP bootstrap and GLM bootstrap models could also be included in the weighting process as long as the simulated results are in the form of random incremental payment streams.

[61] For simplicity, the weights are judgmental and not derived using Bayesian methods.

[62] The "modeled range" in Figure 6.3 is derived using each model that is given at least some weight for any accident year—i.e., if the model is used. In contrast, the "weighted range" is derived using only the models given weight for each accident year, which are highlighted in grey in Figure 6.2 and 6.4.

Figure 6.2. Summary of Mean Results by Model

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Summary of Results by Model

Accident Year	Mean Estimated Unpaid								
	Chain Ladder		Bornhuetter Ferguson		Cape Cod		GLM Bootstrap		Best Est.
	Paid	Incurred	Paid	Incurred	Paid	Incurred	Paid	Incurred	(Weighted)
2006	-	-	-	-	-	-	-	-	-
2007	3	3	2	2	3	3	9	12	3
2008	41	42	28	27	32	33	27	27	41
2009	45	46	37	39	43	45	40	45	46
2010	63	62	60	59	66	71	62	73	64
2011	103	103	96	98	109	115	106	113	103
2012	222	226	169	168	191	199	213	169	224
2013	294	306	327	334	373	385	280	307	335
2014	679	723	722	753	835	871	646	650	752
2015	3,851	3,912	2,660	2,885	3,225	3,430	3,738	4,255	3,742
Totals	5,300	5,422	4,099	4,366	4,878	5,151	5,120	5,650	5,308

Figure 6.3. Summary of Ranges by Accident Year

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Summary of Results by Model

Accident Year	Best Est.	Ranges			
		Weighted		Modeled	
	(Weighted)	Minimum	Maximum	Mininum	Maximum
2006	-				
2007	3	3	3	2	12
2008	41	41	42	27	42
2009	46	45	46	37	46
2010	64	62	63	59	73
2011	103	103	103	96	115
2012	224	222	226	168	226
2013	335	294	385	280	385
2014	752	679	871	646	871
2015	3,742	3,225	4,255	2,660	4,255
Totals	5,308	4,674	5,992	4,099	5,650

estimate. Thus, coefficients of variation by model can be used for this purpose as illustrated in Figure 6.4.

With our focus on the entire distribution, the weights by year were used to randomly sample the specified percentage of iterations from each model. A more complete set of the results for the "weighted" iterations can be created similar to the tables shown in Section 5. The companion "Best Estimate.xlsm" file can be used to weight eight different models together in order to calculate a weighted best estimate. An example for Part A is shown in the table in Figure 6.5.

As one final check of the weighted results it would be common to review the implied IBNR to make sure there are no issues as shown in Figure 6.6. By reviewing this reconciliation, and perhaps also comparing it to deterministic results, additional adjustments could be made to various assumptions. For example, from year 2006 in Figure 6.6 it may be more realistic to revisit the tail factor assumption so that the unpaid estimate is more consistent with the case reserves. Finally, after the interactive process of reviewing results and adjusting assumptions is complete, it may still be

Figure 6.4. Summary of CoV Results by Model

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Summary of Results by Model

Accident	Coefficient of Variation							
	Chain Ladder		Bornhuetter Ferguson		Cape Cod		GLM Bootstrap	
Year	Paid	Incurred	Paid	Incurred	Paid	Incurred	Paid	Incurred
2006								
2007	264.9%	309.9%	310.2%	318.6%	276.2%	326.5%	86.4%	91.5%
2008	74.7%	101.0%	89.2%	109.3%	86.1%	95.6%	177.0%	184.0%
2009	65.5%	93.2%	69.7%	93.5%	69.2%	89.0%	119.3%	118.9%
2010	49.4%	75.6%	52.2%	78.0%	47.2%	72.7%	78.5%	78.1%
2011	34.9%	62.4%	35.7%	64.6%	33.5%	59.5%	51.3%	50.9%
2012	26.1%	49.5%	31.3%	51.4%	28.1%	50.2%	33.6%	41.5%
2013	27.3%	57.5%	26.9%	59.3%	23.3%	56.2%	27.9%	34.9%
2014	18.9%	48.8%	21.8%	51.0%	17.1%	46.7%	20.3%	26.3%
2015	9.2%	39.2%	14.4%	40.5%	8.0%	39.4%	9.0%	16.0%
Totals	8.4%	29.0%	11.1%	28.9%	7.9%	27.5%	8.7%	13.3%

Figure 6.5. Estimated Unpaid Model Results (weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Best Estimate (Weighted)

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	3	9	292.0%	-	173	0	1	17	42
2008	41	37	88.6%	-	391	32	57	111	168
2009	46	37	81.0%	1	522	36	60	114	175
2010	64	41	63.6%	4	537	55	81	139	205
2011	103	50	48.8%	10	636	94	125	193	276
2012	224	89	40.0%	36	917	211	266	382	529
2013	335	148	44.3%	25	1,460	315	401	594	865
2014	752	293	39.0%	106	2,881	725	873	1,265	1,789
2015	3,742	982	26.2%	1,094	10,700	3,654	4,118	5,392	7,059
Totals	5,308	1,044	19.7%	2,116	12,445	5,224	5,758	7,074	8,675
Normal Dist.	5,308	1,044	19.7%			5,308	6,013	7,026	7,738
logNormal Dist.	5,309	1,034	19.5%			5,211	5,935	7,158	8,164
Gamma Dist.	5,308	1,044	19.7%			5,240	5,971	7,135	8,035
TVaR						6,035	6,593	8,140	10,091
Normal TVaR						6,142	6,636	7,463	8,092
logNormal TVaR						6,121	6,691	7,780	8,733
Gamma TVaR						6,137	6,688	7,689	8,516

Figure 6.6. Reconciliation of Total Results (weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Reconciliation of Total Results
Best Estimate (Weighted)

Accident Year	Paid To Date	Incurred To Date	Case Reserves	IBNR	Estimate of Ultimate	Estimate of Unpaid
2006	5,234	5,237	3	(3)	5,234	-
2007	6,470	6,479	9	(6)	6,473	3
2008	7,848	7,867	19	23	7,890	41
2009	7,020	7,046	26	20	7,066	46
2010	7,291	7,341	50	13	7,355	64
2011	8,134	8,225	91	12	8,237	103
2012	10,800	11,085	285	(61)	11,023	224
2013	7,522	7,810	288	46	7,856	335
2014	7,968	8,703	735	17	8,720	752
2015	9,309	12,788	3,478	263	13,051	3,742
Totals	77,596	82,580	4,984	324	82,905	5,308

prudent to make adjustments to the best estimate of the unpaid by shifting the results as noted earlier in this section. For example, since all of the models estimated the unpaid for 2012 to be less than the case reserves, if other studies show that the case reserves are not likely to be redundant then the actuary may decide to shift the unpaid for 2012 so that it is at least 285.

6.1. Additional Useful Output

Three rows of percentile numbers for the normal, lognormal, and gamma distributions, which have been fitted to the total unpaid-claim distribution, may be seen at the bottom of the table in Figure 6.5. The fitted mean, standard deviation, and selected percentiles are in their respective columns; the smoothed results can be used to assess the quality of fit, parameterize a DFA model, or used to smooth the estimate of extreme values,[63] among other applications.

Four rows of numbers indicating the Tail Value at Risk (TVaR), defined as the average of all of the simulated values equal to or greater than the percentile value, may also be seen at the bottom of Figure 6.5. For example, in this table, the 99th percentile value for the total unpaid claims for all accident years combined is 8,675, while the average of all simulated values that are greater than or equal to 8,675 is 10,091. The Normal TVaR, Lognormal TVaR, and Gamma TVaR rows are calculated similarly, except that they use the respective fitted distributions in the calculations rather than actual simulated values from the model.

An analysis of the TVaR values is likely to help clarify a critical issue: if the actual outcome exceeds the X percentile value, by how much will it exceed that value on average? This type of assessment can have important implications related to risk-based capital calculations and other technical aspects of enterprise risk management. But it is worth noting that the purpose of the normal, lognormal, and gamma TVaR numbers is to provide "smoothed" values—that is, that some of the random statistical noise is essentially prevented from distorting the calculations.

6.2. Estimated Cash Flow Results

A model's output may also be reviewed by calendar year (or by future diagonal), as shown in the table in Figure 6.7. A comparison of the values in Figures 6.5 and 6.7 indicates that the total rows are identical, because summing the future payments horizontally or diagonally will produce the same total. Similar diagnostic issues (as discussed in Section 5) may be reviewed in the table in Figure 6.7, with the exception of the relative values of the standard errors and coefficients of variation moving in opposite directions for calendar years compared to accident years. This phenomenon makes sense on an intuitive level when one considers that "final" payments, projected to the furthest point in the future, should actually be the smallest, yet relatively most uncertain.

---

[63] A random instance of an extreme percentile can be quite erratic compared to the same percentile of a distribution fitted to the simulated distribution. This random noise for extreme percentiles could be cause for increasing the number of iterations, but if the same percentiles for the fitted distributions are stable perhaps they can be used in lieu of more iterations. Of course the use of the extreme values assumes that the models are reliable.

Figure 6.7. Estimated Cash Flow (weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Calendar Year Unpaid
Best Estimate (Weighted)

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2016	3,475	754	21.7%	1,297	8,420	3,414	3,797	4,730	5,948
2017	865	208	24.0%	293	2,148	843	982	1,224	1,483
2018	403	118	29.4%	115	1,298	387	467	614	740
2019	204	67	32.7%	56	654	194	240	325	412
2020	140	50	35.9%	40	539	132	165	233	297
2021	90	43	47.4%	12	611	82	112	169	229
2022	70	44	63.2%	6	409	60	91	152	215
2023	51	58	112.2%	-	735	36	75	151	253
2024	10	15	146.5%	-	199	4	15	41	67
Totals	5,308	1,044	19.7%	2,116	12,445	5,224	5,758	7,074	8,675

Figure 6.8. Estimated Loss Ratio (weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Ultimate Loss Ratios
Best Estimate (Weighted)

Accident Year	Mean Loss Ratio	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	67.7%	28.5%	42.1%	0.4%	220.8%	66.1%	71.1%	130.9%	158.2%
2007	79.3%	30.2%	38.1%	8.2%	262.2%	77.8%	83.1%	145.5%	178.5%
2008	90.5%	31.2%	34.5%	16.9%	261.3%	89.0%	94.6%	159.9%	188.9%
2009	72.8%	26.8%	36.7%	10.2%	215.6%	71.4%	76.1%	131.7%	180.4%
2010	65.3%	23.3%	35.7%	10.2%	225.0%	63.8%	68.0%	116.1%	139.7%
2011	64.1%	21.2%	33.1%	13.0%	190.0%	63.2%	67.0%	111.8%	130.5%
2012	80.5%	24.0%	29.9%	25.0%	234.6%	79.0%	83.7%	132.9%	154.6%
2013	54.7%	18.8%	34.4%	9.9%	157.7%	53.9%	57.4%	96.2%	115.1%
2014	58.0%	19.2%	33.0%	13.0%	164.8%	57.1%	60.6%	99.8%	118.8%
2015	88.2%	21.5%	24.4%	30.9%	232.5%	85.5%	92.5%	127.9%	158.7%
Totals	71.3%	7.4%	10.4%	46.6%	112.7%	70.8%	75.7%	84.4%	91.7%

6.3. Estimated Ultimate Loss Ratio Results

Another output table, Figure 6.8, shows the estimated ultimate loss ratios by accident year. Unlike the estimated unpaid and estimated cash-flow tables, the values in this table are calculated using all simulated values, not just the values beyond the end of the historical triangle. Because the simulated sample triangles represent additional possibilities of what could have happened in the past, even as the "squaring of the triangle" and process variance represent what could happen as those same past values are played out into the future, we are in possession of sufficient information to enable us to estimate the variability in the loss ratio from day one until all claims are completely paid and settled for each accident year.[64]

Reviewing the simulated values indicates that the standard errors in Figure 6.8 should be proportionate to the means, while the coefficients of variation should be relatively constant by accident year. In terms of diagnostics, any increases in standard error and coefficient of variation for the most recent years would be consistent with the reasons

[64] If we are only interested in the "remaining" volatility in the loss ratio, then the values in the estimated unpaid table (Figure 6.5) can be added to the cumulative paid values by year and divided by the premiums.

Figure 6.9. Estimated Unpaid Claim Runoff (weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Calendar Year Unpaid Claim Runoff
Best Estimate (Weighted)

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2015	5,308	1,044	19.7%	2,116	12,445	5,224	5,758	7,074	8,675
2016	1,834	365	19.9%	746	4,128	1,797	2,030	2,459	2,957
2017	969	218	22.5%	336	2,316	946	1,088	1,353	1,627
2018	566	146	25.8%	159	1,393	548	647	828	1,004
2019	362	114	31.5%	79	1,171	347	424	565	718
2020	222	92	41.4%	35	956	207	269	386	524
2021	132	76	57.6%	6	863	117	166	268	394
2022	62	59	96.3%	(0)	745	46	84	166	269
2023	10	15	146.5%	(0)	199	4	15	41	67

previously cited in Section 5.4 for the estimated unpaid tables. Risk management-wise, the loss ratio distributions have important implications for projecting pricing risk—the mean loss ratios can be used to view any underwriting cycles and help inform the projected mean for the next few years, while the coefficients of variation can be used to select a standard deviation for the next few years.[65]

6.4. Estimated Unpaid Claim Runoff Results

Figure 6.9, shows the runoff of the total unpaid claim distribution by future calendar year. Like the estimated unpaid and estimated cash-flow tables, the values in this table are calculated using only future simulated values, except that future diagonal results are sequentially removed so that we are left with the remaining unpaid claims at the end of future calendar periods. These results are quite useful for calculating the runoff of the unpaid claim distribution when calculating risk margins using the cost of capital method.

6.5. Distribution Graphs

A final model output to consider is a histogram of the estimated unpaid amounts for the total of all accident years combined, as shown in the graph in Figure 6.10. The histogram is created by counting the number of outcomes within each of 100 "buckets" of equal size spread between the minimum and maximum outcome. To smooth the histogram a kernel density function is often used, which is the green bars in Figure 6.10.

Another useful strategy for graphing the total unpaid distribution may be accomplished by creating a summary of the eight model distributions used to determine the weighted "best estimate" and distribution. An example of this graph using the kernel density functions is shown in Figure 6.11 and dots for the mean estimates, which would represent a traditional range,[66] are also included.

---

[65] The coefficients of variation measure the variability of the loss ratios, given the movements by year. Without this information, it is common to base the future standard deviation on the standard deviation of the historical mean loss ratios, but this is not ideal since the variability of the mean loss ratios is not the same as the possible variation in the actual outcomes given movements in the means.

[66] A traditional range would use deterministic point estimates instead of means of the distributions, but the intent is consistent. While the points would technically have an infinitesimal probability and should therefore sit on the x-axis, they are elevated above the zero probability level purely for illustration purposes.

Figure 6.10. Total Unpaid Claims Distribution

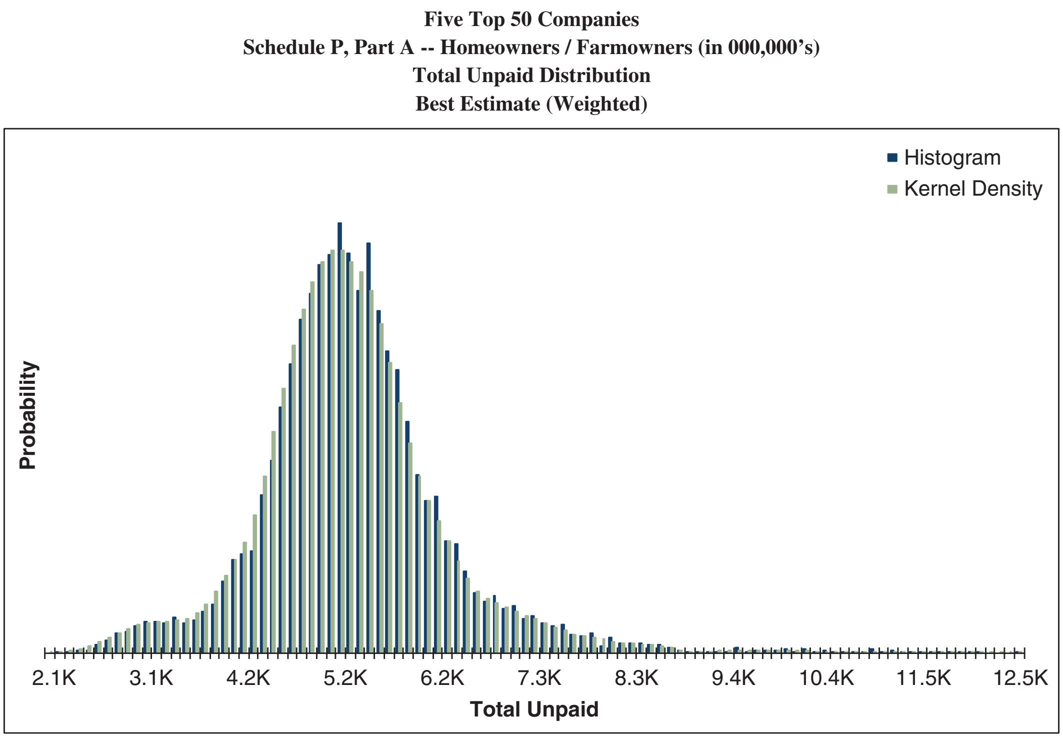
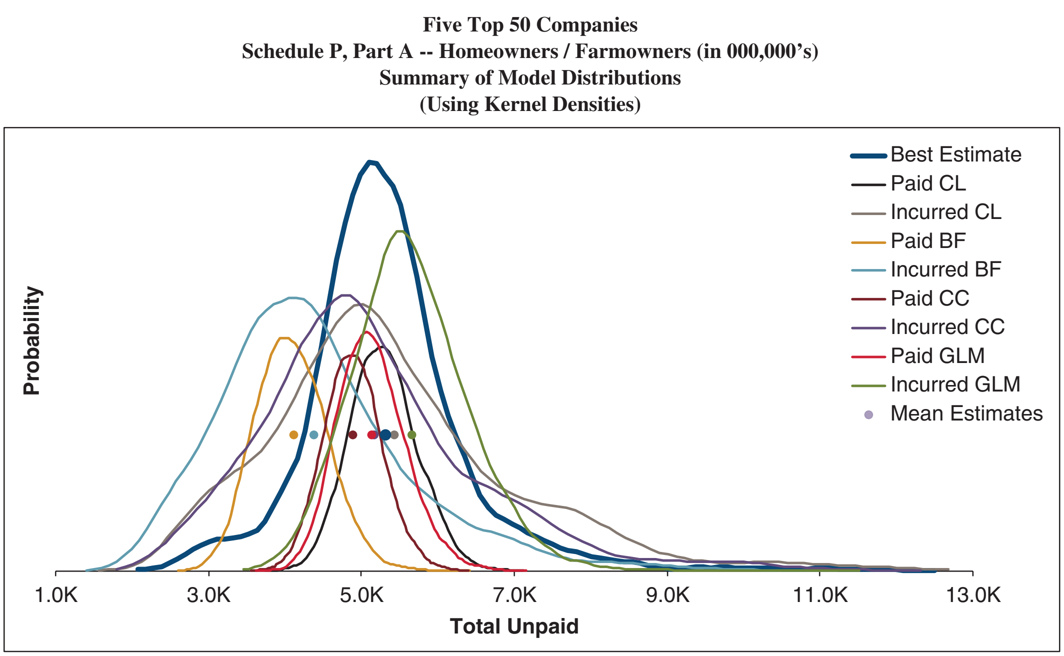


Figure 6.11. Summary of Model Distributions



The corresponding tables and graphs for the Part B and Part C results are shown in Appendices B and C, respectively.[67]

6.6. Correlation

Results for an entire business unit can be estimated, after each business segment has been analyzed and weighted into best estimates, using aggregation. This represents another area where caution is warranted. The procedure is not a simple matter of adding up the distributions for each segment. In order to estimate the distribution of possible outcomes for a company as a whole, a correlation of results between segments must be used.[68]

Simulating correlated variables is commonly accomplished with a multivariate distribution whose parameters and correlations have been previously specified. This type of simulation is most easily applied when distributions are uniformly identical and known in advance (for example, all derived from a multivariate normal distribution). Unfortunately, these conditions do not generally exist for the ODP bootstrap model (or other models), as quite often the modeling process does not allow us to know the characteristics of overall distributions in advance or combining distributions from different types of models is by definition not uniformly identical and known in advance. Indeed, as the shapes of different distributions are usually slightly different, another approach will be needed.[69]

Two useful correlation processes for the ODP bootstrap model are location mapping (or synchronized bootstrapping) and re-sorting.[70]

With location mapping, each iteration will include sampling residuals for the first segment and then going back to note the location in the original residual triangle of each sampled residual.[71] Each of the other segments is sampled using the residuals at the same locations for their respective residual triangles. Thus, the correlation of the original residuals is preserved in the sampling process.

The location-mapping process is easily implemented in Excel and does not require the need to estimate a correlation matrix. There are, however, two drawbacks to this process. First, it requires all of the business segments to use data triangles that are precisely the same size with no missing values or outliers when comparing each location of the residuals.[72] Second, the correlation of the original residuals is used in the model, and no other correlation assumptions can be used for stress testing the aggregate results.

---

[67] For Part B and Part C, tail factors were used to illustrate the results when extrapolated beyond just squaring the triangle. This also flows through to the Aggregate results in Appendix D.

[68] This section assumes the reader is familiar with correlation.

[69] It is possible to use this process with a parametric ODP bootstrap model, as described in Section 4.10, but that is beyond the scope of the monograph.

[70] For a useful reference see Kirschner, et al. (2008).

[71] For example, in the "Bootstrap Models.xlsm" file the locations of the sampled residuals are shown in Step 15, which could be replicated iteration by iteration for each business segment.

[72] It is possible to fill in "missing" residuals in another segment using a randomly selected residual from elsewhere in the triangle, but in order to maintain the same amount of correlation the selection of the other residual would need to account for the correlation between the residuals, which complicates the process.

Figure 6.12. Estimated Correlation and P-values

Rank Correlation of Residuals after Hetero Adjustment - Paid

LOB	1	2	3
1	1.00	0.37	0.19
2	0.37	1.00	0.24
3	0.19	0.24	1.00

P-Values of Rank Correlation of Residuals after Hetero Adjustment - Paid

LOB	1	2	3
1	0.00	0.01	0.17
2	0.01	0.00	0.07
3	0.17	0.07	0.00

The second correlation process, re-sorting, can be accomplished with algorithms such as Iman-Conover[73] or Copulas, among others. The primary advantages of re-sorting include:

- The triangles for each segment may have different shapes and sizes,
- Different correlation assumptions may be employed, and
- Different correlation algorithms may also have other beneficial impacts on the aggregate distribution.

For example, using a *t*-distribution Copula with low degrees of freedom rather than a normal-distribution Copula, will effectively "strengthen" the focus of the correlation in the tail of the distribution, all else being equal. This type of consideration is important for risk-based capital and other risk modeling issues.

To induce correlation among different segments in the ODP bootstrap model, a calculation of the correlation matrix using Spearman's Rank Order and use of re-sorting based on the ranks of the total unpaid claims for all accident years combined may be done. The calculated correlations for Parts A, B, and C based on the paid residuals after hetero adjustments may be seen in the table in Figure 6.12. A second part of Figure 6.12 are the *P*-values for each correlation coefficient, which are an indication of whether a correlation coefficient is significantly different than zero as the *P*-value gets close to zero.[74]

By reviewing the correlation coefficients for each "pair" of segments, along with the *P*-values, from different sets of correlations matrices (e.g., from paid or incurred data before or after the hetero adjustment) judgment can be used to select a correlation matrix assumption. As noted above, caution is warranted as these calculated correlation matrices are limited to the data used in the calculation and the impact of other systemic issues, such as contagion, may also need to be considered.

Using these correlation coefficients, the "Aggregate Estimate.xlsm" file, and the simulation data for Parts A, B, and C, the aggregate results for the three lines of business

[73] For a useful reference see Iman and Conover (1982) or Mildenhall (2006). In the "Aggregate Estimate.xlsm" file the Iman-Conover algorithm is used to "Generate Rank Values" on the Inputs sheet.

[74] While judgment is clearly appropriate, the typical threshold is a *P*-value of 5%—i.e., a *P*-value of 5% or less indicates the correlation is significantly different than zero, while a *P*-value greater than 5% indicates the correlation is not significantly different than zero.

Figure 6.13. Aggregate Estimated Unpaid

Five Top 50 Companies
Aggregate Three Lines of Business
Accident Year Unpaid

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	67	25	37.9%	0	186	66	83	110	130
2007	107	30	28.1%	25	295	105	126	158	185
2008	199	49	24.8%	67	622	194	226	285	342
2009	298	56	18.8%	123	800	293	331	395	457
2010	480	69	14.3%	248	959	475	522	599	668
2011	862	106	12.3%	503	1,561	860	923	1,041	1,135
2012	1,666	187	11.2%	383	2,555	1,662	1,771	1,985	2,148
2013	3,070	333	10.8%	1,808	6,522	3,066	3,249	3,649	3,928
2014	5,632	703	12.5%	2,435	8,555	5,632	6,075	6,801	7,326
2015	13,270	1,788	13.5%	5,217	22,660	13,262	14,348	16,180	18,011
Totals	25,650	2,080	8.1%	16,952	36,085	25,616	26,949	29,088	30,991
Normal Dist.	25,650	2,080	8.1%			25,650	27,053	29,072	30,490
logNormal Dist.	25,650	2,088	8.1%			25,566	27,006	29,222	30,885
Gamma Dist.	25,650	2,080	8.1%			25,594	27,021	29,165	30,736

were calculated and summarized in the table in Figure 6.13. A more complete set of tables for the aggregate results is shown in Appendix D.

Note that using residuals to correlate the lines of business (or other segments), as in the location mapping method, and measuring the correlation between residuals, as in the re-sorting method, both tend to create correlations that are close to zero. For reserve risk, the correlation that is desired is between the total unpaid amounts for two segments. The correlation that is being measured is the correlation between each incremental future loss amount, given the underlying model describing the overall trends in the data. This may or may not be a reasonable approximation.

While not the direct measure we are hoping for, keep in mind that some level of implied correlation between lines of business will naturally occur due to correlations between the model parameters—e.g., similarities in development parameters, so correlation based on the correlation between the remaining random movements in the incremental values given the model parameters (i.e., residuals) may be reasonable. However, an example of an issue not particularly well suited to measurement via residual correlation is contagion between lines of business—i.e., single events that result in claims in multiple lines of business. To account for this, and to add a bit of conservatism, the correlation assumption can be easily changed based on actuarial judgment.

Correlation is often thought of as being much stronger than "close to zero", but in this case the correlation being considered is typically the loss ratio movements by line of business. For pricing risk, the correlation that is desired is between the loss ratio movements by accident year between two segments. This correlation is not as likely to be close to zero, so correlation of loss ratios (e.g., for the data in Figure 6.7) is often done with a different correlation assumption compared to reserving risk.

7. Model Testing

Work on testing stochastic unpaid claim estimation models is still in its infancy. Most papers on stochastic models display results, and some even compare a few different models, but they tend to be void of any statistical evidence regarding how well the model in question predicts the underlying distribution. This is quite understandable since we don't know what the underlying distribution is, so with real data the best we can hope for is to retrospectively test a very old data set to see how well a model predicted the actual outcome.[75]

Testing a few old data sets is better than not, but ideally we would need many similar data sets to perform meaningful tests. One recent paper authored by the General Insurance Reserving Oversight Committee (GI ROC) in their papers for the General Insurance Research Organizing (GIRO) conference in 2007 titled "Best Estimates and Reserving Uncertainty" (ROC/GIRO 2007) and their updated paper in 2008 titled "Reserving Uncertainty" (ROC/GIRO 2008) took a first step in performing more meaningful statistical testing of a variety of models.

A large number of models were reviewed and tested in these studies, but one of the most interesting portions of the studies were done by comparing the unpaid liability distributions created by the Mack and ODP bootstrap model against the "true" artificially generated unpaid loss percentiles. To accomplish these tests, artificial datasets were constructed so that all of the Mack and ODP bootstrap assumptions, respectively, are satisfied. While the artificial datasets were recognized as not necessarily realistic, the "true" results are known so the Working Parties were able to test to see how well each model performed against datasets that could be considered "perfect."

7.1. Bootstrap Model Results

To test the ODP bootstrap model, incremental losses were simulated for a 10 × 10 square of data based on the assumptions of the ODP bootstrap model. For the 30,000 datasets simulated, the upper triangles were used and the OPD bootstrap model from England and Verrall (1999; 2002) were used to estimate the expected results and various percentiles. The proportion of simulated scenarios in which the "true" outcome exceeded the 99th percentile of the ODP Bootstrap method's results was around 2.6–3.1%. For the Mack method, the "true" outcome exceeded the 99th percentile around 8–13%.

[75] For example, data for accident years 1994 to 2004 could be completely settled and all results known as of 2014. Thus, we could use the triangle as it existed at year end 2004 to test how well a model predicted the final results.

Thus, the ODP bootstrap model performed better than the Mack model for "perfect" data, even though the results for both models were somewhat deficient in the sense that they both seem to under-predict the extremes of the "true" distribution. In fairness, it should be noted however, that the ODP bootstrap model that was tested did not include many of the "advancements" described in Section 3.2.

7.2. Future Testing

The testing done for GIRO was a significant improvement over simply looking at results for different models, without knowing anything about the "true" underlying distribution. The next step in the testing process will be to test models against "true" results for realistic data instead of "perfect" data. The CAS Loss Simulation Model Working Party (2011) has created a model that will create datasets from the claim transaction level up. The goal is to create thousands of datasets based on characteristics of real data that can be used for testing various models.

8. Future Research

With testing of stochastic models in its infancy, much work in the area of future research is needed. Only a few such areas are offered here.

- Expand testing of the ODP bootstrap model with realistic data using the CAS loss simulation model.
- Research on how the adjustments to the ODP bootstrap and GLM bootstrap suggested in this monograph perform relative to realistic data—i.e., is there a significant improvement in the predictive power of the model given the different model configurations and adjustments.
- Expand or change the ODP bootstrap model in other ways, for example use of the Munich chain ladder (Quarg and Mack 2008) or Berquist-Sherman (1977) method with an incurred/paid set of triangles, or the use of claim counts and average severities. Other examples could include the use of different residuals, such as deviance or Anscombe residuals noted in Section 3.2.
- Research the use of a Bayesian or other approach to selecting weights for different models by accident year to improve the process of combining multiple models discussed in Section 6.
- Research other risk analysis measures and how the ODP bootstrap model can be used for enterprise risk management.
- Research how the ODP bootstrap model can be used for Solvency II requirements in Europe and the International Accounting Standards.
- Research into the most difficult parameter to estimate: the correlation matrix.

9. Conclusions

While this monograph endeavored to show how the ODP bootstrap model can be used in a variety of practical ways, and to illustrate the diagnostic tools the actuary needs to assess whether the model is working well, it should not be assumed that the ODP bootstrap model is well suited for every data set. However, it is hoped that the ODP bootstrap and GLM bootstrap "toolsets" can become an integral part of the actuaries regular estimation of unpaid claim liabilities, rather than just a "black box" to be used only if necessary or after the deterministic methods have been used to select a point estimate. Finally, the modeling framework allows the actuary to "fit" the model to the data instead of simply accepting the model as is and essentially forcing the data to "fit" the model.

Acknowledgments

The author gratefully acknowledges the many authors listed in the References (and others not listed) that contributed to the foundation of the ODP bootstrap model, without which this research would not have been possible. He also wishes to thank the co-author of the predecessor paper, Jessica Leong, for all her support and the contributions that led to this revised monograph. He would also like to thank all the peer reviewers, Stephen Finch, Roger Hayne, Stephen Lienhard, John Major, Mark Mulvaney and Ben Zehnwirth, who helped to improve the quality of the monograph in a variety of ways. In particular, Stephen Finch is noteworthy for keeping his wits during an intoxicating discussion which led to the creation of the term "heteroecthesious" data. Finally, he is grateful to the CAS referees for their comments which also greatly improved the quality of the monograph.

Supplementary Materials

There are several companion files designed to give the reader a deeper understanding of the concepts discussed in the monograph. T he files are all in the "A Practitioners Guide. zip" file at https://www.casact.org/sites/default/files/2021-02/practitionerssuppl-shaplandmonograph04.zip The files are:

Model Instructions.pdf—this file contains a written description of how to use the primary bootstrap modeling files.

Primary bootstrap modeling iles:

Industry Data.xls—this file contains Schedule P data by line of business for the entire U.S.
industry and five of the top 50 companies, for each LOB that has 10 years of data.

Bootstrap Models.xlsm—this file contains the detailed model steps described in this monograph as well as various modeling options and diagnostic tests. Data can be entered and simulations run and saved for use in calculating a weighted best estimate.

Best Estimate.xlsm—this file can be used to weight the results from eight different models to get a "best estimate" of the distribution of possible outcomes.

Aggregate Estimate.xlsm—this file can be used to correlate the best estimate results from 3 LOBs/segments.

Correlation Ranks.xlsm—this file contains examples of ranks used to correlate results by LOB/segment.

Simple example calculation files:

GLM Framework.xlsm—this file illustrates the calculation of the GLM bootstrap model (framework) and the corresponding ODP bootstrap model for a simple 3×3 triangle using (3.8).

GLM Framework C.xlsm—this file illustrates the calculation of the GLM bootstrap model (framework) and the corresponding ODP bootstrap model for a simple 3×3 triangle using (3.7).

GLM Framework 6.xlsm—this file illustrates the calculation of the GLM bootstrap model (framework) and the corresponding ODP bootstrap model for a simple 6×6 triangle using (3.8).

GLM Framework 6C.xlsm—this file illustrates the calculation of the GLM bootstrap model (framework) and the corresponding ODP bootstrap model for a simple 6×6 triangle using (3.7).

GLM Bootstrap 6 with Outlier.xlsm—this file illustrates how the calculation of the GLM bootstrap for a simple 6 × 6 triangle is adjusted for an outlier. It includes different options for adjusting the ODP bootstrap model to remove an outlier.

GLM Bootstrap 6 with 3yr avg.xlsm—this file illustrates how the calculation of the GLM bootstrap for a simple 6 × 6 triangle is adjusted to only use the equivalent of a three-year average (i.e., the last four diagonals).

GLM Bootstrap 6 with 1 Acc Yr Parameter.xlsm—this file illustrates the calculation of the GLM bootstrap using only one accident year (level) parameter, a development year trend parameter for every year and no calendar year trend parameter for a simple 6 × 6 triangle.

GLM Bootstrap 6 with 1 Dev Yr Parameter.xlsm—this file illustrates the calculation of the GLM bootstrap using only one development year trend parameter, an accident year (level) parameter for every year and no calendar year trend parameter for a simple 6 × 6 triangle.

GLM Bootstrap 6 with 1 Acc Yr & 1 Dev Yr Parameter.xlsm—this file illustrates the calculation of the GLM bootstrap using only one accident year (level) parameter, one development year trend parameter and no calendar year trend parameter for a simple 6 × 6 triangle.

GLM 6 Bootstrap with 1 Acc Yr 1 Dev Yr & 1 Cal Yr Parameter.xlsm—this file illustrates the calculation of the GLM bootstrap using only one accident year (level) parameter, one development year trend parameter and one calendar year trend parameter for a simple 6 × 6 triangle.

Appendices

Appendix A—Schedule P, Part A Results

In this appendix the results for Schedule P, Part A (Homeowners/Farmowners) are shown.

Figure A.1. Estimated Unpaid Model Results (Paid Chain Ladder)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Paid Chain Ladder Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	3	7	264.9%	-	81	0	2	17	33
2008	41	31	74.7%	-	204	35	59	100	131
2009	45	30	65.5%	7	209	38	61	104	137
2010	63	31	49.4%	15	213	56	80	118	161
2011	103	36	34.9%	36	286	96	122	170	213
2012	222	58	26.1%	93	497	216	258	328	376
2013	294	80	27.3%	126	671	285	342	440	513
2014	679	128	18.9%	366	1,190	675	758	894	1,003
2015	3,851	356	9.2%	2,675	5,051	3,831	4,075	4,496	4,790
Totals	5,300	447	8.4%	4,132	6,907	5,282	5,579	6,056	6,421
Normal Dist.	5,300	447	8.4%			5,300	5,602	6,036	6,341
logNormal Dist.	5,300	448	8.4%			5,282	5,591	6,067	6,426
Gamma Dist.	5,300	447	8.4%			5,288	5,595	6,057	6,396

Figure A.2. Total Unpaid Claims Distribution (Paid Chain Ladder)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Total Unpaid Distribution
Paid Chain Ladder Model

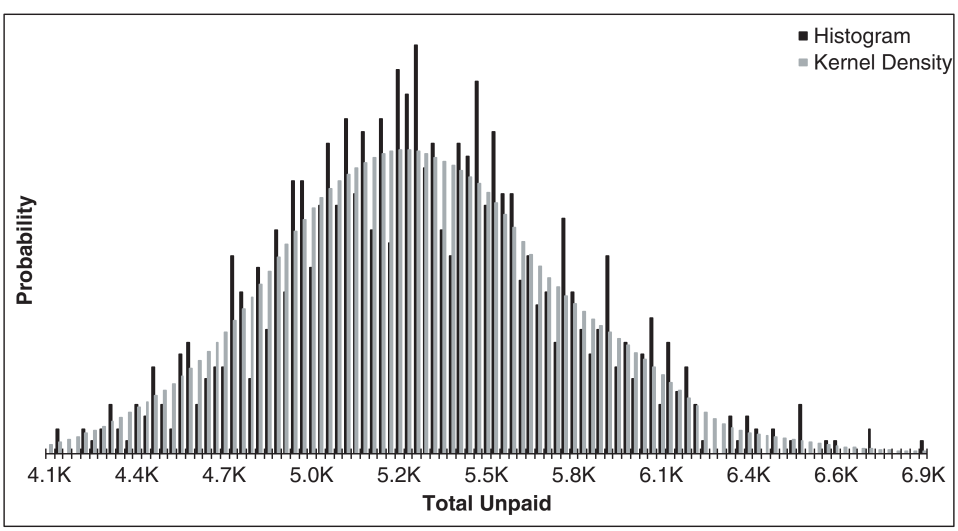


Figure A.3. Estimated Unpaid Model Results (Incurred Chain Ladder)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Incurred Chain Ladder Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	3	9	309.9%	-	93	0	1	17	48
2008	42	42	101.0%	-	306	30	56	126	189
2009	46	42	93.2%	1	325	33	57	135	205
2010	62	47	75.6%	4	355	52	83	149	253
2011	103	64	62.4%	12	473	89	129	231	338
2012	226	112	49.5%	43	984	202	276	435	587
2013	306	176	57.5%	36	1,449	271	384	621	860
2014	723	353	48.8%	109	2,452	664	884	1,418	1,842
2015	3,912	1,534	39.2%	1,306	10,236	3,694	4,523	6,708	9,175
Totals	5,422	1,575	29.0%	1,981	12,631	5,217	6,144	8,197	10,612
Normal Dist.	5,422	1,575	29.0%			5,422	6,485	8,013	9,086
logNormal Dist.	5,423	1,569	28.9%			5,209	6,307	8,305	10,076
Gamma Dist.	5,422	1,575	29.0%			5,271	6,386	8,246	9,741

Figure A.4. Total Unpaid Claims Distribution (Incurred Chain Ladder)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Total Unpaid Distribution
Incurred Chain Ladder Model

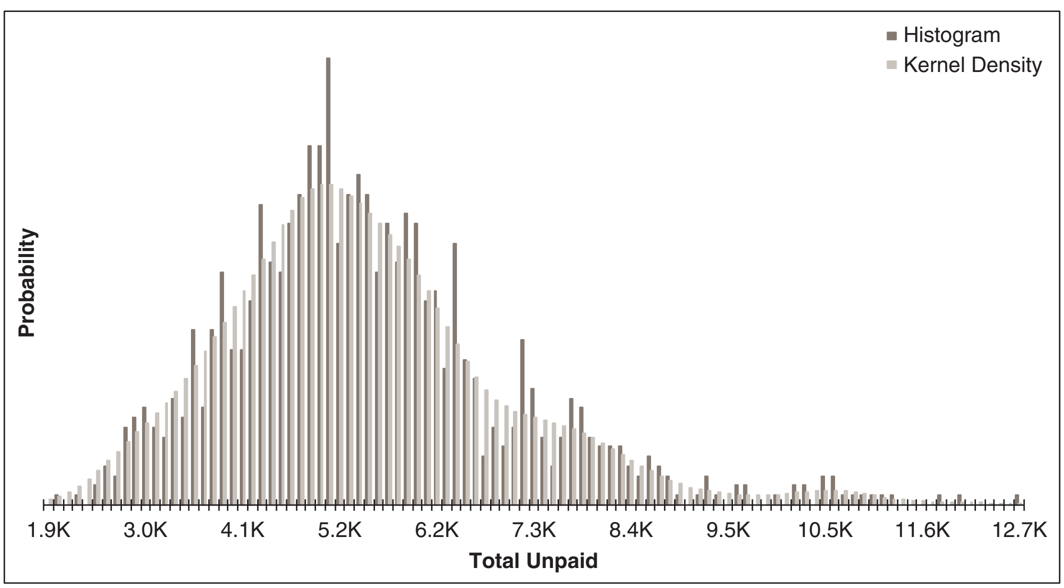


Figure A.5. Estimated Unpaid Model Results (Paid Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Paid Bornhuetter-Ferguson Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	2	6	310.2%	-	48	0	0	10	33
2008	28	25	89.2%	-	188	21	40	71	115
2009	37	26	69.7%	5	152	30	51	87	115
2010	60	31	52.2%	11	186	53	76	127	153
2011	96	34	35.7%	32	274	89	114	163	194
2012	169	53	31.3%	60	367	161	201	269	308
2013	327	88	26.9%	115	804	319	384	483	573
2014	722	157	21.8%	332	1,314	708	826	997	1,129
2015	2,660	383	14.4%	1,689	3,887	2,645	2,908	3,340	3,659
Totals	4,099	456	11.1%	2,835	5,789	4,096	4,392	4,849	5,218
Normal Dist.	4,099	456	11.1%			4,099	4,407	4,850	5,161
logNormal Dist.	4,099	458	11.2%			4,074	4,392	4,894	5,280
Gamma Dist.	4,099	456	11.1%			4,082	4,397	4,877	5,235

Figure A.6. Total Unpaid Claims Distribution (Paid Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Total Unpaid Distribution
Paid Bornhuetter-Ferguson Model

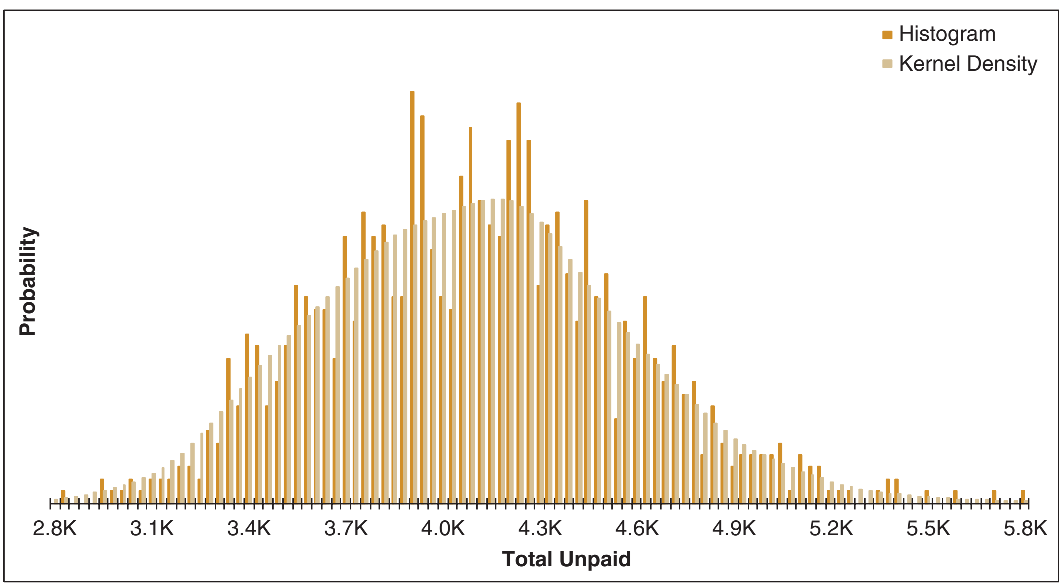


Figure A.7. Estimated Unpaid Model Results (Incurred Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Incurred Bornhuetter-Ferguson Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	2	7	318.6%	-	67	0	1	13	41
2008	27	30	109.3%	-	234	18	37	84	142
2009	39	36	93.5%	1	263	28	50	114	180
2010	59	46	78.0%	4	397	47	78	149	214
2011	98	63	64.6%	9	473	84	123	221	302
2012	168	86	51.4%	30	659	152	210	340	443
2013	334	198	59.3%	34	2,310	304	412	690	972
2014	753	384	51.0%	111	3,131	688	919	1,513	1,883
2015	2,885	1,168	40.5%	921	7,678	2,699	3,449	5,198	6,483
Totals	4,366	1,260	28.9%	1,873	9,804	4,224	5,048	6,860	8,182
Normal Dist.	4,366	1,260	28.9%			4,366	5,216	6,438	7,297
logNormal Dist.	4,367	1,272	29.1%			4,193	5,083	6,704	8,143
Gamma Dist.	4,366	1,260	28.9%			4,246	5,137	6,624	7,817

Figure A.8. Total Unpaid Claims Distribution (Incurred Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Total Unpaid Distribution
Incurred Bornhuetter-Ferguson Model

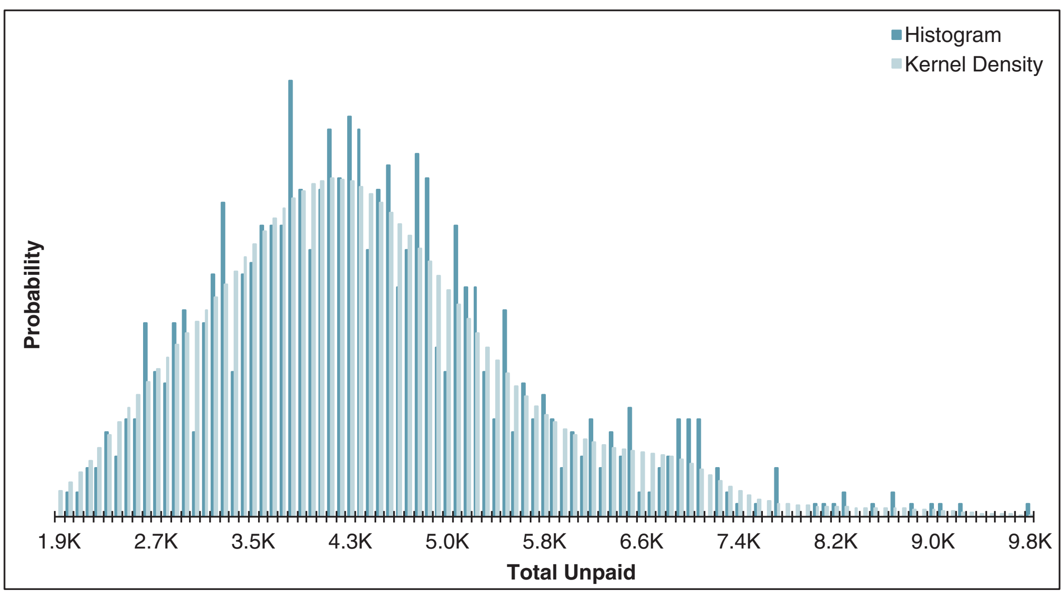


Figure A.9. Estimated Unpaid Model Results (Paid Cape Cod)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Paid Cape Cod Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	3	7	276.2%	-	59	0	1	17	38
2008	32	28	86.1%	-	178	25	45	89	125
2009	43	30	69.2%	6	259	36	59	97	137
2010	66	31	47.2%	16	225	59	85	122	166
2011	109	36	33.5%	43	283	102	130	176	213
2012	191	54	28.1%	74	401	184	226	288	337
2013	373	87	23.3%	156	719	366	424	525	600
2014	835	143	17.1%	407	1,520	832	921	1,082	1,192
2015	3,225	258	8.0%	2,384	4,098	3,227	3,389	3,659	3,855
Totals	4,878	384	7.9%	3,823	6,174	4,871	5,116	5,528	5,836
Normal Dist.	4,878	384	7.9%			4,878	5,137	5,510	5,772
logNormal Dist.	4,878	385	7.9%			4,863	5,128	5,536	5,841
Gamma Dist.	4,878	384	7.9%			4,868	5,132	5,527	5,816

Figure A.10. Total Unpaid Claims Distribution (Paid Cape Cod)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Total Unpaid Distribution
Paid Cape Cod Model

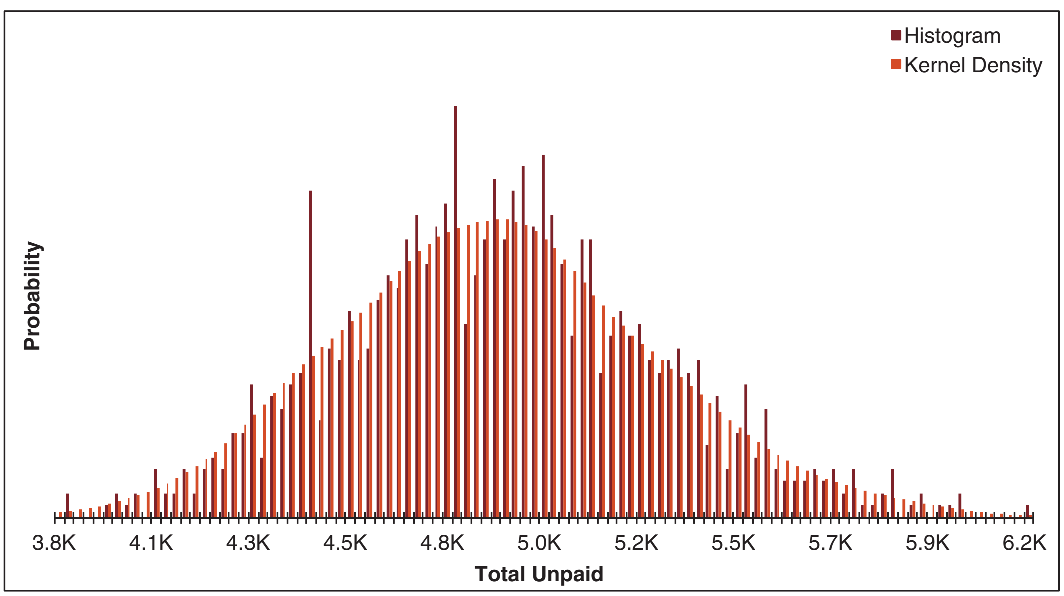


Figure A.11. Estimated Unpaid Model Results (Incurred Cape Cod)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Incurred Cape Cod Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	3	10	326.5%	-	117	0	1	17	50
2008	33	31	95.6%	-	213	24	46	91	148
2009	45	40	89.0%	1	317	33	61	122	184
2010	71	52	72.7%	3	375	58	91	174	251
2011	115	68	59.5%	16	512	102	146	242	366
2012	199	100	50.2%	31	933	181	252	388	499
2013	385	216	56.2%	46	1,629	343	477	812	1,081
2014	871	407	46.7%	132	3,029	802	1,049	1,658	2,191
2015	3,430	1,352	39.4%	1,074	9,190	3,253	3,977	5,946	7,972
Totals	5,151	1,417	27.5%	2,424	11,216	4,972	5,790	7,785	9,512
Normal Dist.	5,151	1,417	27.5%			5,151	6,107	7,482	8,448
logNormal Dist.	5,150	1,404	27.3%			4,969	5,953	7,719	9,264
Gamma Dist.	5,151	1,417	27.5%			5,022	6,023	7,682	9,007

Figure A.12. Total Unpaid Claims Distribution (Incurred Cape Cod)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Total Unpaid Distribution
Incurred Cape Cod Model

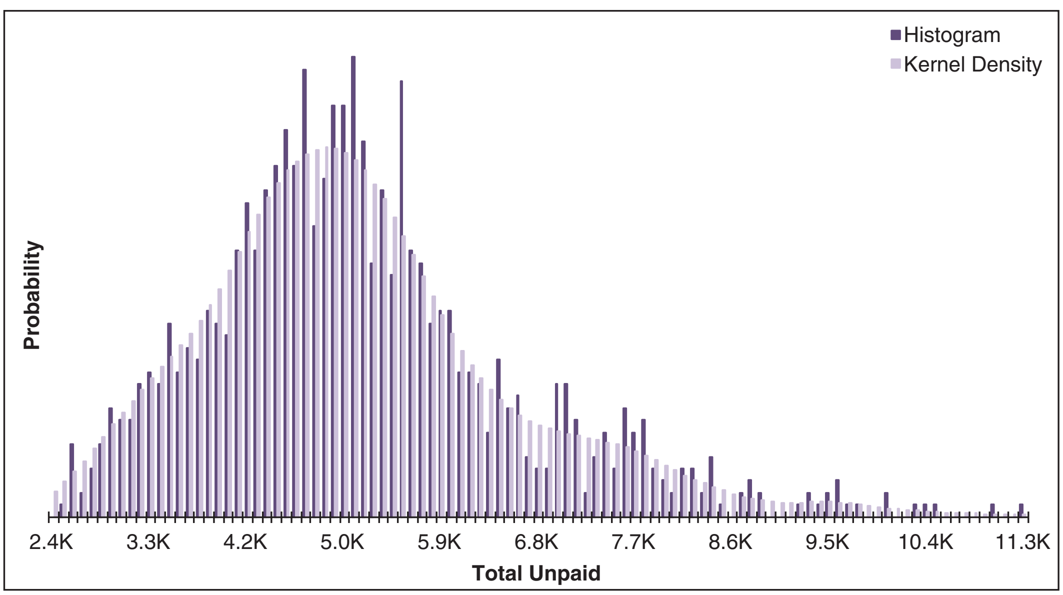


Figure A.13. Estimated Unpaid Model Results (Paid GLM)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Paid GLM Bootstrap Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	9	8	86.4%	0	53	7	13	24	32
2008	27	47	177.0%	0	436	12	24	109	253
2009	40	48	119.3%	2	537	27	44	117	270
2010	62	49	78.5%	11	525	51	69	136	287
2011	106	54	51.3%	31	559	94	117	202	347
2012	213	72	33.6%	79	731	201	242	333	455
2013	280	78	27.9%	100	707	271	325	418	507
2014	646	131	20.3%	337	1,368	634	730	871	979
2015	3,738	335	9.0%	2,696	4,939	3,731	3,953	4,307	4,583
Totals	5,120	447	8.7%	3,766	6,807	5,090	5,411	5,877	6,293
Normal Dist.	5,120	447	8.7%			5,120	5,422	5,856	6,161
logNormal Dist.	5,120	446	8.7%			5,101	5,409	5,886	6,246
Gamma Dist.	5,120	447	8.7%			5,107	5,415	5,878	6,218

Figure A.14. Total Unpaid Claims Distribution (Paid GLM)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Total Unpaid Distribution
Paid GLM Bootstrap Model

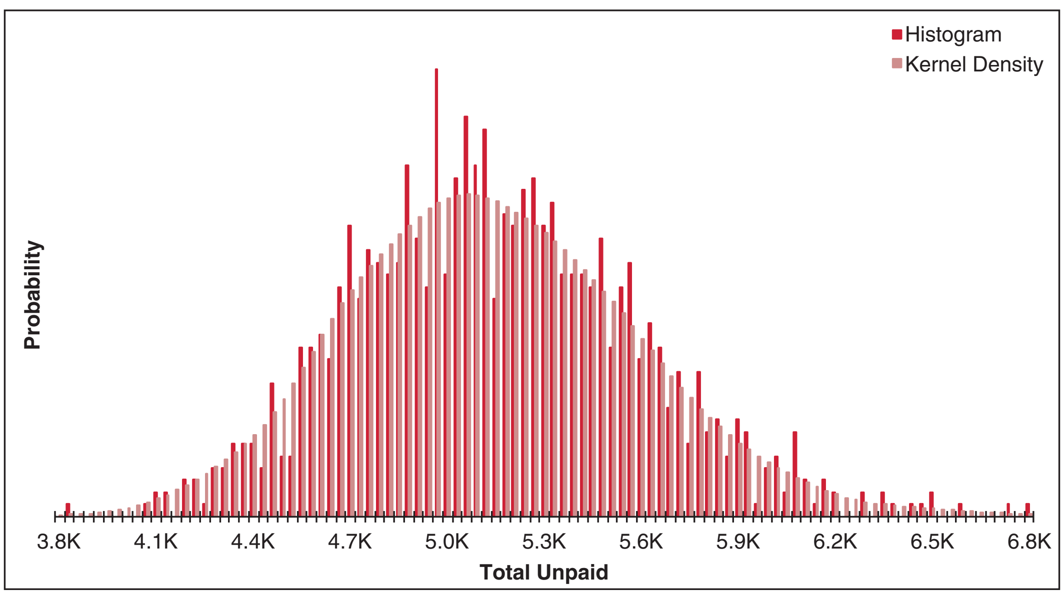


Figure A.15. Estimated Unpaid Model Results (Incurred GLM)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Incurred GLM Bootstrap Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	12	11	91.5%	0	66	8	16	34	48
2008	27	51	184.0%	0	520	12	25	111	262
2009	45	54	118.9%	3	678	31	52	117	268
2010	73	57	78.1%	11	892	59	85	150	301
2011	113	57	50.9%	30	771	101	128	215	360
2012	169	70	41.5%	53	712	153	198	288	415
2013	307	107	34.9%	93	1,550	293	362	491	615
2014	650	171	26.3%	280	2,713	630	743	928	1,057
2015	4,255	682	16.0%	2,581	6,888	4,216	4,670	5,413	6,295
Totals	5,650	751	13.3%	3,707	8,639	5,586	6,137	6,960	7,650
Normal Dist.	5,650	751	13.3%			5,650	6,157	6,886	7,398
logNormal Dist.	5,650	749	13.3%			5,601	6,123	6,960	7,616
Gamma Dist.	5,650	751	13.3%			5,617	6,137	6,940	7,543

Figure A.16. Total Unpaid Claims Distribution (Incurred GLM)

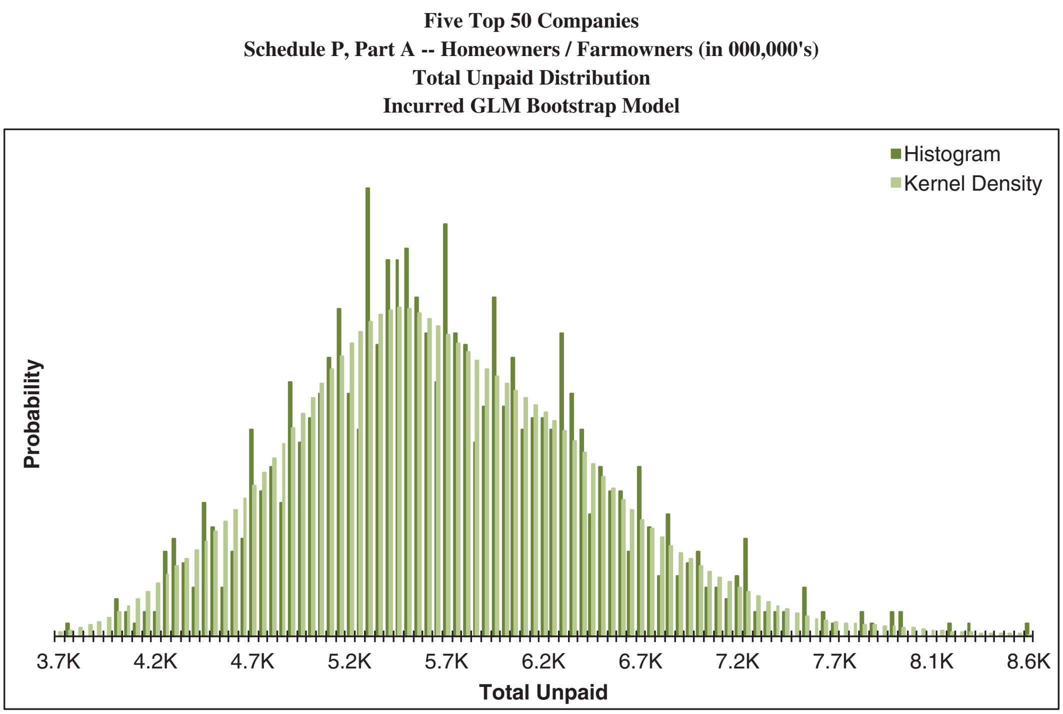


Figure A.17. Model Weights by Accident Year

Accident Year	Model Weights by Accident Year								
	Paid CL	Incd CL	Paid BF	Incd BF	Paid CC	Incd CC	Paid GLM	Incd GLM	TOTAL
2006	50.0%	50.0%							100.0%
2007	50.0%	50.0%							100.0%
2008	50.0%	50.0%							100.0%
2009	50.0%	50.0%							100.0%
2010	50.0%	50.0%							100.0%
2011	50.0%	50.0%							100.0%
2012	50.0%	50.0%							100.0%
2013	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%			100.0%
2014	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%			100.0%
2015	16.7%	16.7%			16.7%	16.7%	16.7%	16.7%	100.0%

Figure A.18. Estimated Mean Unpaid by Model

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Summary of Results by Model

Accident Year	Mean Estimated Unpaid								
	Chain Ladder		Bornhuetter-Ferguson		Cape Cod		GLM Bootstrap		Best Est.
	Paid	Incurred	Paid	Incurred	Paid	Incurred	Paid	Incurred	(Weighted)
2006	-	-	-	-	-	-	-	-	-
2007	3	3	2	2	3	3	9	12	3
2008	41	42	28	27	32	33	27	27	41
2009	45	46	37	39	43	45	40	45	46
2010	63	62	60	59	66	71	62	73	64
2011	103	103	96	98	109	115	106	113	103
2012	222	226	169	168	191	199	213	169	224
2013	294	306	327	334	373	385	280	307	335
2014	679	723	722	753	835	871	646	650	752
2015	3,851	3,912	2,660	2,885	3,225	3,430	3,738	4,255	3,742
Totals	5,300	5,422	4,099	4,366	4,878	5,151	5,120	5,650	5,308

Figure A.19. Estimated Ranges

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Summary of Results by Model

Accident Year	Best Est. (Weighted)	Ranges			
		Weighted		Modeled	
		Minimum	Maximum	Mininum	Maximum
2006	-				
2007	3	3	3	2	12
2008	41	41	42	27	42
2009	46	45	46	37	46
2010	64	62	63	59	73
2011	103	103	103	96	115
2012	224	222	226	168	226
2013	335	294	385	280	385
2014	752	679	871	646	871
2015	3,742	3,225	4,255	2,660	4,255
Totals	5,308	4,674	5,992	4,099	5,650

Figure A.20. Reconciliation of Total Results (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Reconciliation of Total Results
Best Estimate (Weighted)

Accident Year	Paid To Date	Incurred To Date	Case Reserves	IBNR	Estimate of Ultimate	Estimate of Unpaid
2006	5,234	5,237	3	(3)	5,234	-
2007	6,470	6,479	9	(6)	6,473	3
2008	7,848	7,867	19	23	7,890	41
2009	7,020	7,046	26	20	7,066	46
2010	7,291	7,341	50	13	7,355	64
2011	8,134	8,225	91	12	8,237	103
2012	10,800	11,085	285	(61)	11,023	224
2013	7,522	7,810	288	46	7,856	335
2014	7,968	8,703	735	17	8,720	752
2015	9,309	12,788	3,478	263	13,051	3,742
Totals	77,596	82,580	4,984	324	82,905	5,308

Figure A.21. Estimated Unpaid Model Results (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Unpaid
Best Estimate (Weighted)

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	3	9	292.0%	-	173	0	1	17	42
2008	41	37	88.6%	-	391	32	57	111	168
2009	46	37	81.0%	1	522	36	60	114	175
2010	64	41	63.6%	4	537	55	81	139	205
2011	103	50	48.8%	10	636	94	125	193	276
2012	224	89	40.0%	36	917	211	266	382	529
2013	335	148	44.3%	25	1,460	315	401	594	865
2014	752	293	39.0%	106	2,881	725	873	1,265	1,789
2015	3,742	982	26.2%	1,094	10,700	3,654	4,118	5,392	7,059
Totals	5,308	1,044	19.7%	2,116	12,445	5,224	5,758	7,074	8,675
Normal Dist.	5,308	1,044	19.7%			5,308	6,013	7,026	7,738
logNormal Dist.	5,309	1,034	19.5%			5,211	5,935	7,158	8,164
Gamma Dist.	5,308	1,044	19.7%			5,240	5,971	7,135	8,035

Figure A.22. Estimated Cash Flow (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Calendar Year Unpaid
Best Estimate (Weighted)

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2016	3,475	754	21.7%	1,297	8,420	3,414	3,797	4,730	5,948
2017	865	208	24.0%	293	2,148	843	982	1,224	1,483
2018	403	118	29.4%	115	1,298	387	467	614	740
2019	204	67	32.7%	56	654	194	240	325	412
2020	140	50	35.9%	40	539	132	165	233	297
2021	90	43	47.4%	12	611	82	112	169	229
2022	70	44	63.2%	6	409	60	91	152	215
2023	51	58	112.2%	-	735	36	75	151	253
2024	10	15	146.5%	-	199	4	15	41	67
Totals	5,308	1,044	19.7%	2,116	12,445	5,224	5,758	7,074	8,675

Figure A.23. Estimated Loss Ratio (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Ultimate Loss Ratios
Best Estimate (Weighted)

Accident Year	Mean Loss Ratio	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	67.7%	28.5%	42.1%	0.4%	220.8%	66.1%	71.1%	130.9%	158.2%
2007	79.3%	30.2%	38.1%	8.2%	262.2%	77.8%	83.1%	145.5%	178.5%
2008	90.5%	31.2%	34.5%	16.9%	261.3%	89.0%	94.6%	159.9%	188.9%
2009	72.8%	26.8%	36.7%	10.2%	215.6%	71.4%	76.1%	131.7%	180.4%
2010	65.3%	23.3%	35.7%	10.2%	225.0%	63.8%	68.0%	116.1%	139.7%
2011	64.1%	21.2%	33.1%	13.0%	190.0%	63.2%	67.0%	111.8%	130.5%
2012	80.5%	24.0%	29.9%	25.0%	234.6%	79.0%	83.7%	132.9%	154.6%
2013	54.7%	18.8%	34.4%	9.9%	157.7%	53.9%	57.4%	96.2%	115.1%
2014	58.0%	19.2%	33.0%	13.0%	164.8%	57.1%	60.6%	99.8%	118.8%
2015	88.2%	21.5%	24.4%	30.9%	232.5%	85.5%	92.5%	127.9%	158.7%
Totals	71.3%	7.4%	10.4%	46.6%	112.7%	70.8%	75.7%	84.4%	91.7%

Figure A.24. Estimated Unpaid Claim Runoff (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Calendar Year Unpaid Claim Runoff
Best Estimate (Weighted)

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2015	5,308	1,044	19.7%	2,116	12,445	5,224	5,758	7,074	8,675
2016	1,834	365	19.9%	746	4,128	1,797	2,030	2,459	2,957
2017	969	218	22.5%	336	2,316	946	1,088	1,353	1,627
2018	566	146	25.8%	159	1,393	548	647	828	1,004
2019	362	114	31.5%	79	1,171	347	424	565	718
2020	222	92	41.4%	35	956	207	269	386	524
2021	132	76	57.6%	6	863	117	166	268	394
2022	62	59	96.3%	(0)	745	46	84	166	269
2023	10	15	146.5%	(0)	199	4	15	41	67

Figure A.25. Mean Of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident	Mean Values									
Year	12	24	36	48	60	72	84	96	108	120 +
2006	3,776	1,139	218	95	41	21	12	6	25	2
2007	4,635	1,398	268	115	51	25	15	7	31	3
2008	5,647	1,701	327	141	61	31	17	9	38	4
2009	5,065	1,525	294	126	56	28	16	8	34	3
2010	5,318	1,602	307	132	57	29	17	8	36	3
2011	5,882	1,774	340	145	64	32	18	9	40	4
2012	7,909	2,378	457	197	86	43	25	12	53	5
2013	5,589	1,683	323	156	68	35	20	10	42	4
2014	6,197	1,870	392	168	73	37	21	10	46	4
2015	9,615	2,744	521	222	92	53	33	20	47	10

Figure A.26. Standard Deviation of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident Year	Standard Error Values 12	24	36	48	60	72	84	96	108	120 +
2006	1,597	502	119	64	33	11	7	2	23	4
2007	1,779	550	129	68	36	12	8	3	26	9
2008	1,960	603	147	77	38	13	8	3	35	10
2009	1,873	576	139	73	38	13	8	3	34	9
2010	1,906	596	143	75	38	13	9	3	34	9
2011	1,952	610	147	76	40	14	9	3	37	10
2012	2,375	733	173	92	49	17	11	4	44	11
2013	1,938	599	142	88	45	16	10	4	38	10
2014	2,054	639	173	90	47	16	10	4	41	11
2015	2,342	727	178	101	51	20	16	13	57	15

Figure A.27. Coefficient of Variation of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident Year	Coefficients of Variation 12	24	36	48	60	72	84	96	108	120 +
2006	42.3%	44.1%	54.4%	67.8%	80.8%	52.6%	58.8%	43.2%	89.8%	157.5%
2007	38.4%	39.3%	48.2%	59.5%	71.1%	47.7%	52.9%	38.7%	82.4%	292.0%
2008	34.7%	35.5%	44.8%	54.5%	62.5%	43.0%	47.9%	34.9%	92.6%	266.2%
2009	37.0%	37.8%	47.3%	58.1%	68.6%	45.4%	50.3%	37.7%	98.4%	272.2%
2010	35.8%	37.2%	46.5%	56.8%	66.1%	44.8%	52.4%	36.5%	95.5%	279.7%
2011	33.2%	34.4%	43.1%	52.8%	62.4%	42.5%	49.3%	34.0%	92.6%	267.9%
2012	30.0%	30.8%	37.8%	46.6%	57.2%	38.4%	44.6%	30.6%	82.8%	234.2%
2013	34.7%	35.6%	43.9%	56.5%	66.6%	45.8%	51.2%	37.4%	91.1%	250.9%
2014	33.2%	34.2%	44.2%	53.4%	64.1%	43.4%	49.6%	36.0%	88.9%	253.1%
2015	24.4%	26.5%	34.2%	45.3%	55.8%	37.2%	47.4%	66.2%	120.2%	146.5%

Figure A.28. Total Unpaid Claims Distribution (Weighted)

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Total Unpaid Distribution
Best Estimate (Weighted)

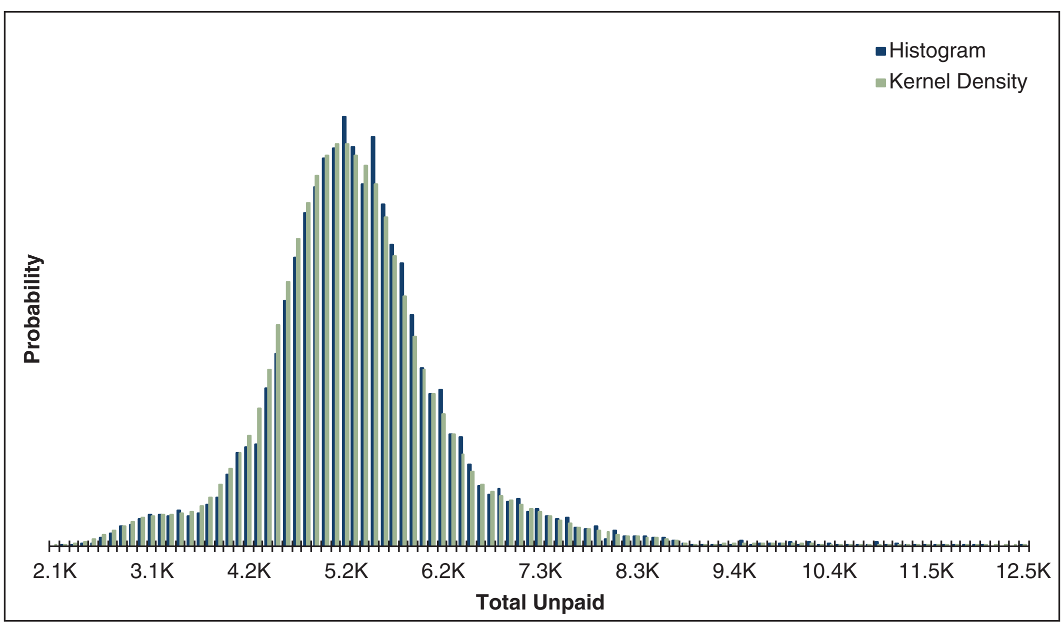
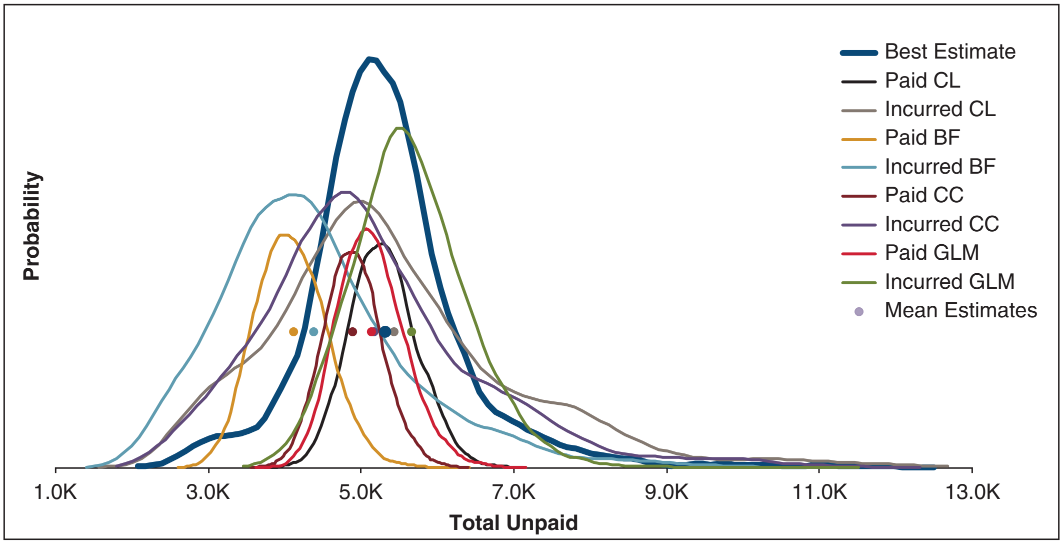


Figure A.29. Summary of Model Distributions

Five Top 50 Companies
Schedule P, Part A -- Homeowners / Farmowners (in 000,000's)
Summary of Model Distributions
(Using Kernel Densities)



Appendix B—Schedule P, Part B Results

In this appendix the results for Schedule P, Part B (Private Passenger Auto Liability) are shown.

Figure B.1. Estimated Unpaid Model Results (Paid Chain Ladder)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Paid Chain Ladder Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	59	23	38.8%	-	125	58	75	97	112
2007	90	25	27.3%	26	164	90	107	131	147
2008	135	27	19.9%	64	217	134	153	178	196
2009	214	32	14.8%	128	322	213	237	265	289
2010	339	31	9.2%	252	443	340	361	390	413
2011	586	38	6.6%	459	707	585	610	651	687
2012	1,109	51	4.6%	949	1,281	1,108	1,144	1,191	1,226
2013	2,089	75	3.6%	1,868	2,329	2,090	2,140	2,211	2,252
2014	3,917	127	3.3%	3,457	4,357	3,919	4,002	4,129	4,203
2015	8,033	219	2.7%	7,335	8,667	8,042	8,175	8,399	8,532
Totals	16,573	385	2.3%	15,252	17,728	16,581	16,842	17,192	17,399
Normal Dist.	16,573	385	2.3%			16,573	16,833	17,207	17,469
logNormal Dist.	16,573	386	2.3%			16,569	16,831	17,216	17,491
Gamma Dist.	16,573	385	2.3%			16,570	16,831	17,212	17,482

Figure B.2. Total Unpaid Claims Distribution (Paid Chain Ladder)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Total Unpaid Distribution
Paid Chain Ladder Model

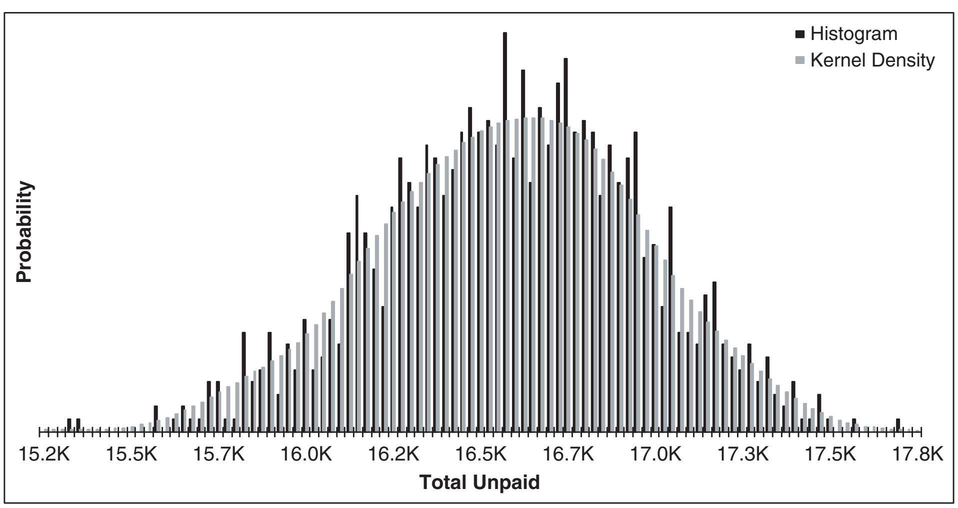


Figure B.3. Estimated Unpaid Model Results (Incurred Chain Ladder)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Incurred Chain Ladder Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	58	27	46.8%	-	156	56	77	103	131
2007	89	31	34.9%	17	212	87	108	146	170
2008	135	37	27.5%	48	278	133	159	196	226
2009	213	46	21.8%	106	397	210	246	290	326
2010	343	63	18.4%	178	560	342	387	445	492
2011	590	106	18.0%	304	886	590	661	764	823
2012	1,125	196	17.4%	610	2,320	1,133	1,265	1,439	1,502
2013	2,133	370	17.4%	1,167	3,115	2,165	2,404	2,722	2,846
2014	4,025	680	16.9%	2,324	5,470	4,078	4,514	5,076	5,298
2015	8,343	1,369	16.4%	4,886	12,352	8,502	9,290	10,413	10,940
Totals	17,054	1,620	9.5%	11,558	21,439	17,111	18,280	19,534	20,583
Normal Dist.	17,054	1,620	9.5%			17,054	18,147	19,719	20,824
logNormal Dist.	17,055	1,653	9.7%			16,976	18,120	19,902	21,257
Gamma Dist.	17,054	1,620	9.5%			17,003	18,117	19,804	21,048

Figure B.4. Total Unpaid Claims Distribution (Incurred Chain Ladder)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Total Unpaid Distribution
Incurred Chain Ladder Model

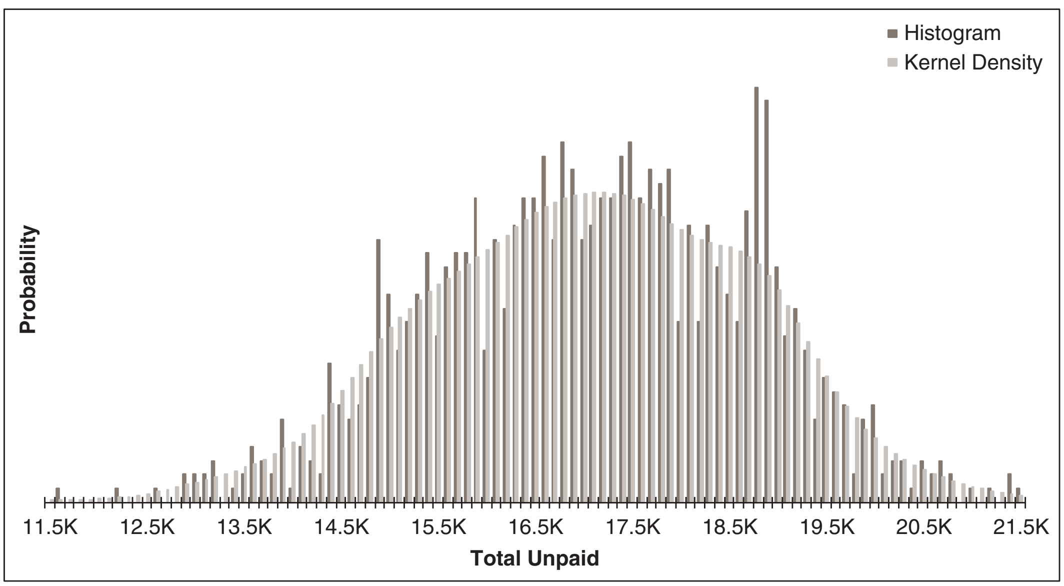


Figure B.5. Estimated Unpaid Model Results (Paid Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Paid Bornhuetter-Ferguson Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	54	22	40.2%	-	126	54	68	91	109
2007	76	22	28.7%	22	157	77	90	112	130
2008	112	24	21.2%	52	189	112	127	154	171
2009	188	30	16.0%	97	295	188	208	238	258
2010	343	36	10.4%	227	472	343	366	404	429
2011	625	50	8.0%	459	819	624	657	709	747
2012	1,162	77	6.7%	910	1,386	1,160	1,212	1,289	1,353
2013	2,217	134	6.1%	1,855	2,666	2,215	2,312	2,450	2,536
2014	3,942	218	5.5%	3,304	4,750	3,937	4,083	4,308	4,444
2015	7,990	441	5.5%	6,885	9,426	7,988	8,271	8,763	9,066
Totals	16,709	562	3.4%	15,239	18,369	16,701	17,096	17,695	18,035
Normal Dist.	16,709	562	3.4%			16,709	17,088	17,633	18,016
logNormal Dist.	16,709	561	3.4%			16,700	17,083	17,648	18,057
Gamma Dist.	16,709	562	3.4%			16,703	17,085	17,644	18,043

Figure B.6. Total Unpaid Claims Distribution (Paid Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Total Unpaid Distribution
Paid Bornhuetter-Ferguson Model

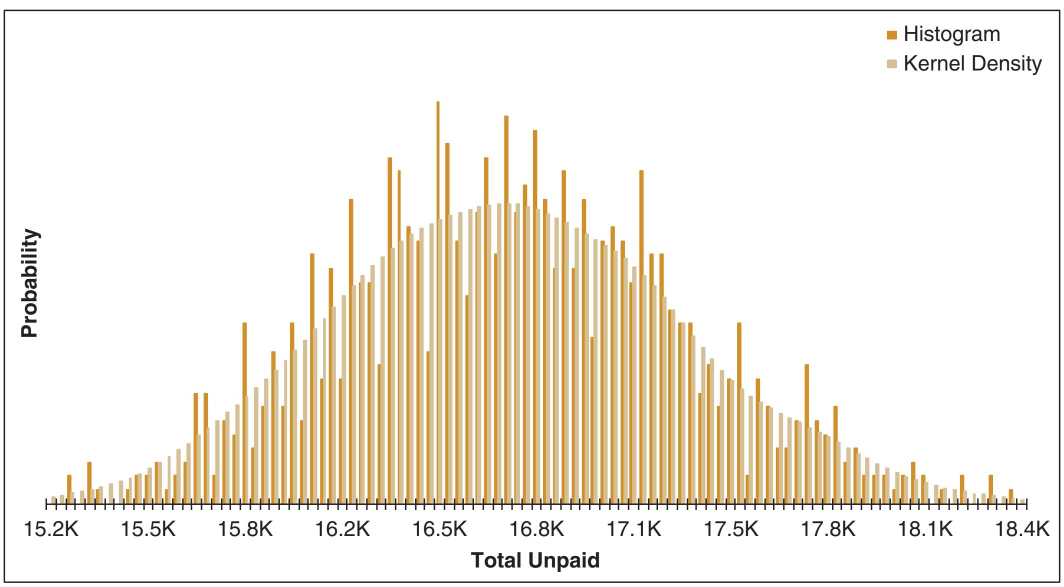


Figure B.7. Estimated Unpaid Model Results (Incurred Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Incurred Bornhuetter-Ferguson Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	54	24	45.4%	-	155	52	68	97	121
2007	76	25	33.1%	13	181	74	92	120	141
2008	111	30	27.2%	42	213	108	132	165	187
2009	188	42	22.5%	78	337	187	215	261	295
2010	344	68	19.7%	142	577	347	391	455	502
2011	627	116	18.5%	319	979	626	709	816	888
2012	1,167	217	18.6%	614	2,121	1,175	1,309	1,517	1,655
2013	2,234	420	18.8%	1,124	5,710	2,270	2,517	2,855	3,060
2014	3,997	689	17.2%	2,017	5,678	4,025	4,470	5,113	5,363
2015	8,289	1,370	16.5%	2,250	11,646	8,398	9,216	10,412	10,925
Totals	17,088	1,617	9.5%	10,942	22,273	17,177	18,198	19,785	20,539
Normal Dist.	17,088	1,617	9.5%			17,088	18,178	19,747	20,849
logNormal Dist.	17,089	1,648	9.6%			17,010	18,150	19,926	21,277
Gamma Dist.	17,088	1,617	9.5%			17,037	18,149	19,831	21,072

Figure B.8. Total Unpaid Claims Distribution (Incurred Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Total Unpaid Distribution
Incurred Bornhuetter-Ferguson Model

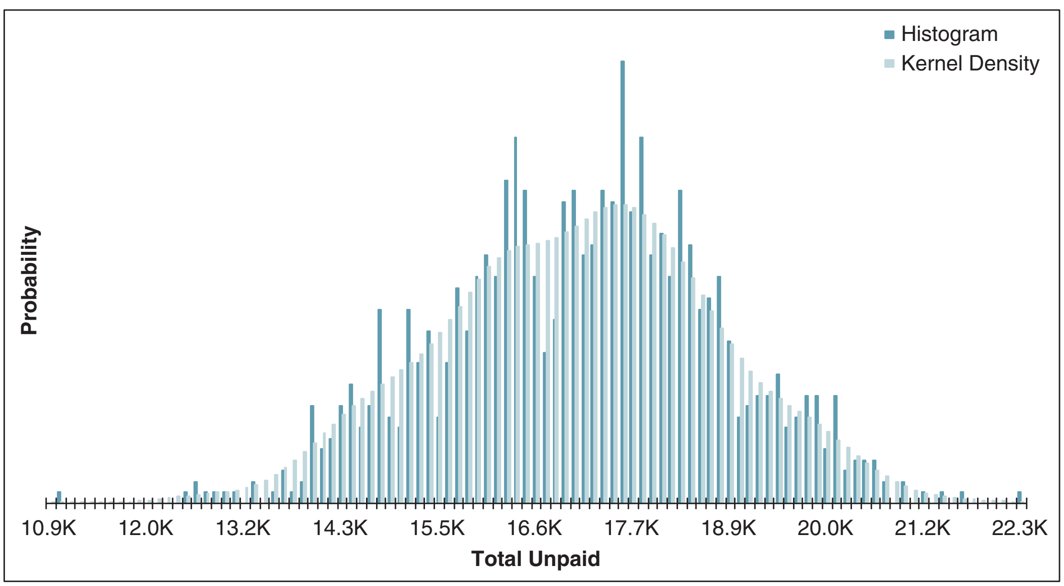


Figure B.9. Estimated Unpaid Model Results (Paid Cape Cod)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Paid Cape Cod Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	55	23	41.2%	-	136	55	70	94	108
2007	80	23	28.9%	23	161	79	95	118	133
2008	117	24	20.9%	57	205	117	134	159	175
2009	196	30	15.5%	116	305	195	216	247	270
2010	354	34	9.5%	263	459	353	377	410	436
2011	642	42	6.5%	513	773	642	670	710	738
2012	1,197	54	4.5%	1,042	1,365	1,198	1,234	1,288	1,331
2013	2,292	80	3.5%	2,045	2,553	2,294	2,345	2,424	2,474
2014	4,145	118	2.9%	3,761	4,502	4,145	4,219	4,345	4,439
2015	8,598	172	2.0%	8,057	9,073	8,596	8,711	8,894	8,987
Totals	17,676	376	2.1%	16,428	18,791	17,675	17,929	18,306	18,488
Normal Dist.	17,676	376	2.1%			17,676	17,930	18,295	18,551
logNormal Dist.	17,676	377	2.1%			17,672	17,928	18,302	18,570
Gamma Dist.	17,676	376	2.1%			17,674	17,929	18,300	18,563

Figure B.10. Total Unpaid Claims Distribution (Paid Cape Cod)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Total Unpaid Distribution
Paid Cape Cod Model

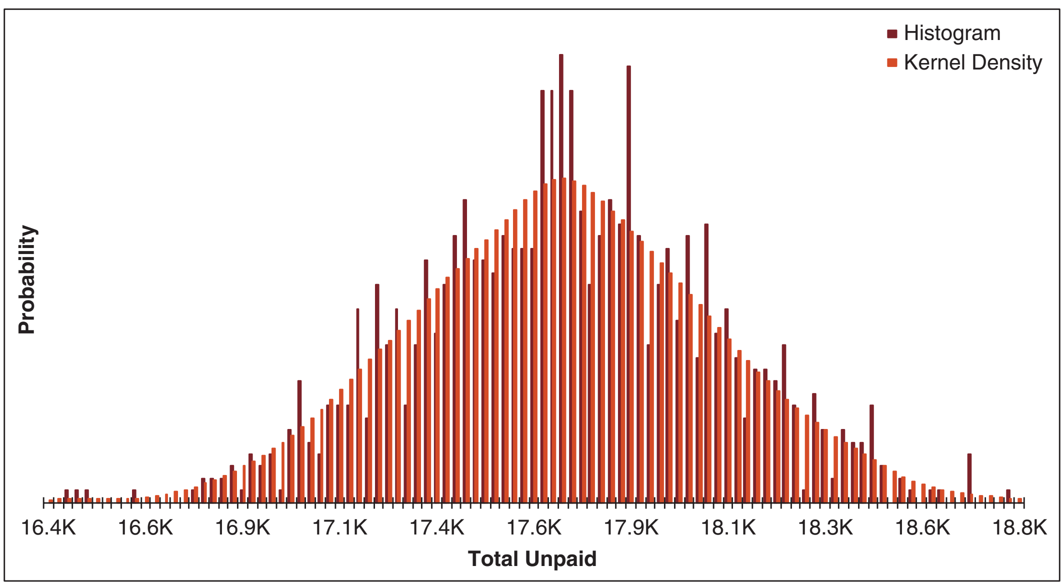


Figure B.11. Estimated Unpaid Model Results (Incurred Cape Cod)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Incurred Cape Cod Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	56	24	43.7%	-	133	54	71	96	114
2007	80	27	33.8%	18	175	79	98	126	147
2008	118	32	27.4%	35	230	116	138	175	203
2009	197	44	22.5%	90	351	194	228	271	309
2010	358	69	19.3%	184	544	358	407	474	509
2011	650	115	17.6%	324	953	647	730	843	900
2012	1,201	213	17.7%	675	1,697	1,224	1,352	1,534	1,632
2013	2,308	388	16.8%	1,247	3,598	2,335	2,579	2,939	3,074
2014	4,178	701	16.8%	2,248	5,709	4,247	4,697	5,271	5,516
2015	8,526	1,424	16.7%	4,605	11,643	8,725	9,508	10,707	11,151
Totals	17,672	1,677	9.5%	12,794	21,955	17,649	18,915	20,366	21,243
Normal Dist.	17,672	1,677	9.5%			17,672	18,803	20,430	21,573
logNormal Dist.	17,673	1,706	9.7%			17,591	18,772	20,610	22,008
Gamma Dist.	17,672	1,677	9.5%			17,619	18,772	20,517	21,805

Figure B.12. Total Unpaid Claims Distribution (Incurred Cape Cod)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Total Unpaid Distribution
Incurred Cape Cod Model

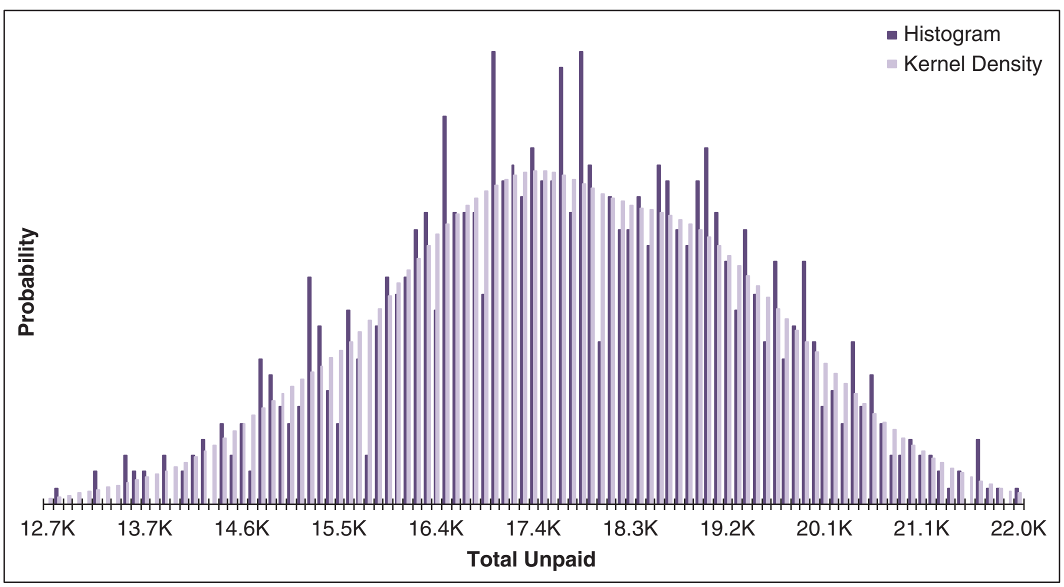


Figure B.13. Estimated Unpaid Model Results (Paid GLM)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Paid GLM Bootstrap Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	29	15	53.7%	2	106	26	37	58	79
2007	56	23	40.9%	7	158	53	69	98	120
2008	99	29	29.6%	29	223	96	116	151	179
2009	177	33	18.5%	99	317	173	198	233	260
2010	302	32	10.7%	200	450	299	324	356	377
2011	552	34	6.2%	465	740	550	573	613	643
2012	1,071	53	5.0%	914	1,288	1,067	1,107	1,162	1,197
2013	2,053	78	3.8%	1,831	2,295	2,052	2,106	2,180	2,244
2014	3,879	118	3.0%	3,525	4,361	3,875	3,955	4,080	4,177
2015	8,004	229	2.9%	7,329	8,746	7,999	8,165	8,380	8,509
Totals	16,222	369	2.3%	15,169	17,945	16,200	16,473	16,833	17,164
Normal Dist.	16,222	369	2.3%			16,222	16,471	16,829	17,080
logNormal Dist.	16,222	367	2.3%			16,218	16,468	16,834	17,096
Gamma Dist.	16,222	369	2.3%			16,220	16,469	16,833	17,092

Figure B.14. Total Unpaid Claims Distribution (Paid GLM)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Total Unpaid Distribution
Paid GLM Bootstrap Model

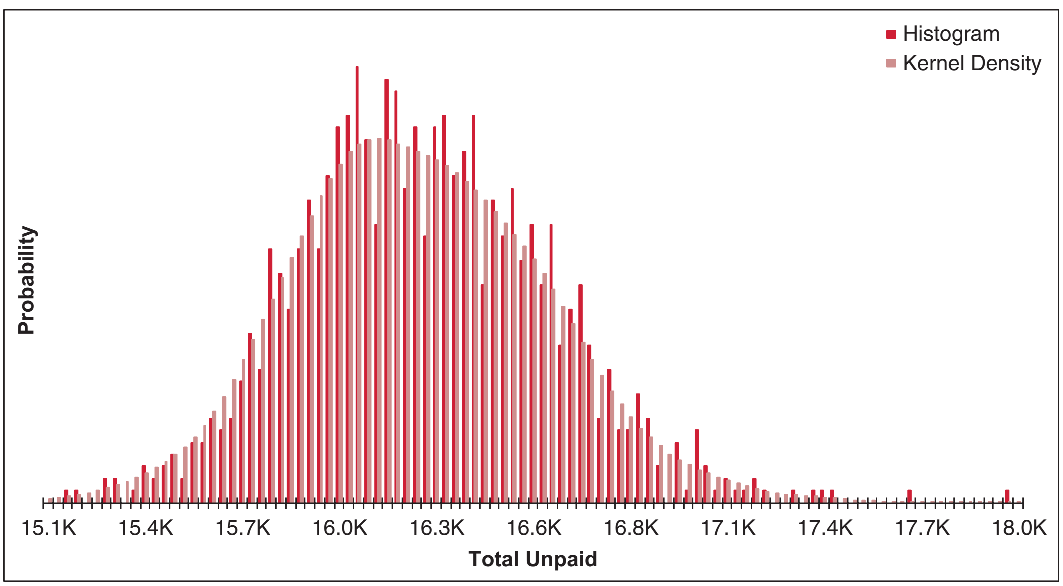


Figure B.15. Estimated Unpaid Model Results (Incurred GLM)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Incurred GLM Bootstrap Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	28	15	55.2%	3	110	25	35	58	76
2007	56	24	42.7%	7	178	53	69	102	138
2008	107	33	30.8%	43	298	101	127	168	200
2009	172	34	19.6%	91	301	169	191	235	263
2010	295	36	12.4%	204	419	290	316	361	394
2011	568	49	8.6%	434	764	565	597	652	702
2012	1,130	90	8.0%	857	1,422	1,126	1,189	1,285	1,332
2013	2,193	168	7.7%	1,738	2,884	2,193	2,307	2,468	2,605
2014	4,058	319	7.9%	3,096	5,040	4,063	4,294	4,573	4,764
2015	8,390	723	8.6%	5,922	10,670	8,375	8,917	9,524	9,986
Totals	16,996	985	5.8%	13,965	19,871	16,965	17,696	18,619	19,079
Normal Dist.	16,996	985	5.8%			16,996	17,660	18,616	19,287
logNormal Dist.	16,996	989	5.8%			16,967	17,645	18,669	19,424
Gamma Dist.	16,996	985	5.8%			16,977	17,649	18,647	19,371

Figure B.16. Total Unpaid Claims Distribution (Incurred GLM)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Total Unpaid Distribution
Incurred GLM Bootstrap Model

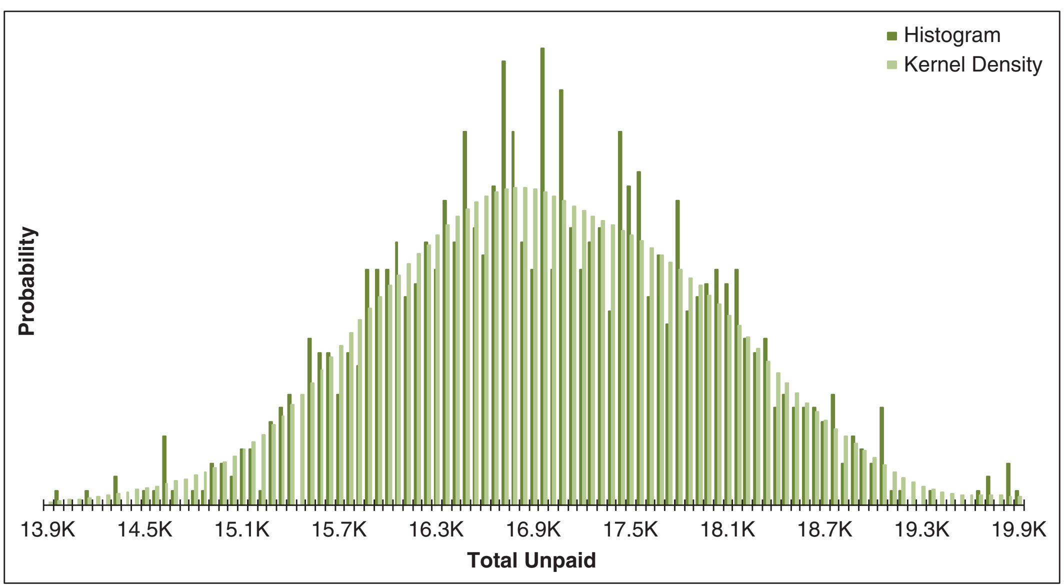


Figure B.17. Model Weights by Accident Year

Accident Year	Model Weights by Accident Year								
	Paid CL	Incd CL	Paid BF	Incd BF	Paid CC	Incd CC	Paid GLM	Incd GLM	TOTAL
2006	50.0%	50.0%							100.0%
2007	50.0%	50.0%							100.0%
2008	50.0%	50.0%							100.0%
2009	50.0%	50.0%							100.0%
2010			25.0%	25.0%	25.0%	25.0%			100.0%
2011			25.0%	25.0%	25.0%	25.0%			100.0%
2012			25.0%	25.0%	25.0%	25.0%			100.0%
2013			25.0%	25.0%	25.0%	25.0%			100.0%
2014		25.0%		25.0%		25.0%		25.0%	100.0%
2015		25.0%		25.0%		25.0%		25.0%	100.0%

Figure B.18. Estimated Mean Unpaid by Model

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Summary of Results by Model

Accident Year	Mean Estimated Unpaid								
	Chain Ladder		Bornhuetter-Ferguson		Cape Cod		GLM Bootstrap		Best Est.
	Paid	Incurred	Paid	Incurred	Paid	Incurred	Paid	Incurred	(Weighted)
2006	59	58	54	54	55	56	29	28	59
2007	90	89	76	76	80	80	56	56	90
2008	135	135	112	111	117	118	99	107	134
2009	214	213	188	188	196	197	177	172	214
2010	339	343	343	344	354	358	302	295	351
2011	586	590	625	627	642	650	552	568	636
2012	1,109	1,125	1,162	1,167	1,197	1,201	1,071	1,130	1,184
2013	2,089	2,133	2,217	2,234	2,292	2,308	2,053	2,193	2,255
2014	3,917	4,025	3,942	3,997	4,145	4,178	3,879	4,058	4,077
2015	8,033	8,343	7,990	8,289	8,598	8,526	8,004	8,390	8,394
Totals	16,573	17,054	16,709	17,088	17,676	17,672	16,222	16,996	17,395

Figure B.19. Estimated Ranges

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Summary of Results by Model

Accident Year	Best Est.	Ranges			
		Weighted		Modeled	
	(Weighted)	Minimum	Maximum	Mininum	Maximum
2006	59	58	59	28	59
2007	90	89	90	56	90
2008	134	135	135	107	135
2009	214	213	214	172	214
2010	351	343	358	295	358
2011	636	625	650	568	650
2012	1,184	1,162	1,201	1,109	1,201
2013	2,255	2,217	2,308	2,089	2,308
2014	4,077	3,997	4,178	3,917	4,178
2015	8,394	8,289	8,526	7,990	8,598
Totals	17,395	17,127	17,720	16,573	17,676

Figure B.20. Reconciliation of Total Results (Weighted)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Reconciliation of Total Results
Best Estimate (Weighted)

Accident Year	Paid To Date	Incurred To Date	Case Reserves	IBNR	Estimate of Ultimate	Estimate of Unpaid
2006	11,816	11,863	47	12	11,875	59
2007	12,679	12,752	72	18	12,770	90
2008	13,631	13,743	112	22	13,765	134
2009	14,472	14,687	216	(1)	14,686	214
2010	13,717	14,079	362	(11)	14,068	351
2011	13,090	13,691	600	36	13,726	636
2012	12,490	13,683	1,193	(9)	13,674	1,184
2013	11,598	13,912	2,313	(58)	13,854	2,255
2014	10,306	14,625	4,319	(243)	14,383	4,077
2015	6,357	15,188	8,830	(437)	14,751	8,394
Totals	120,157	138,223	18,066	(671)	137,551	17,395

Figure B.21. Estimated Unpaid Model Results (Weighted)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Unpaid
Best Estimate (Weighted)

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	59	25	42.2%	-	178	58	75	102	122
2007	90	28	30.8%	17	221	89	109	137	161
2008	134	32	24.0%	41	297	133	156	189	215
2009	214	41	18.9%	73	401	213	240	284	321
2010	351	55	15.6%	160	600	350	383	444	492
2011	636	91	14.2%	314	1,020	636	684	794	867
2012	1,184	157	13.3%	(27)	1,857	1,188	1,260	1,465	1,597
2013	2,255	293	13.0%	1,073	5,710	2,267	2,389	2,781	2,982
2014	4,077	616	15.1%	833	6,049	4,097	4,460	5,120	5,398
2015	8,394	1,234	14.7%	980	12,352	8,468	9,175	10,444	10,911
Totals	17,395	1,428	8.2%	10,057	23,150	17,439	18,375	19,729	20,525
Normal Dist.	17,395	1,428	8.2%			17,395	18,358	19,744	20,717
logNormal Dist.	17,395	1,451	8.3%			17,335	18,336	19,879	21,040
Gamma Dist.	17,395	1,428	8.2%			17,356	18,336	19,809	20,889

Figure B.22. Estimated Cash Flow (Weighted)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Calendar Year Unpaid
Best Estimate (Weighted)

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2016	8,275	715	8.6%	4,501	10,746	8,299	8,761	9,426	9,838
2017	4,072	340	8.4%	2,450	5,608	4,079	4,304	4,621	4,845
2018	2,266	198	8.7%	1,319	3,149	2,267	2,397	2,590	2,718
2019	1,210	109	9.0%	699	1,574	1,210	1,285	1,389	1,461
2020	638	58	9.1%	405	885	638	677	735	778
2021	358	35	9.8%	203	511	358	381	416	439
2022	217	30	13.7%	95	351	216	237	267	291
2023	144	25	17.2%	57	258	144	161	186	205
2024	99	23	23.4%	16	214	98	114	139	157
2025	67	22	33.1%	-	157	66	81	106	124
2026	32	13	40.8%	-	91	32	41	55	66
2027	16	9	57.3%	-	57	15	22	31	38
Totals	17,395	1,428	8.2%	10,057	23,150	17,439	18,375	19,729	20,525

Figure B.23. Estimated Loss Ratio (Weighted)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Ultimate Loss Ratios
Best Estimate (Weighted)

Accident Year	Mean Loss Ratio	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	75.6%	9.7%	12.9%	38.9%	104.5%	75.8%	77.8%	94.6%	99.4%
2007	82.2%	10.3%	12.5%	43.9%	114.0%	82.4%	84.6%	102.2%	107.6%
2008	83.9%	10.2%	12.1%	45.1%	114.4%	83.9%	86.3%	103.7%	108.6%
2009	79.6%	9.2%	11.6%	45.2%	108.3%	79.7%	81.8%	97.8%	102.7%
2010	69.3%	8.2%	11.9%	37.9%	94.6%	69.1%	71.1%	85.3%	90.1%
2011	66.0%	8.1%	12.3%	35.2%	89.7%	66.0%	67.9%	81.7%	85.9%
2012	66.9%	8.1%	12.1%	-1.5%	94.6%	66.9%	68.8%	82.7%	86.8%
2013	66.9%	8.1%	12.2%	35.2%	186.1%	66.9%	68.9%	82.4%	86.3%
2014	71.9%	10.6%	14.7%	14.4%	101.9%	72.7%	78.5%	89.0%	93.5%
2015	73.0%	10.6%	14.5%	8.4%	110.0%	73.9%	79.8%	90.3%	94.2%
Totals	72.9%	3.0%	4.1%	61.6%	90.9%	73.0%	75.0%	77.7%	79.5%

Figure B.24. Estimated Unpaid Claim Runoff (Weighted)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Calendar Year Unpaid Claim Runoff
Best Estimate (Weighted)

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2015	17,395	1,428	8.2%	10,057	23,150	17,439	18,375	19,729	20,525
2016	9,120	739	8.1%	5,556	12,446	9,136	9,623	10,325	10,767
2017	5,048	419	8.3%	3,106	6,838	5,054	5,330	5,738	6,000
2018	2,782	243	8.7%	1,709	3,689	2,781	2,945	3,184	3,360
2019	1,572	157	10.0%	902	2,165	1,570	1,675	1,838	1,951
2020	934	117	12.6%	494	1,387	930	1,011	1,131	1,224
2021	576	94	16.3%	247	988	573	638	733	807
2022	359	75	21.0%	104	687	356	408	488	546
2023	214	59	27.6%	30	467	211	252	317	365
2024	115	41	36.0%	(0)	283	112	142	188	222
2025	48	20	42.4%	(0)	137	47	62	84	101
2026	16	9	57.3%	(0)	57	15	22	31	38
2027	(0)	0	-10502.4%	(0)	0	(0)	0	0	0

Figure B.25. Mean of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident	Mean Values												
Year	12	24	36	48	60	72	84	96	108	120	132	144	156 +
2006	5,232	3,354	1,456	842	457	224	113	58	32	25	30	15	15
2007	5,631	3,608	1,566	907	491	241	121	62	34	27	32	16	16
2008	6,082	3,902	1,691	981	530	261	131	67	37	29	34	17	17
2009	6,480	4,155	1,802	1,043	565	278	139	71	39	31	36	18	18
2010	6,225	3,992	1,732	1,002	543	267	138	71	39	31	36	18	18
2011	6,043	3,876	1,681	974	527	280	141	72	40	31	36	18	18
2012	6,008	3,851	1,671	968	560	274	138	71	39	31	36	18	18
2013	6,046	3,876	1,681	1,051	569	279	140	72	40	31	36	18	18
2014	6,453	4,138	1,821	1,055	572	281	141	72	41	30	32	17	16
2015	6,549	4,261	1,847	1,070	579	284	143	73	41	31	32	17	16

Figure B.26. Standard Deviation of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident	Standard Error Values												
Year	12	24	36	48	60	72	84	96	108	120	132	144	156 +
2006	677	440	199	115	65	35	15	12	4	3	12	6	6
2007	708	460	207	120	68	36	16	13	4	3	13	7	7
2008	742	484	217	126	70	38	16	13	5	4	14	7	7
2009	756	493	220	129	72	39	17	16	5	4	15	8	8
2010	745	485	218	127	71	38	18	16	5	4	15	8	8
2011	747	486	218	128	71	44	19	16	5	4	15	8	8
2012	729	475	213	124	78	42	18	16	5	4	15	8	8
2013	741	483	218	142	79	43	19	16	5	4	15	8	8
2014	955	618	282	165	92	47	22	18	8	7	17	9	9
2015	966	634	280	166	92	48	22	18	9	7	17	9	9

Figure B.27. Coefficient of Variation of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part B -- Private Passenger Auto Liability (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident	Coefficients of Variation												
Year	12	24	36	48	60	72	84	96	108	120	132	144	156 +
2006	12.9%	13.1%	13.6%	13.7%	14.2%	15.5%	13.4%	21.2%	12.9%	12.9%	42.1%	42.2%	42.3%
2007	12.6%	12.7%	13.2%	13.3%	13.8%	15.0%	13.0%	20.6%	12.5%	12.6%	42.0%	42.1%	42.2%
2008	12.2%	12.4%	12.8%	12.8%	13.3%	14.5%	12.6%	19.9%	12.2%	12.3%	42.3%	42.4%	42.5%
2009	11.7%	11.9%	12.2%	12.4%	12.7%	13.9%	12.1%	22.0%	11.7%	11.7%	42.0%	42.1%	42.2%
2010	12.0%	12.2%	12.6%	12.6%	13.1%	14.3%	13.1%	22.5%	12.5%	12.6%	42.5%	42.6%	42.7%
2011	12.4%	12.5%	13.0%	13.1%	13.5%	15.6%	13.4%	22.5%	12.9%	12.9%	42.5%	42.7%	42.8%
2012	12.1%	12.3%	12.8%	12.9%	13.9%	15.3%	13.2%	22.4%	12.6%	12.7%	42.3%	42.5%	42.5%
2013	12.3%	12.5%	13.0%	13.5%	13.9%	15.4%	13.2%	22.3%	12.7%	12.7%	42.0%	42.2%	42.2%
2014	14.8%	14.9%	15.5%	15.7%	16.1%	16.8%	15.4%	24.5%	20.8%	23.7%	52.6%	51.2%	57.5%
2015	14.7%	14.9%	15.2%	15.5%	15.9%	16.7%	15.2%	24.4%	20.6%	23.4%	52.1%	51.1%	57.3%

Figure B.28. Total Unpaid Claims Distribution (Weighted)

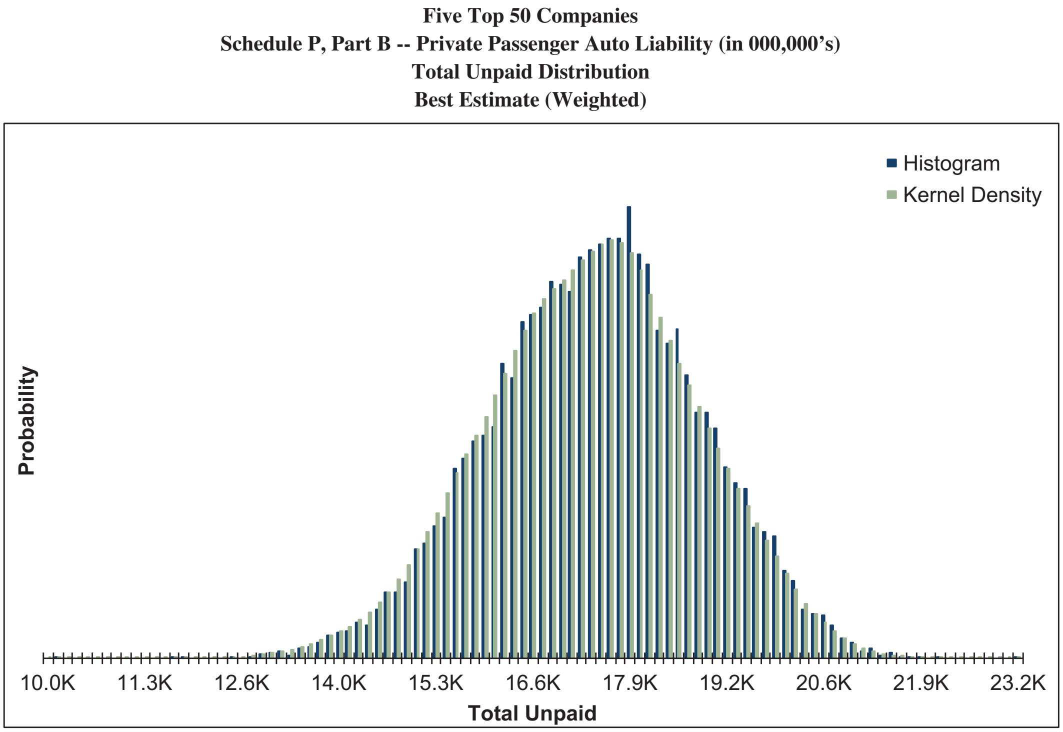
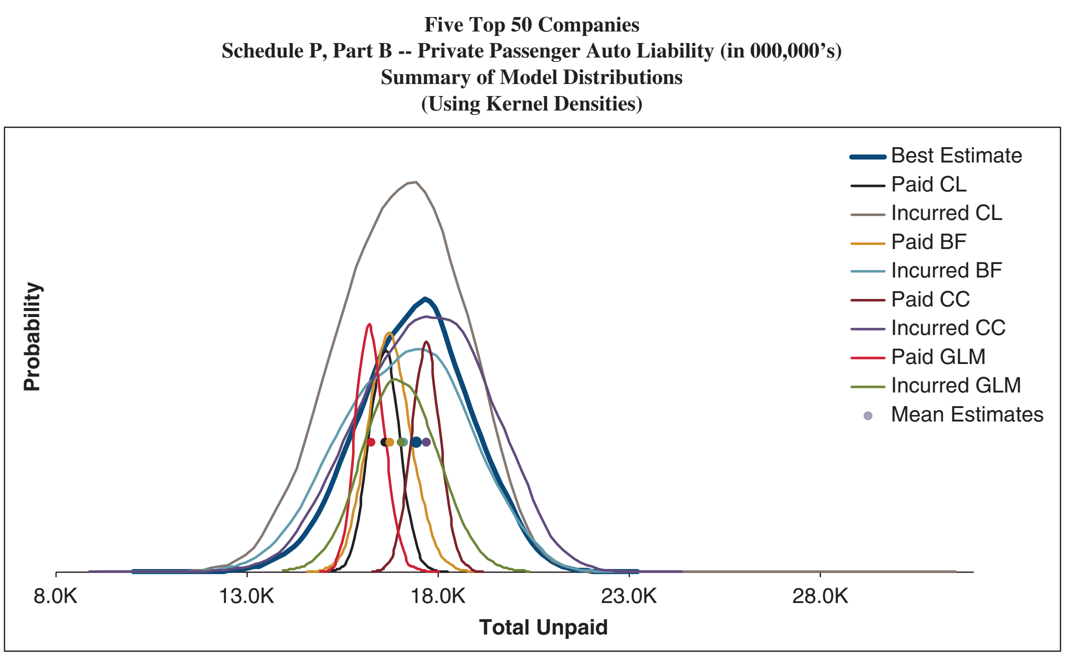


Figure B.29. Summary of Model Distributions



Appendix C—Schedule P, Part C Results

In this appendix the results for Schedule P, Part C (Commercial Auto Liability) are shown.

Figure C.1. Estimated Unpaid Model Results (Paid Chain Ladder)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Paid Chain Ladder Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	8	4	50.6%	-	22	8	10	15	19
2007	11	4	39.9%	(0)	28	10	13	18	22
2008	21	5	24.3%	7	43	21	24	29	34
2009	35	6	18.3%	18	66	34	39	46	51
2010	61	10	16.6%	34	97	60	67	80	87
2011	110	22	20.0%	57	195	107	124	150	173
2012	216	33	15.4%	111	359	215	237	273	296
2013	410	39	9.4%	294	550	408	434	474	513
2014	773	52	6.7%	610	946	770	806	863	901
2015	1,103	75	6.8%	872	1,345	1,100	1,152	1,232	1,285
Totals	2,746	122	4.4%	2,357	3,171	2,741	2,830	2,951	3,019
Normal Dist.	2,746	122	4.4%			2,746	2,828	2,946	3,029
logNormal Dist.	2,746	122	4.4%			2,743	2,827	2,951	3,041
Gamma Dist.	2,746	122	4.4%			2,744	2,827	2,949	3,037

Figure C.2. Total Unpaid Claims Distribution (Paid Chain Ladder)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Total Unpaid Distribution
Paid Chain Ladder Model

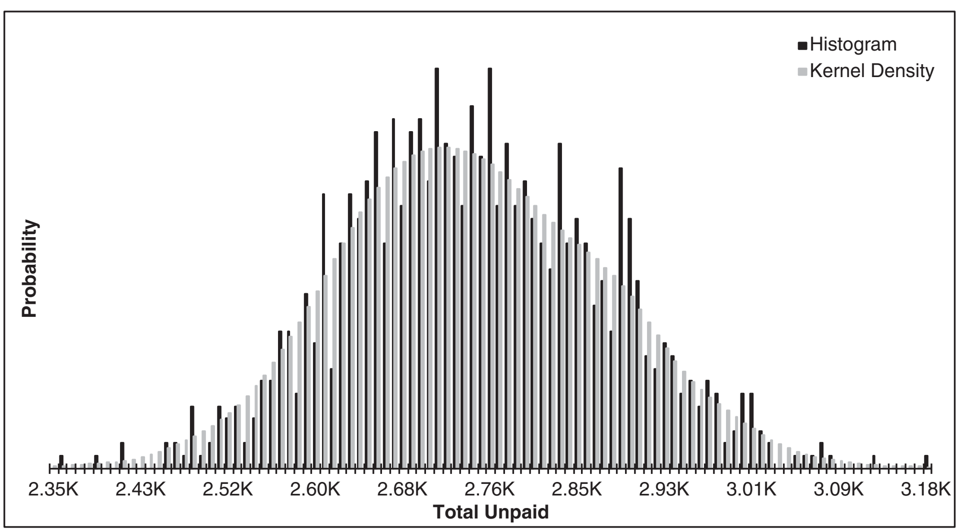


Figure C.3. Estimated Unpaid Model Results (Incurred Chain Ladder)

Five Top 50 Companies
Schedule P, Part C -- Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Incurred Chain Ladder Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	11	12	108.0%	-	74	7	16	35	48
2007	15	16	110.1%	0	157	9	22	46	66
2008	31	33	105.0%	-	354	23	47	91	127
2009	53	54	102.3%	-	533	41	86	144	200
2010	92	103	111.1%	-	1,654	69	145	258	369
2011	168	176	104.5%	-	1,625	127	264	498	681
2012	328	372	113.3%	-	4,031	217	528	963	1,307
2013	623	615	98.7%	-	3,767	484	1,049	1,782	2,238
2014	1,223	1,415	115.7%	-	21,802	1,019	2,010	3,319	4,335
2015	1,513	1,618	107.0%	-	13,830	1,062	2,546	4,356	5,798
Totals	4,056	2,421	59.7%	146	30,092	3,725	5,273	7,786	10,983
Normal Dist.	4,056	2,421	59.7%			4,056	5,689	8,038	9,687
logNormal Dist.	4,168	2,899	69.6%			3,422	5,227	9,616	14,755
Gamma Dist.	4,056	2,421	59.7%			3,586	5,328	8,677	11,670

Figure C.4. Total Unpaid Claims Distribution (Incurred Chain Ladder)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Total Unpaid Distribution
Incurred Chain Ladder Model

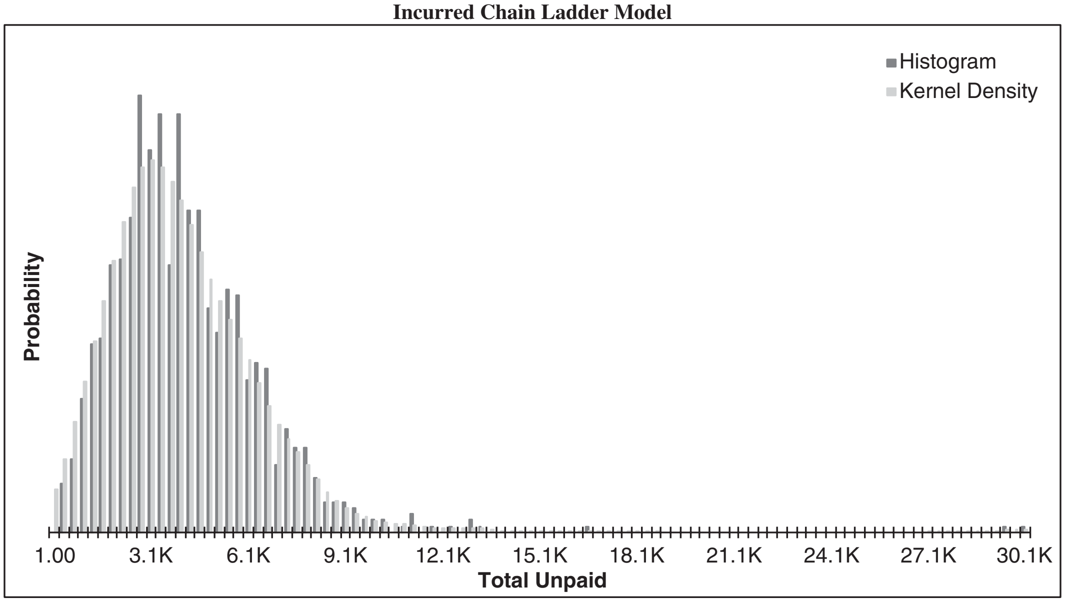


Figure C.5. Estimated Unpaid Model Results (Paid Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part C -- Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Paid Bornhuetter-Ferguson Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	5	3	54.4%	-	16	5	7	10	13
2007	8	3	42.3%	0	22	8	10	14	17
2008	17	4	26.7%	5	32	17	20	25	29
2009	35	7	19.3%	13	64	34	39	46	52
2010	65	11	17.1%	38	110	65	73	84	94
2011	123	25	20.5%	44	211	121	140	167	197
2012	259	40	15.6%	145	420	256	287	327	353
2013	481	52	10.8%	315	658	477	517	565	607
2014	812	76	9.3%	590	1,078	811	860	936	996
2015	1,132	100	8.9%	857	1,480	1,127	1,198	1,300	1,369
Totals	2,936	153	5.2%	2,472	3,474	2,939	3,040	3,180	3,313
Normal Dist.	2,936	153	5.2%			2,936	3,040	3,188	3,293
logNormal Dist.	2,936	154	5.2%			2,932	3,038	3,196	3,312
Gamma Dist.	2,936	153	5.2%			2,934	3,038	3,193	3,305

Figure C.6. Total Unpaid Claims Distribution (Paid Bornhuetter-Ferguson)

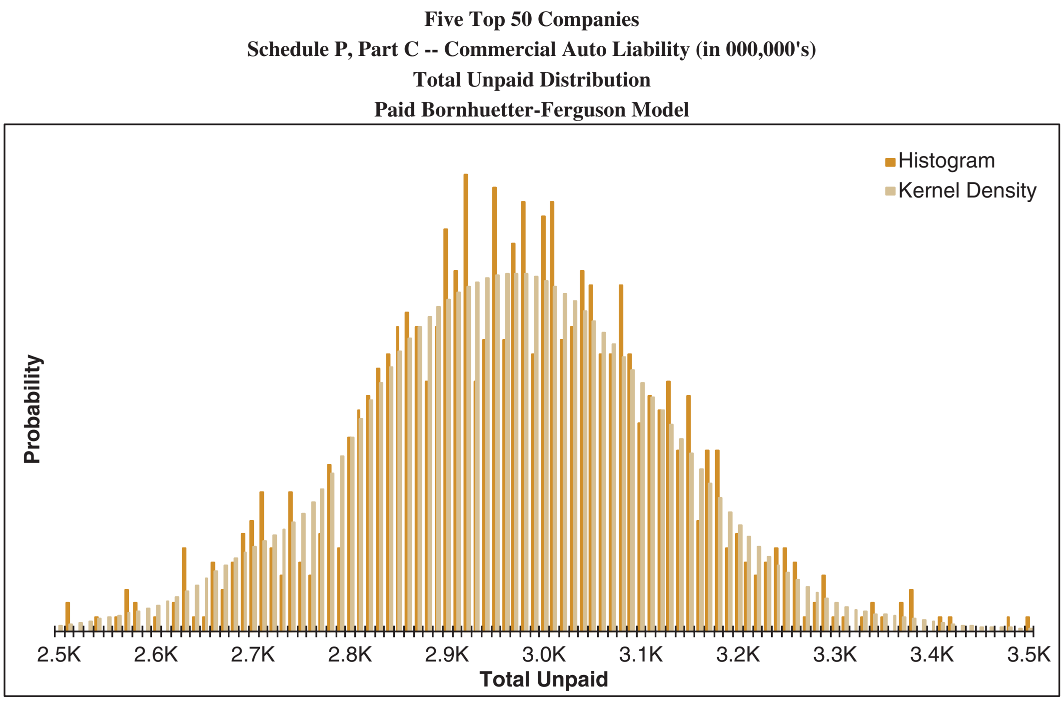


Figure C.7. Estimated Unpaid Model Results (Incurred Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Incurred Bornhuetter-Ferguson Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	7	8	116.0%	-	48	4	10	24	34
2007	11	12	110.5%	-	61	7	15	36	52
2008	24	23	96.0%	-	124	18	37	68	93
2009	49	45	92.9%	-	216	38	80	139	165
2010	99	88	88.8%	0	375	82	162	265	318
2011	176	164	93.3%	0	821	134	279	505	630
2012	362	338	93.5%	0	1,547	296	584	1,005	1,228
2013	642	597	93.1%	1	2,344	502	1,066	1,792	2,119
2014	1,118	996	89.1%	0	4,243	980	1,862	2,919	3,447
2015	1,554	1,409	90.6%	0	5,956	1,371	2,626	4,146	4,729
Totals	4,040	1,873	46.4%	387	10,575	3,901	5,304	7,418	8,445
Normal Dist.	4,040	1,873	46.4%			4,040	5,303	7,120	8,397
logNormal Dist.	4,116	2,390	58.1%			3,560	5,120	8,638	12,472
Gamma Dist.	4,040	1,873	46.4%			3,755	5,099	7,530	9,612

Figure C.8. Total Unpaid Claims Distribution (Incurred Bornhuetter-Ferguson)

Five Top 50 Companies
Schedule P, Part C -- Commercial Auto Liability (in 000,000's)
Total Unpaid Distribution
Incurred Bornhuetter-Ferguson Model

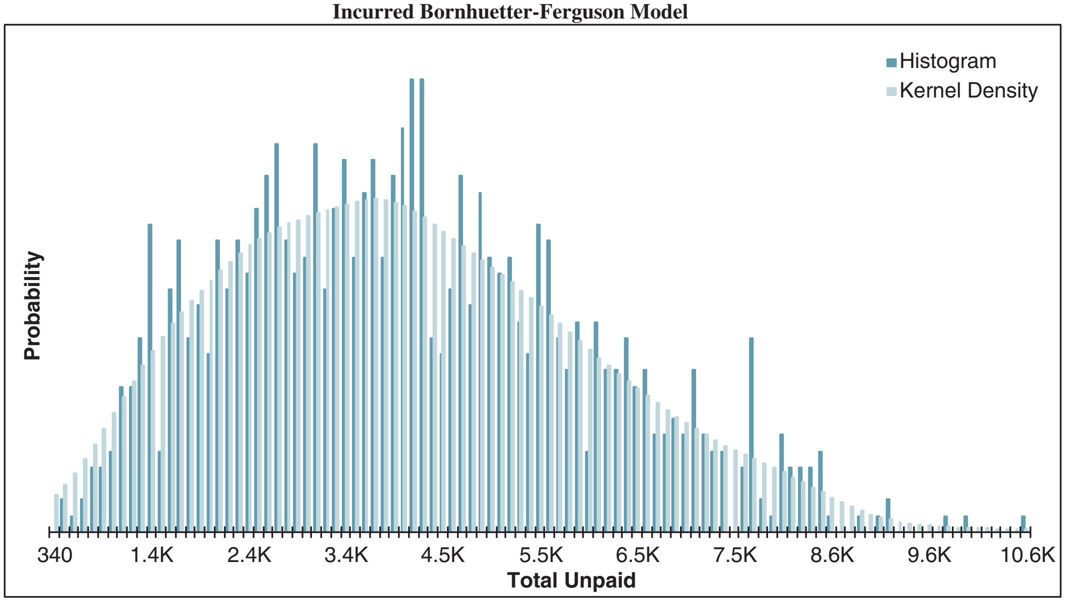


Figure C.9. Estimated Unpaid Model Results (Paid Cape Cod)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Paid Cape Cod Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	6	3	52.3%	-	17	6	8	12	14
2007	9	4	41.0%	0	26	9	11	15	19
2008	18	5	26.1%	7	34	18	22	27	31
2009	36	7	17.9%	20	59	36	41	48	52
2010	67	11	16.1%	39	101	66	74	86	94
2011	124	23	18.8%	67	245	122	138	163	192
2012	258	38	14.8%	166	416	255	283	323	359
2013	481	40	8.4%	363	629	478	509	548	583
2014	827	50	6.0%	684	975	827	858	915	948
2015	1,178	53	4.5%	990	1,348	1,176	1,212	1,268	1,308
Totals	3,004	122	4.0%	2,559	3,428	3,001	3,088	3,204	3,297
Normal Dist.	3,004	122	4.0%			3,004	3,086	3,204	3,286
logNormal Dist.	3,004	121	4.0%			3,001	3,084	3,208	3,297
Gamma Dist.	3,004	122	4.0%			3,002	3,085	3,206	3,294

Figure C.10. Total Unpaid Claims Distribution (Paid Cape Cod)

Five Top 50 Companies
Schedule P, Part C -- Commercial Auto Liability (in 000,000's)
Total Unpaid Distribution
Paid Cape Cod Model

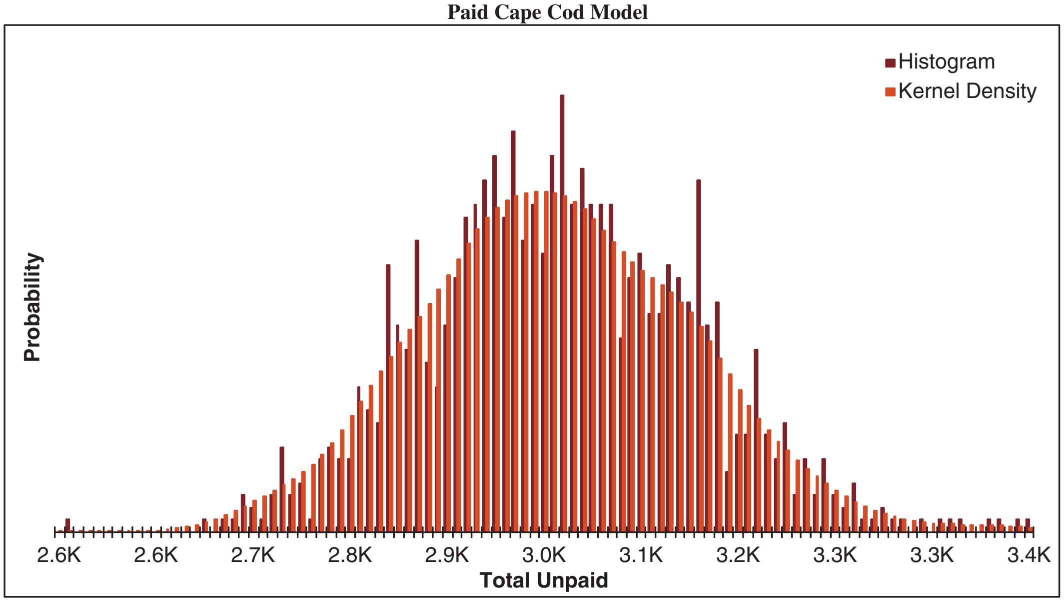


Figure C.11. Estimated Unpaid Model Results (Incurred Cape Cod)

Five Top 50 Companies
Schedule P, Part C -- Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Incurred Cape Cod Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	8	9	110.7%	-	62	5	12	25	36
2007	13	14	108.2%	-	98	9	19	43	60
2008	25	25	98.6%	0	185	18	40	76	99
2009	52	51	98.2%	0	481	37	82	145	201
2010	101	98	97.9%	0	1,082	81	160	267	339
2011	183	199	108.7%	0	3,031	140	282	515	644
2012	403	410	101.7%	0	4,350	320	637	1,106	1,514
2013	696	747	107.4%	0	11,739	577	1,110	1,930	2,405
2014	1,287	1,239	96.3%	0	20,322	1,121	2,045	3,306	4,162
2015	1,647	1,748	106.1%	1	31,078	1,408	2,694	4,317	5,401
Totals	4,415	3,174	71.9%	372	72,036	4,089	5,685	8,357	11,920
Normal Dist.	4,415	3,174	71.9%			4,415	6,555	9,635	11,797
logNormal Dist.	4,465	2,906	65.1%			3,743	5,588	9,947	14,914
Gamma Dist.	4,415	3,174	71.9%			3,682	5,956	10,581	14,867

Figure C.12. Total Unpaid Claims Distribution (Incurred Cape Cod)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Total Unpaid Distribution
Incurred Cape Cod Model

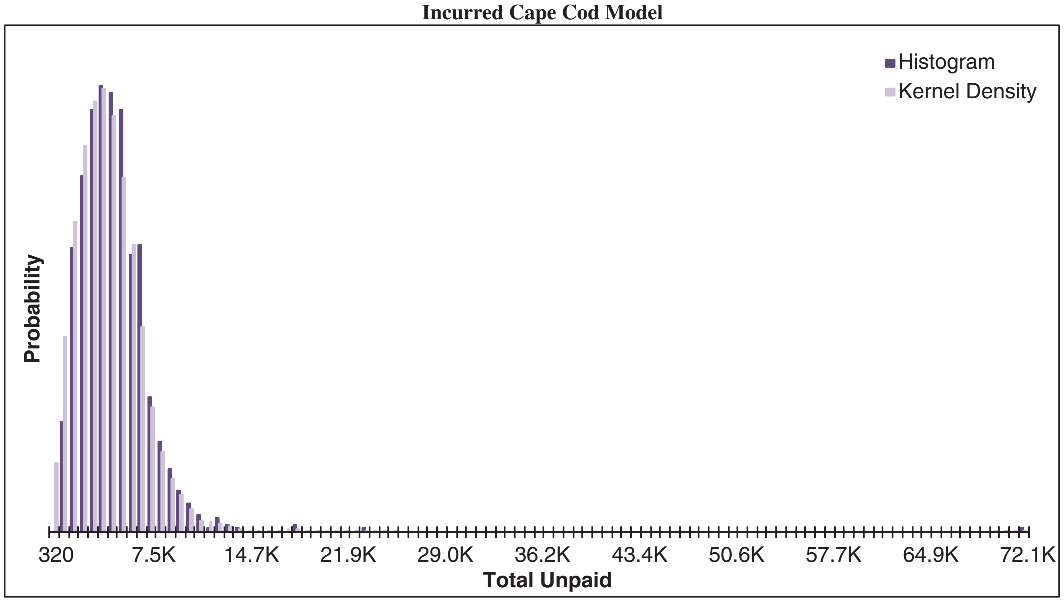


Figure C.13. Estimated Unpaid Model Results (Paid GLM)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Paid GLM Bootstrap Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	8	5	63.7%	(5)	33	7	10	17	23
2007	14	7	52.9%	(3)	52	12	18	27	33
2008	23	9	39.9%	(1)	72	22	29	39	49
2009	38	12	30.2%	8	90	38	45	58	70
2010	64	13	20.8%	27	112	64	73	88	100
2011	123	17	13.8%	81	178	122	135	152	162
2012	244	25	10.4%	169	331	243	261	286	305
2013	457	37	8.1%	361	577	455	480	520	543
2014	747	53	7.1%	597	926	749	784	831	870
2015	1,063	77	7.3%	851	1,346	1,060	1,112	1,192	1,259
Totals	2,781	188	6.8%	2,234	3,480	2,775	2,904	3,097	3,251
Normal Dist.	2,781	188	6.8%			2,781	2,907	3,090	3,218
logNormal Dist.	2,781	188	6.8%			2,774	2,903	3,100	3,246
Gamma Dist.	2,781	188	6.8%			2,776	2,905	3,097	3,237

Figure C.14. Total Unpaid Claims Distribution (Paid GLM)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Total Unpaid Distribution
Paid GLM Bootstrap Model

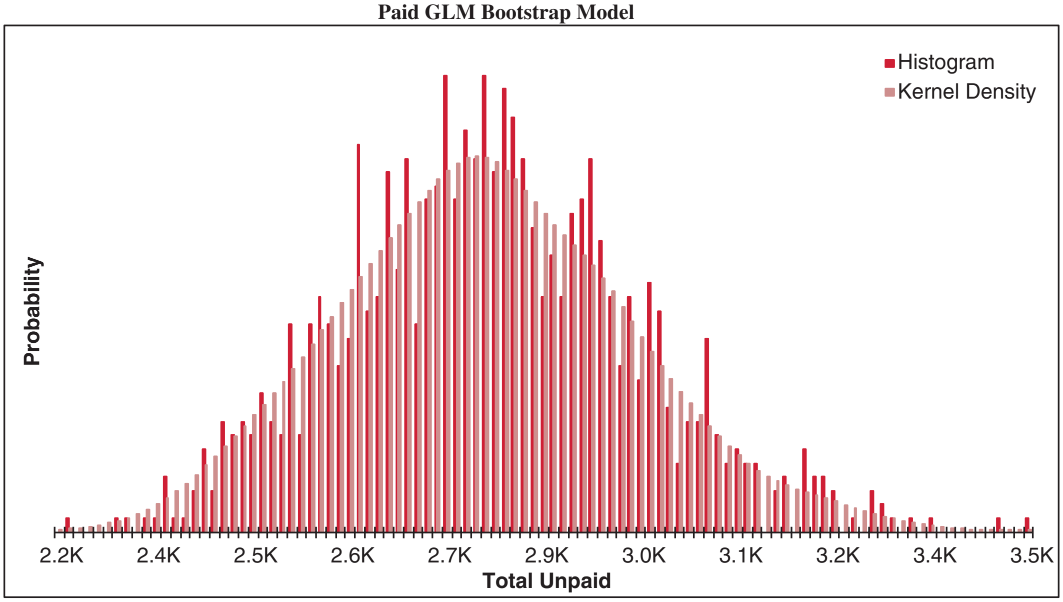


Figure C.15. Estimated Unpaid Model Results (Incurred GLM)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Incurred GLM Bootstrap Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	10	8	81.9%	(9)	57	8	13	25	39
2007	17	11	62.4%	(5)	65	15	22	39	53
2008	31	17	53.2%	(0)	104	29	40	63	82
2009	54	23	43.3%	7	177	50	67	97	119
2010	92	35	38.2%	17	251	88	113	153	184
2011	174	63	36.1%	23	378	171	217	278	333
2012	363	119	32.8%	76	773	360	443	572	648
2013	682	224	32.9%	100	1,490	666	833	1,078	1,211
2014	1,097	366	33.3%	267	2,346	1,084	1,334	1,716	2,055
2015	1,567	555	35.4%	452	4,027	1,515	1,899	2,536	3,071
Totals	4,087	760	18.6%	2,190	6,754	4,018	4,584	5,485	6,034
Normal Dist.	4,087	760	18.6%			4,087	4,599	5,336	5,854
logNormal Dist.	4,087	769	18.8%			4,017	4,555	5,460	6,200
Gamma Dist.	4,087	760	18.6%			4,040	4,570	5,411	6,058

Figure C.16. Total Unpaid Claims Distribution (Incurred GLM)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Total Unpaid Distribution
Incurred GLM Bootstrap Model

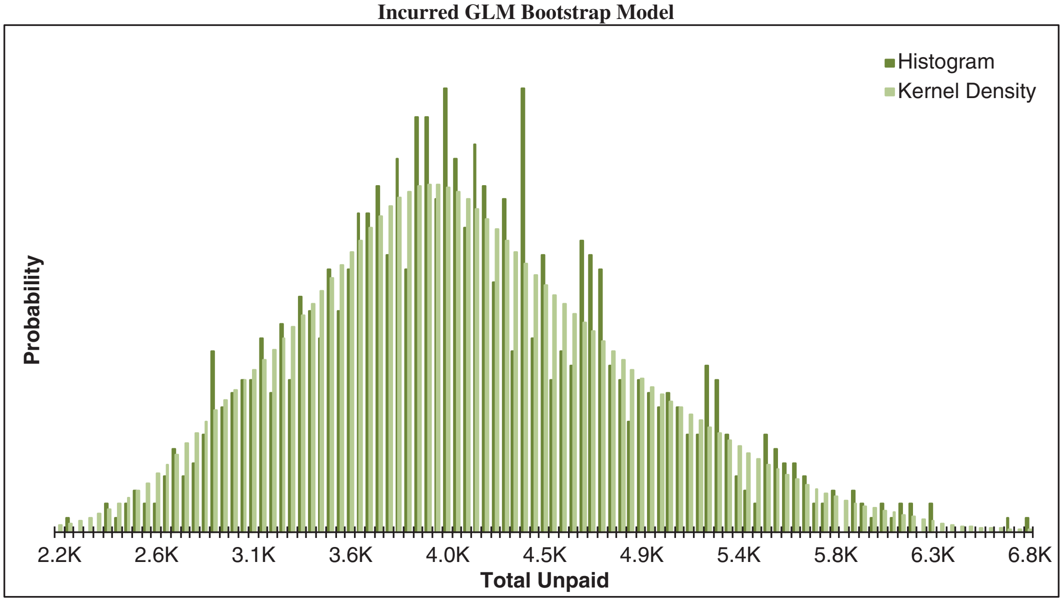


Figure C.17. Model Weights By Accident Year

Accident	Model Weights by Accident Year								
Year	Paid CL	Incd CL	Paid BF	Incd BF	Paid CC	Incd CC	Paid GLM	Incd GLM	TOTAL
2006							100.0%		100.0%
2007							100.0%		100.0%
2008							100.0%		100.0%
2009							100.0%		100.0%
2010			33.3%		33.3%		33.3%		100.0%
2011			33.3%		33.3%		33.3%		100.0%
2012			50.0%		50.0%				100.0%
2013			50.0%		50.0%				100.0%
2014	33.3%		33.3%		33.3%				100.0%
2015	33.3%		33.3%		33.3%				100.0%

Figure C.18. Estimated Mean Unpaid By Model

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Summary of Results by Model

Accident	Mean Estimated Unpaid								
	Chain Ladder		Bornhuetter-Ferguson		Cape Cod		GLM Bootstrap		Best Est.
Year	Paid	Incurred	Paid	Incurred	Paid	Incurred	Paid	Incurred	(Weighted)
2006	8	11	5	7	6	8	8	10	8
2007	11	15	8	11	9	13	14	17	13
2008	21	31	17	24	18	25	23	31	23
2009	35	53	35	49	36	52	38	54	38
2010	61	92	65	99	67	101	64	92	66
2011	110	168	123	176	124	183	123	174	124
2012	216	328	259	362	258	403	244	363	258
2013	410	623	481	642	481	696	457	682	480
2014	773	1,223	812	1,118	827	1,287	747	1,097	803
2015	1,103	1,513	1,132	1,554	1,178	1,647	1,063	1,567	1,134
Totals	2,746	4,056	2,936	4,040	3,004	4,415	2,781	4,087	2,947

Figure C.19. Estimated Ranges

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Summary of Results by Model

Accident		Ranges			
	Best Est.	Weighted		Modeled	
Year	(Weighted)	Minimum	Maximum	Mininum	Maximum
2006	8	8	8	5	8
2007	13	14	14	8	14
2008	23	23	23	17	23
2009	38	38	38	35	38
2010	66	64	67	61	67
2011	124	123	124	110	124
2012	258	258	259	216	259
2013	480	481	481	410	481
2014	803	773	827	747	827
2015	1,134	1,103	1,178	1,063	1,178
Totals	2,947	2,884	3,018	2,746	3,004

Figure C.20. Reconciliation of Total Results (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Reconciliation of Total Results
Best Estimate (Weighted)

Accident Year	Paid To Date	Incurred To Date	Case Reserves	IBNR	Estimate of Ultimate	Estimate of Unpaid
2006	1,563	1,577	14	(6)	1,571	8
2007	1,469	1,505	36	(23)	1,482	13
2008	1,387	1,436	49	(26)	1,410	23
2009	1,350	1,417	67	(29)	1,388	38
2010	1,342	1,445	102	(37)	1,408	66
2011	1,198	1,345	147	(24)	1,321	124
2012	1,061	1,339	278	(20)	1,318	258
2013	853	1,327	474	6	1,333	480
2014	645	1,442	797	6	1,448	803
2015	294	1,422	1,128	6	1,428	1,134
Totals	11,162	14,255	3,093	(146)	14,109	2,947

Figure C.21. Estimated Unpaid Model Results (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Unpaid
Best Estimate (Weighted)

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	8	5	65.8%	(8)	35	7	11	18	24
2007	13	7	51.3%	(7)	52	13	17	26	33
2008	23	9	39.4%	(5)	72	22	28	39	48
2009	38	11	28.9%	7	92	37	45	58	68
2010	66	12	17.8%	30	130	65	73	86	96
2011	124	22	17.6%	59	247	122	137	161	182
2012	258	40	15.4%	140	485	255	284	326	359
2013	480	47	9.8%	311	737	478	509	559	604
2014	803	65	8.1%	580	1,151	802	845	912	967
2015	1,134	83	7.3%	800	1,569	1,138	1,189	1,266	1,327
Totals	2,947	132	4.5%	2,471	3,532	2,947	3,036	3,162	3,257
Normal Dist.	2,947	132	4.5%			2,947	3,036	3,164	3,254
logNormal Dist.	2,947	132	4.5%			2,944	3,035	3,170	3,268
Gamma Dist.	2,947	132	4.5%			2,945	3,035	3,168	3,263

Figure C.22. Estimated Cash Flow (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Calendar Year Unpaid
Best Estimate (Weighted)

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2016	1,156	58	5.0%	937	1,378	1,155	1,194	1,254	1,299
2017	796	53	6.7%	611	993	795	832	886	927
2018	475	42	8.9%	332	668	474	503	547	580
2019	248	38	15.3%	129	410	246	273	315	342
2020	125	23	18.6%	60	260	123	139	165	187
2021	64	11	16.6%	25	110	63	71	82	91
2022	37	6	17.2%	15	71	37	41	48	53
2023	22	5	23.7%	5	52	21	25	30	35
2024	11	4	35.5%	(1)	31	10	13	17	21
2025	7	3	43.3%	-	28	7	9	13	16
2026	4	2	53.2%	-	17	3	5	7	9
2027	2	1	69.8%	-	11	2	2	4	6
2028	1	1	95.7%	-	9	1	1	3	4
Totals	2,947	132	4.5%	2,471	3,532	2,947	3,036	3,162	3,257

Figure C.23. Estimated Loss Ratio (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Ultimate Loss Ratios
Best Estimate (Weighted)

Accident Year	Mean Loss Ratio	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	88.5%	2.7%	3.0%	79.6%	98.3%	88.5%	90.3%	92.9%	94.7%
2007	82.9%	2.5%	3.0%	73.9%	92.3%	82.9%	84.6%	87.0%	88.6%
2008	74.9%	2.3%	3.1%	65.8%	83.0%	74.9%	76.5%	78.7%	80.3%
2009	60.3%	1.9%	3.2%	52.4%	67.3%	60.4%	61.7%	63.5%	64.7%
2010	55.0%	2.1%	3.9%	47.5%	62.5%	55.0%	56.5%	58.4%	59.7%
2011	54.3%	1.9%	3.5%	46.8%	62.7%	54.4%	55.7%	57.5%	58.7%
2012	51.8%	2.0%	3.9%	44.0%	61.8%	51.8%	53.1%	55.2%	56.7%
2013	54.1%	2.3%	4.2%	46.9%	64.8%	54.1%	55.6%	57.9%	59.8%
2014	58.3%	2.8%	4.9%	48.6%	72.0%	58.3%	60.1%	63.0%	65.1%
2015	59.9%	3.6%	6.1%	45.7%	77.7%	60.1%	62.4%	65.7%	68.3%
Totals	62.4%	0.8%	1.3%	59.1%	65.8%	62.4%	63.0%	63.7%	64.3%

Figure C.24. Estimated Unpaid Claim Runoff (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Calendar Year Unpaid Claim Runoff
Best Estimate (Weighted)

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2015	2,947	132	4.5%	2,471	3,532	2,947	3,036	3,162	3,257
2016	1,791	101	5.6%	1,449	2,233	1,790	1,859	1,959	2,027
2017	995	73	7.3%	739	1,286	993	1,043	1,117	1,170
2018	520	52	10.0%	345	712	518	554	608	649
2019	271	31	11.5%	161	415	270	291	325	352
2020	147	18	12.6%	65	246	146	159	178	193
2021	83	13	16.0%	31	156	82	91	106	116
2022	46	10	22.2%	11	97	45	53	63	71
2023	24	7	30.7%	1	65	24	29	37	44
2024	14	5	37.9%	(0)	42	13	17	23	27
2025	6	3	45.2%	(0)	24	6	8	11	14
2026	3	2	60.2%	(0)	13	2	4	6	8
2027	1	1	95.7%	(0)	9	1	1	3	4
2028	0	0	19936.0%	(0)	0	0	0	0	0

Figure C.25. Mean of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident	Mean Values													
Year	12	24	36	48	60	72	84	96	108	120	132	144	156	168 +
2006	326	384	355	237	124	55	27	16	10	6	4	2	1	1
2007	328	388	326	218	114	61	27	16	9	6	3	2	1	1
2008	331	356	299	200	125	60	27	16	9	6	3	2	1	1
2009	303	327	274	219	124	60	27	16	9	6	3	2	1	1
2010	290	328	306	218	121	58	27	16	11	4	4	2	1	1
2011	269	323	291	207	115	60	27	15	10	4	4	2	1	1
2012	269	312	281	198	130	62	28	16	11	3	4	2	1	1
2013	266	308	278	229	126	60	27	16	11	3	4	2	1	1
2014	299	346	325	228	126	60	27	15	11	3	4	2	1	1
2015	294	351	317	223	123	59	26	15	11	3	4	2	1	1

Figure C.26. Standard Deviation of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident	Standard Error Values													
Year	12	24	36	48	60	72	84	96	108	120	132	144	156	168 +
2006	21	23	22	18	13	9	6	5	4	3	3	2	2	1
2007	21	23	21	17	13	9	6	5	4	3	3	2	2	1
2008	21	22	20	17	13	9	6	5	4	3	3	2	2	1
2009	21	21	19	17	13	9	6	5	4	3	3	2	2	1
2010	18	26	23	14	22	15	8	4	4	3	2	2	1	1
2011	18	17	22	15	22	18	8	4	4	3	2	2	1	1
2012	13	14	22	11	30	21	8	4	3	2	2	1	1	1
2013	13	14	22	18	31	21	8	4	3	2	2	1	1	1
2014	13	14	33	18	30	21	8	4	3	2	2	1	1	1
2015	13	26	32	19	30	21	8	4	3	2	2	1	1	1

Figure C.27. Coefficient of Variation of Incremental Values (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Accident Year Incremental Values by Development Period
Best Estimate (Weighted)

Accident	Coefficients of Variation													
Year	12	24	36	48	60	72	84	96	108	120	132	144	156	168 +
2006	6.5%	6.0%	6.2%	7.6%	10.4%	15.9%	22.6%	29.2%	37.6%	49.0%	74.6%	95.5%	122.0%	156.9%
2007	6.4%	5.9%	6.5%	8.0%	11.0%	15.0%	22.4%	28.9%	38.1%	56.2%	73.4%	94.7%	123.1%	157.6%
2008	6.5%	6.2%	6.8%	8.3%	10.4%	15.0%	22.6%	29.3%	41.7%	56.2%	73.4%	95.5%	122.4%	160.6%
2009	6.8%	6.4%	7.1%	7.9%	10.6%	15.2%	22.7%	31.5%	42.2%	56.9%	74.2%	96.8%	121.7%	162.3%
2010	6.1%	7.9%	7.3%	6.3%	18.2%	25.8%	28.6%	26.8%	35.0%	65.1%	63.8%	81.4%	111.5%	113.7%
2011	6.6%	5.4%	7.6%	7.2%	18.7%	30.1%	29.0%	27.1%	34.8%	66.8%	64.0%	82.4%	113.2%	115.9%
2012	4.8%	4.5%	7.8%	5.6%	23.4%	34.1%	30.0%	24.4%	30.1%	60.7%	57.6%	71.7%	93.4%	94.2%
2013	4.8%	4.4%	7.9%	7.9%	24.2%	34.5%	30.4%	24.4%	30.2%	61.4%	58.2%	72.2%	94.5%	94.4%
2014	4.5%	4.2%	10.0%	8.1%	23.6%	34.6%	30.2%	24.7%	30.3%	62.0%	59.4%	73.2%	94.7%	95.6%
2015	4.6%	7.4%	10.0%	8.3%	24.3%	35.0%	31.0%	24.6%	30.6%	61.4%	59.9%	73.5%	97.0%	95.7%

Figure C.28. Total Unpaid Claims Distribution (Weighted)

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Total Unpaid Distribution
Best Estimate (Weighted)

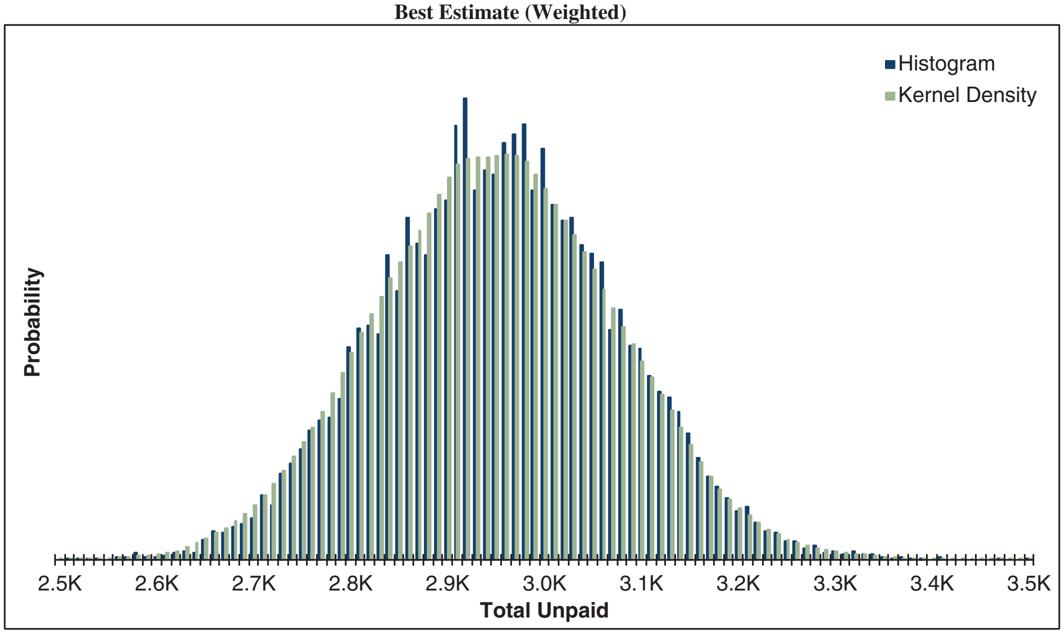
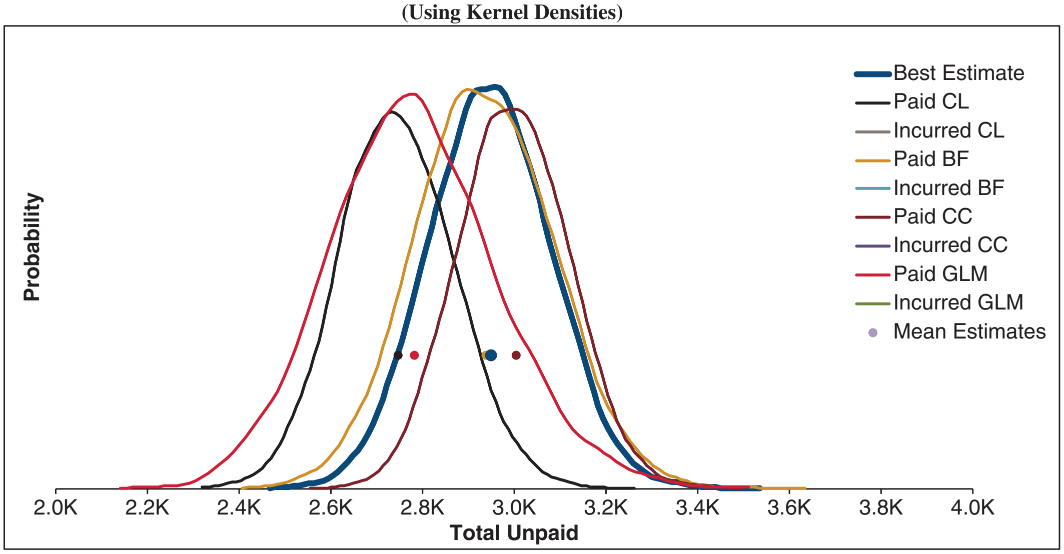


Figure C.29. Summary of Model Distributions

Five Top 50 Companies
Schedule P, Part C – Commercial Auto Liability (in 000,000's)
Summary of Model Distributions
(Using Kernel Densities)



Appendix D—Aggregate Results

In this appendix the results for the correlated aggregate of the three Schedule P lines of business (Parts A, B, and C) are shown, using the correlation calculated from the paid data after adjustment for heteroscedasticity.

Figure D.1. Estimated Unpaid Model Results

Five Top 50 Companies
Aggregate Three Lines of Business
Accident Year Unpaid

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	67	25	37.9%	0	186	66	83	110	130
2007	107	30	28.1%	25	295	105	126	158	185
2008	199	49	24.8%	67	622	194	226	285	342
2009	298	56	18.8%	123	800	293	331	395	457
2010	480	69	14.3%	248	959	475	522	599	668
2011	862	106	12.3%	503	1,561	860	923	1,041	1,135
2012	1,666	187	11.2%	383	2,555	1,662	1,771	1,985	2,148
2013	3,070	333	10.8%	1,808	6,522	3,066	3,249	3,649	3,928
2014	5,632	703	12.5%	2,435	8,555	5,632	6,075	6,801	7,326
2015	13,270	1,788	13.5%	5,217	22,660	13,262	14,348	16,180	18,011
Totals	25,650	2,080	8.1%	16,952	36,085	25,616	26,949	29,088	30,991
Normal Dist.	25,650	2,080	8.1%			25,650	27,053	29,072	30,490
logNormal Dist.	25,650	2,088	8.1%			25,566	27,006	29,222	30,885
Gamma Dist.	25,650	2,080	8.1%			25,594	27,021	29,165	30,736

Figure D.2. Estimated Cash Flow

Five Top 50 Companies
Aggregate Three Lines of Business
Calendar Year Unpaid

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2016	12,906	1,209	9.4%	8,242	19,475	12,869	13,611	14,897	16,182
2017	5,733	453	7.9%	3,991	7,589	5,727	6,024	6,488	6,836
2018	3,144	257	8.2%	2,132	4,373	3,137	3,310	3,573	3,781
2019	1,663	144	8.6%	1,163	2,415	1,657	1,757	1,906	2,018
2020	903	86	9.5%	617	1,331	900	958	1,050	1,122
2021	512	59	11.5%	319	1,064	508	546	613	678
2022	324	55	16.9%	140	699	317	353	423	484
2023	217	64	29.4%	86	931	205	245	328	431
2024	120	28	23.7%	21	308	118	137	170	197
2025	74	22	30.1%	7	165	73	89	113	131
2026	36	13	37.2%	2	94	35	45	59	70
2027	18	9	51.9%	0	58	17	24	33	41
2028	1	1	95.7%	-	9	1	1	3	4
Totals	25,650	2,080	8.1%	16,952	36,085	25,616	26,949	29,088	30,991

Figure D.3. Estimated Loss Ratio

Five Top 50 Companies
Aggregate Three Lines of Business
Accident Year Ultimate Loss Ratios

Accident Year	Mean Loss Ratio	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	74.0%	10.7%	14.5%	33.5%	132.5%	73.7%	77.5%	93.7%	109.6%
2007	81.3%	11.5%	14.2%	38.3%	147.1%	81.0%	85.0%	102.0%	121.0%
2008	85.4%	11.8%	13.8%	39.5%	153.1%	85.0%	89.2%	107.7%	123.9%
2009	76.0%	10.2%	13.4%	36.8%	131.0%	75.6%	79.4%	94.7%	111.2%
2010	66.9%	9.3%	13.9%	31.0%	119.9%	66.3%	70.1%	84.1%	97.9%
2011	64.5%	8.9%	13.8%	30.1%	117.2%	64.2%	67.5%	81.1%	91.4%
2012	71.0%	10.1%	14.3%	31.6%	129.3%	70.5%	74.0%	90.5%	104.6%
2013	61.4%	8.5%	13.9%	29.3%	125.5%	61.1%	64.2%	77.3%	88.8%
2014	65.4%	9.7%	14.8%	31.3%	115.9%	65.2%	70.3%	82.2%	94.9%
2015	78.2%	11.5%	14.7%	39.0%	143.2%	77.8%	83.8%	97.6%	113.8%
Totals	71.6%	3.3%	4.6%	59.5%	88.1%	71.5%	73.7%	77.3%	80.3%

Figure D.4. Estimated Unpaid Claim Runoff

Five Top 50 Companies
Aggregate Three Lines of Business
Calendar Year Unpaid Claim Runoff

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2015	25,650	2,080	8.1%	16,952	36,085	25,616	26,949	29,088	30,991
2016	12,744	944	7.4%	8,710	17,043	12,733	13,373	14,296	15,047
2017	7,012	536	7.6%	4,664	9,551	7,000	7,368	7,905	8,324
2018	3,868	319	8.2%	2,512	5,388	3,861	4,075	4,406	4,671
2019	2,205	213	9.7%	1,348	3,259	2,196	2,340	2,567	2,762
2020	1,302	158	12.1%	730	2,266	1,292	1,400	1,574	1,733
2021	790	126	15.9%	401	1,697	781	864	1,003	1,145
2022	466	99	21.2%	166	1,272	458	524	636	746
2023	249	62	24.9%	45	533	245	289	359	403
2024	129	42	32.4%	13	294	126	156	202	236
2025	55	21	37.9%	3	141	53	68	90	107
2026	19	9	49.6%	0	60	18	25	34	42
2027	1	1	95.7%	(0)	9	1	1	3	4

Figure D.5. Mean of Incremental Values

Five Top 50 Companies
Aggregate Three Lines of Business
Accident Year Incremental Values by Development Period

Accident	Mean Values													
Year	12	24	36	48	60	72	84	96	108	120	132	144	156	168 +
2006	9,334	4,878	2,029	1,175	621	300	151	79	67	33	33	17	16	1
2007	10,595	5,394	2,159	1,239	655	327	163	85	75	35	35	18	17	1
2008	12,060	5,959	2,317	1,321	716	352	175	91	84	38	38	19	18	1
2009	11,848	6,007	2,371	1,389	745	365	182	95	83	40	40	20	20	1
2010	11,834	5,923	2,345	1,351	721	354	182	95	85	38	40	20	19	1
2011	12,195	5,972	2,312	1,326	707	372	185	96	90	39	40	20	19	1
2012	14,186	6,541	2,409	1,362	775	380	191	99	103	39	40	20	19	1
2013	11,901	5,868	2,282	1,436	763	374	187	97	93	38	40	20	19	1
2014	12,949	6,354	2,538	1,451	771	378	189	98	98	38	36	19	17	1
2015	16,458	7,356	2,685	1,515	794	395	202	108	99	44	36	19	17	1

Figure D.6. Standard Deviation of Incremental Values

Five Top 50 Companies
Aggregate Three Lines of Business
Accident Year Incremental Values by Development Period

Accident	Standard Deviation Values													
Year	12	24	36	48	60	72	84	96	108	120	132	144	156	168 +
2006	1,735	668	233	134	74	37	18	13	23	6	13	7	6	6
2007	1,909	712	244	140	77	39	18	14	26	10	14	7	7	10
2008	2,085	768	264	147	81	41	20	14	35	11	15	8	7	11
2009	2,010	754	260	149	82	41	19	17	34	10	15	8	8	10
2010	2,059	775	264	148	84	43	22	17	35	11	15	8	8	11
2011	2,085	777	261	150	84	49	22	17	37	11	16	8	8	11
2012	2,492	875	277	155	98	50	23	17	44	12	15	8	8	12
2013	2,078	767	261	169	97	51	23	17	39	11	15	8	8	11
2014	2,300	907	341	192	109	55	26	19	42	13	17	9	9	13
2015	2,728	1,087	365	210	116	59	30	23	58	17	17	9	9	17

Figure D.7. Coefficient of Variation of Incremental Values

Five Top 50 Companies
Aggregate Three Lines of Business
Accident Year Incremental Values by Development Period

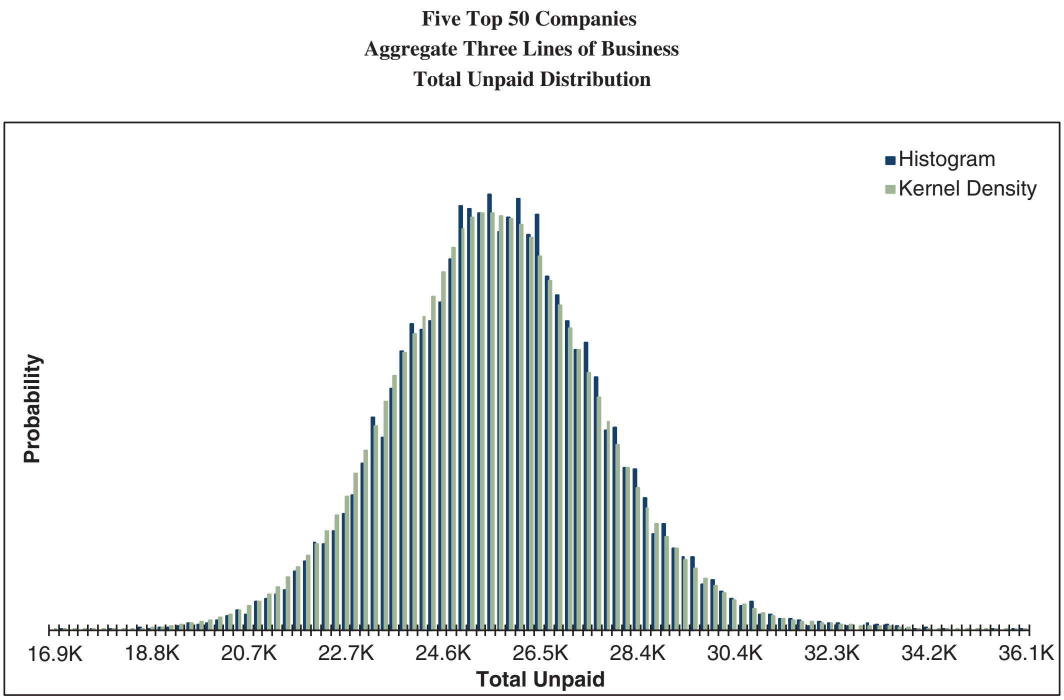
Accident	Coefficients of Variation													
Year	12	24	36	48	60	72	84	96	108	120	132	144	156	168 +
2006	18.6%	13.7%	11.5%	11.4%	11.9%	12.5%	11.7%	16.8%	35.0%	17.0%	38.4%	38.5%	40.0%	702.6%
2007	18.0%	13.2%	11.3%	11.3%	11.8%	12.0%	11.4%	16.3%	35.1%	27.8%	38.5%	38.6%	40.0%	1216.6%
2008	17.3%	12.9%	11.4%	11.1%	11.3%	11.7%	11.2%	15.7%	42.1%	28.0%	39.1%	39.2%	40.5%	1356.5%
2009	17.0%	12.6%	11.0%	10.7%	11.0%	11.3%	10.7%	17.7%	41.4%	26.2%	38.8%	39.0%	40.2%	1287.1%
2010	17.4%	13.1%	11.3%	11.0%	11.6%	12.2%	11.9%	17.7%	40.5%	27.6%	39.0%	39.4%	40.8%	1164.5%
2011	17.1%	13.0%	11.3%	11.3%	11.9%	13.3%	11.9%	17.5%	41.5%	28.2%	39.2%	39.5%	40.9%	1219.1%
2012	17.6%	13.4%	11.5%	11.4%	12.6%	13.2%	12.0%	16.9%	42.8%	31.6%	38.6%	39.0%	40.6%	1268.9%
2013	17.5%	13.1%	11.4%	11.8%	12.6%	13.6%	12.1%	17.3%	41.9%	29.1%	38.5%	38.9%	40.4%	1214.5%
2014	17.8%	14.3%	13.5%	13.3%	14.2%	14.5%	13.8%	19.0%	43.0%	34.8%	47.6%	46.7%	54.6%	1429.6%
2015	16.6%	14.8%	13.6%	13.8%	14.6%	15.0%	14.9%	21.5%	58.1%	38.8%	47.3%	46.7%	54.4%	1901.0%

Figure D.8. Calculation of Risk Based Capital

Five Top 50 Companies
Aggregate Three Lines of Business
Indicated Unpaid Claim Risk Portion of Required Capital

LOB / Segment	Earned Premium	Mean Unpaid	99.0% Unpaid	Value at Risk Capital	Allocated Capital	Unpaid Ratio	Premium Ratio
Schedule P, Part A	15,148	5,308	8,675	3,367	2,642	49.8%	17.4%
Schedule P, Part B	20,467	17,395	20,525	3,130	2,456	14.1%	12.0%
Schedule P, Part C	2,383	2,947	3,257	310	243	8.3%	10.2%
Total	37,997	25,650	32,457	6,807			
Aggregate	37,997	25,650	30,991	5,341	5,341	20.8%	14.1%

Figure D.9. Total Unpaid Claims Distribution



Appendix E—GLM Bootstrap Results

In this appendix the results for the GLM Bootstrap model, as illustrated in Figures 5.9 through 5.12 using the Taylor and Ashe (1983) data, are shown.

Figure E.1. Estimated Unpaid Model Results

Taylor and Ashe Data
Accident Year Unpaid
Paid GLM Bootstrap Model

Accident Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	-	-		-	-	-	-	-	-
2007	201,062	86,944	43.2%	13,857	542,484	186,940	254,238	361,288	438,224
2008	438,222	193,377	44.1%	48,640	1,570,379	405,070	547,131	798,395	996,074
2009	701,223	229,176	32.7%	192,462	1,747,698	679,682	831,657	1,122,868	1,320,964
2010	1,024,913	264,752	25.8%	405,036	2,286,536	1,009,377	1,186,714	1,467,758	1,825,411
2011	1,452,650	315,901	21.7%	619,534	2,544,116	1,424,030	1,660,714	1,996,927	2,261,272
2012	2,181,115	481,962	22.1%	916,307	4,248,064	2,136,166	2,480,213	3,027,607	3,396,995
2013	3,468,030	603,268	17.4%	1,751,033	5,598,537	3,424,738	3,862,292	4,553,992	4,965,982
2014	4,568,990	695,194	15.2%	2,331,572	6,824,685	4,526,036	5,039,460	5,731,706	6,408,694
2015	5,672,877	744,661	13.1%	3,681,244	8,333,062	5,657,952	6,171,074	6,954,411	7,414,615
Totals	19,709,081	2,176,864	11.0%	13,360,401	27,429,908	19,594,207	21,069,822	23,354,466	24,752,422
Normal Dist.	19,709,081	2,176,864	11.0%			19,709,081	21,177,353	23,289,703	24,773,224
logNormal Dist.	19,709,844	2,194,514	11.1%			19,588,799	21,111,651	23,512,537	25,360,134
Gamma Dist.	19,709,081	2,176,864	11.0%			19,628,994	21,130,455	23,421,097	25,123,713

Figure E.2. Estimated Cash Flow

Taylor and Ashe Data
Calendar Year Unpaid
Paid GLM Bootstrap Model

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2016	5,367,217	639,639	11.9%	3,363,863	7,428,225	5,343,203	5,770,597	6,447,544	6,986,539
2017	4,312,360	599,300	13.9%	2,363,704	6,455,658	4,279,059	4,673,264	5,338,534	5,922,511
2018	3,310,498	539,509	16.3%	1,993,107	5,419,760	3,288,209	3,657,889	4,209,239	4,690,515
2019	2,245,627	417,764	18.6%	1,078,000	4,088,770	2,221,086	2,510,176	2,948,019	3,475,039
2020	1,676,436	369,916	22.1%	619,943	3,157,564	1,644,779	1,921,249	2,318,054	2,614,635
2021	1,224,109	326,624	26.7%	444,913	2,352,525	1,202,484	1,436,029	1,782,066	2,085,204
2022	838,442	264,751	31.6%	226,969	2,477,444	803,316	991,076	1,302,125	1,532,640
2023	507,334	211,762	41.7%	104,873	1,268,302	480,233	635,243	889,537	1,135,405
2024	227,058	93,270	41.1%	32,667	711,619	213,471	277,710	403,483	498,676
2025	-	-		-	-	-	-	-	-
Totals	19,709,081	2,176,864	11.0%	13,360,401	27,429,908	19,594,207	21,069,822	23,354,466	24,752,422

Figure E.3. Estimated Loss Ratio

Taylor and Ashe Data
Accident Year Ultimate Loss Ratios
Paid GLM Bootstrap Model

Accident Year	Mean Loss Ratio	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2006	54.8%	6.4%	11.7%	38.1%	74.0%	54.7%	59.0%	65.7%	70.4%
2007	65.0%	6.4%	9.8%	48.1%	84.1%	65.0%	68.9%	75.7%	80.7%
2008	63.1%	6.4%	10.1%	42.6%	82.0%	63.1%	67.3%	73.4%	78.6%
2009	56.0%	6.2%	11.0%	38.0%	76.4%	55.9%	60.0%	66.2%	71.6%
2010	53.1%	5.9%	11.0%	34.7%	74.7%	52.8%	57.1%	63.1%	66.6%
2011	50.5%	5.6%	11.1%	33.9%	70.0%	50.2%	54.2%	60.0%	63.9%
2012	53.8%	7.6%	14.2%	31.3%	81.3%	53.1%	59.1%	66.8%	72.8%
2013	55.3%	6.9%	12.5%	34.6%	78.4%	55.1%	59.5%	66.9%	73.4%
2014	52.9%	6.8%	12.8%	31.7%	74.5%	52.4%	57.2%	64.5%	70.1%
2015	50.7%	6.5%	12.8%	33.0%	72.4%	50.6%	55.1%	61.6%	65.8%
Totals	55.1%	2.9%	5.2%	46.9%	63.6%	55.1%	57.1%	60.0%	61.7%

Figure E.4. Estimated Unpaid Claim Runoff

Taylor and Ashe Data
Calendar Year Unpaid Claim Runoff
Paid GLM Bootstrap Model

Calendar Year	Mean Unpaid	Standard Error	Coefficient of Variation	Minimum	Maximum	50.0% Percentile	75.0% Percentile	95.0% Percentile	99.0% Percentile
2015	19,709,081	2,176,864	11.0%	13,360,401	27,429,908	19,594,207	21,069,822	23,354,466	24,752,422
2016	14,341,864	1,839,659	12.8%	8,990,374	21,139,070	14,231,008	15,525,987	17,412,102	19,106,264
2017	10,029,504	1,499,062	14.9%	5,923,686	15,623,104	9,926,619	10,979,472	12,605,655	13,627,923
2018	6,719,006	1,188,158	17.7%	3,317,118	11,201,515	6,612,903	7,438,758	8,841,160	9,734,081
2019	4,473,380	922,335	20.6%	1,884,408	7,436,971	4,366,371	5,040,244	6,143,079	6,968,601
2020	2,796,943	678,192	24.2%	1,137,743	5,050,304	2,740,868	3,192,138	4,018,580	4,623,373
2021	1,572,834	443,756	28.2%	595,162	3,523,942	1,524,022	1,852,397	2,369,545	2,820,528
2022	734,392	257,467	35.1%	204,545	1,654,724	708,577	888,204	1,167,670	1,463,534
2023	227,058	93,270	41.1%	32,667	711,619	213,471	277,710	403,483	498,676
2024	0	0	4017.8%	(0)	0	-	0	0	0

Figure E.5. Mean of Incremental Values

Taylor and Ashe Data
Accident Year Incremental Values by Development Period
Paid GLM Bootstrap Model

Accident	Mean Values									
Year	12	24	36	48	60	72	84	96	108	120+
2006	260,293	698,693	688,850	704,606	388,809	311,880	258,794	214,532	169,749	142,707
2007	353,111	978,505	972,391	972,627	539,441	447,302	359,572	300,611	234,076	201,062
2008	355,598	975,396	989,087	971,633	541,986	440,002	357,470	297,335	237,981	200,241
2009	343,575	914,108	911,442	913,681	502,676	421,801	335,888	282,854	231,129	187,240
2010	341,295	923,102	914,709	919,809	500,195	420,057	337,719	275,372	224,883	186,939
2011	336,529	924,119	917,372	913,328	503,784	409,092	338,662	284,360	234,436	186,099
2012	381,818	1,028,561	1,036,624	1,025,187	578,558	451,767	374,253	312,453	251,461	212,623
2013	402,258	1,107,072	1,108,427	1,111,762	614,292	501,712	410,170	332,713	265,392	231,989
2014	408,511	1,104,124	1,109,649	1,096,598	616,324	491,960	408,977	338,285	274,715	232,482
2015	406,207	1,098,540	1,104,298	1,121,727	609,668	497,186	407,810	331,738	274,852	227,058

Figure E.6. Standard Deviation of Incremental Values

Taylor and Ashe Data
Accident Year Incremental Values by Development Period
Paid GLM Bootstrap Model

Accident Year	Standard Error Values									
	12	24	36	48	60	72	84	96	108	120+
2006	108,496	120,663	181,091	248,062	129,788	119,862	108,476	67,654	126,590	56,408
2007	131,381	142,390	209,358	306,437	159,961	138,486	133,743	80,122	152,143	86,944
2008	127,448	146,072	215,044	306,874	152,841	142,207	122,350	78,132	159,664	86,683
2009	125,340	137,368	201,409	295,530	154,057	138,440	121,788	90,057	156,660	80,901
2010	127,558	139,764	193,891	297,664	152,539	136,999	129,441	88,860	139,515	78,776
2011	125,839	139,522	196,494	285,649	156,869	139,339	128,102	92,988	160,838	81,152
2012	137,400	150,449	208,476	321,223	187,435	156,077	150,736	103,377	165,336	93,178
2013	137,189	150,565	221,025	338,863	195,056	173,394	151,060	103,542	169,356	96,165
2014	132,459	159,062	254,892	329,254	195,407	162,115	149,531	106,739	171,923	94,787
2015	135,172	183,619	247,413	336,959	177,810	163,745	147,122	102,400	167,873	93,270

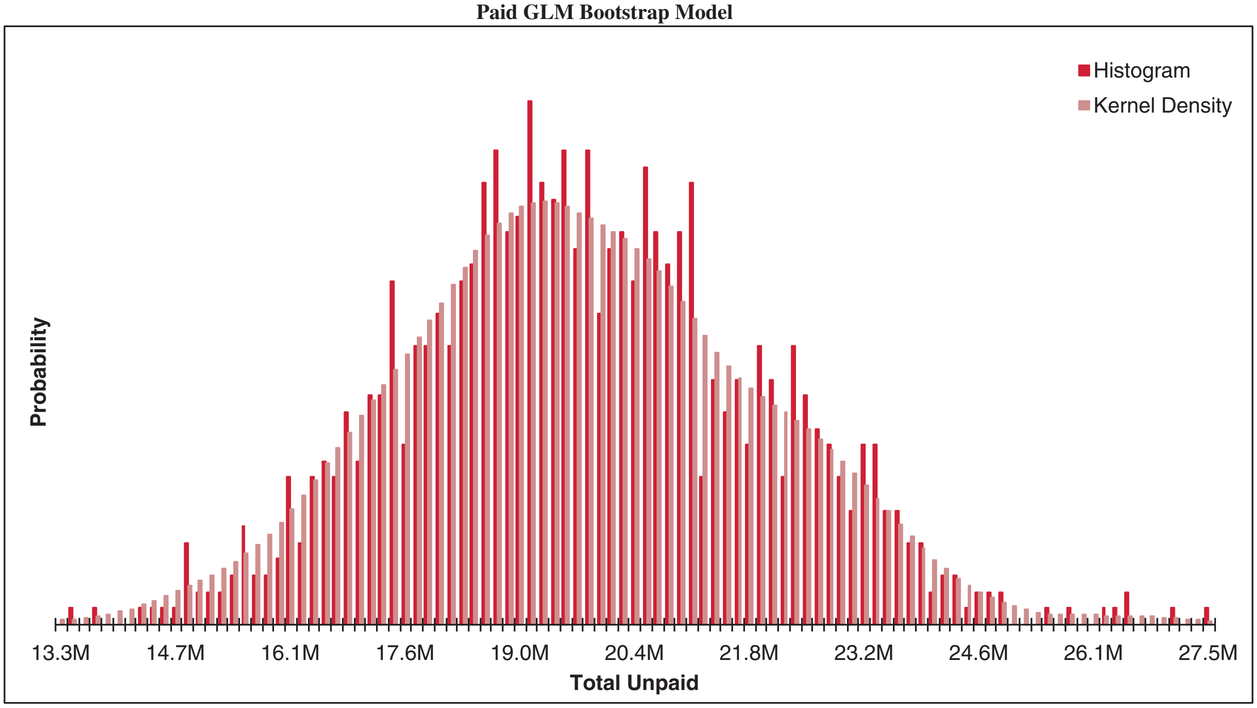
Figure E.7. Coefficient of Variation of Incremental Values

Taylor and Ashe Data
Accident Year Incremental Values by Development Period
Paid GLM Bootstrap Model

Accident Year	Coefficient of Variation Values									
	12	24	36	48	60	72	84	96	108	120+
2006	41.7%	17.3%	26.3%	35.2%	33.4%	38.4%	41.9%	31.5%	74.6%	39.5%
2007	37.2%	14.6%	21.5%	31.5%	29.7%	31.0%	37.2%	26.7%	65.0%	43.2%
2008	35.8%	15.0%	21.7%	31.6%	28.2%	32.3%	34.2%	26.3%	67.1%	43.3%
2009	36.5%	15.0%	22.1%	32.3%	30.6%	32.8%	36.3%	31.8%	67.8%	43.2%
2010	37.4%	15.1%	21.2%	32.4%	30.5%	32.6%	38.3%	32.3%	62.0%	42.1%
2011	37.4%	15.1%	21.4%	31.3%	31.1%	34.1%	37.8%	32.7%	68.6%	43.6%
2012	36.0%	14.6%	20.1%	31.3%	32.4%	34.5%	40.3%	33.1%	65.8%	43.8%
2013	34.1%	13.6%	19.9%	30.5%	31.8%	34.6%	36.8%	31.1%	63.8%	41.5%
2014	32.4%	14.4%	23.0%	30.0%	31.7%	33.0%	36.6%	31.6%	62.6%	40.8%
2015	33.3%	16.7%	22.4%	30.0%	29.2%	32.9%	36.1%	30.9%	61.1%	41.1%

Figure E.8. Total Unpaid Claims Distribution

Taylor and Ashe Data
Total Unpaid Distribution
Paid GLM Bootstrap Model



References

Anderson, Duncan, Sholom Feldblum, Claudine Modlin, Doris Schirmacher, Ernesto Schirmacher, and Neeza Thandi. 2007. "A Practitioner's Guide to Generalized Linear Models," *CAS Exam Study Note,* 3rd Edition: 1–116.

Ashe, Frank. 1986. "An Essay at Measuring the Variance of Estimates of Outstanding Claim Payments." *ASTIN Bulletin* 16:S: 99–113.

Barnett, Glen, and Ben Zehnwirth. 2000. "Best Estimates for Reserves." *Proceedings of the Casualty Actuarial Society* 87, 2: 245–321.

Berquist, James R., and Richard E. Sherman. 1977. "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach." *Proceedings of the Casualty Actuarial Society* 64: 123–184.

Bornhuetter, Ronald, and Ronald Ferguson. 1972. "The Actuary and IBNR." *Proceedings of the Casualty Actuarial Society* 59: 181–195.

CAS Loss Simulation Model Working Party Summary Report. 2011. "Modeling Loss Emergence and Settlement Processes." *Casualty Actuarial Society Forum* (Winter) 1: 1–124.

CAS Working Party on Quantifying Variability in Reserve Estimates. 2005. "The Analysis and Estimation of Loss & ALAE Variability: A Summary Report." *Casualty Actuarial Society Forum* (Fall): 29–146.

CAS Tail Factor Working Party. 2013. "The Estimation of Loss Development Tail Factors: A Summary Report." *Casualty Actuarial Society E-Forum* (Fall): 1–111.

Christofides, S. 1990. "Regression Models Based on Log-Incremental Payments." *Claims Reserving Manual,* vol. 2. Institute of Actuaries, London.

Efron, Bradley. 1979. "Bootstrap Methods: Another Look at the Jackknife." *The Annals of Statistics* 7-1: 1–26.

England, Peter D., and Richard J. Verrall. 1999. "Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving." *Insurance: Mathematics and Economics* 25: 281–293.

England, Peter D., and Richard J. Verrall. 2002. "Stochastic Claims Reserving in General Insurance." *British Actuarial Journal* 8-3: 443–544.

England, Peter D., and Richard J. Verrall. 2006. "Predictive Distributions of Outstanding Liabilities in General Insurance." *The Annals of Actuarial Science* 1, 2: 221–270.

Foundations of Casualty Actuarial Science, 4th ed. 2001. Arlington, Va.: Casualty Actuarial Society.

Iman, R., and W. Conover. 1982. "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables." *Communications in Statistics—Simulation and Computation* 11(3): 311–334.

IAA (International Actuarial Association). 2010. "Stochastic Modeling—Theory and Reality from an Actuarial Perspective." Available from www.actuaries.org/stochastic.

Kirschner, Gerald S., Colin Kerley, and Belinda Isaacs. 2008. "Two Approaches to Calculating Correlated Reserve Indications Across Multiple Lines of Business." *Variance* 1: 15–38.

Kremer, E. 1982. "IBNR Claims and the Two Way Model of ANOVA" *Scandinavian Actuarial Journal:* 47–55.

Liu, H., and R. Verrall. 2010. "Bootstrap Estimation of the Predictive Distributions of Reserves Using Paid and Incurred Claims." *Variance* 4: 125–135.

McCullagh, P., and J. Nelder. 1989. *Generalized Linear Models,* 2nd ed. Chapman and Hall.

Mildenhall, Stephen J. 2006. "Correlation and Aggregate Loss Distributions with an Emphasis on the Iman-Conover Method." *Casualty Actuarial Society E-Forum* (Winter): 103–204.

Milliman. 2014. "Using the Milliman Arius Reserving Model." Version 2.1.

Pinheiro, Paulo J. R., João Manuel Andrade e Silva, and Maria de Lourdes Centeno. 2001. "Bootstrap Methodology in Claim Reserving." *ASTIN Colloquium:* 1–13.

Pinheiro, Paulo J. R., João Manuel Andrade e Silva, and Maria de Lourdes Centeno. 2003. "Bootstrap Methodology in Claim Reserving." *Journal of Risk and Insurance* 70: 701–714.

Quarg, Gerhard, and Thomas Mack. 2008. "Munich Chain Ladder: A Reserving Method that Reduces the Gap between IBNR Projections Based on Paid Losses and IBNR Projections Based on Incurred Losses." *Variance* 2: 266–299.

ROC/GIRO Working Party. 2007. "Best Estimates and Reserving Uncertainty." Institute of Actuaries.

ROC/GIRO Working Party. 2008. "Reserving Uncertainty." Institute of Actuaries.

Renshaw, A. E., 1989. "Chain Ladder and Interactive Modelling (Claims Reserving and GLIM)." *Journal of the Institute of Actuaries* 116 (III): 559–587.

Renshaw, A. E., and R. J. Verrall. 1994. "A Stochastic Model Underlying the Chain Ladder Technique." Proceedings XXV *ASTIN Colloquium,* Cannes.

Struzzieri, Paul J., and Paul R. Hussian. 1998. "Using Best Practices to Determine a Best Reserve Estimate." *Casualty Actuarial Society Forum* (Fall): 353–413.

Taylor, Greg, and Frank Ashe. 1983. "Second Moments of Estimates of Outstanding Claims." *Journal of Econometrics* 23-1: 37–61.

Venter, Gary G. 1998. "Testing the Assumptions of Age-to-Age Factors." *Proceedings of the Casualty Actuarial Society* 85: 807–47.

Verrall, Richard J. 1991. "On the Estimation of Reserves from Loglinear Models." *Insurance: Mathematics and Economics* 10: 75–80.

Verrall, Richard J. 2004. "A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving." *North American Actuarial Journal* 8-3: 67–89.

Zehnwirth, Ben, 1989. "The Chain Ladder Technique—A Stochastic Model." *Claims Reserving Manual* vol. 2. Institute of Actuaries, London.

Zehnwirth, Ben. 1994. "Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals and Risk Based Capital." *Casualty Actuarial Society Forum* (Spring), 2: 447–606.

Selected Bibliography

Björkwall, Susanna. 2009. "Bootstrapping for Claims Reserve Uncertainty in General Insurance." Mathematical Statistics, Stockholm University. Research Report 2009:3, Licenciate thesis. http://www2.math.su.se/matstat/reports/seriea/2009/rep3/report.pdf.

Björkwall, Susanna, Ola Hössjer, and Esbjörn Ohlsson. 2009. "Non-parametric and Parametric Bootstrap Techniques for Age-to-Age Development Factor Methods in Stochastic Claims Reserving." *Scandinavian Actuarial Journal* 4: 306–331.

Freedman, D. A. 1981. "Bootstrapping Regression Models." *The Annals of Statistics* 9-6: 1218–1228.

Hayne, Roger M. 2008. "A Stochastic Framework for Incremental Average Reserve Models." *Casualty Actuarial Society E-Forum* (Fall): 174–195.

Mack, Thomas. 1993. "Distribution Free Calculation of the Standard Error of Chain Ladder Reserve Estimates." *ASTIN Bulletin* 23-2: 213–225.

Mack, Thomas. 1999. "The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor." *ASTIN Bulletin* 29-2: 361–366.

Mack, Thomas, and Gary Venter. 2000. "A Comparison of Stochastic Models that Reproduce Chain Ladder Reserve Estimates." *Insurance: Mathematics and Economics* 26: 101–107.

Merz, Michael, and Mario V. Wüthrich. 2008. "Modeling the Claims Development Result For Solvency Purposes." *Casualty Actuarial Society E-Forum* (Fall): 542–568.

Moulton, Lawrence H., and Scott L. Zeger. 1991. "Bootstrapping Generalized Linear Models." *Computational Statistics and Data Analysis* 11: 53–63.

Murphy, Daniel M. 1994. "Unbiased Loss Development Factors." *Proceedings of the Casualty Actuarial Society* 81: 154–222.

Ruhm, David L., and Donald F. Mango. 2003. A Method of Implementing Myers-Read Capital Allocation in Simulation. *Casualty Actuarial Society Forum* (Fall): 451–458.

Shapland, Mark R. 2007. "Loss Reserve Estimates: A Statistical Approach for Determining 'Reasonableness'." *Variance* 1: 120–148.

Venter, Gary G. 2003. "A Survey of Capital Allocation Methods with Commentary Topic 3: Risk Control." *ASTIN Colloquium.*

Abbreviations and Notations

AIC: Akaike Information Criterion
APD: Automobile Physical Damage
BIC: Bayesian Information Criterion
BF: Bornhuetter-Ferguson
CC: Cape Cod
CL: Chain Ladder
CoV: Coefficient of Variation
DFA: Dynamic Financial Analysis
ELR: Expected Loss Ratio
ERM: Enterprise Risk Management
GLM: Generalized Linear Models
MLE: Maximum Likelihood Estimate
ODP: Over-Dispersed Poisson
OLS: Ordinary Least Squares
RSS: Residual Sum Squared
SSE: Sum of Squared Errors

About the Author

Mark R. Shapland is Senior Consulting Actuary in Milliman's Dubai office where he is responsible for various reserving and pricing projects for a variety of clients and was previously the lead actuary for the Property & Casualty Insurance Software (PCIS) development team. He has a B.S. degree in Integrated Studies (Actuarial Science) from the University of Nebraska-Lincoln. He is a Fellow of the Casualty Actuarial Society, a Fellow of the Society of Actuaries and a Member of the American Academy of Actuaries. He was the leader of Section 3 of the Reserve Variability Working Party, the Chair of the CAS Committee on Reserves, co-chair of the Tail Factor Working Party, and co-chair of the Loss Simulation Model Working Party. He is also a co-developer and co-presenter of the CAS Reserve Variability Limited Attendance Seminar and has spoken frequently on this subject both within the CAS and internationally. He can be contacted at mark.shapland@milliman.com.

ABOUT THE SERIES:

CAS monographs are authoritative, peer-reviewed, in-depth works focusing on important topics within property and casualty actuarial practice. For more information on the CAS Monograph Series, visit the CAS website at www.casact.org.

Expertise. Insight.
Solutions.

www.casact.org

A Model for Reserving Workers Compensation High Deductibles

Jerome J. Siewert, FCAS

A Model for Reserving Workers Compensation High Deductibles

Jerome J. Siewert, FCAS, MAAA

Abstract

Several approaches for estimating liabilities under a high deductible program are described. Included is a proposal for a more sophisticated approach relying upon a loss distribution model. Additionally, the discussion addresses several related issues dealing with deductible size and mix, absence of long-term histories, as well as the determination of consistent loss development factors among deductible limits. Lastly, approaches are proposed for estimating aggregate loss limit charges, if any, and the asset value for associated servicing revenue.

Biography

Jerry Siewert is an Assistant Vice President and Actuary with Wausau Insurance Companies. He currently manages the reserving unit of the actuarial department. His previous experience includes managing several pricing units and serving on various industry ratemaking committees. Prior to joining Wausau, he taught mathematics for four years at the secondary level.

Jerry is a Member of the American Academy of Actuaries and became a Fellow of the Casualty Actuarial Society in 1985. He graduated from the University of Wisconsin - Stevens Point (1972) earning bachelor degrees in both mathematics and psychology.

Acknowledgments

Special thanks to Jim Golz and Terry Knull of Wausau Insurance and Roger Bovard of T.I.G. Insurance for their very helpful comments on earlier drafts of this paper.

A Model for Reserving Workers Compensation High Deductibles

1. Abstract

Several approaches for estimating liabilities under a high deductible program are described. Included is a proposal for a more sophisticated approach relying upon a loss distribution model. Additionally, the discussion addresses several related issues dealing with deductible size and mix, absence of long-term histories, as well as the determination of consistent loss development factors among deductible limits. Lastly, approaches are proposed for estimating aggregate loss limit charges, if any, and the asset value for associated servicing revenue.

2. Introduction

With the advent of the high deductible program in the early '90s, actuarial efforts focused principally on pricing issues. Insurers initially developed this program to provide both themselves and insureds many advantages, including:

1. achieving price flexibility while passing additional risk to larger insureds in what was considered at that time an unprofitable line of business,
2. ameliorating onerous residual market charges and premium taxes in some states,
3. realizing cash flow advantages similar to those of the closely related product - the paid loss retro,
4. providing insureds with another vehicle to control losses while protecting them against random, large losses, and
5. allowing "self-insurance" without submitting insureds to sometimes demanding state requirements.

Now as the program matures, the focus shifts to the liability side. Questions are being asked as to what losses are actually emerging and, more importantly, what will they ultimately cost insurers. For the actuary, the question is how best to estimate these liabilities when losses are not expected to emerge above deductible limits for many years. Many issues need to be addressed:

1. In the absence of long-term development histories under a deductible program, how can the actuary construct reasonable development factors?
2. How can the actuary determine development patterns that reflect the diversity of both deductible size and mix?
3. How should the actuary determine consistent development factors between limited and excess values?
4. What is a reasonable approach for the indexing of deductible limits over time?
5. How can the actuary estimate the liability associated with aggregate loss limits, if any?

6. Is there a sound way to determine the proper asset value for associated service revenue?[1]

In the remainder of this paper I describe possible approaches dealing with those issues.

3. Development Approaches

Overview

The development approach presented relies heavily upon my company's extensive history of full coverage workers compensation claim experience. In effect, I create deductible/excess development patterns as needed. Of course, this approach poses problems if credible histories of full coverage losses are not readily available.

Once I establish the appropriate development factors, I apply them at the account level and determine the overall aggregate reserve by summarizing estimated ultimates for each account. I argue this is a reasonable approach, if you view each account as belonging to a cohort of policies with similar limit characteristics. Determining the overall reserve in such a fashion allows me to address the issue of deductible mix by reflecting each account's unique limits.

Later I describe the possible use of a loss distribution model to enforce consistent results between deductible/excess development factors. Once the parameters of the distribution are set, it is possible to determine development factors, as needed, for any deductible size. Perhaps, the use of such a model may even provide an alternative approach for determining tail factors through the projection of the distribution parameters.

Loss Ratio

In the absence of credible development histories, a common approach for determining liabilities is to apply loss ratios to premiums arising from the exposures. Historically, as that element was required to first price the product, loss ratios for the various accounts written should be readily available. For immature years, where data is sparse, applying loss ratios is probably the most practical approach to take. Given the long-tailed nature of this business, actual experience over deductible limits emerges slowly over time . Also the expected experience is readily converted to an accident year basis based upon a pro rata earnings of the policy year exposures.

The loss ratio approach requires a database of individual accounts and pricing elements. The database should include an estimate of the full coverage loss ratio. From a pricing standpoint, that number can come from a variety of sources. One approach would be to use company experience by state, reflecting the individual account's premium distribution. Possibly, that experience to the extent credible could be blended with industry experience. As with other

[1] Similar in usage to a loss conversion factor in retro rating, loss multipliers are applied to deductible losses to capture expenses that vary with loss.

pricing efforts, that experience ought to be developed to ultimate, brought on level, and trended to the appropriate exposure period.

Besides an estimate of the full coverage loss ratio, the database should include estimates of excess losses for both occurrence and aggregate limits. For the occurrence limit, several approaches are possible including estimating excess ratios based upon company experience. A potentially more credible approach uses excess loss pure premium ratios underlying industry-based excess loss factors used in retro rating. Besides their availability by multiple limits, excess loss factors can easily be adjusted to a loss basis and reflect hazard groups with differing severity potential. Utilizing account-based excess ratios reflecting unique state and hazard group characteristics should lead to reasonable estimates of per occurrence excess losses:

(3.1) $P \cdot E \cdot \chi$

where P = premium, E = expected loss ratio, and χ = per occurrence charge

Regarding the aggregate loss charge, if any, an approach I prefer uses a process similar to that for determining insurance charges in a retro rating program. Those charges would, in turn, rely on the National Council on Compensation Insurance's (NCCI) Table M. I refer the interested reader to the Retrospective Rating Plan [1] for further details. The process reflects the size of the account, deductible, state severity relativities, prospective rating period, and appropriate rating plan parameters:

(3.2) $P \cdot E \cdot (1-\chi) \cdot \phi$

where P = premium, E = expected loss ratio, χ = per occurrence charge, and ϕ = per aggregate charge

Applying this procedure to each account and aggregating leads to an estimate of ultimate accident year losses. I show in Table 1 a hypothetical case of how to apply those factors to determine the ultimate liabilities. Incurred But Not Reported (IBNR) amounts are easily determined by subtracting known losses from the ultimate estimate.

Again, this approach is particularly useful when no data is available or the data is so immature as to be virtually useless. Obviously, loss ratio estimates can be consistently tied to pricing programs, at least at the outset. This procedure also benefits from its reliance on a more credible pool of company and/or industry experience. On the negative side, a loss ratio approach ignores actual emerging experience, which in some circumstances may differ significantly from estimated ultimate losses. For this reason alone, the loss ratio approach is not particularly useful after several years of development. Another shortcoming of this method is that it may not properly reflect account characteristics, as development may emerge differently due to the exposures written.

Table 1
Countrywide Insurance Enterprise
Account: Widget, Inc.
Expected Deductible/Aggregate Loss Charges

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			(2) x (3)		(4) x (5)		[(4) - (6)] x (7)
					Deductible	Aggre-	Aggregate
			Expected	Excess	Loss	gate	Loss
State	Premium	ELR	Loss	Ratio	Charge	Ratio	Charge
Arkansas	9,084	.567	5,151	.062	319	.02	97
Illinois	573,066	.532	304,871	.105	32,011	.02	5,457
Iowa	373,072	.588	219,366	.096	21,059	.02	3,966
Kansas	70,549	.644	45,434	.071	3,226	.02	844
Minnesota	1,012,622	.457	462,768	.143	66,176	.02	7,932
South Carolina	22,980	.522	11,996	.048	576	.02	228
South Dakota	94,401	.697	65,797	.211	13,883	.02	1,038
Total	2,155,774	.517	1,115,383	.123	137,250	.02	19,562

Implied Development

There are many ways to incorporate actual emergence in high deductible reserve estimates. Determining excess development implicitly is one possibility. By implied development, I mean an approach that works as follows:

1. Develop full coverage losses to ultimate.
2. Next, develop deductible losses to ultimate by applying development factors reflecting various inflation indexed limits.
3. Finally, determine ultimate excess losses by differencing the full coverage ultimate losses and the limited ultimate losses.

A variety of the usual development techniques could be applied to determine full coverage losses. Those methods include paid and incurred techniques designed consistently with the company's reserving procedures for full coverage workers compensation. However, care should be exercised in determining a full coverage tail factor consistent with the limited loss tail factors. In particular, the actuary should avoid developing limited losses beyond unlimited losses, or even losses for lower limits beyond those of higher limits.

When calculating development factors for the various deductibles, it is appropriate to index the limits for inflationary effects. Adjusting the deductible by indexing keeps the proportion of deductible/excess losses constant about the limit from year to year, at least, in theory. For example, if inflationary forces drive claim costs ten percent higher each year, the percentage of losses over a $100,000 deductible for one year equate to those of a $110,000 deductible in the next. Indexing of deductible limits allows for the possibility of combining differing experience years in the determination of development factors.

There is really no set method for determining the indexing value. One approach would be to determine that index by fitting a line to average severities over a long-term history. Another simpler approach might be to use an index that reflects the movement in annual severity changes. In any event, the actuary needs to be cognizant that a constant deductible over time usually implies increasing excess losses.

An advantage of the implied development approach is that it provides an estimate of excess losses at early maturities even when excess losses have not emerged. Also, the development factors for limited losses are more stable than those determined for losses above the deductible. This procedure also provides an important byproduct in the estimation of assets under the high deductible program. Specifically, estimating deductible losses helps determine the asset represented by revenue collected from the application of a loss multiplier to future losses. Despite these advantages, this approach does appear to have its focus misplaced, as one would like to explicitly recognize excess loss development.

Direct Development

This approach explicitly focuses on excess development, though it relies upon development factors implicit from the previous technique. That is, given development factors for limited as well as full coverage losses, excess loss development factors are fixed. It is important to recognize here that excess development is part of overall development, and the actuary should strive to determine excess factors in conjunction with limited development factors that balance back to full coverage development. That is not to say that reserve indications from implicit and explicit methods necessarily will be the same, but only that the underlying loss development factors should be.

Again, a variety of paid and incurred techniques are applicable here. I see several disadvantages to directly determining excess development factors and applying them to excess losses. Those factors tend to be quite leveraged and extremely volatile, making them difficult to select. Additionally, if excess losses have not actually emerged at any particular stage of development, it is not possible to get an estimate of the required liability.

Credibility Weighting Techniques/Bornhuetter-Ferguson

Given the significant drawbacks mentioned for the previous approaches to determining excess liabilities for the deductible product, the next approach described offers greater promise. It relies on credibility weighting indications based upon actual experience with expected values, preferably based on pricing estimates. This method requires that the actuary determine a suitable set of weights or credibilities. The Bornhuetter-Ferguson [2] technique offers one possible approach for determining credibilities that are specified as reciprocals of loss development factors.

(3.3) $L = O_t \cdot LDF_t \cdot Z + E \cdot (1 - Z)$ (Credibility viewpoint)

where L = ultimate loss estimate, O_t = observed loss at time t, LDF_t = age to ultimate development factor, Z = credibility, and E = expected ultimate loss

Letting $Z = \frac{1}{LDF_t}$ leads to:

(3.4) $L = O_t + E \cdot \left(\frac{LDF_t - 1}{LDF_t} \right)$ (Bornhuetter-Ferguson viewpoint)

Using the Bornhuetter-Ferguson approach allows the actuary to determine liabilities either directly or indirectly. This procedure affords the ability to tie into pricing estimates for recent years where excess losses have yet to emerge. Also, it provides more stable estimates over time, rather than the volatility arising from erratic emergence or leveraged development factors. Hopefully, a credibility weighting approach like this provides better estimators of ultimate liabilities as well. Of course, a disadvantage of this technique is that it ignores actual experience to the extent of the complement of credibility. That drawback suggests finding alternative weights or credibilities that may be more responsive to the actual experience as desired.

4. Development Model

This section deals more specifically with a number of the issues I described at the outset. How best can the actuary determine development factors in the absence of a long-term history under the deductible program? How can the actuary determine development patterns that reflect the diversity of both deductible size and mix? What is a reasonable approach for indexing deductible limits over time? How best should the process relate development for various limits consistently? Determining development factors for a high deductible program is really an exercise in partitioning development about the deductible limit. Is it possible to develop consistent tail factors among the limits the company is exposed to?

Some Possible Approaches

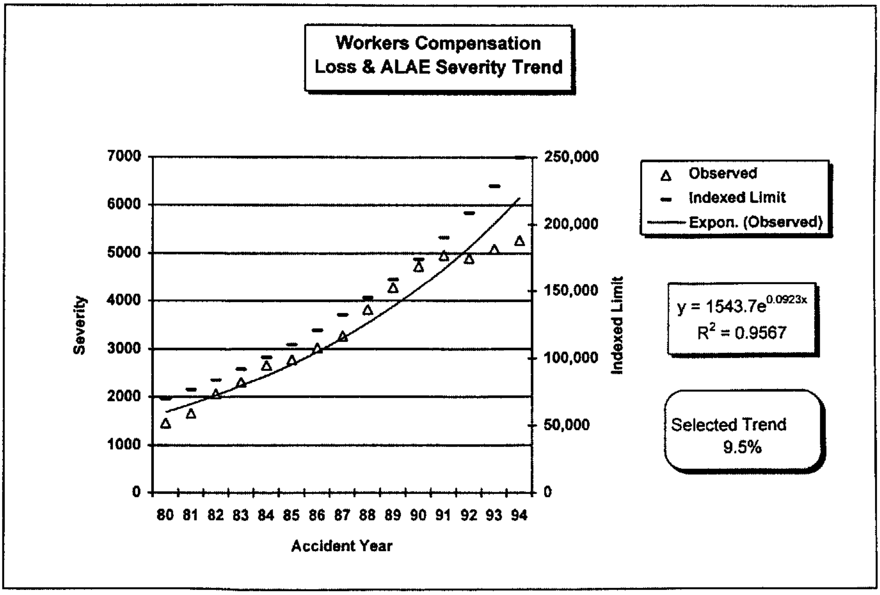
As I stated earlier, in the absence of long-term experience under the deductible program, I suggest making extensive use of a company's history of full coverage workers compensation claims, if available. It is also appropriate to apply an indexed limit to the claims in order to determine a series of accident year loss development histories by limit. In some of the analyses I performed, I looked at selected limits ranging from \$50,000 to \$1,000,000. I focused, however, on the more common deductible sizes in the neighborhood of \$250,000. I used case losses that included indemnity, medical, and any subject allocated claim expense. The histories I reviewed ran out for 25 years but were not further separated by account, injury, or state. That suggests eventually creating alternative development patterns that do reflect those types of break-out. I show in Table 2, age-to-age development factors by indexed limit resulting from my preliminary studies.

Table 2
Workers Compensation - High Deductibles
Limited Loss & ALAE Age-to-Age Development Factors
by Indexed Limit (Middle 6 of Last 8)

Limit	12:24 Months	24:36 Months	36:48 months	48:60 Months	60:72 Months
$50,000	1.5031	1.0418	1.0038	1.0025	1.0020
$100,000	1.6225	1.0727	1.0151	1.0063	1.0080
$250,000	1.6791	1.1300	1.0451	1.0207	1.0060
$500,000	1.6827	1.1393	1.0684	1.0322	1.0170
$750,000	1.6816	1.1408	1.0720	1.0359	1.0214
$1,000,000	1.6811	1.1411	1.0728	1.0371	1.0229
Unlimited	1.6876	1.1430	1.0749	1.0391	1.0196

In order to determine those development factors, I combined several years of experience based upon indexed limits. For example, for the most recent year, limits were used as stated. But for the first prior year, I adjusted limits downward by an indexing factor of 1.095. For the current year, I assumed a limit of $250,000 was the equivalent of a limit of $228,311 for the first prior year. Each limit was adjusted by the same index, back for as many years as needed, to generate the desired development factors.

Chart 1



I simply based the selected indexing factor of 1.095 upon a long-term severity history. As I alluded to earlier, other approaches may be better. Possibly varying the indexing factor by year or adjusting for the distorting effects of larger claims are but a couple of examples of improvements that could be explored. I show in Chart 1 the exponential trend line fit through known data points determining the long-term indexing factor of 1.095. Also depicted is the indexed $250,000 loss limit.

The approach I recommend requires separating claim count development from severity development. In my work to date I focused on the counts for full coverage losses rather than worrying about emergence of claims over a specific deductible limit. I feel it is much easier to recognize development in this fashion, as there is generally very little true claim count IBNR after about three years. This is true even for the larger claims, as they will be reported early on just like the other claims, but their true severity will not be known for some time.

Table 3
Workers Compensation
Age-to-Age Development Factors
Full Coverage Claim Count

Accident Year	12:24 Months	24:36 Months	36:48 months	48:60 Months
1988	-	-	-	0.9999
1989	-	-	0.9999	0.9994
1990	-	1.0026	0.9999	1.0001
1991	1.1111	1.0022	1.0002	-
1992	1.1305	1.0017	-	-
1993	1.1283	-	-	-
Last 3	1.1233	1.0022	1.0000	0.9998
Selected	1.1250	1.0025	1.0000	1.0000
Age to Ultimate	1.1278	1.0025	1.0000	1.0000

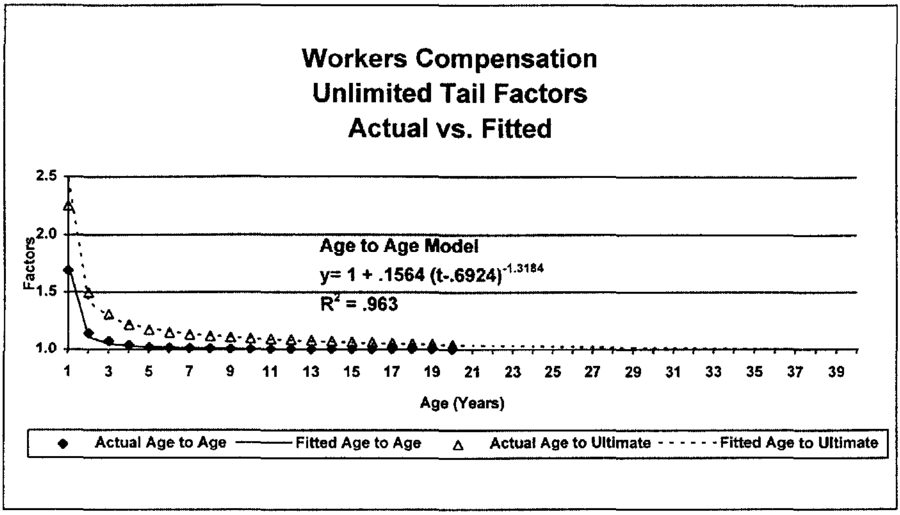
In order to handle the issue of how to develop limited losses to ultimate, I relied upon an inverse power curve recommended by Richard Sherman [3] to model the development arising in the tail. Specifically, I used a three parameter version of the curve depicted as follows:

(4.1) $y = 1 + a \cdot (t + c)^{-b}$

Again, my concern was to determine consistent tail factors by limit. Starting with the unlimited loss development and fitting an inverse power curve to known age-to-age factors allowed me to project ultimate unlimited losses. As the inverse power curve continues indefinitely, there is a need to select a time at which the projection should end. At this point I tied this approach to a similar method used for determining our full coverage tail factor that relies upon extended development triangles. That procedure suggested that I could get an equivalent result from the inverse power curve model by stopping its projected age-to-age development factors at 40 years. Compounding the age-to-age factors from the fitted curve leads to the desired completion or tail factors.

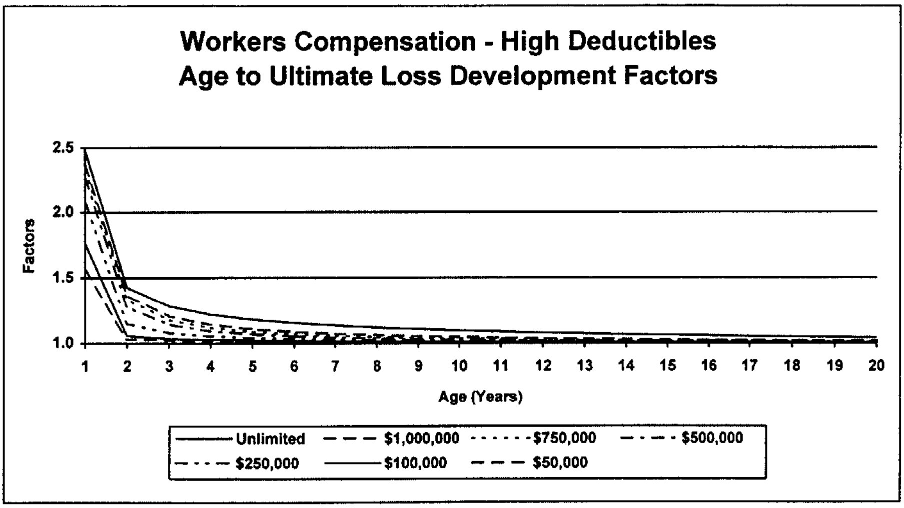
Once I set the ultimate age, I fit the inverse power curves to age-to-age factors for the various deductible limits under review and extended to that common maturity. Though this approach utilizes a consistent technique and generates uniformly decreasing tail factors, it is still an open issue whether the bias in extending all curves to a common maturity is significant or not. (At lower limits, development likely ceases well before forty years.) Chart 2 depicts the age-to-age model determined for the unlimited loss development.

Chart 2



In Chart 3 I show the pattern of age-to-ultimate limited loss development factors resulting from the inverse power curve model.

Chart 3



Another issue the actuary needs to be sensitive to is the relationship between loss development factors and limited severity relativities.[2] In some of my earlier efforts I attempted to uniquely develop losses by limit without regard to how they might relate to one another. This led to inconsistencies in development factors where completion factors for smaller deductibles, for example, sometimes exceeded factors for larger deductibles. Upon closer inspection, I found that any attempts to determine deductible development factors need to address the relationship between the full coverage loss development and severity relativities. The following formulas show the relationship between limited and excess development factors with the unlimited loss development and severity relativities.

(4.2) $$LDF^L = \frac{C}{C_t} \cdot \frac{S}{S_t} \cdot \frac{R^L}{R_t^L}$$

where L = Deductible Limit, C = Counts, S = Severity, R = Severity Relativity, and t = age

(4.3) $$XSLDF^L = \frac{C}{C_t} \cdot \frac{S}{S_t} \cdot \frac{\left(1-R^L\right)}{\left(1-R_t^L\right)}$$

where L = Deductible Limit, C = Counts, S = Severity, R = Severity Relativity, and t = age

[2] Limited severity relativities are defined simply as the ratio of the limited to unlimited severity.

Motivation for Relationship

$$(4.4)\quad LDF_t = R_t^L \cdot LDF^L + \left(1 - R_t^L\right) \cdot XSLDF^L$$

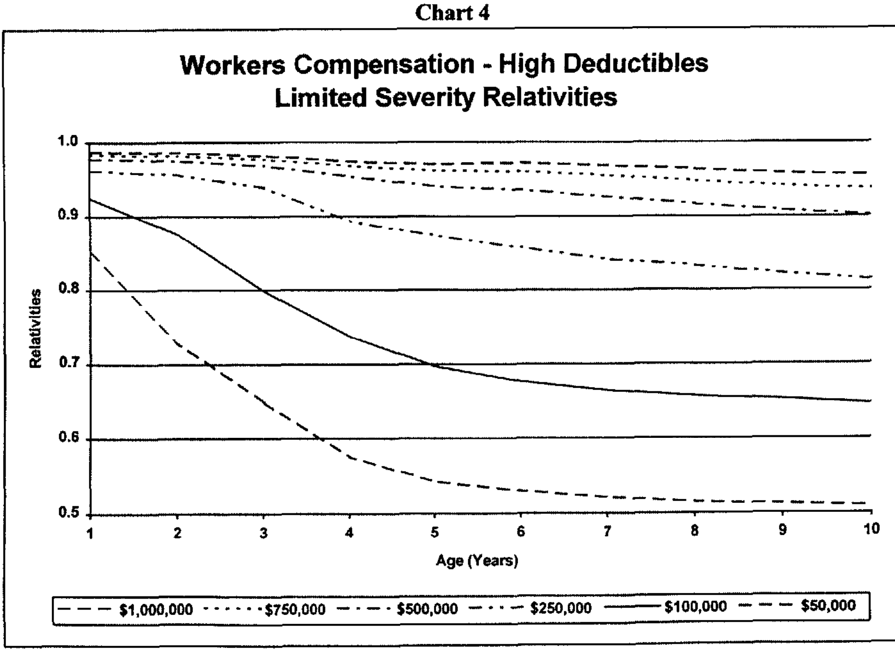
$$(4.5)\quad LDF_t = R_t^L \cdot \frac{C}{C_t} \cdot \frac{S}{S_t} \cdot \frac{R^L}{R_t^L} + (1 - R_t^L) \cdot \frac{C}{C_t} \cdot \frac{S}{S_t} \cdot \frac{\left(1 - R^L\right)}{\left(1 - R_t^L\right)}$$

$$(4.6)\quad LDF_t = \frac{C}{C_t} \cdot \frac{S}{S_t} \cdot R^L + \frac{C}{C_t} \cdot \frac{S}{S_t} \cdot (1 - R^L)$$

$$(4.7)\quad LDF_t = \frac{C}{C_t} \cdot \frac{S}{S_t}$$

The motivation for these relationships results from the desire to partition total loss development in a consistent fashion between limited and excess development. I show in Chart 4 how the historical limited severity relativities ought to relate to one another and change over time.

Chart 4



In Table 4 I show age-to-age development about a $250,000 deductible limit.

Table 4
Workers Compensation
High Deductibles

Age-to-Age Loss & ALAE Development Factors
(Unlimited)

Accident Year	12:24	24:36	36:48	48:60	60:72
1989	1.7063	1.1756	1.0929	1.0359	1.0273
1990	1.8219	1.1574	1.0744	1.0387	-
1991	1.7724	1.1506	1.0737	-	-
1992	1.6912	1.1398	-	-	-
1993	1.6044	-	-	-	-
Average	1.7192	1.1559	1.0803	1.0373	1.0273

Age-to-Age Loss & ALAE Development Factors
($250,000 Deductible)

Accident Year	12:24	24:36	36:48	48:60	60:72
1989	1.7077	1.1598	1.0657	1.0221	1.0120
1990	1.7755	1.1509	1.0550	1.0247	-
1991	1.7734	1.1461	1.0643	-	-
1992	1.6750	1.1363	-	-	-
1993	1.6229	-	-	-	-
Average	1.7109	1.1483	1.0617	1.0234	1.0120

Age-to-Age Loss & ALAE Development Factors
(Excess of $250,000 Deductible)

Accident Year	12:24	24:36	36:48	48:60	60:72
1989	1.6646	1.6582	1.6742	1.1927	1.2011
1990	4.4890	1.3049	1.3151	1.2411	-
1991	1.7373	1.3115	1.3675	-	-
1992	2.2474	1.2291	-	-	-
1993	1.1684	-	-	-	-
Average	2.2613	1.3759	1.4523	1.2169	1.2011

In Table 5 I show relativities and their changes for the selected deductible limit.

Table 5
Workers Compensation
High Deductibles

Limited Severity Relativities
($250,000 Deductible)

Accident Year	12 Months	24 Months	36 Months	48 Months	60 Months	72 Months
1989	0.9675	0.9683	0.9553	0.9315	0.9191	0.9053
1990	0.9829	0.9578	0.9524	0.9353	0.9227	-
1991	0.9723	0.9728	0.9690	0.9605	-	-
1992	0.9717	0.9623	0.9594	-	-	-
1993	0.9593	0.9704	-	-	-	-
Average	0.9707	0.9663	0.9590	0.9424	0.9209	0.9053

Changes in Limited Severity Relativities
($250,000 Deductible)

Accident Year	12:24	24:36	36:48	48:60	60:72
1989	1.0008	0.9866	0.9751	0.9867	0.9850
1990	0.9745	0.9944	0.9820	0.9865	-
1991	1.0005	0.9961	0.9912	-	-
1992	0.9903	0.9970	-	-	-
1993	1.0116	-	-	-	-
Average	0.9955	0.9935	0.9828	0.9866	0.9850

Note how the change in limited loss development relates to the unlimited loss development. Also note how actual case loss development does not always conform to expectations, as the limited loss development factor sometimes exceeds the unlimited.

(4.8) $LDF^{L} = LDF \cdot \Delta R^{L}$

For example, for accident year 1993, moving from 12 to 24 months, a limited factor of 1.6229 is observed. That is equivalent to the unlimited loss development factor of 1.6044 compounded with the change in severity relativities for the same time period of 1.0116.

Note also the relationship of the excess development to the unlimited loss development for the same year.

$$(4.9)\quad XSLDF^{L} = LDF \cdot \Delta\left(1 - R^{L}\right)$$

There the accident year 1993 excess development factor of 1.1684 is equivalent to the unlimited development factor compounded with the ratio of the complements of the severity relativities moving from 12 to 24 months. (1.1684 = (1.6044) (1 - 0.9704) / (1 - 0.9593))

And, as desired, the weighted average of the limited and excess development factors using the relativity as the weight leads to the unlimited loss development factor.

$$(4.10)\quad LDF_{t} = R_{t}^{L} \cdot LDF_{t}^{L} + \left(1 - R_{t}^{L}\right) \cdot XSLDF_{t}^{L}$$

(Accident Year 1993: 1.6044 = (0.9704) (1.6229) + (1 - 0.9704) (1.1684))

Distributional Model - A More Promising Approach

Because of the concepts just described, this whole approach seems ideally suited for the application of some form of loss distribution model. That model helps to tie the relativities to the severities and consequently provides consistent loss development factors. Not only that, a distributional model easily allows for interpolation among limits and years, as needed.

The approach I propose models the development process by determining parameters of a distribution that vary over time. Once the distribution and its parameters are specified, it is possible to calculate the desired limited/excess severities. Comparing those severities over time leads to the needed development factors. Of course, care has to be exercised to recognize claim count development at earlier maturities.

For my work, I relied upon a Weibull distribution to specify the workers compensation claim loss distribution. That distribution has been commonly used for workers compensation claims and is familiar to actuaries working with distributional models. It is ideally suited for this type of work, as it gives a reasonable depiction of the loss distributions and is easy to work with.

Of course, the most difficult aspect of working with distributional models is estimating the parameters involved. There are various approaches that can be used, including Method of Moments as well as Maximum Likelihood. I tried an alternative approach that optimizes the fit between actual and theoretical severity relativities around the $250,000 deductible size. Specifically, I minimized the chi-square between actual and expected severity relativities to determine the needed parameters. I made use of a solver routine incorporated in Microsoft Excel's spreadsheet application, which allowed me to constrain the optimization routine in such a fashion that the parameters generated produced the actual unlimited severity at the specified maturity.

I show in Table 6 an example of results used to determine age-to-ultimate loss development factors by limit from 48 months to ultimate. I selected 48 months in order to focus attention on changes in severity rather than changes in total claim counts assuming no IBNR count development after 36 months. (Please see Appendix I for details.)

Table 6
Workers Compensation High Deductibles
Actual Versus Fitted Limited/Excess Development Factors (@ 48 Months)
(using a Weibull Loss Distribution)

Ultimate

Limit	Unlimited	$1,000,000	$750,000	$500,000	$250,000	$100,000	$50,000
			Observed				
Limited Severity	6,846.4	6,159.2	5,980.4	5,714.4	5,094.8	3,939.6	3,036.5
Relativity	1.0000	0.8996	0.8735	0.8347	0.7442	0.5754	0.4435
Excess Severity	0.0	687.2	866.0	1,132.0	1,751.6	2,906.8	3,809.9
			Fitted				
Limited Severity	6,846.4	6,295.2	6,106.5	5,778.7	5,064.4	3,926.7	3,043.8
Relativity	1.0000	0.9195	0.8919	0.8440	0.7397	0.5735	0.4446
Excess Severity	0.0	551.2	739.9	1,067.7	1,782.0	2,919.7	3,802.6

Weibull Parameters Scale = 180.0 Shape = .2326
Mean = 6,846.4 Coefficient of Variation = 10.07

48 Months

Limit	Unlimited	$1,000,000	$750,000	$500,000	$250,000	$100,000	$50,000
			Observed				
Limited Severity	5,530.2	5,346.6	5,288.5	5,182.3	4,824.0	3,807.5	2,937.1
Relativity	1.0000	0.9668	0.9563	0.9371	0.8723	0.6885	0.5311
Limited LDF	1.2380	1.1520	1.1308	1.1027	1.0561	1.0347	1.0338
Excess Severity	0.0	183.6	241.7	347.9	706.2	1,722.7	2,593.1
Excess LDF	-	3.7429	3.5830	3.2538	2.4803	1.6874	1.4692
			Fitted				
Limited Severity	5,530.2	5,380.5	5,301.4	5,142.5	4,722.4	3,894.0	3,144.1
Relativity	1.0000	0.9729	0.9586	0.9299	0.8539	0.7041	0.5685
Limited LDF	1.2380	1.1700	1.1519	1.1237	1.0724	1.0084	0.9681
Excess Severity	0.0	149.7	228.8	387.7	807.8	1,636.2	2,386.1
Excess LDF	-	3.6820	3.2338	2.7539	2.2060	1.7844	1.5936

Weibull Parameters Scale = 305.7 Shape = .2625
Mean = 5,530.2 Coefficient of Variation = 7.35

Lastly, the following formulation shows how expected development can be partitioned about the deductible limit.

$$(4.11)\quad \text{Expected Development} = 1 - \frac{1}{LDF_t}$$

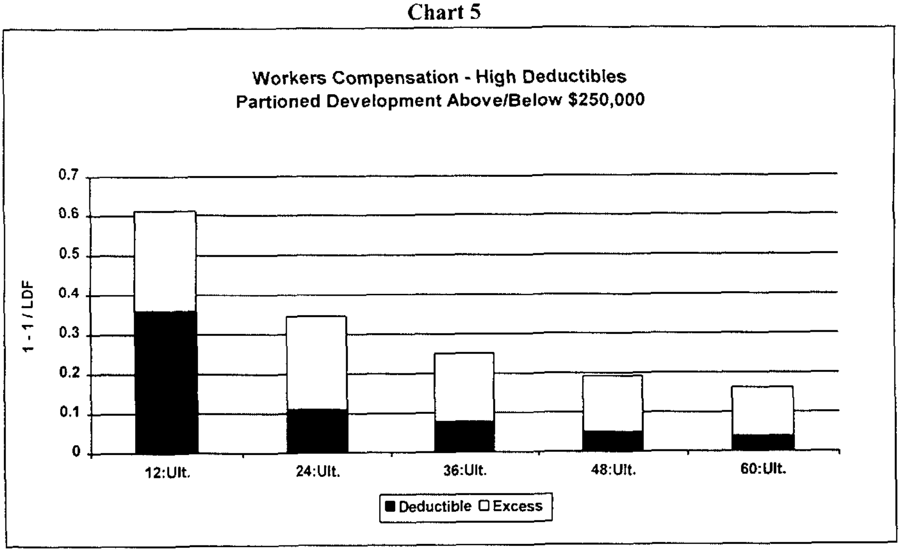
$$(4.12)\quad = \frac{LDF_t - 1}{LDF_t}$$

$$(4.13)\quad = \frac{R_t^L \cdot LDF_t^L + \left(1 - R_t^L\right) \cdot XSLDF_t^L - 1}{R_t^L \cdot LDF_t^L + \left(1 - R_t^L\right) \cdot XSLDF_t^L}$$

$$(4.14)\quad = \frac{R_t^L \cdot \left(LDF_t^L - 1\right) + \left(1 - R_t^L\right) \cdot \left(XSLDF_t^L - 1\right)}{R_t^L \cdot LDF_t^L + \left(1 - R_t^L\right) \cdot XSLDF_t^L}$$

I show graphically in Chart 5 partitioned development for a selected \$250,000 deductible limit based upon the previously described Weibull loss distribution model. Note the changing proportions of development over time. Not unexpectedly, excess development represents the vast majority of development with increasing age.

Chart 5



5. Other Elements

Several other elements associated with high deductible plans call for further discussion: aggregate limits, service revenue and allocated claim expense. Determining sound estimates for those items involves a fair amount of complexity. In the following discussion I recommend using advanced collective risk modeling techniques to estimate losses excess of aggregate limits. I also suggest an alternative procedure using the NCCI Table M, if collective risk modeling is not considered practical. The asset for service revenue, though not as difficult to determine, however, depends upon prior estimates of losses for deductible/aggregate limits. Treating allocated claim expense in a similar fashion to loss simplifies the estimation process for that liability, but separating the two pieces is problematic.

Aggregate Limits

Some risks, besides choosing to limit their per occurrence losses, desire to limit all losses that they will pay under a high deductible program. Insurers satisfy that need by providing aggregate loss limits. Those limits are conceptually similar to maximum premium limitations used in retro rating plans.

Determining loss development factors for losses excess of aggregate limits is more complicated than for per occurrence limitations. However, the obligations arising from those aggregate limits are generally less significant than for per occurrence limits. Besides the additional complexity, the data needed to determine development factors for these limits is generally sparse and not likely to be very credible. Outside of actually attempting to gather data for development factors of this sort, I suggest making use of collective risk modeling techniques to determine the needed loss development factors. Such a model could utilize the loss distributions just described for the deductible limits in conjunction with selected claim frequency distributions.

I used a collective risk model described by Heckman and Meyers [4] to determine development factors for losses excess of aggregate limits. I show in Table 7 selected development factors using the same Weibull loss distribution I used previously to determine deductible development factors. I assumed a Poisson claim count distribution to model frequency. Though I did not incorporate any parameter risk in determining the development factors, the model does allow for that possibility. I refer the interested reader to a discussion by Meyers and Schenker [5] describing how to incorporate parameter risk into the collective risk model.

The sampling of development factors I calculated shows that development for losses excess of aggregate limits decreases more rapidly over time with smaller deductibles than larger ones. That is not unexpected as most of the later development occurs in the layers of loss above the deductible limits, which is not covered by the aggregate. Also, not unexpectedly, development is more leveraged for larger aggregate limits. There is one additional point the reader should note in reviewing Table 7. Though I show hypothetical results for risks of $1 million and $2.5 million in expected loss size, the limited expectations are considerably smaller.

Table 7
Workers Compensation High Deductibles
Development Factors for Losses Excess of Aggregate Limits
(Collective Risk Model Utilizing Weibull Loss Distribution)

Expected Unlimited Losses of $1,000,000

Aggregate Limit = 500,000

Deductible	12 Months Excess Loss	12 Months LDF	48 Months Excess Loss	48 Months LDF	Ultimate Excess Loss
$100,000	9,253.6	13.024	114,646.0	1.051	120,523.3
$250,000	22,882.5	12.007	228,070.7	1.205	274,761.6
$500,000	28,653.6	13.255	289,389.2	1.312	379,794.3

Aggregate Limit = 750,000

Deductible	12 Months Excess Loss	12 Months LDF	48 Months Excess Loss	48 Months LDF	Ultimate Excess Loss
$100,000	155.1	136.451	18,005.9	1.175	21,163.6
$250,000	1,844.9	63.845	84,475.1	1.394	117,788.5
$500,000	4,257.2	49.763	138,526.3	1.529	211,851.8

Aggregate Limit = 1,000,000

Deductible	12 Months Excess Loss	12 Months LDF	48 Months Excess Loss	48 Months LDF	Ultimate Excess Loss
$100,000	.8	2,242.750	1,274.7	1.408	1,794.2
$250,000	94.5	418.531	23,343.1	1.694	39,551.2
$500,000	494.5	213.275	57,471.2	1.835	105,464.6

Expected Unlimited Losses of $2,500,000

Aggregate Limit = 1,000,000

Deductible	12 Months Excess Loss	12 Months LDF	48 Months Excess Loss	48 Months LDF	Ultimate Excess Loss
$100,000	39,703.2	11.761	456,498.9	1.023	466,934.1
$250,000	81,084.7	10.876	759,354.4	1.161	881,844.0
$500,000	95,069.6	12.021	912,976.1	1.252	1,142,866.6

Aggregate Limit = 1,250,000

Deductible	12 Months Excess Loss	12 Months LDF	48 Months Excess Loss	48 Months LDF	Ultimate Excess Loss
$100,000	3,829.0	64.779	236,271.2	1.050	248,037.5
$250,000	17,740.7	36.191	522,364.3	1.229	642,046.5
$500,000	26,520.1	33.986	674,759.3	1.336	901,315.4

Aggregate Limit = 1,500,000

Deductible	12 Months Excess Loss	12 Months LDF	48 Months Excess Loss	48 Months LDF	Ultimate Excess Loss
$100,000	173.5	564.077	87,988.1	1.112	97,867.3
$250,000	2,693.1	158.522	318,464.5	1.341	426,916.3
$500,000	6,001.8	112.833	463,359.8	1.461	677,200.3

Given the volatility of losses excess of aggregate limits, I recommend using a Bornhuetter-Ferguson method to smooth out indications of ultimate liability. The example I show in Table 8 makes use of expected aggregate loss charges as well as expected development factors based upon the previously described collective risk modeling approach. The final indication adds together known losses excess of aggregate limits and IBNR based upon the modeled development patterns.

Table 8
Countrywide Insurance Enterprise
Workers Compensation High Deductibles
Estimated Ultimate Aggregate Excess of Loss
(Utilizing Bornhuetter-Ferguson Methodology)

	Known Loss (@ 48 Months)			Aggregate Excess of Loss		
Account	Deductible	Deductible	Excess of Aggregate	Expected	LDF	Indicated
Expected Unlimited Loss - 1,000,000; Aggregate Limit - 750,000						
A	100,000	581,252	-	21,164	1.175	3,152
B	250,000	703,027	-	117,789	1.394	33,292
C	500,000	764,493	14,493	211,852	1.529	87,789
Expected Unlimited Loss - 2,500,000; Aggregate Limit - 1,250,000						
X	100,000	1,453,169	203,169	248,038	1.050	214,980
Y	250,000	1,757,616	507,616	642,047	1.229	627,248
Z	500,000	1,911,285	661,285	901,315	1.336	887,963

An alternative approach for determining IBNR estimates for aggregate excess of loss coverage merits consideration. That procedure utilizes the NCCI methodology [1] for determining insurance charges in retrospective rating plans. I consider it a more practical approach than collective risk modeling, but its accuracy hinges upon determining the proper insurance charge table.

Essentially the IBNR is determined by subtracting insurance charges at different maturities. The process used to determine the ultimate insurance charge would be the same as that used for pricing purposes. The key to the NCCI procedure is the adjustment of expected losses reflecting loss limits. That adjustment increases expected losses used in determining the appropriate insurance charge table by use of the following formula:

$$(5.1)\quad \text{Adjustment Factor} = \frac{(1+0.8\cdot\chi)}{(1-\chi)}$$

where χ = per occurrence charge

The intent of increasing expected losses for the use of a per occurrence limit is to utilize a less dispersed loss ratio distribution and, consequently, a smaller insurance charge. Though this adjustment for a loss limit moves the selection of an insurance charge table in the right direction, the question remains whether it does so in an appropriate manner. Additionally, the procedure generates smaller insurance charges by the use of limited losses in the entry ratio calculation.

In order to calculate the insurance charge at earlier maturities I suggest determining the per occurrence charge used in the NCCI procedure by relating undeveloped, limited losses to ultimate, unlimited losses. For example, using the fitted results depicted in Table 6 for a 250,000 deductible leads to a per occurrence charge of 31 percent (1 - 4722.4 / 6846.4) at 48 months. Besides reflecting the impact of the limit, this approach also captures the effects of development. Again, the issue remains whether or not the adjustment for both the limit and development is appropriate.

I show in Table 9 a comparison of IBNR estimates determined using the NCCI Table M with estimates from the previously described collective risk modeling approach depicted in Table 8. I further detail IBNR estimates from the NCCI Table M in Appendix II.

Table 9
A Comparison of Aggregate Excess of Loss IBNR Estimates (@ 48 Months)
Collective Risk Model Versus NCCI Table M

Account	Deductible	Collective Risk Model	NCCI Table M
	Expected Unlimited Loss - 1,000,000; Aggregate Limit - 750,000		
A	100,000	3,152	1,809
B	250,000	33,292	38,500
C	500,000	73,296	68,811
	Expected Unlimited Loss - 2,500,000; Aggregate Limit - 1,250,000		
X	100,000	11,811	9,959
Y	250,000	119,633	103,000
Z	500,000	226,678	222,168

Service Revenue

A significant element that ought to be reflected on the asset side of the balance sheet is the revenue associated with servicing claims under a high deductible program. As I noted earlier, service revenue is generated in an analogous fashion to the use of a loss conversion factor in a retro rating plan. Generally, a factor is applied to deductible losses, limited by any applicable aggregate, to cover expenses that vary with those losses. In practice, however, other elements are captured by the loss multiplier, reflecting the desire of the individual accounts to fund the cost of the program as losses emerge. The service revenue is often collected as losses are paid, but it may also be gathered as a function of case incurred losses.

I propose determining the asset in the following fashion:

1. Determine ultimate deductible losses at the account level.
2. Subtract ultimate losses excess of aggregate limits from ultimate deductible losses.
3. Apply the selected loss multiplier to the difference determined in step 2 to determine ultimate recoverables.
4. Determine the total asset by subtracting any known recoveries from the estimated ultimate recoverables and aggregate results for all accounts.

Table 10 shows an example of how in practice the asset for the service revenue might be determined.

Table 10
Countrywide Insurance Enterprise
Workers Compensation - High Deductibles
Estimated Ultimate Service Revenue

Expected Unlimited Loss - 2,500,000; Aggregate Limit - 1,250,000; Loss Multiplier - 10%

	Ultimate Loss					
Account	Deductible	Excess of Aggregate	Net of Aggregate	Multiplier Revenue	Known Recoveries	Asset
X	1,465,376	214,980	1,250,396	125,040	96,960	28,080
Y	1,884,867	627,248	1,257,619	125,762	102,712	23,050
Z	2,147,711	887,963	1,259,748	125,975	106,912	19,063
Total	5,497,954	1,730,191	3,767,763	376,777	306,584	70,193

Allocated Claim Expense

There are two principal means of handling allocated claim expense under a high deductible program. Either the account manages this expense itself or it is treated as loss and subjected to applicable limits. In the first instance development patterns reflecting loss only would be appropriate for determining liabilities, while a combination of loss and expense is appropriate for the second case. For this discussion I determined development factors combining loss and expense components assuming expenses were equivalent to additional loss dollars. Though different development patterns are likely for loss and expense versus loss only, the gain in precision is likely not worth the effort.

A remaining issue is how best to split loss and allocated claim expense for financial reporting purposes. Though splitting them proportionately based upon their full coverage counterparts is expeditious, other more actuarially sound approaches, even if available, may not be cost justifiable.

6. Conclusion

I intended with this discussion to suggest some possible approaches for estimating liabilities under a high deductible program. As with many actuarial procedures, much work and improvement are still needed. I hope my suggestions provoke further discussion as to how to better estimate these liabilities.

Although the reader probably has many ideas to improve upon the suggestions I have made, I feel several stand out including:

- Obtain longer histories of experience under the program better reflecting risk characteristics.
- Derive (Select) parameters (distributions) that provide better fits to the actual data.
- Determine better tail factors and/or parameters of the utilized loss distribution.
- Develop more advanced approaches to index loss limits.

None of these are really unknown issues for actuaries, who have long been confronted with developing either limited or excess losses. The availability of more comprehensive data in a workers compensation program allows for the application of more sophisticated loss distributional approaches that affords greater consistency to all of the pieces involved. To that end I hope this paper provides a few steps toward developing sounder actuarial techniques for analyzing workers compensation high deductible loss development.

REFERENCES

[1] National Council on Compensation Insurance, *Retrospective Rating Plan Manual for Workers' Compensation and Employers Liability Insurance - Appendix*

[2] Bornhuetter, Ronald L., and Ferguson, Ronald E., "The Actuary and IBNR," *PCAS* LIX, 1972, p. 181

[3] Sherman, Richard E., "Extrapolating, Smoothing, and Interpolating Development Factors," *PCAS* LXXI, 1984, p. 122

[4] Heckman, Philip E., and Meyers, Glenn G., "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *PCAS* LXX, 1983, p. 22

[5] Meyers, Glenn G., and Schenker, Nathaniel, "Parameter Uncertainty in the Collective Risk Model," *PCAS* LXX, 1983, p. 111

Appendix I

Weibull Distribution

1. Cumulative Distribution Function $F(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$; where $x > 0, \beta > 0, \alpha > 0$

2. Probability Density Function $f(x) = \frac{\alpha \cdot x^{\alpha - 1}}{\beta^{\alpha}} \cdot e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$

3. $E(x) = \beta \cdot \Gamma\left(\frac{1}{\alpha} + 1\right)$; where $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha - 1} e^{-x} dx$

4. $LEV(x) = \beta \cdot \Gamma\left(\frac{1}{\alpha} + 1\right) \cdot \Gamma\left(\frac{1}{\alpha} + 1; \left(\frac{x}{\beta}\right)^{\alpha}\right) + x \cdot e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$

LDF calculations about $250,000 deductible limit

Severities at ultimate
$\beta = 180.0; \alpha = .2326$

$$E(x) = 180.0 \cdot \Gamma\left(\frac{1}{.2326} + 1\right) = 6846$$

$$LEV(x) = 6846 \cdot \Gamma\left[\frac{1}{.2326} + 1; \left(\frac{250000}{180}\right)^{.2326}\right] + 250000 \cdot \left(e^{-\left(\frac{250000}{180}\right)^{.2326}}\right) = 5064$$

$$E(x) - LEV(x) = 6846 - 5064 = 1782$$

Severities at 48 Months
$\beta = 305.7; \alpha = .2625$

$$E(x) = 305.7 \cdot \Gamma\left(\frac{1}{.2625} + 1\right) = 5530$$

$$LEV(x) = 5530 \cdot \Gamma\left[\frac{1}{.2625} + 1; \left(\frac{250000}{305.7}\right)^{.2625}\right] + 250000 \cdot \left(e^{-\left(\frac{250000}{305.7}\right)^{.2625}}\right) = 4722$$

$$E(x) - LEV(x) = 5530 - 4722 = 808$$

$$LDF_{48} = \frac{6846}{5530} = 1.238$$

$$LDF_{48}^{250000} = \frac{5064}{4722} = 1.072$$

$$XSLDF_{48}^{250000} = \frac{1782}{808} = 2.205$$

Appendix II

Determination of IBNR for an Aggregate Excess of 1,250,000
Risk Characteristics: Expected Unlimited Loss - 2,500,000;
Severity - 6846.4; Frequency - 365.2

	48 Months	Ultimate
a. Severity: Deductible = 250,000	4722.4	5064.4
b. Frequency	365.2	365.2
c. Limited Loss: a • b	1,724,620.5	1,849,518.9
d. Entry Ratio: Aggregate / c	0.72	0.68
e. Loss Excess of Deductible: 1 - LEV(x) / E(x)	0.310	0.260
f. Adjustment for Limit: (1 + .8 • e) / (1 - e)	1.810	1.633
g. Adjusted Limited Loss: Expected Unlimited Loss • f	4,525,000	4,082,500
h. 1994 Expected Loss Group	19	20
i. Insurance Charge Ratio	.336	.369
j. Insurance Charge Amount: c • i	579,472	682,472
k. IBNR	682,472 - 579,472 = 103,000	

CAS MONOGRAPH SERIES
NUMBER 3

STOCHASTIC LOSS RESERVING USING GENERALIZED LINEAR MODELS

Greg Taylor and Gráinne McGuire

CASUALTY ACTUARIAL SOCIETY

The purpose of the monograph is to provide access to generalized linear models for loss reserving but initially with strong emphasis on the chain ladder. The chain ladder is formulated in a GLM context, as is the statistical distribution of the loss reserve. This structure is then used to test the need for departure from the chain ladder model and to formulate any required model extensions.

The chain ladder is by far the most widely used method for loss reserving. The chain ladder algorithm itself is non-stochastic, but Mack (1993) defined a stochastic version of the model and showed how a mean square error of prediction may be associated with any loss reserve obtained from this model.

There are, however, two families of stochastic model which generate the chain ladder algorithm for the estimation of loss reserve, as discussed in Taylor (2011). They require differing treatments for the estimation of mean square error of prediction. Both families of model may be formulated as generalized linear models. This is not widely appreciated of the Mack model. The monograph commences with the identification of these two families and their respective GLM formulations.

GLM formulation naturally invites the use of a bootstrap to estimate prediction error. The bootstrap estimates the entire distribution of loss reserve rather than just the mean square error of prediction obtainable from Mack's algorithm. The monograph discusses both parametric and semi-parametric forms of the GLM bootstrap.

Emphasis is placed on the use of statistical software to implement the GLM formulation. This formulation and the associated software provide the diagnostics for testing the validity of the model. This aspect is covered by the existing literature but the monograph reviews this material in view of its importance.

Practical applications of the chain ladder often depart from the strict model. There are a number of causes but prominent among these are:

- the need to smooth the age-to-age factor tail;
- the need to give greater weight to more recent data than to older.

These two matters are considered within the GLM context. The subject of smoothing leads to a discussion of generalized additive models.

As regards the second point, the GLM structure is used to test whether or not data are time-homogeneous (as is required by the strict chain ladder model) and, if not, to suggest a procedure for recognising and accommodating time-heterogeneity in the data. This may lead to the common practice of discarding all but the last m diagonals of the claim triangle, but more general approaches are also be considered.

As time-heterogeneity is not consistent with the chain ladder model, it amounts to model failure, and is recognizable from the diagnostics introduced above. Various forms of model failure are considered and, in each case, a model extension constructed to deal with it.

Finally, extension to several models that go beyond the scope of generalized linear models is discussed.

STOCHASTIC LOSS RESERVING USING GENERALIZED LINEAR MODELS

Greg Taylor and Gráinne McGuire

Casualty Actuarial Society
4350 North Fairfax Drive, Suite 250
Arlington, Virginia 22203
www.casact.org
(703) 276-3100

Stochastic Loss Reserving Using Generalized Linear Models
By Greg Taylor and Gráinne McGuire

Copyright 2016 by the Casualty Actuarial Society

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. For information on obtaining permission for use of material in this work, please submit a written request to the Casualty Actuarial Society.

Library of Congress Cataloging-in-Publication Data
Taylor, Greg
McGuire, Gráinne
Stochastic Loss Reserving Using Generalized Linear Models / Greg Taylor and Gráinne McGuire
ISBN 978-0-9968897-0-4 (print edition)
ISBN 978-0-9968897-1-1 (electronic edition)
1. Actuarial science. 2. Loss reserving 3. Insurance—mathematical models.
I. Taylor, Greg. II. McGuire, Gráinne.

Contents

About the Authors

Greg Taylor holds an honorary professorial position in Risk and Actuarial Studies at the University of New South Wales. He previously spent 44 years in commercial actuarial practice and eight years as an actuarial academic. Taylor has published two books on loss reserving and numerous articles in mathematics, statistics, and actuarial science. He is an Officer of the Order of Australia, and holds a Gold Medal from the Australian Actuaries Institute and a Silver Medal from the United Kingdom Institute and Faculty of Actuaries.

Gráinne McGuire is a consulting actuary with Taylor Fry, with over 12 years' experience in that role, and has specialized in complex claim modeling and loss reserving using GLMs. She has co-authored a number of papers on various topics in loss reserving including the use of GLMs, bootstrapping and automatic reserving methods (of which one paper, co-authored with Greg Taylor, was awarded the Michelbacher prize in 2004). She has also been a member of GIRO working parties examining stochastic loss reserving and uncertainty measurement. She holds a Ph.D. in statistics and is a Fellow of the Actuaries Institute, Australia.

2015 CAS Monograph Editorial Board

C. K. “Stan” Khury, Editor in Chief
Emmanuel Bardis
Craig Davis
Richard Fein
Jesse Groman
Ali Ishaq
Leslie Marlo
Sholom Feldblum, consultant
Glenn Meyers, consultant
Katya Prell, consultant

Foreword

The oral tradition of the CAS has it that what we know today as the "chain-ladder method" was first used in the 1950s at a small Midwestern insurance company. In fact, the method during those early years was named after that small Midwestern mutual insurance company. Since that time, its name has evolved variously to the "loss development method," the "chain-ladder method," and the "link ratio" method. Since those early days, its use spread to other companies and ultimately became the most widely used actuarial methodology for estimating ultimate losses. This monograph begins at that same point, a point that, in effect, saw the chain-ladder method as a heuristic.

In this work, Taylor and McGuire note the evolution of the chain-ladder method through its various developmental stages: from the first time the estimate produced by the chain-ladder method was recognized as a maximum likelihood estimate of a stochastic model (Hachemeister and Stanard, 1975), through the development of a non-parametric model that recognized variance in the observations (Mack, 1993), and then the development of a collection of models that fit this description (Taylor, 2011), and on to the recent demonstration that all these models may be represented by generalized linear models (Taylor, 2015).

In addition to describing the various formal models for which the chain ladder algorithm provides a maximum likelihood estimate of ultimate losses, the authors show how the generalized linear model outputs may be used to estimate the associated prediction error and thus test whether the chain ladder is a reasonable representation of the claim data. The authors also show how adjustments to recognize eccentricities in the data could be made within a GLM formulation. The authors introduce two variations of the chain-ladder method that could not be contemplated within the conventional chain-ladder framework.

The authors conclude by introducing a series of model extensions that deal with a variety of conditions that are faced in the daily work of an actuary.

The authors make use of two devices that facilitate the assimilation of the content of this monograph: one is that each chapter begins with a brief abstract that describes the contents in direct simple terms and the other is that a single data set is used throughout the monograph to illustrate the results of various models and their variations. To this end, the reader is able to compare outputs and points of sensitivity among the various model presentations.

This monograph in effect covers the chain-ladder method from its humble beginnings through all the layers that ultimately identify its stochastic parent distributions in their most generalized form. It makes for a complete presentation that practicing actuaries can put to good use. The Monograph Editorial Board is grateful to the authors for a valuable contribution to the casualty actuarial literature.

C. K. "Stan" Khury
Chairperson
Monograph Editorial Board

1. The Chain Ladder Algorithm

Chapter summary. The claims triangle, and its generalization to arrays of other shapes, is introduced, along with notation and a few basic concepts such as that of outstanding losses. A data set to be used consistently through a number of numerical examples is also introduced.

Next the chain ladder algorithm is introduced, and illustrated by application to the example data set. The Bornhuetter-Ferguson and Cape Cod extensions of the chain ladder are described.

1.1. Introduction

The chain ladder is the most ubiquitous of loss reserving models. For much of its life it existed as an algorithm rather than a model. Here "algorithm" implies a mere calculation procedure, not necessarily subject to any rigorous theoretical foundation.

This was remedied by Hachemeister and Stanard (1975) who defined a stochastic model of claims data for which chain ladder estimation was found to be maximum likelihood (**"ML"**). Subsequently, the collection of models that fit this description was extended, as discussed by Taylor (2011).

It was further shown (Taylor 2015) that all of these models could be represented as generalized linear models (**"GLMs"**), enabling their parameter estimation by means of statistical software. The use of this software also returns a good deal of additional information about the model, particularly the dispersion of the parameter estimates. This may be used as the basis for estimation of the prediction error associated with the model.

The purpose of this monograph is to provide a brief account of these matters, specifically:

- to describe the various formal models for which the chain ladder algorithm provides an ML forecast of loss reserve;
- to discuss how these models may be used to estimate the associated prediction error;
- to discuss how the output of GLM software may be used to test whether the chain ladder is indeed a reasonable representation of the claims data; and
- to consider some natural extensions of the chain ladder that are well adapted to the GLM framework.

A prior knowledge of the chain ladder as a heuristic loss reserving algorithm, though not its theoretical properties, is assumed. Some of the latter will be discussed in Chapter 3.

Although the essentials of GLMs are reviewed, a nodding acquaintance of the reader with them would be distinct advantage.

In any event, the purpose of the monograph is **not** to provide a primer on either the chain ladder or GLMs, but rather to show that the former may be placed within the context of the latter with many beneficial results. The intention is to provide this in tight, minimalist mathematical form.

To venture into a more discursive approach to the intuition of the modeling would expand this work considerably, perhaps beyond monograph length. The reader interested in a more intuitive approach to GLMs might consult Lindsey (1997).

1.2. Framework and Notation

It will be convenient to follow the framework and notation of Buchwalder, Bühlmann, Merz and Wüthrich (2006). They consider a $K \times J$ rectangle of claims observations Y_{kj} with:

- accident periods represented by rows and labelled $k = 1, 2, \ldots, K$;
- development periods represented by columns and labelled by $j = 1, 2, \ldots, J \le K$.

Within the rectangle they identify a **development trapezoid** of **past** observations

$$\mathfrak{D}_K = \left\{ Y_{kj} \colon 1 \le k \le K \text{ and } 1 \le j \le \min\left(J, K - k + 1\right) \right\}$$

The complement of this subset, representing **future** observations is

$$\begin{aligned} \mathfrak{D}_K^c &= \left\{ Y_{kj} \colon 1 \le k \le K \text{ and } \min\left(J, K - k + 1\right) < j \le J \right\} \\ &= \left\{ Y_{kj} \colon K - J + 1 < k \le K \text{ and } K - k + 1 < j \le J \right\} \end{aligned}$$

Also let

$$\mathfrak{D}_K^+ = \mathfrak{D}_K \cup \mathfrak{D}_K^c$$

On the d-th diagonal of $\mathfrak{D}_K$, $k + j - 1 = d$, and so the diagonal represents claims experience from the d-th **calendar period** contained in the trapezoid. Diagonals will be referred to as **experience periods**. The final diagonal of $\mathfrak{D}_K$ is the K-th diagonal, consisting of observations $Y_{k,K-k+1}$, $k = K - J + 1, \ldots, K$.

In general, the problem is to predict $\mathfrak{D}_K^c$ on the basis of observed $\mathfrak{D}_K$.

At this stage the nature of the observations Y_{kj} will be left unspecified. They might be defined to be paid losses, reported claim counts, etc. The mathematical structure of the chain ladder model does not require stipulation of this.

The usual case in the literature (though often not in practice) is that in which $J = K$, so that the trapezoid becomes a triangle. The more general trapezoid will be retained throughout the present monograph.

Define the **cumulative row sums**

$$X_{kj} = \sum_{i=1}^{j} Y_{ki} \tag{1-1}$$

Let $\Sigma^{\mathcal{R}(k)}$ denote summation over the entire row k of $\mathfrak{D}_K$, i.e., $\sum_{j=1}^{\min(J,K-k+1)}$ for fixed k.

Similarly, let $\Sigma^{\mathcal{C}(j)}$ denote summation over the entire column j of $\mathfrak{D}_K$, i.e., $\sum_{k=1}^{K-j+1}$ for fixed j.

Also define, for $k = K - J + 2, \ldots, K$,

$$R_k = \sum_{j=K-k+2}^{J} Y_{kj} = X_{kJ} - X_{k,K-k+1} \tag{1-2}$$

$$R = \sum_{k=2}^{K} R_k \tag{1-3}$$

Note that R is the sum of the (future) observations in $\mathfrak{D}_K^c$. It will be referred to as the total amount of **outstanding losses**. Likewise, R_k denotes the amount of outstanding losses in respect of accident period k. The objective stated earlier is to forecast the R_k and R.

1.3. Data for Numerical Examples

A number of the developments described in subsequent chapters will be illustrated by numerical example. It will be convenient to relate all examples to the same data set. The chosen data set appears as Table 1-1. It will be referred to henceforth as **"the example data set"**.

It is seen that the generic "observations" Y_{kj} of Section 1.2 have now been particularized as incremental paid losses.

The triangle has been obtained from the data base of Meyers and Shi (2011). It is in fact the workers compensation triangle of the New Jersey Manufacturers Group.

Table 1-1. Triangle of Incremental Paid Losses for Numerical Examples

		Incremental Paid Losses in Development Year ($000)									
Accident Year		1	2	3	4	5	6	7	8	9	10
1988	1	41821	34729	20147	15965	11285	5924	4775	3742	3435	2958
1989	2	48167	39495	24444	18178	10840	7379	5683	4758	3959	
1990	3	52058	47459	27359	17916	11448	8846	5869	5391		
1991	4	57251	49510	27036	20871	14304	10552	7742			
1992	5	59213	54129	29566	22484	14114	10000				
1993	6	59475	52076	26836	22332	14756					
1994	7	65607	44648	27062	22655						
1995	8	56748	39315	26748							
1996	9	52212	40030								
1997	10	43962									

The "Accident year" column shows the actual accident year, and then its translated version in which the earliest accident year has been re-labelled "1", as in the general framework set out in Section 1.2. This dual notation will be retained through subsequent chapters.

Although remaining chapters will be concerned with just this one type of triangle (an **"incremental paid loss triangle"**), it should be understood that there are many other commonly used types, namely:

- **"cumulative paid loss triangles"**, in which each entry is equal to total payments up to and including the relevant development year of the row concerned, i.e., the entry in the (k, j) cell is X_{kj} instead of Y_{kj} as in the above example;
- "incurred loss triangles", in which the entry in the (k, j) cell is the insurer's estimate, as at the end of development year j, of the total claim cost incurred in accident year k, i.e., X_{kj} plus the insurer's estimate of the claim cost remaining unpaid at the end of development year j.

The incurred loss triangles might reasonably be referred to as **"cumulative incurred loss triangles"**, and one might define **"incremental incurred loss triangles"**, obtained by differencing rows of the cumulative incurred loss triangles.

There are yet other triangles. These include triangles of **claim counts**, instead of claim amounts. These might contain, for example, counts of:

- Reported claims;
- Finalized claims;
- Unfinalized claims.

These data are required by the models explored in Chapter 4 of Taylor (2000).

1.4. The Chain Ladder Algorithm

This section will give a statement of the chain ladder algorithm as it has been used in years past. The description below is taken largely from Mack (1993).

Define the following **age-to-age factors**:

$$\hat{f}_{kj} = X_{k,j+1} \big/ X_{kj}, \quad k = 1, 2, \ldots, K-1;\ j = 1, 2, \ldots, \min(J-1, K-k) \tag{1-4}$$

and the **weighted average age-to-age factors**:

$$\hat{f}_j = \sum_{k=1}^{K-j} w_{kj} \hat{f}_{kj}, \quad j = 1, 2, \ldots, J-1 \tag{1-5}$$

where, for each fixed j, $\{w_{kj}, k = 1, 2, \ldots, K-j\}$ is some set of **weights**, i.e., $w_{kj} \geq 0$ and

$$\sum_{k=1}^{K-j} w_{kj} = 1 \tag{1-6}$$

Suppose the weights are chosen as

$$w_{kj} = X_{kj} \Big/ \sum_{k=1}^{K-j} X_{kj} \tag{1-7}$$

Then the weighted average age-to-age factors in (1-5) become

$$\hat{f}_j = \sum_{k=1}^{K-j} X_{k,j+1} \Big/ \sum_{k=1}^{K-j} X_{kj} \tag{1-8}$$

Now define the following forecasts of the X_{kj} corresponding to the $Y_{kj} \in \mathfrak{D}_K^c$:

$$\hat{X}_{kj} = X_{k,K-k+1}\hat{f}_{K-k+1} \ldots \hat{f}_{j-1} \tag{1-9}$$

whence, by (1-4), the forecasts of the Y_{kj} are:

$$\hat{Y}_{kj} = X_{k,K-k+1}\hat{f}_{K-k+1} \ldots \hat{f}_{j-2}\left(\hat{f}_{j-1} - 1\right) \tag{1-10}$$

It follows from (1-5) that outstanding losses R_k are estimated by

$$\hat{R}_k = \hat{X}_{kJ} - X_{k,K-k+1} = X_{k,K-k+1}\left(\hat{f}_{K-k+1} \ldots \hat{f}_{J-1} - 1\right) \tag{1-11}$$

Finally, denote total (over all accident years) outstanding losses by R and their estimate by

$$\hat{R} = \sum_{k=1}^{K-1} \hat{R}_k \tag{1-12}$$

As the heading of the current section indicates, the estimation schema (1-8) to (1-12) is only an algorithm, not a model. No model has yet been formulated in the sense of expressing the observations in terms of a set of parameters. This will be addressed in Chapter 3.

1.5. Numerical Example

The development in Section 1.4 provides the necessary background for an explanation of the choice of data set in Table 1-1. That triangle has been chosen purposefully rather than at random. The reasons for the choice can be seen in Table 1-3. This is constructed from Table 1-2, which is the table of cumulative observations X_{kj} in the notation of Section 1.2. The X_{kj} are obtained from Table 1-1.

Then Table 1-3 is the table of $\hat{f}_{kj}$ in the notation of Section 1.4. In this table the age-to-age factor labelled as belonging to development year j is $\hat{f}_{kj}$, defined in (1-4) as relating development years j and $j + 1$.

The averaging of age-to-age factors over a column in (1-5) and (1-8) suggests an implicit assumption of random variation of the f_{kj} about a constant parameter for fixed j.

Table 1-2. Triangle of Cumulative Paid Losses

Accident Year		Cumulative Paid Losses to and Including Development Year ($000)									
		1	2	3	4	5	6	7	8	9	10
1988	1	41821	76550	96697	112662	123947	129871	134646	138388	141823	144781
1989	2	48167	87662	112106	130284	141124	148503	154186	158944	162903	
1990	3	52058	99517	126876	144792	156240	165086	170955	176346		
1991	4	57251	106761	133797	154668	168972	179524	187266			
1992	5	59213	113342	142908	165392	179506	189506				
1993	6	59475	111551	138387	160719	175475					
1994	7	65607	110255	137317	159972						
1995	8	56748	96063	122811							
1996	9	52212	92242								
1997	10	43962									

This assumption will be made explicit in the model formulation of Section 3.3.1. In the meantime, the approximate constancy of the $\hat{f}_{kj}$ for fixed j in Table 1-3 may be noted.

As a consequence, the chosen data set will be compatible with the formal chain ladder models formulated in Chapter 3. The data set has been selected for this reason as it is to be used for numerical illustration of various aspects of the chain ladder.

1.6. Common Chain Ladder Extensions

There are a couple of extensions to the chain ladder forecast just described that will not be discussed further in this monograph but are integral to loss reserving practices

Table 1-3. Triangle of Age-to-Age Factors

Accident Year		Age-to-Age Factor for Development Year								
		1	2	3	4	5	6	7	8	9
1988	1	1.830	1.263	1.165	1.100	1.048	1.037	1.028	1.025	1.021
1989	2	1.820	1.279	1.162	1.083	1.052	1.038	1.031	1.025	
1990	3	1.912	1.275	1.141	1.079	1.057	1.036	1.032		
1991	4	1.865	1.253	1.156	1.092	1.062	1.043			
1992	5	1.914	1.261	1.157	1.085	1.056				
1993	6	1.876	1.241	1.161	1.092					
1994	7	1.681	1.245	1.165						
1995	8	1.693	1.278							
1996	9	1.767								

to the extent that they will be related here. Their origins lie in the fact that the chain ladder algorithm, at least in its incremental paid loss form, is highly sensitive to the amount of claim payments to date.

Note that, by (1-10), all forecasts in respect of accident year k are directly proportional to $X_{k,K-k+1}$, the total paid losses to date for that accident year. This sensitivity can be particularly acute in the case of the more recent accident years. For example, forecasts for the most recent accident year K will be directly proportional to the single observation Y_{kj} ($= X_{kj}$).

Some variations of the chain ladder algorithm seek to reduce this sensitivity by relating the estimate ultimate claim cost of an accident year to some kind of budget (i.e., prior-to-data estimate) cost.

Let B_k denote a budget ultimate claim cost for accident year k. An estimate of the portion of this paid in the future (i.e., after development year $K - k + 1$), based on the age-to age factors (1-8) is obtained by inversion of (1-11) thus:

$$\hat{R}_k^{(B)} = B_k - \hat{X}_{k,K-k+1} = B_k \left[1 - \frac{1}{\hat{f}_{K-k+1} \ldots \hat{f}_{J-1}} \right] \tag{1-13}$$

There are two common forms of this forecast used in practice, involving different budget claim costs:

- **Bornhuetter-Ferguson** forecast (Bornhuetter and Ferguson, 1972): $B_k = P_k \pi_k$, where P_k denotes earned premium for accident year k, and π_k budget loss ratio for the accident year; and
- **Cape Cod** forecast (Straub, 1988): $B_k = P_k \Sigma_{i=1}^{K} \omega_i \left[(X_{i,K-i+1} + \hat{R}_i) / P_i \right] / \Sigma_{i=1}^{K} \omega_i$, with $\omega_i = 1 / \hat{f}_{K-i+1} \ldots \hat{f}_{J-1}$.

The Bornhuetter-Ferguson forecast uses a budget ultimate claim cost calculated according to the budget loss ratio for the relevant accident year. The Cape Cod forecast is similar but uses the same budget loss ratio for each accident year. This single loss ratio is a weighted average of the loss ratios forecast by the chain ladder for the individual accident years.

2. Stochastic Models

Chapter summary. This chapter provides the theoretical background for GLMs. A GLM assumes observations to be subject to a distribution drawn from the Exponential Dispersion Family. This family, and its properties, are introduced. Important sub-families, namely the Tweedie sub-family, and the over-dispersed Poisson (nested within Tweedie), are identified.

A GLM is then defined and explained. The two types of covariate, categorical and continuous, are discussed. A number of aspects of goodness-of-fit of a GLM are discussed, including deviance and residuals. The use of weights to control heteroscedasticity, and to deal with outlying observations, is explained. The use of a GLM to generate forecasts is also discussed.

2.1. Exponential Dispersion Family

Subsequent chapters will present the chain ladder models in terms of GLMs, which will be defined in Section 2.2. GLMs rest on the family of distributions called the exponential dispersion family (**"EDF"**), which is defined in the present subsection.

2.1.1. The Exponential Dispersion Family in General

The EDF was introduced by Nelder and Wedderburn (1972), and discussed in detail in McCullagh and Nelder (1989). It is the family of distributions with probability density function (**"pdf"**) $\pi(y; \theta, \phi)$ of the form

$$ln\, \pi(y; \theta, \phi) = \frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi) \tag{2-1}$$

where

y is the value of an observation Y;
θ is a location parameter called the **canonical parameter**;
ϕ is a **dispersion parameter**, sometimes called the **scale parameter**;
$b(.)$ is called the **cumulant function**, and determines the shape of the distribution;
$exp\; c(y, \phi)$ is a **normalizing factor** producing unit total mass for the distribution.

It is assumed that the functions a, b, c are continuous and that b is one-one and twice differentiable with first derivative also one-one.

A family of distributions is specified by the selection of $a, b, c,$ and members of this family are then characterised by the parameters θ, ϕ. A specific member of this family will be denoted $EDF(\theta, \phi; a, b, c)$.

Table 2-1. Examples of Distributions from the EDF

Distribution	$b(\theta)$	$a(\phi)$	$c(y, \phi)$
Normal	$\frac{1}{2}\theta^2$	ϕ	$-\frac{1}{2}[y^2/\phi + ln(2\pi\phi)]$
Poisson	$exp\ \theta$	1	$-ln\ y!$
Binomial	$ln\ (1 + e^\theta)$	n^{-1}	$ln\binom{n}{ny}$
Gamma	$-ln\ (-\theta)$	v^{-1}	$v\ ln(vy) - ln\ y - ln\ (\Gamma v)$
Inverse Gaussian	$-(-2\theta)^{-1/2}$	ϕ	$-\frac{1}{2}[ln\ (2\pi\phi y^3 + 1/\phi y)]$

The form (2-1) is one which includes a number of the well-known distributions, as illustrated in Table 2-1.

The selection of an EDF distribution from this table to be assumed within a model will depend on the subject of the model and its properties. For example, the Poisson and binomial cases might be suitable for a model of counts; the other cases for amounts.

It may be shown that, when Y is distributed according to (2-1),

$$E[Y] = b'(\theta) \tag{2-2}$$

$$Var[Y] = a(\phi)b''(\theta) \tag{2-3}$$

If $E[Y]$ is denoted by μ, then (2-2) establishes a connection between μ and θ:

$$\theta = (b')^{-1}(\mu) \tag{2-4}$$

which justifies the above description of θ as a location parameter.

The relation (2-4) is one-one and so, with just a slight abuse of notation, one may write the pdf of y as $p(y; \mu, \phi)$, as an alternative to $p(y; \theta, \phi)$.

Use of (2-2) converts (2-3) to the form:

$$Var[Y] = \alpha(\phi)V(\mu) \tag{2-5}$$

where

$$V(\mu) = b''\left((b')^{-1}(\mu)\right) \tag{2-6}$$

and $V(\mu)$ is called the **variance function**.

Note that the somewhat confusingly named variance function is not equal to the variance. In fact, (2-5) decomposes the variance into factors that depend on the mean and the dispersion parameter respectively. The variance function is the factor dependent on the mean.

For all practical purposes, it is sufficient to restrict (2-1) to the special case

$$a(\phi) = \phi/w \tag{2-7}$$

for some constant w, and this restriction will be assumed henceforth. Variation of w from one observation to another creates any required variation in $a(\phi)$, as will be explained in Section 2.2.1. However, unless otherwise stated in the following, it will be assumed that $w = 1$.

2.1.2. The Tweedie Sub-Family

The Tweedie sub-family of the EDF was introduced by Tweedie (1984). It is obtained from the EDF by restriction of the variance function as follows:

$$V(\mu) = \mu^p,\ p \le 0 \text{ or } p \ge 1 \tag{2-8}$$

So, according to (2-5) and (2-7), $Var[Y] = \phi\mu^p$ and variance is proportional to a power of the mean.

It may be shown that this form of variance function implies that the cumulant function takes the form

$$b(\theta) = (2-p)^{-1}\left[(1-p)\theta\right]^{\frac{2-p}{1-p}} \tag{2-9}$$

and this in turn implies

$$\mu = \left[(1-p)\theta\right]^{\frac{1}{1-p}} \tag{2-10}$$

$$ln\,\pi(y;\mu,\phi) = \left[\frac{y\mu^{1-p}}{(1-p)} - \frac{\mu^{2-p}}{(2-p)}\right]\Big/\phi + c(y,\phi) \tag{2-11}$$

Note that several of the example distributions appearing in Table 2-1 are characterized by a cumulant function of the form (2-9). In fact all distributions in that table other than binomial satisfy this condition, or at least a limiting version of it, when it is recognized that

$$\lim_{p \downarrow 1}\left[(1-p)\theta\right]^{\frac{1}{1-p}} = exp\,\theta \tag{2-12}$$

$$\lim_{p \uparrow 2}(2-p)^{-1}\theta^{2-p} = ln\,\theta \tag{2-13}$$

The Tweedie sub-family, which will be denoted $Tw(\mu, \phi; p)$, thus contains these distributions, as set out in Table 2-2. It also contains the over-dispersed version of the Poisson distribution. The final column here omits the term $c(y, \phi)$.

It follows from (2-8) that the tail heaviness of Tweedie distributions increases with increasing p. The choice of Tweedie member for a model may therefore depend on the heaviness of tail indicated by the data. If, for example, a model based on index p generates more widely dispersed residuals than are consistent with that model,

Table 2-2. Some Well-Known Members of the Tweedie Family

Distribution	p	$b(\theta)$	μ	$\ln \pi(y; \mu, \phi)$
Normal	0	$\frac{1}{2}\theta^2$	θ	$[y\mu - \frac{1}{2}\mu^2]/\phi$
Over-dispersed Poisson	1	$\exp\theta$	$\exp\theta$	$[y \ln \mu - \mu]/\phi$
Gamma	2	$\ln(-\theta)$	$-1/\theta$	$[-y/\mu - \ln \mu]/\phi$
Inverse Gaussian	3	$-(-2\theta)^{1/2}$	$(-2\theta)^{-1/2}$	$[-(y/2\mu^2) + 1/\mu]/\phi$

then consideration might be given to increasing p. This matter is discussed further in Section 6.6.

Moreover, it has been shown (Jorgensen and Paes de Souza, 1994) that the cases $1 \le p < 2$ can be identified as compound Poisson distributions with gamma severity distributions.

2.1.3. The Over-Dispersed Poisson Sub-Family

The over-dispersed Poisson (**"ODP"**) distribution will play a central role in some subsequent chapters, and so is discussed a little further here.

As noted in Table 2-2, it is the Tweedie case $p = 1$. It may be represented, as a family, by $Tw(\mu, \phi; 1)$, which will be abbreviated to $ODP(\mu, \phi)$. From the last column of that table, its pdf is

$$\pi(y; \mu, \phi) = \mu^{y/\phi} exp[-\mu/\phi + c(y, \phi)],\ y = 0, \phi, 2\phi, \text{etc.} \tag{2-14}$$

with $\mu = e^{\theta}$.

It may be checked that a unit total probability mass is obtained if

$$exp\ c(y, \phi) = [(y/\phi)!]^{-1} \tag{2-15}$$

Substitution of (2-15) in (2-14) yields

$$\pi(y; \mu, \phi) = \frac{\mu^{y/\phi} exp(-\mu/\phi)}{(y/\phi)!},\ y = 0, \phi, 2\phi, \text{etc.} \tag{2-16}$$

and this is recognizable as the Poisson distribution

$$Y/\phi \sim Poiss(\mu/\phi) \tag{2-17}$$

From this it follows that

$$E[Y] = \phi E[Y/\phi] = \mu \tag{2-18}$$

$$Var[Y] = \phi^2 Var[Y/\phi] = \phi\mu \tag{2-19}$$

Note that (2-18) checks with the definition of μ, and (2-19) checks with (2-5), (2-7) and (2-8). Note also that, in the case $\phi = 1$, (2-17) reduces to the simple Poisson

$$Y \sim Poiss(\mu) \tag{2-20}$$

Thus, by (2-17)–(2-19), the ODP variate is similar to a Poisson variate but with the relation between variance and mean changed by the dispersion parameter ϕ.

An ODP assumption is often a convenient one when little is known of the subject distribution. As a simple modification of the Poisson distribution, it retains much of the simplicity of that case, but its 2-parameter nature endows it with much more flexibility. Nonetheless, as in the case of any other distributional assumption, it requires validation by reference to the data (see Section 6.6). Its major relevance to this monograph will become apparent in Section 3.3.

2.2. Generalized Linear Models (GLMs)

2.2.1. Definition

For the purpose of the current sub-section, let $\pi(.; \mu, \phi)$ denote a member of the EDF, fixed except that the parameters μ, ϕ remain variable.

Consider a sample of observations Y_i, $i = 1, 2, \ldots, n$. Suppose that each Y_i is associated with a known q-vector $(x_{i1}, x_{i2}, \ldots, x_{iq})$ of **predictors** (or **covariates**). Let the transpose of this vector be denoted x_i. Suppose also that these observations satisfy the following conditions:

(1) $Y_i \sim \pi(.; \mu_i, \phi_i)$ with the μ_i being unknown parameters.
(2) $h(\mu_i) = x_i^T\beta$, where $h(.)$, known as the **link function**, is a given one-one function with range $(-\infty, +\infty)$, β is a q-vector of unknown parameters, and the upper T denotes vector or matrix transposition.
(3) The observations Y_i are stochastically independent.

The structure defined by conditions (1)–(3) is called a generalized linear model ("GLM"), discussed in depth by McCullagh and Nelder (1989). The variate Y_i is called the **response** and the linear expression $x_i^T\beta$ is called the **linear response**. The choice of link function must be such as to transform the mean of each observation into a linear function of the parameter vector β. An example will be given in Section 3.3.2.

The dispersion parameters ϕ_i may be known but more commonly it is assumed that

$$\phi_i = \phi / w_i \tag{2-21}$$

with ϕ unknown but the w_i (called **weights**) known.

The GLM is a regression model. Note that, if $\pi(.; \mu_i, \phi_i) = n(.; \mu_i, \phi_i)$, the normal density, and $h = identity$, then conditions (1) and (2) may be expressed in the form

$$Y_i = x_i^T\beta + \varepsilon_i \text{ with } \varepsilon_i \sim N(0, \phi_i) \tag{2-22}$$

This is recognizable as a **weighted linear regression** model. Thus a GLM may be regarded as a generalization of linear regression in which:

- The relation between observations and covariates may be non-linear;
- Error terms may be non-normal.

It will sometimes be useful to represent condition (2) in vector and matrix notation. Let Y denote the vector whose i-th component is Y_i, μ denote the vector whose i-th component is μ_i, and let X denote the matrix whose i-th row is x_i^T. The matrix X is called the **design matrix** of the regression. Then condition (2) is written as

$$\mu = h^{-1}(X\beta) \tag{2-23}$$

where h^{-1} is understood to operate componentwise on its vector argument.

The parameter vector β is related to the canonical parameters θ of (2-1) through (2-2) and (2-23). Within the GLM, there will be an n-vector $(\theta_1, \ldots, \theta_n)$ of canonical parameters, one corresponding to each observation. Let this vector henceforth be denoted by θ. Then

$$b'(\theta_i) = E[Y_i] = \mu_i = h^{-1}(x_i^T\beta) \tag{2-24}$$

It is evident from (2-8), (2-24) and the discussion surrounding Table 2-2 that selection of a GLM consists of:

- selection of a cumulant function, controlling the model's assumed error distribution;
- as part of this, selection of index p, which controls the relation between the model mean and variance;
- selection of the covariates x_i^T, those explanatory variables considered to influence the cell mean μ_i;
- selection of a link function, which specifies the functional relation between the cell mean μ_i and the associated covariates.

Chapter 6 discusses in some detail how diagnostics derived from the data might be used to guide these selections.

One way in which the parameters of the GLM may be estimated from data is by maximum likelihood estimation (**"MLE"**). Usually, the MLE solutions are not expressible in closed form, and numerical solution is required. The numerical solution is non-trivial, and specialist software is required.

Well known GLM software packages are SAS, R and Emblem. These use MLE, and this form of estimation will be assumed for the remainder of this monograph.

Sections 2.2.2 to 2.2.6 discuss a number of aspects of a GLM that are essential to its meaningful formulation. As part of the present chapter, which establishes the theoretical background, these sections are abstract in nature. However, many of the features discussed are illustrated numerically in Chapter 6.

2.2.2. Categorical and Continuous Covariates

Some covariates are discrete by nature, possibly non-numerical (e.g., gender). Such covariates are usually referred to as **categorical** in the regression context. Other covariates are **continuous** by nature (e.g., age).

Consider a categorical variate with m possible values (often referred to as **levels** of the variate), denoted $\xi_1, \ldots, \xi_m$. This is represented in the GLM as m distinct 0–1 variates $x_{k+1}, \ldots, x_{k+m}$, where $x_1, \ldots, x_k, x_{k+m+1}, \ldots$ denote the other regression covariates. The 0–1 variates are defined as

$$\begin{aligned} x_{k+r} &= 1 \text{ if the categorical variate assumes the value } \xi_r \\ &= 0 \text{ otherwise} \end{aligned} \tag{2-25}$$

Note that

$$\sum_{r=1}^{m} x_{k+r} = 1 \tag{2-26}$$

For example, if one wished to include development year as a covariate in a model, this might be done by treatment of development year as a categorical variate ξ with J levels $\xi = j$, $j = 1, \ldots, J$, where the associated 0–1 variates are defined as:

$$\begin{aligned} x_{k+j} &= 1 \text{ if } \xi = j \\ &= 0 \text{ otherwise} \end{aligned}$$

This treatment of categorical variates can sometimes lead to the introduction of redundant parameters. This will be illustrated, and the remedy given, in Sections 3.2 and 3.3.2, where representation of development year as a categorical variate will be pursued further.

A continuous variate on the other hand assumes numerical values in a continuous range (e.g., age). Such a variate may be represented in a regression as simply itself. Alternatively, it may be represented as some transformation of itself.

For example, the function

$$L_{mM}(x) = min\left[M - m, max(0, x - m)\right] \text{ with } m < M \tag{2-27}$$

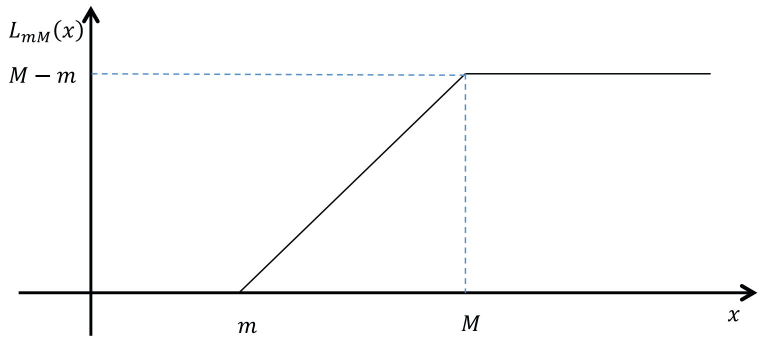
is linear with unit gradient between m and M, and constant outside this range, as illustrated in Figure 2-1.

Functions of this sort may be used to incorporate **linear splines** (piecewise linear functions) in the regression. For example, the function

$$\sum_{k=1}^{K} \beta_k L_{m_k m_{k+1}}(x) \tag{2-28}$$

is a linear spline with knots at $x = m_1, \ldots, m_{K+1}$ and gradient β_k for $x \in [m_k, m_{k+1}]$.

Figure 2-1. Illustration of the Function $L_{mM}(x)$



The functions $L_{m_k m_{k+1}}(x)$ are called **basis functions** since the spline may be constructed as a linear combination of them. If these basis functions are included as covariates in a regression, then the regression will return estimates of the gradients β_k. Splines of higher degree (e.g., cubic splines) may be similarly incorporated in the regression model by means of appropriately defined basis functions. Basis functions will be central to the development of generalized additive models in Section 7.2.

2.2.3. Goodness-of-Fit and Deviance

Let $\hat{\beta}$ denote the MLE of β. The vector

$$\hat{Y} = h^{-1}\left(X\hat{\beta}\right) \tag{2-29}$$

is the MLE of μ and is referred to as the vector of **fitted values** (c.f. (2-23)).

The principal measure of goodness-of-fit of a GLM is its **scaled deviance**, defined as

$$\begin{aligned} D\left(Y, \hat{Y}\right) &= 2\left[ln\, \pi\left(Y; \hat{\theta}^{(s)}, \phi\right) - ln\, \pi\left(Y; \hat{\theta}, \phi\right)\right] \qquad (2\text{-}30) \\ &= 2\sum_{i=1}^{n}\left[ln\, \pi\left(Y_i; \hat{\theta}^{(s)}, \phi\right) - ln\, \pi\left(Y_i; \hat{\theta}, \phi\right)\right] \end{aligned}$$

where θ is the vector of canonical parameters introduced just before (2-24), $\hat{\theta}$ is the MLE of θ, and $\hat{\theta}^{(s)}$ is the estimate of θ in the **saturated model**, a model with a parameter for every observation so that $\hat{Y} = Y$.

It should be noted that nomenclature differs between authors. For example, McCullagh and Nelder refer to (2-30) as the **scaled deviance**, as is done here, whereas other authors refer to just the deviance.

It is evident from a comparison of (2-30) with (2-1) that maximization of likelihood is equivalent to minimization of deviance. A smaller scaled deviance indicates improved goodness-of-fit. The minimum achievable deviance is zero, when there is no difference between observations and fitted values (as in the saturated model).

Calculation of the scaled deviance (2-30) requires computation of a value for ϕ. However, it is evident from (2-1) that ϕ will factor out of any minimisation of scaled deviance, whence its value is irrelevant to MLE of parameters.

For this reason it is common to define an unscaled version of the deviance, referred to subsequently as just the **deviance**, as follows:

$$D^*\left(Y,\hat{Y}\right)=2\sum_{i=1}^{n}\left[ln\ \pi\left(Y_i;\hat{\theta}^{(s)},1\right)-ln\ \pi\left(Y_i;\hat{\theta},1\right)\right] \tag{2-31}$$

which, in effect, ignores ϕ. MLE is then carried out by minimization of $D^*(Y, \hat{Y})$ with respect to $\hat{\theta}$, equivalently $\hat{\beta}$.

The deviance can be viewed as the logarithm of a likelihood ratio and, by an application of Wilks' theorem, it is asymptotically χ^2 distributed with $n - p$ as the number of **degrees of freedom**. The usual estimate of the scale parameter ϕ is therefore

$$\hat{\phi}=D^*\left(Y,\hat{Y}\right)/(n-p) \tag{2-32}$$

2.2.4. Residuals

Pearson Residuals

Define the **standardized Pearson residual** associated with observation Y_i as

$$R_i^P=\left(Y_i-\hat{Y}_i\right)/\hat{\sigma}_i \tag{2-33}$$

where $\hat{\sigma}_i^2$ is an estimator of $\sigma_i^2 = Var[Y_i]$.

If it may be assumed that $\hat{Y}_i$ is approximately unbiased as an estimator of μ_i, and that $Var[Y_i - \hat{Y}_i]$ differs little from $Var[Y_i]$ (these assumptions are often reasonable), then approximately

$$E\left[R_i^P\right]=0 \text{ and } Var\left[R_i^P\right]=1 \tag{2-34}$$

It is in fact possible to correct (2-33) with a further scalar multiplier in order to ensure that $Var[R_i^P] = 1$ but details are not given here.

In this case a plot of the Y_i against i will produce a scatter of residuals evenly about zero (unbiasedness) and with uniform dispersion as one reads from left to right (**homoscedasticity**). An example appears as Figure 2-2.

In fact the homoscedasticity of Figure 2-2 is only approximate, as is indicated by Figure 2-3. This plots the standard deviation of residuals by age group (right-hand scale). The standard deviation varies from about 0.8 to about 1.1, indicating mild heteroscedasticity. The same figure plots the lower quartile ("p_25") and upper quartile ("p_75") of the residuals in each age group (left-hand scale).

Routine **model validation** includes the examination of a separate residual plot against each covariate (e.g., age), checking for unbiasedness and homoscedasticity. The reason that unbiasedness is sought is obvious. The reason for the requirement of homoscedasticity will be discussed in Section 2.2.5.

Figure 2-2. Example of Unbiased Approximately Homoscedastic Residual Plot

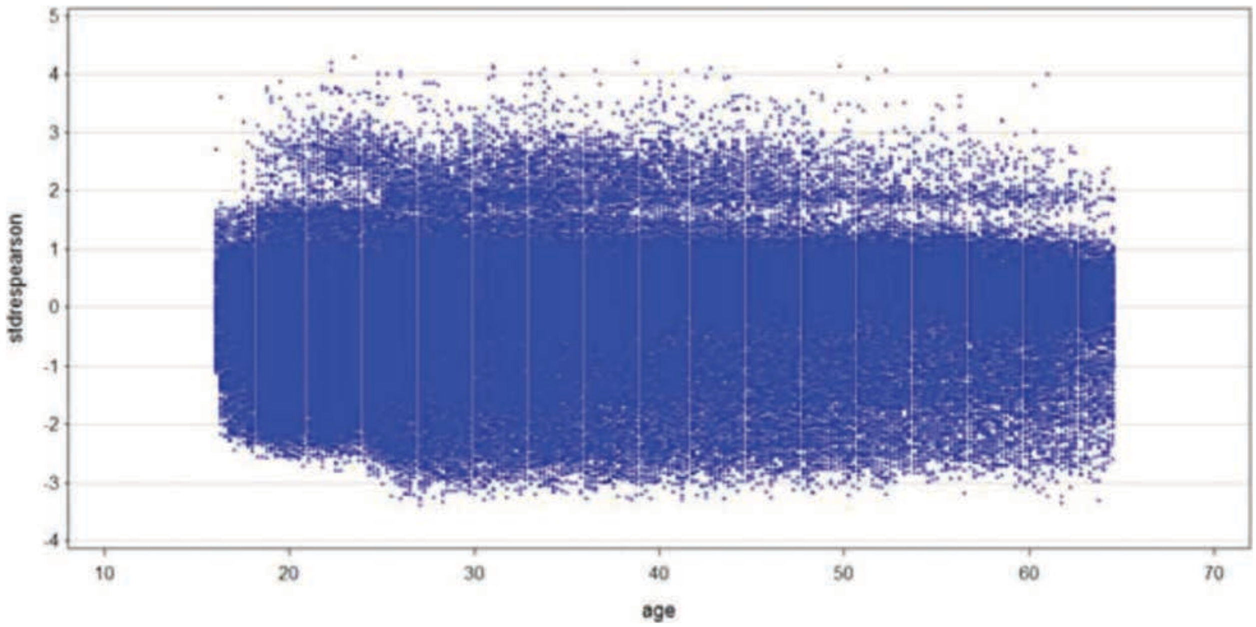
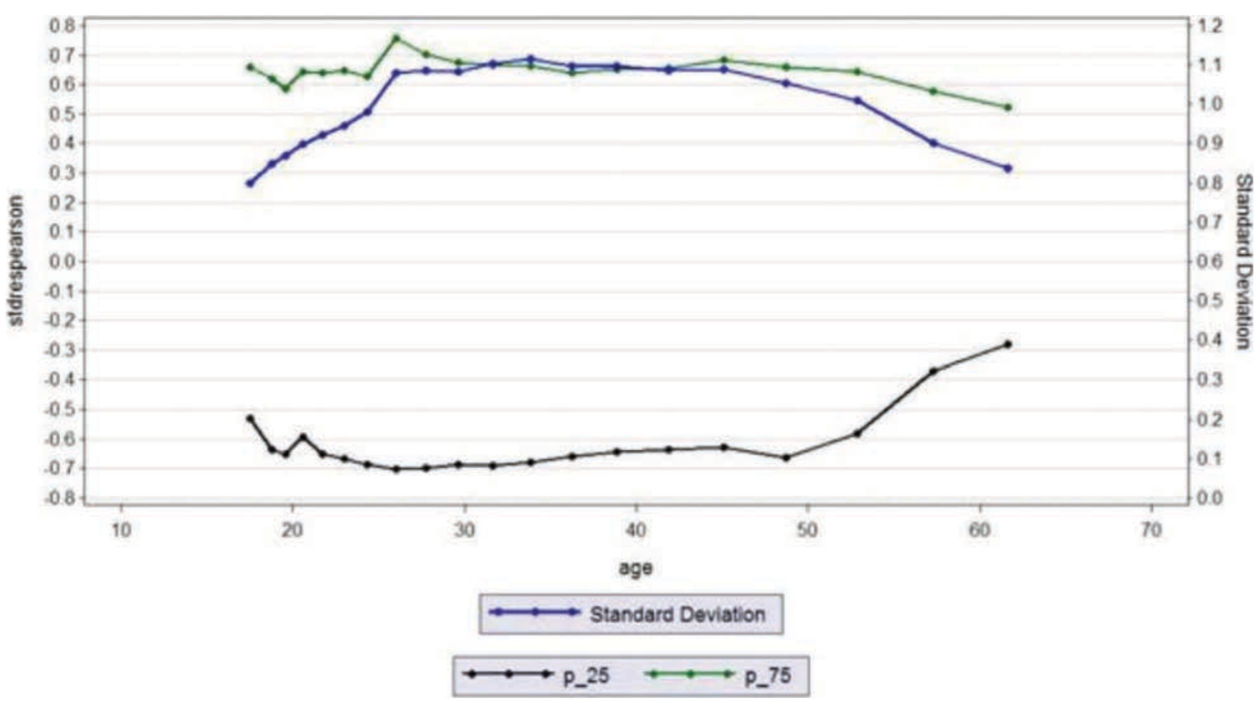


Figure 2-3. Example of Biased Homoscedastic Residual Plot



Deviance Residuals

Although Pearson residuals have a simple intuitive interpretation, they are linear transformations of the observations and will reproduce any non-normality that exists in them. For this reason it is common to use a different form of residual in the assessment of a GLM. This is the **standardized deviance residual,** defined in relation to the observation Y_i as

$$R_i^D = sgn\left(Y_i - \hat{Y}_i\right)\left(d_i / \hat{\phi}\right)^{1/2} \tag{2-35}$$

where d_i is the contribution of the i-th observation to the deviance $D^*(Y, \hat{Y})$.

As was the case with Pearson residuals, it is possible to correct (2-35) with a further scalar multiplier in order to ensure that $Var[R_i^D] = 1$ but again details are not given here.

Pierce and Schafer (1986) showed that deviance residuals are normally distributed with error of order $m^{-1/2}$, where m is a certain index derived from the specific member of the EDF associated with the GLM. As a result of this property, deviance residuals often remove much of the non-normality present in Pearson residuals and, in consequence, are often more useful.

An example of this is given in Figure 2-4 and Figure 2-5, which plot histograms of residuals from a model of individual auto bodily injury claims in one Australian state. Individual claims are modeled as gamma distributed with mean value depending on various claim characteristics but constant (and large) coefficient of variation, 1.16.

Figure 2-4. Histogram of Standardized Pearson Residuals

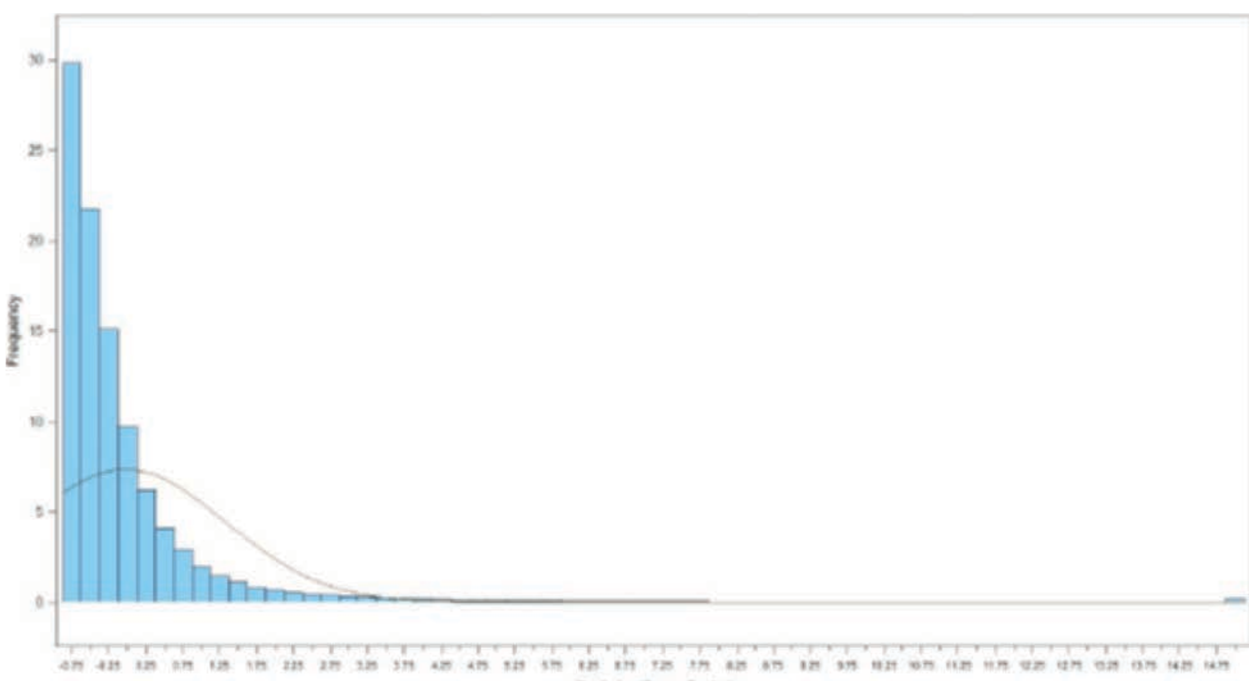


Figure 2-5. Histogram of Standardized Deviance Residuals

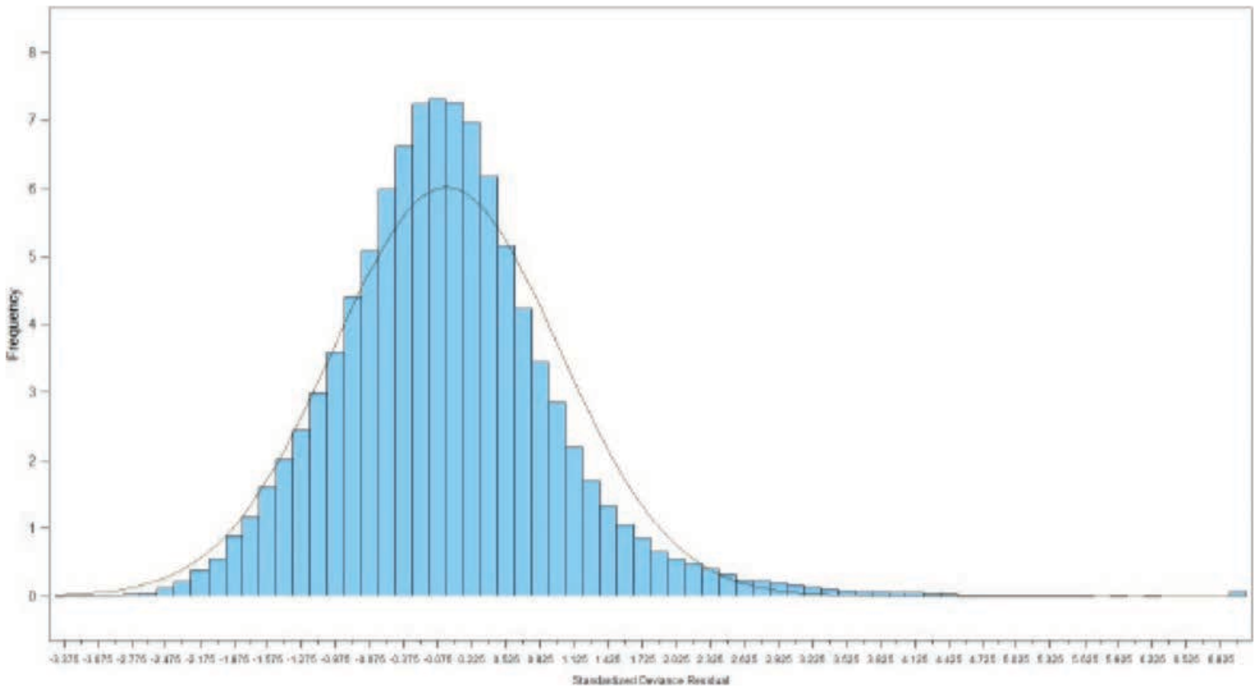


Figure 2-4 plots standardized Pearson residuals, and Figure 2-5 plots standardized deviance residuals.

In each case the best normal approximation to the histogram, calculated by the method of moments, is also shown. The Pearson residuals are seen to be highly skew and poorly fit by the normal approximation. The deviance residuals, while still exhibiting some degree of non-normality, are seen to be much closer to normal.

2.2.5. Outliers and the Use of Weights

The need for homoscedasticity was discussed in Section 2.2.4. The reason for this will be discussed below. However, first a short digression on the use of **variance weights** (or simply **weights**).

If a residual plot reveals heteroscedasticity, correction may be made by means of weights. Consider the following example that is rather exaggerated but illustrates the point nonetheless. Suppose a GLM has been formulated on the assumption of homoscedasticity, specifically that (see (2-5) and (2-7))

$$Var[Y_i] = \phi V(\mu_i) \tag{2-36}$$

with ϕ independent of i.

Suppose that standardized Pearson residuals (2-33) have been plotted by age and it appears that residuals above age 55 have double the standard deviation of those below age 55.

First use (2-5) and (2-7) to express (2-33) in the form

$$R_i^P = \left(Y_i - \hat{Y}_i\right) / \left[\hat{\phi} V(\hat{\mu})\right]^{1/2} \tag{2-37}$$

Then the observed heteroscedasticity indicates that the value of ϕ for ages above 55 is in fact about four times that for lower ages. The heteroscedasticity would be removed if the model were adjusted to reflect this variation in ϕ over age. This may be achieved by the use of weights. By (2-21) the required result may be achieved by setting

w_i = 1 if the *i*-th observation involves an age below 55
= ¼ if the *i*-th observation involves an age above 55

In the default case in which there is no explicit introduction of weights (Section 2.2.5), all observations will be equally weighted in parameter estimation. This is appropriate if all observations are subject to the same ϕ, but undesirable otherwise. It is intuitively obvious that observations of larger variance than this should receive lesser weight than those of smaller variance.

Indeed, it can be shown that estimation efficiency will be optimized if each observation is assigned a weight that is inversely proportional to its ϕ. As noted above, the relative values of ϕ for different observations are reflected in the variance of their standardized residuals.

Thus, in general, if a residual plot displays heteroscedasticity, one adjusts weights roughly in inverse proportion with variance of the residuals. A specific example of the use of weights in this way is given in Section 6.6 (see particularly Figure 6-15, Figure 6-17 and associated text).

A residual plot might also identify isolated observations with very large residuals. These are referred to as **outliers**. Such observations can influence the regression unduly by shifting the fitted values away from the main body of observations in favor of the outliers, as illustrated in Figure 2-6.

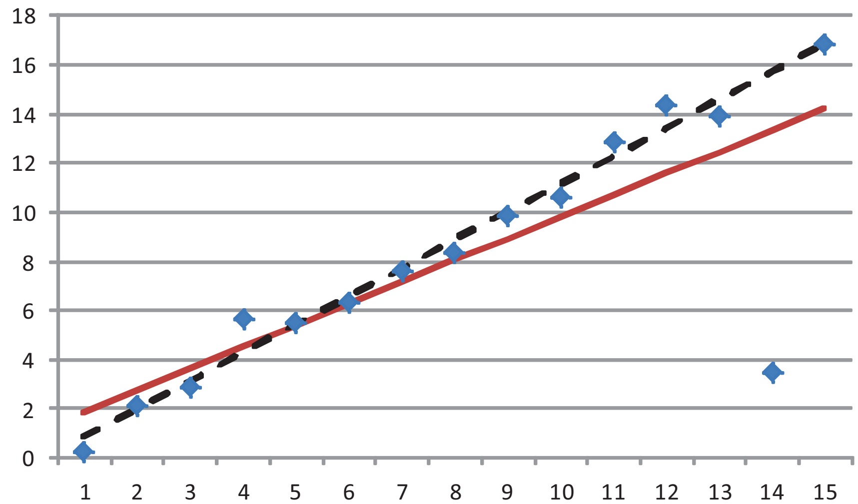
The solid line in the diagram is the result of linear regression using all observations, including the outlier at $x = 14$, whereas the dotted line is the result of linear regression excluding this observation.

In the event that a specific observation is identified as an outlier, and its inclusion in the regression considered distorting, it may be excluded by assigning it zero weight.

Care must be taken in the exclusion of any data points. For example, if the outlier represented a major natural event, whereas the other observations represented attritional events, the exclusion of the former from the regression may be appropriate but the cost of major natural events would need to be accounted for somewhere.

Moreover, the exclusion of selected observations from parameter estimation will have consequences for the estimation of prediction error, as discussed in Sections 5.3.1 and 5.3.2.

Figure 2-6. Illustration of Distortion of Regression by Outlier



2.2.6. Forecasts

Recall from Section 2.2.1 that

$$E[Y_i] = \mu_i = h^{-1}\left(x_i^T \beta\right) \tag{2-38}$$

When the GLM is to be used for forecasting, as in loss reserving, the covariate vectors x_i will typically include information on the time of measurement. They may also contain other information. For example, in the case of workers compensation claims, the x_i may include the type of claim (income replacement, medical only, etc.).

When the model is applied to forecast future observations, those observations will be characterized by their own covariate vectors x_i. These will be distinct from those in the data set in that, to the extent that they include time variates, their values will all relate to the future.

It will be convenient to distinguish future observations from the past Y_i by the notation Y_i^*, characterised by the covariate vector x_i^*. In general, the addition of a star to a symbol will indicate future values of the variate represented by the corresponding unstarred symbol. Thus, for example, Y^* will denote the vector of target random quantities Y_i^* to be forecast, and the relation (2-23) is extended to future values as follows:

$$\mu^* = h^{-1}\left(X^* \beta\right) \tag{2-39}$$

where X^* is the matrix whose rows are the $(x_i^*)^T$ discussed above and may be referred to as the **forecast design matrix**.

A reasonable forecast of Y^* is then

$$\hat{Y}^* = \hat{\mu}^* = h^{-1}\left(X^* \hat{\beta}\right) \tag{2-40}$$

3. Stochastic Models Supporting the Chain Ladder

Chapter summary. This chapter is concerned with the fact that the chain ladder algorithm of Chapter 1, known to many actuaries as merely a heuristic device, in fact provides the maximum likelihood forecasts of outstanding claims for a couple of formal models. Several formal chain ladder models from the literature are surveyed.

Two distinctly different stochastic models are defined whose MLEs of future claims experience are the same as the predictions of the heuristic algorithm. Moreover, these MLEs are also seen to possess certain minimum variance properties. These results are summarized in three theorems.

It is shown that these formal stochastic models are expressible as GLMs, and therefore estimates and forecasts from these GLMs will match the chain ladder estimates and forecasts. This is illustrated by numerical example.

Practical applications of the chain ladder often incorporate various *ad hoc* adjustments, such as omission of older diagonals from the claims triangle or omission of isolated observations that are considered rogue. It is shown that such adjustments can be accommodated within the GLM formulation, thus maintaining a formal model structure in their presence.

3.1. Mack Models

3.1.1. Non-Parametric Mack Model

Mack (1993) introduced a stochastic chain ladder model that has subsequently become known as the **Mack model**. It satisfies the following conditions:

(M1) Accident years are stochastically independent, i.e., $Y_{k_1 j_1}$, $Y_{k_2 j_2}$ are independent if $k_1 \neq k_2$.

(M2) For each $k = 1, 2, \ldots, K$, the X_{kj} (j varying) form a Markov chain.

(M3) For each $k = 1, 2, \ldots, K$ and $j = 1, 2, \ldots, J-1$,

(a) $E[X_{k,j+1}|X_{kj}] = f_j X_{kj}$ for some parameter $f_j > 0$;

(b) $Var[X_{k,j+1}|X_{kj}] = \sigma_j^2 X_{kj}$ for some parameter $\sigma_j > 0$.

The model was stochastic in the sense that it considered not only expected values but also variances of observations. However, it was non-parametric in the sense that it did not consider the distribution of observations.

Mack derived a number of results from this model, including the following:

Result 1: The conventional chain ladder estimators $\hat{f}_j$ of f_j according to (1-8) are:
(a) unbiased; and
(b) minimum variance among estimators that are unbiased linear combinations of the $\hat{f}_{kj}$ defined by (1-4).

Result 2: The conventional chain ladder estimator $\hat{R}_k$ of R_k given by (1-11) is unbiased.

3.1.2. Parametric Mack Models

A parametric version of the Mack model requires that assumption (M3) be supplemented by a distributional assumption. Parametric versions of the Mack model were studied by Taylor (2011). The observations $Y_{k,j+1}|X_{kj}$ were assigned distributions according to a member of the EDF, creating the **EDF Mack model**, defined as follows.

(EDFM1) As for (M1).
(EDFM2) As for (M2).
(EDFM3) For each k = 1, 2, . . . , K and $j = 1, 2, \ldots, J-1$,
(a) $Y_{k,j+1}|X_{kj} \sim EDF(\theta_{kj}, \phi_{kj}; a, b, c)$; and
(b) As for (M3a).

Assumption (EDFM3a) provides the required distributional assumption, with a general requirement that conditional observations be distributed according to some specific member of the EDF. Assumption (EDFM3b) retains the same form of conditional expectation as in the Mack model. No assumption about variance has been made other than that inherent in the selected EDF member. So the form of variance allowed in the EDF Mack model is more general than in the non-parametric Mack model.

Taylor (2011) also considered the following sub-families of the EDF Mack models:

Tweedie Mack model, in which (EDFM3a) is replaced by $Y_{k,j+1}|X_{kj} \sim Tw(\mu_{kj}, \phi_{kj}; p)$.
ODP Mack model, in which (EDFM3a) is replaced by $Y_{k,j+1}|X_{kj} \sim ODP(\mu_{kj}, \phi_{kj})$.

Taylor derived the following result.

Theorem 3.1. Suppose that the data array $\mathfrak{D}_K$ is a triangle ($J = K$) with observations subject to the EDF Mack model defined by assumptions (EDFM1-3).

(a) If assumption (M3b) also holds, then the model's MLEs of the f_j are the conventional chain ladder estimators $\hat{f}_j$ from (1-8). These are in turn unbiased estimators in the Mack model (see Result 1 of Section 3.1.1).
(b) If the EDF Mack model is restricted to an ODP Mack model in Assumption (EDFM3a), and if in addition the dispersion parameters ϕ_{kj} are just column dependent ($\phi_{kj} = \phi_j$) (the condition (M3b) automatically holds in this case), then the $\hat{f}_j$ from (1-8) are minimum variance unbiased estimators ("**MVUEs**") of the f_j.
(c) Under the same conditions as in (b), the predictors $\hat{X}_{kj}$, $\hat{R}_k$ defined by (1-9) and (1-11) are also MVUEs of X_{kj}, R_k. ■

The results of the theorem were also shown to extend to certain cases in which the distributions of the Y_{kj} were binomial or negative binomial.

The theorem is remarkable because it shows that estimates and forecasts that had been introduced to the actuarial literature many years earlier on an entirely heuristic basis turn out to be optimal estimators in the MLE and MVUE sense.

This MVUE result is much stronger than that of Mack referred to in Section 3.1.1 as the estimators here are minimum variance out of **all unbiased estimators**, not just out of the linear combinations of the $\hat{f}_{kj}$.

3.2. Cross-Classified Models

Consider a model of $\mathfrak{D}_K^+$ defined by the following conditions:

(EDFCC1) The random variables $Y_{kj} \in \mathfrak{D}_K^+$ are stochastically independent.
(EDFCC2) For each $k = 1, 2, \ldots, K$ and $j = 1, 2, \ldots, J$,
(a) $Y_{kj} \sim EDF(\theta_{kj}, \phi_{kj}; a, b, c)$;
(b) $E[Y_{kj}] = \alpha_k \beta_j$ for some parameters $\alpha_k, \beta_j > 0$; and
(c) $\sum_{j=1}^{J} \beta_j = 1$.

Models subject to (EDFCC2b) are variously referred to in the literature as cross-classified, ANOVA, or non-recursive. This model will be referred to here as the **EDF cross-classified model**.

The condition (EDFCC2c) merely removes redundancy from the model's parameter set. If it were absent, all α's could be doubled and all β's halved without any substantive change to the model. A single restriction on the parameters is required to render their values unique. Condition (EDFCC2c) is widely used for this purpose but other constraints would serve equally well, e.g., $\beta_1 = 1$ or $\alpha_1 = 1$.

It is noteworthy that the parameters of the EDF cross-classified model consist of both row and column parameters α_k and β_j respectively, whereas the only parameters contained in the Mack models are the column parameters f_j. This appears to imply that the EDF cross-classified structure is more general.

There was considerable discussion of this around the turn of the century (e.g., Mack and Venter, 2000; Verrall, 2000) in which it was pointed out that, although the Mack model contains no explicit row parameters, its conditioning on prior observations (see (M3a)) in effect plays the same role. The accumulated experience $X_{k,J-k+1}$ of row k serves as a row parameter in the forecast of future experience of that row.

Just as for the EDF Mack model of Section 3.1.2, Tweedie and ODP sub-families of the EDF cross-classified family may be identified. These will be referred to as the **Tweedie cross-classified family** and **ODP cross-classified family** respectively.

Let $\hat{\alpha}_k, \hat{\beta}_j$ denote MLEs of α_k, β_j and let $\hat{Y}_{kj} = \hat{\alpha}_k \hat{\beta}_j$ denote the fitted value associated with $Y_{kj} \in \mathfrak{D}_K$ or the forecast of $Y_{kj} \in \mathfrak{D}_K^c$. The following result was obtained by England & Verrall (2002).

Theorem 3.2. Suppose that the data array $\mathfrak{D}_K$ is a triangle ($J = K$) with observations subject to the ODP cross-classified model defined by assumptions (EDFCC1-2) and the following additional conditions:

(EDFCC3a) In (EDFCC2a) Y_{kj} is restricted to an ODP distribution;
(EDFCC3b) The dispersion parameters ϕ_{kj} are identical for all cells in $\mathfrak{D}_K^+$ (i.e., $\phi_{kj} = \phi$).

Then the MLE fitted values and forecasts $\hat{Y}_{kj}$ are the same as those given by the conventional chain ladder forecasts from (1-10). ■

The same result had been obtained earlier for the special case of the simple Poisson distribution ($\phi = 1$) by Hachemeister and Stanard (1975) and Renshaw and Verrall (1998).

The same results are not true for EDF distributions more general than ODP. In fact, the explicit (and different) ML equations for the Tweedie case are given by Peters, Shevchenko and Wüthrich (2009) and by Taylor (2009), and for the general EDF case by Taylor (2011).

The MLEs $\hat{Y}_{kj}$ will not be unbiased in general. However, Taylor (2011) obtained the following result.

Theorem 3.3. Suppose that the data array $\mathfrak{D}_K^+$ is subject to the same conditions as in Theorem 3.2. Suppose also that the fitted values and forecasts $\hat{Y}_{kj}$ and $\hat{R}_k$ are corrected for bias. Then they are MVUEs of Y_{kj} and R_k respectively. ■

Theorems 3.2 and 3.3 together parallel Theorem 3.1 but are even more remarkable. First, they state that the forecasts obtained from the ODP Mack and ODP cross-classified models are identical (and equal to those obtained from the conventional chain ladder) despite the very different formulations of the models. Moreover, notwithstanding that the cross-classified model is formulated in terms of parameters α_k, β_j, one may obtain forecasts without any consideration of them, but working as if the model were ODP Mack.

Numerical Example

It is instructive to illustrate this by reference to the data set in Table 1-1. It is worthy of note at the outset that the Mack models apply to cumulative data, whereas the cross-classified models apply to incremental data.

Commence by applying the chain ladder algorithm of Section 1.4 to the data. Average age-to-age factors are obtained by the application of (1-8), yielding the results in Table 3-1.

Forecasts are obtained by means of (1-9). For example, the first cell requiring forecast for accident year 1996 is that relating to development year 3. The forecast is $\hat{X}_{1996,3} = X_{1996,2}\hat{f}_2 = 92242 \times 1.261 = 116312$. Hence $\hat{Y}_{1996,3} = 116312 - 92242 = 24070$.

The full set of forecasts is given in Table 3-2, where the bold-face diagonal is merely transferred from Table 1-2, and then subsequent cells contain forecasts according to (1-9). The final column of the table contains the amounts of estimated outstanding losses $\hat{R}_k$, obtained by means of (1-11).

Table 3-1. Average Age-to-Age Factors

Average Age-to-Age Factor for Development Year								
1	2	3	4	5	6	7	8	9
1.815	1.261	1.158	1.088	1.055	1.039	1.030	1.025	1.021

Table 3-2. Estimation of Outstanding Losses

Accident Year		Forecast Cumulative Paid Losses to and Including Development Year ($000)										Estimated Outstanding Claims ($000)
		1	2	3	4	5	6	7	8	9	10	
1988	1										**144781**	
1989	2									**162903**	166301	3398
1990	3								**176346**	180731	184501	8155
1991	4							**187266**	192924	197721	201845	14579
1992	5						**189506**	196828	202774	207817	212151	22645
1993	6					**175475**	185209	192364	198176	203104	207340	31865
1994	7				**159972**	174108	183766	190866	196632	201522	205725	45753
1995	8			**122811**	142227	154795	163381	169693	174820	179168	182904	60093
1996	9		**92242**	116312	134700	146603	154735	160713	165569	169686	173225	80983
1997	10	**43962**	79788	100608	116513	126809	133843	139014	143214	146775	149836	105874
												373346

Now consider MLE within the ODP cross-classified model. The ML equations are well known (see any of the authors listed earlier in the present sub-section). They are merely **marginal sum estimation** equations (Schmidt and Wünsche, 1998), which means that they equate each row sum of observations with the corresponding sum of MLEs, and similarly for column sums. That is,

$$\sum^{\mathcal{R}(k)} Y_{kj} = \sum^{\mathcal{R}(k)} \hat{\alpha}_k \hat{\beta}_j = \hat{\alpha}_k \sum^{\mathcal{R}(k)} \hat{\beta}_j = \hat{\alpha}_k \sum_{j=1}^{J-k+1} \hat{\beta}_j = \hat{\alpha}_k \left[1 - \sum_{j=J-k+2}^{J} \hat{\beta}_j \right] \tag{3-1}$$

the last equality following from (EDFCC2c). Also

$$\sum^{\mathcal{C}(j)} Y_{kj} = \sum^{\mathcal{C}(j)} \hat{\alpha}_k \hat{\beta}_j = \hat{\beta}_j \sum^{\mathcal{C}(j)} \hat{\alpha}_k \tag{3-2}$$

It is further known that, for a triangular data set such as in Table 1-1, these equations are simply solved in the following order: (3-1) for $k = 1$, (3-2) for $j = J$, (3-1) for $k = 2$, (3-2) for $j = J - 1$, etc. and with repeated use of the constraint (EDFCC2c).

The first step in this procedure yields

$$144781 = \sum\nolimits^{\mathcal{R}(1)} Y_{kj} = \hat{\alpha}_1 \sum\nolimits^{\mathcal{R}(1)} \hat{\beta}_j = \hat{\alpha}_1$$

whence $\hat{\alpha}_1 = 144781$.

The second step yields

$$2958 = \sum\nolimits^{\mathcal{C}(10)} Y_{kj} = \hat{\beta}_{10} \sum\nolimits^{\mathcal{C}(10)} \hat{\alpha}_k = \hat{\beta}_{10} \hat{\alpha}_1$$

whence $\hat{\beta}_{10} = 2958/\hat{\alpha}_1 = 0.020$.

And so on, resulting in Table 3-3.

Table 3-3. Parameter Estimates for ODP Cross-Classified Model

j or k	$\hat{\alpha}_k$	$\hat{\beta}_j$
1	144781	0.293
2	166301	0.239
3	184501	0.139
4	201845	0.106
5	212151	0.069
6	207340	0.047
7	205725	0.035
8	182904	0.028
9	173225	0.024
10	149836	0.020

From these results, one may calculate $\hat{Y}_{1996,3} = \hat{\alpha}_9\hat{\beta}_3 = 173225 \times 0.139 = 24070$, in agreement with the estimate from the ODP Mack model. Similarly, all forecasts $\hat{Y}_{kj} \in \mathfrak{D}_K^c$ may be shown to reconcile with the ODP Mack model, indicating that it and the ODP cross-classified model yield the same estimates of outstanding losses (see Table 3-2).

Indeed, it follows from the identical forecasts of the ODP Mack and ODP cross-classified models that one may translate between the two by means of one-one relation. This relation, proven by Verrall (2000) using a Bayesian argument, is

$$\hat{f}_j = \sum_{i=1}^{j+1} \hat{\beta}_i \Big/ \sum_{i=1}^{j} \hat{\beta}_i \tag{3-3}$$

or its inverse

$$\hat{\beta}_{j+1} = \left(\hat{f}_j - 1\right) \prod_{r=1}^{j-1} \hat{f}_r \Big/ \prod_{r=1}^{J-1} \hat{f}_r \tag{3-4}$$

subject to the convention that $\Pi_{r=1}^{0} \hat{f}_r = 1$. Table 3-1 and Table 3-3 may be reconciled by this correspondence.

3.3. GLM Representation of Chain Ladder Models

3.3.1. ODP Mack Model

Consider the ODP Mack model of Section 3.1.2, and particularly the conditions (EDFM3a), modified to its ODP form, and (EDFM3b). Together these conditions amount to the following:

$$Y_{k,j+1} | X_{kj} \sim ODP\left(\left(f_j - 1\right) X_{kj}, \phi_{kj}\right) \tag{3-5}$$

Add the condition

$$\phi_{kj} = \phi_j, \text{ independent of } k \tag{3-6}$$

which was a pre-requisite in Section 3.1.2 for the ODP Mack model to yield the conventional chain ladder estimators as MLEs. Then

$$Y_{k,j+1} | X_{kj} \sim ODP\left(\left(f_j - 1\right) X_{kj}, \phi_j\right) \tag{3-7}$$

Now replace $Y_{k,j+1}$ here by $\hat{f}_{kj} - 1 = Y_{k,j+1}/X_{kj}$ from (1-4). It may be checked that

$$E\left[\hat{f}_{kj} - 1 | X_{kj}\right] = f_j - 1 \tag{3-8}$$

$$Var\left[\hat{f}_{kj} - 1 | X_{kj}\right] = Var\left[Y_{k,j+1} | X_{kj}\right] \Big/ X_{kj}^2 = \phi_j\left(f_j - 1\right) / X_{kj} \tag{3-9}$$

The ODP family is known to be closed under scaling, i.e., an ODP variate, divided by a constant, produces another ODP variate. Combining this fact with (3-8) and (3-9) yields

$$\hat{f}_{kj} - 1 | X_{kj} \sim ODP(f_j - 1, \phi_j / X_{kj}) \tag{3-10}$$

This may be formulated as a (rather trivial) GLM by comparison of (3-10) with the definition of a GLM in Section 2.2.1. The response vector of the GLM consists of the observations $\hat{f}_{kj} - 1 | X_{kj}$, listed in some convenient order such as dictionary order. The link function is the identity.

The parameter vector β consists of the parameters $f_1, \ldots f_9$, and the row of the design matrix X corresponding to $\hat{f}_{kj} - 1 | X_{kj}$ is the co-ordinate 9-vector e_j, which has unity in the j-th position and zeros elsewhere. In the terms usually required by GLM software for the specification of a model, this amounts to:

- Specification of development year j (= 1, 2, . . . , 9) as a categorical variate (referred to in some software systems as a **class variate**).
- Specification of the "model", i.e., the expected value, of each observation as

$$E\left[\hat{f}_{kj} - 1 | X_{kj}\right] = \sum_{i=1}^{9} (f_i - 1)\delta_{ji} \tag{3-11}$$

where δ_{ji} is the Kronecker delta, and the 9 delta functions are the 0–1 variates associated with the categorical variate development year, as mentioned in Section 2.2.2.

It is also necessary that the model include the variance structure set out in (3-10), and, by (2-21), this requires that observation $\hat{f}_{kj} - 1$ be assigned weight X_{kj}/ϕ_j. The values of ϕ_j are unknown, but the following argument will show that knowledge of their values is not required.

Consider MLE of the f_j. Commence with the log-likelihood of the claims trapezoid $\mathfrak{D}_K$:

$$\begin{aligned}\ell(\mathfrak{D}_K) &= \sum_{\mathfrak{D}_K, j\neq 1} \ell\left(\hat{f}_{kj} - 1\right) \\ &= \sum_{\mathfrak{D}_K, j\neq 1} \left\{ \frac{(Y_{kj}/X_{k,j-1})\, ln(f_{j-1} - 1) - (f_{j-1} - 1)}{\phi_{j-1}/X_{k,j-1}} - ln\left[(Y_{kj}/\phi_{j-1})!\right] \right\} \end{aligned} \tag{3-12}$$

where $\ell(\hat{f}_{kj} - 1)$ has been evaluated by substitution of (3-7)–(3-9) into (2-16).

The MLE of f_{j-1} for a specific j, say $j = i$, is obtained by differentiating (3-12) with respect to f_{i-1} and setting the result to zero. On differentiation:

- The final member within the braces is eliminated since it does not depend on f_{i-1}.
- The summation over $\mathfrak{D}_K$ is reduced to a summation over only $\mathcal{C}(i)$ since only this column depends on f_{i-1}.

The result is as follows:

$$\frac{\partial \ell(\mathfrak{D}_K)}{\partial f_{i-1}} = \frac{1}{\phi_{i-1}} \sum_{(k,i)\in\mathcal{C}(i)} X_{k,i-1} \frac{\partial}{\partial f_{i-1}} \left\{ (Y_{ij}/X_{k,i-1})\, ln(f_{i-1} - 1) - (f_{i-1} - 1) \right\} = 0 \tag{3-13}$$

The interested reader may complete the calculation to obtain the conventional chain ladder estimator (1-8) as the MLE, verifying the result cited in Section 3.1.2. However, all that is necessary for present purposes is to note that ϕ_{i-1} may be factored out of (3-13), in which case it does not enter into the MLE.

This means that the value of ϕ_{i-1} is arbitrary for the purpose of estimation of f_{i-1}, and so it may conveniently be set to unity. This lengthy digression thus shows that the above requirement of a weight X_{kj}/ϕ_j (ϕ_j unknown) to be associated with observation $\hat{f}_{kj} - 1$ in the GLM is reduced to a requirement of the simpler weight X_{kj}.

The ODP Mack model is now fully specified as a GLM. It may therefore be written in the general form of a GLM, as set out in Section 2.2.1. Specifically, the response vector Y now consists of all observations $Y_{k,j+1}/X_{kj}$ for all $Y_{k,j+1}$ in $\mathfrak{D}_K$ other than its first column, and written in some convenient order. The order is unimportant, but dictionary order is obvious and convenient: $\hat{f}_{11}, \ldots, \hat{f}_{1,J-1}, \hat{f}_{21}, \hat{f}_{22}, \ldots, \hat{f}_{K-2,1}, \hat{f}_{K-2,2}, \hat{f}_{K-1,1}$, and this will be assumed for the purpose of illustration.

Let μ denote the vector of μ_{kj}, also in dictionary order, and express it in the GLM form (2-23):

$$\mu = h^{-1}(X\beta) \tag{3-14}$$

where h, X and β can be determined by reference to (3-11):

$$h = identity$$
$$\beta = (f_1, f_2, \ldots, f_9)^T$$

$$X = \left[\begin{array}{ccc} 1 & & \\ & 1 & \\ & & \ddots \\ 1 & & \\ & 1 & \\ & & \ddots \\ & \vdots & \\ 1 & & \\ & 1 & \\ 1 & & \end{array}\right] \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{9 rows} \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{8 rows} \\ \\ \left.\begin{array}{c} \\ \\ \end{array}\right\} \text{2 rows} \\ \left.\begin{array}{c} \\ \end{array}\right\} \text{1 row} \end{array}$$

3.3.2. ODP Cross-Classified Model

Consider the ODP cross-classified model of Section 3.2, and particularly the conditions (EDFCC2a), modified to its ODP form, and (EDFCC2b). Together these conditions amount to the following:

$$Y_{kj} \sim ODP(\alpha_k \beta_j, \phi_{kj}) \tag{3-15}$$

Add the further condition

$$\phi_{kj} = \phi \tag{3-16}$$

which was seen in Section 3.2 to be a pre-requisite for ODP cross-classified model to yield the conventional chain ladder estimators as MLEs. Then

$$Y_{kj} \sim ODP(\alpha_k \beta_j, \phi) = ODP(\mu_{kj}, \phi) \tag{3-17}$$

where

$$\mu_{kj} = exp(ln\,\alpha_k + ln\,\beta_j) \tag{3-18}$$

The final equality here expresses the mean of the (k, j) cell as the exponential of a linear function of $ln\,\alpha_k$ and $ln\,\beta_j$. Thus (3-17) may be formulated as GLM in which the response vector consists of the observations Y_{kj}, the error distribution is ODP, the link function is the natural logarithm and the parameter vector takes the form $(ln\,\alpha_1, \ldots, ln\,\alpha_{10}, ln\,\beta_1, \ldots, ln\,\beta_{10})$. The scale parameter is unknown but will be estimated by the GLM software. Note how the logarithmic link function is pre-ordained by the multiplicative form of the assumption (EDFCC2b).

Just as in Section 3.3.1, the model may be expressed in the GLM form (2-23). If the components of Y are again written in dictionary order, then the design matrix is

$$X = \begin{bmatrix} 1\,0\ldots0 & 1\,0\ldots0 \\ 1\,0\ldots0 & 0\,1\ldots0 \\ & \vdots \\ 1\,0\ldots0 & 0\,0\ldots1 \\ 0\,1\ldots0 & 1\,0\ldots0 \\ 0\,1\ldots0 & 0\,1\ldots0 \\ & \vdots \\ & \vdots \\ 0\,0\ldots1 & 1\,0\ldots0 \end{bmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} \text{10 rows} \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} \text{9 rows} \\ \vphantom{a} \\ \} \text{ 1 row} \end{array}$$

Section 3.2 noted that the full parameter vector $(\alpha_1, \ldots, \alpha_{10}, \beta_1, \ldots, \beta_{10})$ contained one degree of redundancy, which was removed by the addition of the constraint (EDFCC2c). Likewise, the full parameter vector $(ln\,\alpha_1, \ldots, ln\,\alpha_{10}, ln\,\beta_1, \ldots, ln\,\beta_{10})$ of the GLM will contain a degree of redundancy.

In fact, this is no impediment to the fitting of the GLM for most GLM software. Most such software will remove redundancy by setting one or more (just one in the present case) parameters to zero. These parameters are said to be **aliased**.

Generally, this will lead to parameter estimates that differ from those obtained under condition (EDFCC2c), though the two GLMs are equivalent, simply stated differently. This is illustrated as follows.

Suppose that the GLM software chooses to set $ln\,\beta_1 = 0$, i.e., $\beta_1 = 1$. Simply replace each estimate $\hat{\beta}_j$ by $\hat{\beta}_j / \sum_{i=1}^{10} \hat{\beta}_i$ in order to satisfy (EDFCC2c). To compensate for this change, replace each $\hat{\alpha}_k$ by $\hat{\alpha}_k \sum_{i=1}^{10} \hat{\beta}_i$. With these replacements, the fitted value associated with Y_{kj} is

$$\hat{Y}_{kj} = \left[\hat{\alpha}_k \sum_{i=1}^{10} \hat{\beta}_i\right]\left[\hat{\beta}_j \Big/ \sum_{i=1}^{10} \hat{\beta}_i\right] = \hat{\alpha}_k \hat{\beta}_j \tag{3-19}$$

In other words, the model fitted values are unaltered by this re-scaling of the parameters α_k, β_j. Similarly for forecasts. In this sense, the alternative statements of the GLM are equivalent.

The forecast design matrix, as defined in (2-39), takes the form

$$X^* = \begin{bmatrix} 0\ 1\ 0 \ldots 0 & 0\ 0 \ldots 0\ 1 \\ 0\ 0\ 1 \ldots 0 & 0\ 0 \ldots 1\ 0 \\ 0\ 0\ 1 \ldots 0 & 0\ 0 \ldots 0\ 1 \\ & \vdots \\ 0\ 0 \ldots 0\ 1 & 0\ 1\ 0 \ldots 0 \\ & \vdots \\ 0\ 0 \ldots 0\ 1 & 0\ 0 \ldots 0\ 1 \end{bmatrix} \begin{matrix} \} \text{ 1 row} \\ \left.\begin{matrix} \\ \\ \end{matrix}\right\} \text{ 2 rows} \\ \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \text{ 9 rows} \end{matrix}$$

3.3.3. Numerical Example

The discussion in Sections 3.3.1 and 3.3.2 is illustrated by reference to the example data set. This data set is submitted to the GLM procedure GENMOD in SAS software according to both ODP Mack and ODP cross-classified models.

ODP Mack Model

The GLM formulation of the ODP Mack model, as described at the end of Section 3.3.1, has been applied to the example data set with the results displayed in Table 3-4. These results are seen to accord with those obtained by application of the chain ladder algorithm and set out in Table 3-1.

Table 3-4. GLM Parameter Estimates for ODP Mack Model

j	$\hat{f}_j - 1$
1	0.815
2	0.261
3	0.158
4	0.088
5	0.055
6	0.039
7	0.030
8	0.025
9	0.021

ODP Cross-Classified Model

The GLM formulation of the ODP cross-classified model, as set out in (3-17) and (3-18), has been applied to the example data set with the results displayed in Table 3-5. The parameter estimates in the columns headed $ln\ \alpha_k$ and $ln\ \beta_j$ have been extracted directly from the GLM output. In the next two columns they have been exponentiated, and in final two columns re-scaled as described in the paragraph preceding (3-19) so that the $\sum_{j=1}^{10} \hat{\beta}_j = 1$. The results are seen to agree with those found in Table 3-3 (subject to a couple of microscopic differences).

Table 3-5. GLM Parameter Estimates for ODP Cross-Classified Model

	Estimated Directly from GLM				Re-normalised	
j or k	$\ln \hat{\alpha}_k$	$\ln \hat{\beta}_j$	$\hat{\alpha}_k$	$\hat{\beta}_j$	$\hat{\alpha}_k$	$\hat{\beta}_j$
1	10.657	0.000	42479	1.000	144781	0.293
2	10.795	–0.205	48793	0.815	166301	0.239
3	10.899	–0.747	54133	0.474	184501	0.139
4	10.989	–1.017	59221	0.362	201845	0.106
5	11.039	–1.452	62245	0.234	212151	0.069
6	11.016	–1.833	60834	0.160	207341	0.047
7	11.008	–2.140	60360	0.118	205726	0.035
8	10.891	–2.348	53664	0.096	182905	0.028
9	10.836	–2.513	50824	0.081	173225	0.024
10	10.691	–2.664	43962	0.070	149837	0.020
Total				3.408		1.000

3.4. Minor Variations of Chain Ladder

Hitherto the chain ladder model has been presented as containing no flexibility; as the non-parametric Mack model, the EDF Mack model, or one of the other variations defined earlier in this chapter, but in each case fully defined without any scope for variation by the user. In practice, many variations occur. This section will consider a few of the common variations and show that they may be easily incorporated in a GLM.

3.4.1. Reliance on Only Recent Experience Years

It is common to view only the most recent m experience years as relevant to parameter estimation. This would mean in the ODP Mack model (Section 3.3.1), for example, that the only observations used would be $\hat{f}_{kj} - 1|X_{kj}$, $k = 1, \ldots, K-1$, $j = 1, \ldots, J-1$, $K+1-m \le k+j \le K$.

This restriction is easily implemented within the GLM defined in Section 3.3.1 by simply setting the weight of each observation other than those above to zero, i.e., the weight X_{kj} assigned to observation $\hat{f}_{kj} - 1|X_{kj}$ at the end of Section 3.3.1 is modified to the following:

$$w_{kj} = X_{kj} I\left(K+1-m \le k+j \le K\right) \tag{3-20}$$

where $I(.)$ is the indicator function:

$$\begin{aligned} I(c) &= 1 \text{ if the logical condition } c \text{ is true} \\ &= 0 \text{ otherwise} \end{aligned} \tag{3-21}$$

Consider the likelihood (3-12), omitting the final member within the braces since it was seen in Section 3.1.2 to vanish in the likelihood maximization, and set weights according to (3-20). The weights are included in the model by means of (2-21). Thus log-likelihood (3-12) becomes:

$$\ell(\mathfrak{D}_K) = \sum_{\mathfrak{D}_{K}, j \neq 1} I(K+1-m \leq k+j \leq K) \times \left\{ \frac{(Y_{kj}/X_{k,j-1}) ln(f_{j-1}-1) - (f_{j-1}-1)}{\phi_{j-1}/X_{k,j-1}} \right\} \tag{3-22}$$

and the indicator function has the effect of simply selecting the Y_{kj} from the last m experience years for inclusion in the log-likelihood.

3.4.2. Outlier Observations

The argument leading to the last result has been phrased in terms specific to the ODP Mack model. However, it may be generalized to **any model** with the conclusion that setting the weight of any observation to zero causes that observation, in effect, to be deleted from the data set.

It follows that **outlier observations** may be excluded from the model fitting simply by the assignment of zero weights to them.

4. Prediction Error

Chapter summary. This chapter is concerned with the error contained in a forecast derived from a GLM in accordance with Chapter 2, as compared with the actual value of the predictand when ultimately observed. This error is decomposed into its components: parameter error, process error, and model error.

The properties of parameter and process errors follow from the model, whereas the properties of model error do not. For the main part, the chapter deals with the more tractable parameter and process errors.

Mean square error of prediction is discussed as a measure of forecast error, and it is noted that increased goodness-of-fit of a model does not necessarily imply reduced forecast error. Information criteria are introduced as simple rough proxies for forecast error to assist in the evaluation of competing models.

The literature on model error is scant, but the subject receives some discussion at the end of the chapter.

4.1. Parameter Error and Process Error

4.1.1. Individual Observations

For the purpose of the current chapter the model used for the forecast of outstanding losses will not be limited to the chain ladder. The model will be loosely specified as follows:

$$Y_{kj} = u(k, j; \theta) + \varepsilon_{kj} \text{ for } Y_{kj} \in \mathfrak{D}_K^+ \tag{4-1}$$

for some function u, dependent on a parameter vector θ, and **centered** stochastic error ε_{kj}, i.e.,

$$E[\varepsilon_{kj}] = 0 \tag{4-2}$$

It will be supposed that this model has been calibrated against that data set $\mathfrak{D}_K$. The means of calibration is left unspecified. It yields parameter estimates $\hat{\theta}$. Now define

$$\hat{Y}_{kj} = u(k, j; \hat{\theta}) \text{ for } Y_{kj} \in \mathfrak{D}_K^+ \tag{4-3}$$

The $\hat{Y}_{kj}$ associated with $Y_{kj} \in \mathfrak{D}_K$ are fitted values, as in (2-29). The $\hat{Y}_{kj}$ associated with $Y_{kj} \in \mathfrak{D}_K^c$ are forecasts.

The **prediction error** associated with the forecast $\hat{Y}_{kj}$ is

$$e_{kj} = Y_{kj} - \hat{Y}_{kj} = \left[u(k, j; \theta) - u\left(k, j; \hat{\theta}\right)\right] + \varepsilon_{kj} \tag{4-4}$$

where the second equality follows from (4-1) and (4-3).

It may be noted from (4-1) and (4-2) that

$$E\left[Y_{kj}\right] = u(k, j; \theta) \tag{4-5}$$

and so (4-4) may be represented in the alternative form

$$e_{kj} = \left[\mu_{kj} - \hat{Y}_{kj}\right] + \varepsilon_{kj} \tag{4-6}$$

where μ_{kj} denotes $E[Y_{kj}]$.

The square bracketed term in (4-6) (or (4-4)) is the difference between the true (but unknown) mean of the future observation and its forecast, and is referred to as the **parameter error** associated with forecast $\hat{Y}_{kj}$. The remaining term ε_{kj} is noise or, as it is usually referred to, **process error**. It reflects the fact that, even if the model had been perfectly calibrated (zero parameter error), a prediction error would still arise from the stochastic nature of future observation.

Typically, parameter error and process error may be shown to be stochastically independent. Note that $\hat{Y}_{kj}$, on which parameter error depends, is necessarily some function of **past data** $\mathfrak{D}_K$, whereas the ε_{kj} are components of the **future data** $\mathfrak{D}_K^c$. If the model formulation is such that the past Y_{kj} and the future ε_{kj} are independent, then so are the parameter and process errors.

This follows very simply in any model, such as the EDF cross-classified model of Section 3.2, which specifies that all observations are independent.

The above argument is subject to a substantial qualification that will not be pursued in the present volume. The relation (4-5) may indeed be consistent with (4-1), but both assume that the model u has been correctly specified.

In fact, it is unlikely that the precise functional form of u will have been correctly chosen. As a result, a further component of prediction error arises in practice. This is the difference between $E[Y_{kj}]$, as specified by (4-5), and its correct specification, usually referred to as **model error**. It is discussed in greater detail in Section 4.5.

Model error, by its nature, lacks amenability to rigorous statistical treatment. For this reason, it is regarded as outside the scope of this monograph. This is by no means to suggest that it is insignificant. Indeed, its magnitude may in some cases exceed the total of parametric and process errors. The interested reader might consult O'Dowd, Smith and Hardy (2005) for a suggested treatment of model error.

4.1.2. Loss Reserves

For notational brevity, it will be convenient to represent the above prediction errors in vector terms. Let Y denote the observations $Y_{kj} \in \mathfrak{D}_K$, assembled into a vector, and let Y^* denote the observations $Y_{kj} \in \mathfrak{D}_K^c$, similarly assembled into a vector. The ordering of the components of these vectors is immaterial for present purposes.

Similarly, assemble any other quantity that depends on k, j into a vector and denote that vector by the original quantity's symbol with k and j omitted. Add a star to the symbol if it refers to the future. Again, the ordering of components is immaterial, but it must be consistent between different vectors, e.g., the ordering of cells of $\mathfrak{D}_K^c$ must be the same in Y^* and $\hat{Y}^*$.

In this notation, (4-6) becomes

$$e^* = \left[\mu^* - \hat{Y}^*\right] + \varepsilon^* \tag{4-7}$$

Now consider any linear combination of the components of Y^*, represented by $r^T Y^*$, where r denotes some vector and the upper T denotes vector transposition. For example, the total amount of outstanding claims is equal to $1^T Y^*$, where the vector 1 has all components equal to unity. As a second example, the amount of outstanding claims in respect of just accident year k is equal to $r_k^T Y^*$, where the vector r_k contains unity in those components that refer to accident year k, and zero for all other components.

The prediction error associated with $r^T Y^*$ will be denoted $e^*_{(r)}$ and, by (4-7), is

$$e^*_{(r)} = r^T e^* = \left[r^T \mu^* - r^T \hat{Y}^*\right] + r^T \varepsilon^* \tag{4-8}$$

where the members on the right can be recognized as follows:

$r^T \mu^*$ is the statistical expectation of outstanding losses;
$r^T \hat{Y}^*$ is the forecast of the quantum of these losses;
$r^T \varepsilon^*$ is the process error associated with this quantum.

The square-bracketed term in (4-8) can be identified as the parameter error associated with the forecast of outstanding losses. If Y and ε^* are stochastically independent, then, by the same argument as in Section 4.1.1, parameter error and process error will be independent.

4.2. Mean Square Error of Prediction

4.2.1. Definition

A useful summary measure of the magnitude of prediction error $e^*_{(r)}$ is its **mean square error of prediction**, abbreviated to **MSEP** and denoted $MSEP[e^*_{(r)}]$. It is defined as

$$MSEP\left[e^*_{(r)}\right] = E\left\{\left[e^*_{(r)}\right]^2\right\} \tag{4-9}$$

In the case where parameter and process errors can be established to be stochastically independent, substitution of (4-8) into (4-9) yields

$$MSEP\left[e^*_{(r)}\right] = E\left\{\left[e^*_{(r)param}\right]^2\right\} + E\left\{\left[e^*_{(r)proc}\right]^2\right\} \tag{4-10}$$

where the following notation has been introduced:

$$e^*_{(r)param} = r^T \mu^* - r^T \hat{Y}^* = \text{parameter error} \tag{4-11}$$

$$e^*_{(r)proc} = r^T \varepsilon^* = \text{process error} \tag{4-12}$$

4.2.2. Goodness-of-Fit and Prediction Error

The MSEP estimates the tightness of a forecast around its target. A model generating a smaller MSEP is generally to be preferred over one generating a larger MSEP.

It is to be noted, however, that improving the goodness-of-fit of a model to a data set does not necessarily improve its MSEP. It is evident that an effective model requires some degree of goodness-of-fit, but the achievement of this by the inclusion of an excessive number of parameters in the model will in fact increase the MSEP.

In short, the inclusion of too many parameters in a model amounts to over-fitting, and destabilizes the model's predictions. The situation is summarized by Figure 4-1 (see, e.g., Hastie, Tibshirani and Friedman (2009, pp. 219–223)). The figure considers the effect of increased model complexity (number of model parameters) on the model's predictive value.

It is supposed that the available data set is divided into two subsets, a **training set** and a **test (or holdout) set**. The model is fitted to the training set. Some form of error in the fit ("model error" in the figure) of the model to the data, such as squared error, deviance, etc., is selected and plotted against model complexity. The fit of the model to the data is seen to improve monotonically as model complexity is increased.

However, the value of the model as a predictor of unseen data does not improve in the same way. The model error when the model is used to generate fitted values corresponding to the test set is also plotted in the figure. It is seen that a model with very few parameters produces a poor fit; it represents a weak attempt to extract the main characteristics of the training data set.

As complexity is added to the model, it not only fits the training data set better, but also predicts the test set better. Beyond a certain point, however, additional complexity detracts from the model; its performance in the prediction of the test set begins to deteriorate.

This indicates over-fitting. The model is beginning to parameterize the noise in the data, of no value for prediction. In the extreme case in which the model contains as many parameters as the training data set contains observations, the model will fit the data perfectly (zero error). However, this cannot be regarded as a model at all in the usual sense. It has no predictive value.

The minimum point on the "Test" curve of Figure 4-1 represents the optimum model complexity. It is the model with greatest predictive value.

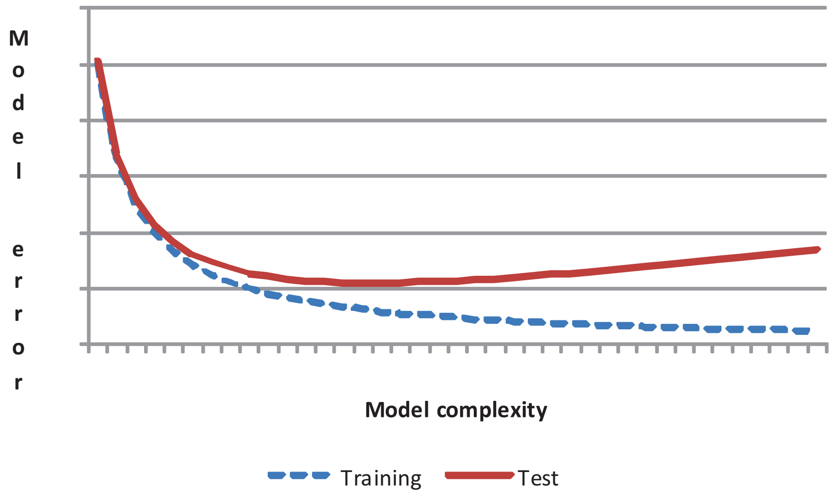
4.3. Information Criteria

There exist statistics which function as proxies for measurement of model predictive error relative to a test data set. These are called **information criteria**, and take the general form:

$$\begin{aligned} \textit{information criterion} &= \textit{measure of model fit error} \left(\textit{relative to training data set}\right) \\ &\quad + \textit{penalty for number of parameters} \end{aligned} \tag{4-13}$$

As model complexity increases the error in the fit of the model decreases but the penalty for number of parameters increases. The information criterion behaves in a manner similar to the model error relative to a test data set, as in Figure 4-1.

Figure 4-1. Goodness-of-Fit and Prediction Error



For a GLM, a convenient form of (4-13) for a model based on data Y and producing fitted values $\hat{Y}$ is:

$$IC\left(Y,\hat{Y}\right)=D\left(Y,\hat{Y}\right)+f\left(p\right) \tag{4-14}$$

where

$IC(Y, \hat{Y})$ denotes the information criterion;
$D(Y, \hat{Y})$ denotes the scaled deviance, defined by (2-30);
p denotes the number of model parameters; and
$f(.)$ is some monotonically increasing function.

The two most common forms of information criterion are defined by the penalty functions set out in Table 4-1, where n denotes the dimensionality of Y, i.e., the number of observations used in the fitting of the model.

The penalty functions of both criteria are linear in p, but the BIC applies the heavier penalty.

There is a modified form of the AIC, called AICc, that contains a correction for finite sample size n. In this case, $f(p) = 2p[1 + (p + 1)/(n - p - 1)] \to 2p$ as $n/p \to \infty$.

The information criteria are used for the comparison of different models of the same data set. All models involve some loss of information contained in the data. If the AIC (say) assumes a lower value for Model 1 than for Model 2, then Model 1 is indicated as the more likely of the two to have minimized the information loss, and Model 1 would be selected in preference to Model 2.

Table 4-1. Information Criteria

Information Criterion	Function $f(p)$
Akaike Information Criterion (AIC)	$2p$
Bayes Information Criterion (BIC)	$p \ln n$

4.4. Generalized Cross-Validation

Cross validation is a frequently used method for estimating prediction error, being easily applicable to regression and non-regression models alike. For example, in K-fold cross-validation, the data is split into K equal sized parts, with the model fitted on K-1 parts and tested on the final K^{th} part. A common choice for K is *n,* i.e., one point is left out of the fit for each iteration of the calculation. This is also referred to as leave-one-out cross-validation.

For linear models, where the fitted value may be expressed as $\hat{y} = Hy$, it may be shown that an approximation to leave-one-out validation is given by the **generalized cross-validation** (**"GCV"**) measure:

$$GCV = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n\left[1 - trace\left(H\right)/n\right]^2} \tag{4-15}$$

where:

Y_i is the i^{th} observed value
$\hat{Y}_i$ is the i^{th} fitted value
n is the number of observations

H is often referred to as the hat matrix. The trace of the hat matrix, trace(H), is defined as the **effective number of parameters** in a model.

Further discussion of all these points is given in Hastie, Tibshirani and Friedman (2009, pp. 232–233 and 241–245), who also note that the GCV measure is related to likelihood based measures such as AIC and BIC. As with those measures, it is composed of two parts: the first relating to the measure of model fit error (the residual sum of squares in this case, i.e., $\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2$) and the second being a penalty for the number of parameters (the remainder of (4-15)).

4.5. Model Error

Re-consider the decomposition of prediction error into parameter and process error in Section 4.1. Recall (4-1), where the non-stochastic part of each observation is $u(k, j; \theta)$. Now note that the fitted values of (4-3) are assumed to take the form $u(k, j; \hat{\theta})$, i.e., the same parametric form with unknown parameters replaced by their estimates.

There is a tacit pre-supposition here that the function $u(k, j; \theta)$ underlying the data can be accurately identified for modeling purposes. Thus was useful for didactic reasons in Section 4.1, but in fact this function will be unknown, and essentially unknowable. Not even the parameter set on which it depends will be identifiable in practice.

To recognize this, continue to suppose that (4-1) holds, but now suppose that, in ignorance of this parametric form, one has supposed for modeling purposes that

$$Y_{kj} = v\left(k, j; \xi\right) + \varepsilon_{kj} \text{ for } Y_{kj} \in \mathfrak{D}_K^+ \tag{4-16}$$

for some different approximation function $v(.)$ with a different parameter set.

The fitted values from this model will be

$$\hat{Y}_{kj} = v\left(k,\, j;\, \hat{\xi}\right) \text{ for } Y_{kj} \in \mathfrak{D}_K^+ \tag{4-17}$$

and the decomposition of prediction error corresponding to (4-4) is now

$$e_{kj} = Y_{kj} - \hat{Y}_{kj} = \underbrace{\left[v(k,\, j;\, \xi) - v\left(k,\, j;\, \hat{\xi}\right)\right]}_{\text{Parameter error}} + \underbrace{\varepsilon_{kj}}_{\text{Process error}} + \underbrace{\left[u(k,\, j;\, \theta) - v(k,\, j;\, \xi)\right]}_{\text{Model error}} \tag{4-18}$$

The decomposition contains parameter error and process error terms as in (4-4), but now includes an additional term that has been labelled **model error**. This is the term $[u(k, j; \theta) - v(k, j; \xi)]$, which measures the difference between the parametric form assumed for the model and the true but unknown parametric form, i.e., the error introduced by the choice of model.

Since model error involves the form $u(k, j; \theta)$, that has already been pronounced unknowable, its quantification is difficult. There is no known procedure for its estimation by reference just to the data relied on by the modeling.

There have, however, been one or two attempts to estimate model error from data and/or opinions external to the data set. Notable in this respect is the contribution by O'Dowd, Smith and Hardy (2005), which sets out:

- to identify the major potential causes of model error;
- to score each subjectively for its likely magnitude in the model under assessment;
- to map the scores to quantitative measures of error (e.g., coefficient of variation);
- to combine these measures with those for parameter and process error, with due allowance for any dependencies (also subjectively assessed) between the various components of model error.

This monograph is, as its title indicates, concerned with the application of GLMs to loss reserving. The assessment of model error will address the GLM used but, as can be seen from the description of O'Dowd, Smith and Hardy (2005), will not be carried out within the framework of that GLM. It will therefore not be discussed further here.

This is not, however, to minimize the importance of model error and the need to address it. In many cases it will represent a material, possibly even a dominant, proportion of total prediction error. For example, in the case of one large insurer, model error was assessed as representing about three-quarters of total prediction error.

5. The Bootstrap

Chapter summary. This chapter is concerned with the estimation of the prediction error associated with outstanding losses, excluding the contribution of model error (as explained in the summary of Chapter 4). Two approaches are taken: the delta method, and the bootstrap.

Although the delta method is relatively simple computationally, its accuracy in any particular application is unknown, and may be dubious in some cases. Further, although it provides an estimate of MSEP, it provides no information on the distributional properties of prediction error, e.g., quantiles.

The bootstrap, while computationally more demanding, remedies both shortcomings. This is a device that generates many synthetic data sets with the same stochastic properties as the original one, and produces an estimate of outstanding losses from each dataset. It thus estimates the full distribution of prediction error and, with sufficient computation, its accuracy can be increased arbitrarily. Two forms of the bootstrap are examined.

The chapter concludes with numerical examples of both the delta method and the bootstrap.

5.1. Background

A chain ladder forecast was carried out in Table 3-2 on the basis of the chain ladder algorithm. The algorithm was merely heuristic and so the stochastic properties of the forecast were undetermined.

However, it was shown in Chapter 3 that the same algorithm, and so the same forecast, emerged from two different stochastic models. In each of those cases, the stochastic properties of the forecast follow, at least in principle.

The two stochastic chain ladder models were formulated in the form of GLMs in Section 3.3, whose parameter estimates were reported in Table 3-4 and Table 3-5. Although only the estimates themselves were reported there, the GLM software in fact also provides estimates of the associated standard errors, as in Table 5-1.

The parameter $ln\ \beta_1$ has been aliased here in the manner described in Section 3.3.2. Since this amounts to selecting a zero (deterministic) value for this parameter, the associated standard error is zero.

The estimated correlations between parameter estimates are also provided by the GLM software. These are displayed in Table 5-2. Only the lower triangle of the correlation

Table 5-1. GLM Parameter Estimates and Standard Errors for ODP Cross-Classified Model

	$ln\ \hat{\alpha}_k$		$ln\ \hat{\beta}_j$	
j or *k*	Estimate	Standard Error	Estimate	Standard Error
1	10.657	0.0316	0.000	
2	10.795	0.0299	–0.205	0.0228
3	10.899	0.0289	–0.747	0.0282
4	10.989	0.0281	–1.017	0.0328
5	11.039	0.0278	–1.452	0.0421
6	11.016	0.0285	–1.833	0.0547
7	11.008	0.0295	–2.140	0.0715
8	10.891	0.0327	–2.348	0.0931
9	10.836	0.0367	–2.513	0.1267
10	10.691	0.0510	–2.664	0.1993

Table 5-2. Estimated Correlation Matrix of GLM Parameter Estimates for ODP Cross-Classified Model

	Parameter									
Parameter	$ln\ \hat{\alpha}_1$	$ln\ \hat{\alpha}_2$	$ln\ \hat{\alpha}_3$	$ln\ \hat{\alpha}_4$	$ln\ \hat{\alpha}_5$	$ln\ \hat{\alpha}_6$	$ln\ \hat{\alpha}_7$	$ln\ \hat{\alpha}_8$	$ln\ \hat{\alpha}_9$	$ln\ \hat{\alpha}_{10}$
$ln\ \hat{\alpha}_1$	1.00									
$ln\ \hat{\alpha}_2$	0.20	1.00								
$ln\ \hat{\alpha}_3$	0.20	0.21	1.00							
$ln\ \hat{\alpha}_4$	0.20	0.21	0.22	1.00						
$ln\ \hat{\alpha}_5$	0.19	0.20	0.21	0.22	1.00					
$ln\ \hat{\alpha}_6$	0.18	0.19	0.20	0.20	0.20	1.00				
$ln\ \hat{\alpha}_7$	0.16	0.17	0.18	0.18	0.18	0.18	1.00			
$ln\ \hat{\alpha}_8$	0.13	0.14	0.14	0.15	0.15	0.14	0.14	1.00		
$ln\ \hat{\alpha}_9$	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.09	1.00	
$ln\ \hat{\alpha}_{10}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

(continued on next page)

Table 5-2. Estimated Correlation Matrix of GLM Parameter Estimates for ODP Cross-Classified Model *(continued)*

	Parameter									
Parameter	$ln\ \hat{\alpha}_1$	$ln\ \hat{\alpha}_2$	$ln\ \hat{\alpha}_3$	$ln\ \hat{\alpha}_4$	$ln\ \hat{\alpha}_5$	$ln\ \hat{\alpha}_6$	$ln\ \hat{\alpha}_7$	$ln\ \hat{\alpha}_8$	$ln\ \hat{\alpha}_9$	$ln\ \hat{\alpha}_{10}$
$ln\ \hat{\beta}_2$	–0.32	–0.34	–0.35	–0.36	–0.37	–0.36	–0.35	–0.31	–0.28	0.00
$ln\ \hat{\beta}_3$	–0.28	–0.29	–0.30	–0.31	–0.32	–0.31	–0.30	–0.27	–0.10	0.00
$ln\ \hat{\beta}_4$	–0.25	–0.27	–0.28	–0.29	–0.29	–0.28	–0.27	–0.12	–0.09	0.00
$ln\ \hat{\beta}_5$	–0.21	–0.22	–0.23	–0.24	–0.24	–0.24	–0.12	–0.10	–0.07	0.00
$ln\ \hat{\beta}_6$	–0.18	–0.19	–0.20	–0.20	–0.20	–0.10	–0.09	–0.07	–0.05	0.00
$ln\ \hat{\beta}_7$	–0.16	–0.17	–0.17	–0.18	–0.09	–0.08	–0.07	–0.06	–0.04	0.00
$ln\ \hat{\beta}_8$	–0.14	–0.15	–0.16	–0.07	–0.07	–0.06	–0.06	–0.04	–0.03	0.00
$ln\ \hat{\beta}_9$	–0.14	–0.15	–0.05	–0.05	–0.05	–0.04	–0.04	–0.03	–0.02	0.00
$ln\ \hat{\beta}_{10}$	–0.16	–0.03	–0.03	–0.03	–0.03	–0.03	–0.03	–0.02	–0.01	0.00

	Parameter								
Parameter	$ln\ \hat{\beta}_2$	$ln\ \hat{\beta}_3$	$ln\ \hat{\beta}_4$	$ln\ \hat{\beta}_5$	$ln\ \hat{\beta}_6$	$ln\ \hat{\beta}_7$	$ln\ \hat{\beta}_8$	$ln\ \hat{\beta}_9$	$ln\ \hat{\beta}_{10}$
$ln\ \hat{\beta}_2$	1.00								
$ln\ \hat{\beta}_3$	0.36	1.00							
$ln\ \hat{\beta}_4$	0.31	0.27	1.00						
$ln\ \hat{\beta}_5$	0.24	0.21	0.19	1.00					
$ln\ \hat{\beta}_6$	0.19	0.16	0.15	0.12	1.00				
$ln\ \hat{\beta}_7$	0.14	0.12	0.11	0.09	0.08	1.00			
$ln\ \hat{\beta}_8$	0.11	0.09	0.09	0.07	0.06	0.05	1.00		
$ln\ \hat{\beta}_9$	0.08	0.07	0.06	0.05	0.04	0.04	0.04	1.00	
$ln\ \hat{\beta}_{10}$	0.05	0.04	0.04	0.03	0.03	0.02	0.02	0.02	1.00

matrix is displayed, the upper triangle being given by symmetry. Since parameter $ln\ \beta_1$ has been fixed at zero, it is non-stochastic and does not appear in the matrix.

5.2. Delta Method

From Table 5-1 and Table 5-2 all estimated second order moments of the parameter estimates are available. This is sufficient for an approximate estimate of the second moments of the estimated total outstanding losses $\hat{R}$. This is done using the so-called **delta method** (Kendall and Stuart, 1977).

5.2.1. Uni-Dimensional

This method is most easily understood for a single-dimensional variate. Here the purpose is to calculate the variance of a transformed variate when the variance of the untransformed variate is known.

In the interest of simplicity, the following notation will apply just to the present sub-section. It is unrelated to the notation introduced in Section 1.2.

Let X denote a random variate with $E[X] = \mu$, $Var[X] = \sigma^2$, and let f denote a differentiable one-one transformation of X. The quantity $Var[f(X)]$ is required.

Take the Taylor series expansion of $f(X)$ to second order about $X = \mu$:

$$f(X) = f(\mu) + (X-\mu) f'(\mu) + \tfrac{1}{2}(X-\mu)^2 f''(\mu) + \cdots \qquad (5\text{-}1)$$

where the primes denote differentiation.

Take expectations with respect to X throughout (5-1):

$$E[f(X)] = f(\mu) + \tfrac{1}{2}E\left[(X-\mu)^2\right] f''(\mu) + \cdots = f(\mu) + \tfrac{1}{2}\sigma^2 f''(\mu) + \cdots \qquad (5\text{-}2)$$

as a second order approximation of $E[f(X)]$, where $E[X - \mu]$ is seen to vanish in the first equation.

Now consider $Var[f(X)] = E\{[f(X) - E[f(X)]]^2\}$. For a second order approximation of this quantity, a first order approximation of $f(X) - E[f(X)]$ is required. This is obtained from (5-1) and (5-2) as

$$f(X) - E[f(X)] = (X-\mu) f'(\mu) + \cdots \qquad (5\text{-}3)$$

from which

$$Var[f(X)] = E\left[(X-\mu)^2 [f'(\mu)]^2 + \cdots\right] = \sigma^2 [f'(\mu)]^2 + \cdots \qquad (5\text{-}4)$$

This provides an easily calculated second order approximation of $Var[f(X)]$.

5.2.2. Multi-Dimensional

With Section 5.2.1 for guidance, extend to the case in which $Y = f(X)$ with X now a column n-vector, and with $f: \Re \rightarrow \Re$ acting on X componentwise (just as h^{-1} did in (2-23)). Let the components of X, Y be denoted X_i, Y_i respectively. In parallel with (5-3),

$$Y_i - E[Y_i] = f(X_i) - E[f(X_i)] = (X_i - \mu_i) f'(\mu_i) + \cdots \qquad (5\text{-}5)$$

with $\mu_i = E[X_i]$.

Then second order approximations of covariances may be obtained as

$$Cov[Y_i, Y_j] = E\{[Y_i - E[Y_i]][Y_j - E[Y_j]]\} = Cov[X_i, X_j] f'(\mu_i) f'(\mu_j) \qquad (5\text{-}6)$$

This may be conveniently expressed in matrix form, thus:

$$Var[Y] = D\,Var[X]\,D \qquad (5\text{-}7)$$

where $Var[Y]$ now denotes the entire variance-covariance matrix of vector Y, similarly for $Var[X]$, and $D = diag[f'(\mu_1), \ldots, f'(\mu_n)]$.

Table 5-3. Replacements in Section 5.2.2 for Estimation of Forecast Error

Quantity from Section 5.2.2	Replaced By
Y	$\hat{Y}^*$
X	$X^*\hat{\beta}$
μ	$X^*\beta$
f	h^{-1}

5.2.3. Application to Loss Reserving

Now replace Y of Section 5.2.2 by the forecast $\hat{Y}^*$, defined by (2-40), in order to estimate the variance of that forecast due to variation in $\hat{\beta}$, i.e., parameter error as defined in Section 4.1. It will be assumed that the components of $\hat{Y}^*$ appear in dictionary order, as was illustrated in Section 3.3.2. Other quantities from Section 5.2.2 also require replacement by those relevant to (2-40). Table 5-3 lists the required replacements.

With the replacements in the table, supplemented by this last one, (5-7) becomes

$$Var_{param}\left[\hat{Y}^*\right] = DVar\left[X^*\hat{\beta}\right]D = DX^*Var\left[\hat{\beta}\right]\left(X^*\right)^T D \tag{5-8}$$

where $Var_{param}[\hat{Y}^*]$ has been written instead of $Var[\hat{Y}^*]$ as a reminder that only parameter error is being estimated, $Var[\hat{\beta}]$ is estimated by the GLM software and

$$D = diag\left[\left(h^{-1}\right)'\left(h\left(\hat{Y}^*_{2,10}\right)\right), \ldots, \left(h^{-1}\right)'\left(h\left(\hat{Y}^*_{10,10}\right)\right)\right] \tag{5-9}$$

where the vector $X^*\beta$ of the innermost arguments has been replaced by $X^*\hat{\beta} = h(\hat{Y}^*)$.

Finally, the full prediction error of $\hat{Y}^*$, other than model error, may be estimated by adding process error (see (4-10)) where, for the case of the ODP distribution, process error is given by (2-19). Translation of this to the present context yields an estimated process error of

$$Var_{proc}\left[\hat{Y}^*\right] = \hat{\phi}DIAG\left[\hat{Y}^*\right] \tag{5-10}$$

where, for a vector v, $DIAG[v]$ denotes the diagonal matrix with the components of v along its diagonal, and the estimate $\hat{\phi}$ of scale parameter ϕ is provided by the GLM software.

The MSEP of the forecast $\hat{Y}^*$ is now obtainable by combination of (5-8) and (5-10) thus:

$$\begin{aligned} MSEP\left[\hat{Y}^*\right] &= DVar\left[X^*\hat{\beta}\right]D + \hat{\phi}DIAG\left[\hat{Y}^*\right] \\ &= DX^*Var\left[\hat{\beta}\right]\left(X^*\right)^T D + \hat{\phi}DIAG\left[\hat{Y}^*\right] \end{aligned} \tag{5-11}$$

Estimates of the prediction error of outstanding loss amounts R_k and R, or for that matter sums over any other subset of $\hat{Y}^*$, can be obtained by the use of vectors consisting of just 0-1 components, selecting out the relevant components of $\hat{Y}^*$.

For example,

$$MSEP\left[\hat{R}_k\right] = MSEP\left[1_k \hat{Y}^*\right] = 1_k^T\, MSEP\left[\hat{Y}^*\right] 1_k \tag{5-12}$$

where $1_k^T = (0, \ldots, 0, 1, \ldots, 1, 0, \ldots, 0)$ with the 1's so placed as to select the components $\hat{Y}^*_{k,J-k+2}, \ldots, \hat{Y}^*_{kJ}$ of $\hat{Y}^*$.

Similarly

$$Var\left[\hat{R}\right] = 1^T\, MSEP\left[\hat{Y}^*\right] 1 \tag{5-13}$$

where 1 is a vector consisting entirely of unit components.

A numerical example will be given in Section 5.4.1.

5.3. The Bootstrap

The delta method presents two difficulties.

First, since it is a second order approximation to covariance, it leaves an unquantified third order error. It is evident from the development in (5-1)–(5-4) that the error depends on the magnitudes of the higher derivatives $f^{(m)}$ (equivalently $(h^{-1})^{(m)}$ in Section 5.2.3), and especially on the convexity f'' (equivalently $(h^{-1})''$). This knowledge may be insufficient, however, for the formation of a clear view of the magnitude of error.

Second, even a relatively accurate estimation of second order moments provides little distributional information. It may be necessary to estimate quantiles of $\hat{R}$ for loss reserving purposes. For example, some regulators require the loss reserve to be equal to the estimated amount of outstanding losses with $100p\%$ ($p > 50$) probability of adequacy ("**PoA**"). If this amount is denoted $\hat{R}_p$, then it is defined as follows:

$$Prob\left[R < \hat{R}_p\right] = p \tag{5-14}$$

It is evident that estimation of $\hat{R}_p$ requires knowledge of the distribution of R. The delta method does not provide this. It is possible, of course, to assume some distribution. Often this is done in practice, where the lognormal distribution is often assumed for R. In fact, the lognormal often appears to perform quite well, but there is no guarantee of this and the procedure is at risk of producing erroneous PoA loss reserves, particularly for high p.

The bootstrap is a procedure which estimates the entire distribution of the estimand. It is a particularly convenient computational device since it does this without the need for any algebraic development such as in Section 5.2. Naturally, since it estimates an entire distribution, it also generates an estimate of variance (for that matter, any other moment or functional of the distribution).

There are many different approaches to the bootstrap. Shibata (1997) provides a useful classification of some of these into "non-parametric," "semi-parametric," and "parametric," with the terminology indicating the level of reliance on model and distributional assumptions. For any specific case, it is useful to consider the estimation of parameter and process separately and which of non-parametric or parametric methods are best suited to the problem in hand.

Some possible approaches to bootstrapping claims data are discussed below, following the terminology of Shibata where appropriate.

5.3.1. Semi-Parametric Bootstrap

The original form of the bootstrap was introduced by Efron (1979). It is a procedure for estimation of the properties of a defined statistic, particularly when analytical computation of those properties would be complex. It falls within the general family of **re-sampling methods**, since it involves repeated sampling from the available data.

For regression models, Efron (1979) proposed a procedure that involved resampling residuals and constructing pseudo datasets from these and fitted values. This type of procedure is outlined here. Consider an n-dimensional data vector Y. For the moment this is a general vector, and the bootstrap will be described in a general context. Later it will be specialized to the loss reserving context. Suppose that a model has been fitted to the data vector and a prediction $\hat{Y}^*$ of some vector Y^* of future observations made.

Suppose the target prediction is some function $R(Y^*)$ of Y^*, and it has been estimated by $R(\hat{Y}^*)$. The objective now is to estimate the distribution of the prediction $R(\hat{Y}^*)$.

Let $\hat{Y}$ denote the model's vector of fitted values corresponding to Y, and let $S(Y; \hat{Y})$ denote the vector of standardized residuals associated with Y. Residuals may be Pearson, deviance or any other for which the inverse $S^{-1}(.; \hat{Y})$ exists.

For example, in the case of Pearson residuals, the i-th component of $S(Y; \hat{Y})$ is

$$S_i\left(Y;\hat{Y}\right)=\left(Y_i-\hat{Y}_i\right)\big/\hat{\sigma}_i \tag{5-15}$$

where $\hat{\sigma}_i^2$ is an estimator of $Var[Y_i]$. In this case

$$Y_i=S^{-1}\left(S_i;\hat{Y}\right)=\hat{Y}_i+\hat{\sigma}_i S_i \tag{5-16}$$

Now suppose that the S_i are iid. In fact, the residuals from a regression will be dependent, and so the requirement is actually that the S_i be approximately iid. The requirement of identical distribution is an essential one, as will be explained further below, and the most egregious results can be obtained if it is violated.

Now draw a random n-sample from $S(Y; \hat{Y})$. The sampling can be without replacement (in which case the sample will be simply a permutation of Y), or with replacement. Let the members of the sample be denoted $\tilde{S}_i$, $i = 1, \ldots, n$, and arrange these in a vector denoted $\tilde{S}$. This is the process of **data re-sampling** referred to earlier.

Form the vector $\tilde{Y}$ with i-th component

$$\tilde{Y}_i=S^{-1}\left(\tilde{S}_i;\hat{Y}\right) \tag{5-17}$$

and let $\tilde{Y}$ denote the vector with components $\tilde{Y}_i$, ordered in the same way as the $\tilde{S}_i$ in S.

Since the S_i were iid, S and $\tilde{S}$ have the same stochastic properties, and then, by (5-16) and (5-17), Y and $\tilde{Y}$ have the same stochastic properties. That is, $\tilde{Y}$ may be viewed as an alternative data set with the same stochastic properties as the original one. It is in fact called a **pseudo-data set**.

In the case in which the residuals S_i are Pearson residuals (see (5-15) and (5-16)), the construction of the pseudo-data (5-17) takes the form

$$\tilde{Y}_i = S^{-1}\left(\tilde{S}_i; \hat{Y}\right) = \hat{Y}_i + \hat{\sigma}_i \tilde{S}_i \tag{5-18}$$

It is possible to draw many pseudo-data sets. The number of possibilities is $n!$ if sampling without replacement is used, and n^n if with replacement. These are very large numbers even for n of moderate size.

So draw some large number r of pseudo-data sets, denoted $\tilde{Y}_{(1)}, \tilde{Y}_{(2)}, \ldots, \tilde{Y}_{(r)}$, and model each of them, using precisely the same model as was applied to Y originally. Here "precisely the same model" means having precisely the same algebraic structure. Obviously, the parameters will change as the data inputs change. Call the model $\mathcal{M}$.

For each pseudo-data set, form the same forecasts as for the original data set. Thus, let $\hat{\beta}_{(j)}$ denote the vector of parameter estimates (**"pseudo-estimates"**) associated with the pseudo-data set $\tilde{Y}_{(j)}$, and let $\hat{\tilde{Y}}^*_{(j)}$ denote the forecast of Y^* using the j-th pseudo-data set, and let $R(\hat{\tilde{Y}}^*_{(j)})$ denote the associated forecast of the target $R(Y^*)$. This is a **pseudo-forecast** of $R(Y^*)$, and there are now r pseudo-forecasts $R(\hat{\tilde{Y}}^*_{(j)})$, $j = 1, \ldots, r$.

The set of pseudo-forecasts has the same stochastic properties as an r-sample of forecasts of $R(Y^*)$, obtained by application of model $\mathcal{M}$ to an r-sample of data sets. The variation between the pseudo-forecasts reflects parameter error introduced in Section 4.1, the error arising from the fact that the application of the same model to randomly varying data sets produces variation in the model parameter estimates.

As was also noted in Section 4.1, forecast error also needs to take account of the process error, or noise, contained in $R(Y^*)$ (see (4-6)). This may also be achieved by re-sampling, as follows.

Let the process error associated with the i-th component of Y^* be denoted

$$\varepsilon^*_i = Y^*_i - E\left[Y^*_i\right] \tag{5-19}$$

or, equivalently,

$$Y^*_i = E\left[Y^*_i\right] + \varepsilon^*_i \tag{5-20}$$

Now, in the j-th replication (also referred to as a **replicate**) $E[Y^*_i]$ is estimated by the i-th component of $\hat{\tilde{Y}}^*_{(j)}$. To obtain a set of random drawings with the same properties as the collection $\{\varepsilon^*_i\}$, draw a second vector $\tilde{S}_{proc}$ in the same way as $\tilde{S}$ was drawn, form the pseudo-observation vector $\tilde{Y}_{proc}$ in parallel with (5-17), and then define the vector

$$\varepsilon^*_{proc} = \hat{Y}_{proc} - \hat{Y} \tag{5-21}$$

The components of ε^*_{proc} then have the same properties as the collection $\{\varepsilon^*_i\}$. The procedure can be repeated to obtain r replicates $\varepsilon^*_{proc(j)}$ of ε^*_{proc}.

In the case of Pearson residuals, (5-17) is specialized to (5-18) in this process, and (5-21) simplifies to

$$\varepsilon^*_{proc,i} = \hat{\sigma}_i \tilde{S}_{proc,i} \tag{5-22}$$

where $\varepsilon^*_{proc,i}$ and $\tilde{S}_{proc,i}$ are the i-th components of ε^*_{proc} and $\tilde{S}_{proc}$ respectively.

Replace $E[Y^*_i]$ and ε^*_i in (5-20) by the estimators just formed to define

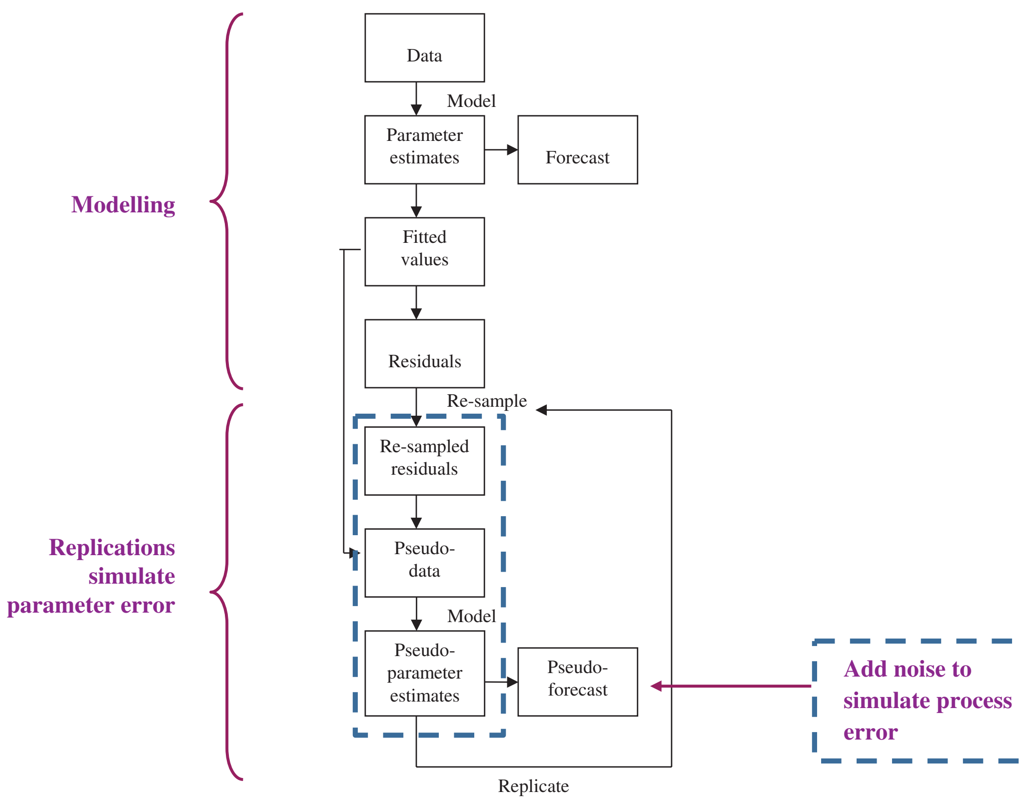
$$\left(\hat{\tilde{Y}}^*_{(j)}\right)^+ = \hat{\tilde{Y}}^*_{(j)} + \varepsilon^*_{proc(j)} \tag{5-23}$$

whereupon $(\hat{\tilde{Y}}^*_{(j)})^+$ becomes a pseudo-forecast, augmented to include process error. Pseudo-forecasts of $R(Y^*)$, also including process error, can now be obtained as simply $R((\hat{\tilde{Y}}^*_{(j)})^+), j = 1, \ldots, r$.

These are iid drawings with the same distribution as $R(Y^*)$, and so the r replicates form an empirical distribution of $R(Y^*)$. Any stochastic property of $R(Y^*)$, e.g., MSEP, may then be estimated from the distribution.

The bootstrap process just described may be represented diagrammatically as in Figure 5-1. The dashed rectangles are marked for discussion in Section 5.3.2.

Figure 5-1. Diagrammatic Representation of the Semi-Parametric Bootstrap



The version of the bootstrap just described is called semi-parametric here and in Shibata 1997 (though elsewhere in the actuarial literature, it is often referred to as non-parametric bootstrapping) because the generation of the pseudo-data sets by means of the re-sampling procedure (5-17) or (5-18) makes no distributional assumption. However, it does rely on a fitted model from which to calculate predicted values and residuals. The distribution of the pseudo-data $\tilde{Y}_{(j)}$ is determined entirely by that of the residuals *S*. Similarly in the addition of process error in (5-23).

By contrast, the non-parametric bootstrap (terminology as per Shibata, 1997) does not require a fitted model prior to resampling. It simply generates a large number of pseudo-samples by repeatedly sampling the observed data with replacement. Clearly this is inappropriate for aggregated insurance loss data where the magnitude differs from one development period to the next. The use of the term "semi-parametric" for the residual resampling approach may be helpful to distinguish the two types of bootstrap, which were both proposed in Efron (1979).

It is evident from the re-sampling basis of the bootstrap that the exclusion of any outlying observations, as discussed in Section 2.2.5, will have ramifications not only for model parameter estimation (as remarked in that sub-section) but will also reduce any bootstrap estimate of dispersion. Once again, one would need to consider whether adjustment of that estimated dispersion might be required. Such adjustments are beyond the scope of this volume.

5.3.2. Parametric Bootstrap

Parametric bootstrapping as defined in Shibata (1997) is functionally very similar to the semi-parametric method described above, but based on theoretical rather than empirical residuals. Thus for models such as GLMs, where the standardized deviance residuals are asymptotically normal, resampling of the actual residuals may be replaced by sampling from a normal distribution with the appropriate variance.

There are other possible ways to make use of the GLM assumptions to generate a distribution of reserves, including the approach described below which simplifies the area of Figure 5-1 in the dotted box, in which replicates of parameter estimates are obtained, and also simplifies the generation of process error. With some abuse of terminology, this is also referred to as parametric bootstrapping in this monograph.

Parameter Estimates

It is supposed that the original parameter estimates $\hat{\beta}$ (the second box in the figure) are MLEs, as is usually the case for GLMs. It is known that an MLE is an asymptotically normal unbiased estimator for indefinitely increasing sample size in the presence of some technical conditions (Cox and Hinckley, 1974). In symbolic terms,

$$\hat{\beta} \sim N\left(\beta, Var\left[\hat{\beta}\right]\right) \text{ asymptotically} \tag{5-24}$$

If this asymptotic relation is assumed to hold precisely for the finite data sample under consideration, then one may assume that

$$\hat{\beta} \sim N\left(\beta, \hat{C}\right) \tag{5-25}$$

where $Var[\hat{\beta}]$ has been denoted by C, and $\hat{C}$ denotes the estimate of C provided by the GLM software (as already mentioned just prior to (5-9)). The parameter estimate replicates $\hat{\beta}_{(j)}$ may then be sampled from the multi-normal $N(\hat{\beta}, \hat{C})$.

The sampling requires care in view of the correlations contained in $\hat{C}$. The usual sampling process consists of the following steps:

- apply a linear transformation M to $\hat{\beta}$ such that the components of $M\hat{\beta}$ are uncorrelated;
- sample the each of these components from a univariate normal distribution to obtain a random vector γ;
- apply the inversion of M to the sampled vector γ to obtain the required sampling from $N(\hat{\beta}, \hat{C})$.

In mathematical terms, find M such that $Var[M\hat{\beta}] = \Lambda$, diagonal, i.e.,

$$M\hat{C}M^T = \Lambda = diag\left(\lambda_1, \ldots, \lambda_p\right) \tag{5-26}$$

Now make random drawings

$$\gamma_i \sim N\left(\left(M\hat{\beta}\right)_i, \lambda_i\right), i = 1, 2, \ldots, p \tag{5-27}$$

where $(M\hat{\beta})_i$ denotes the i-th component of $M\hat{\beta}$.

Finally, construct replicates of $\hat{\beta}_{(j)}$ as

$$\hat{\beta}_{(j)} = M^{-1}\gamma \tag{5-28}$$

where $\gamma = (\gamma_1, \ldots, \gamma_p)^T$.

To check that $\hat{\beta}_{(j)} \sim N(\hat{\beta}, \hat{C})$, note that

$$E\left[\hat{\beta}_{(j)}\right] = M^{-1}E[\gamma] = M^{-1}M\hat{\beta} = \hat{\beta} \tag{5-29}$$

and

$$Var\left[\hat{\beta}_{(j)}\right] = M^{-1}Var[\gamma]\left(M^{-1}\right)^T = M^{-1}\left[M\hat{C}M^T\right]\left(M^{-1}\right)^T = \hat{C} \tag{5-30}$$

Central to the above sampling process is the identification of the required matrix M in (5-26). this may be achieved by either **Cholesky decomposition** or **spectral decomposition** of $\hat{C}$, both of which will be available from conventional statistical software.

Cholesky decomposition expresses $\hat{C}$ in the form

$$\hat{C} = LL^T \tag{5-31}$$

with L a lower triangular matrix. This is equivalent to (5-26) with $M = L^{-1}$ and $\Lambda = I$.

Spectral decomposition expresses $\hat{C}$ in the form

$$\hat{C} = P\Lambda P^T \tag{5-32}$$

with P an orthogonal matrix and $\gamma_1, \ldots, \gamma_p$ the eigenvalues of $\hat{C}$. This is equivalent to (5-26) with $M = P^{-1} = P^T$.

Process Error

The addition of process error is indicated in the bottom right box of Figure 5-1 and is described in (5-21) to (5-23). Now $\tilde{Y}_{proc}$ in (5-21) is a replicate of Y, which the GLM will have assumed subject to some particular distribution. Hence $\tilde{Y}_{proc}$ may be obtained simply as a random drawing from that distribution.

For example, if the assumed distribution of Y_i is ODP, the i-th component of $\tilde{Y}_{proc}$ may be obtained as a random drawing from a ODP distribution with mean $\tilde{Y}_i$ and scale parameter $\hat{\phi}/w_i$, where this last quantity is the GLM's estimate of (2-21).

Discussion

The parametric version of the bootstrap is so called because it makes use of assumed parametric forms: the normal distribution for parameter error, and the GLM's chosen distributional form for process error.

Its implementation is somewhat simpler than that of the semi-parametric form with shorter computational times, considerably so for larger data sets. Evidently, however, its validity is dependent on the assumptions just stated, and will become more dubious as:

- the sample size n declines to the point where reliance cannot be placed upon the asymptotic result (5-24); and/or
- the error structure assumed within the GLM becomes a poor representation of the data.

The commentary at the end of Section 5.3.1 on the exclusion of isolated observations from the bootstrap applies equally to the parametric bootstrap.

5.4. Numerical Examples

5.4.1. Delta Method

Table 3-2 obtained the chain ladder forecasts associated with the data triangle of Table 1-1. These were obtained by application of the conventional chain ladder, but it was noted in Section 3.2 that the ODP cross-classified model produces the same forecasts.

The delta method, as described in Section 5.2.3, is now applied to estimate the forecast error associated with the ODP cross-classified model forecasts. Note that, although the

ODP Mack and ODP cross-classified models produce the same forecasts, they are different models and do not produce the same estimates of forecast error.

The forecast error required here is estimated by application of (5-11) to (5-13), where it was noted in Section 5.2.3. that values of $Var[\hat{\beta}]$ and $\hat{\phi}$ are provided by the GLM software. These formulas required the evaluation of D, defined by (5-9). It is noted that, for the ODP cross-classified model, the link function is $h = ln$, and so $(h^{-1})'(h(.)) = identity$. Thus, (5-9) simplifies to

$$D = diag\left(\hat{\mu}^*_{2,10}, \hat{\mu}^*_{3,9}, \ldots, \hat{\mu}^*_{10,10}\right) \tag{5-33}$$

The results are displayed in Table 5-4. The table contains the **root mean square errors of prediction** (**"RMSEP"**) and **coefficient of variation of prediction** (**"CVP"**). The first of these is simply the square root of the MSEP, and the second is defined as

$$CVP = \frac{RMSEP}{Forecast}$$

5.4.2. Bootstrap

The parametric bootstrap, as described in Section 5.3.2, has been applied to estimate the forecast error associated with the ODP cross-classified model forecasts.

Table 5-4. Chain Ladder Forecast Error

Accident Year	Outstanding Losses		
	Forecast	RMSEP	CVP
	$000	$000	%
1989	3,398	924	27.2
1990	8,155	1,363	16.7
1991	14,579	1,775	12.2
1992	22,645	2,169	9.6
1993	31,865	2,523	7.9
1994	45,753	3,036	6.6
1995	60,093	3,577	6.0
1996	80,983	4,538	5.6
1997	105,874	6,786	6.4
Total	373,346	14,076	3.8

It may be noted that the table reveals positive correlation between (at least some) accident years. If accident years were independent, then the MSEP of the total forecast would be simply the sum of the accident year MSEPs, and the RMSEP of the total forecast would be 10,275 ($000), substantially less than the actual result of 14,076. The difference is accounted for by positive correlation.

Table 5-5. Parametric Bootstrap Estimates of Chain Ladder Forecast Error

Accident Year	Outstanding Losses		
	Forecast	RMSEP	CVP
	$000	$000	%
1989	3,476	937	27.0
1990	8,269	1,366	16.5
1991	14,738	1,794	12.2
1992	22,776	2,186	9.6
1993	32,043	2,525	7.9
1994	45,963	3,057	6.7
1995	60,273	3,608	6.0
1996	81,249	4,589	5.6
1997	106,204	6,831	6.4
Total	374,992	14,286	3.8

The information required for this consists of that in Table 5-1, together with the GLM estimate of the scale parameter, which is $\phi = 114.5$.

The results of 10,000 bootstrap replications are contained in Table 5-5, in which:

- "Forecast" is taken as the arithmetic mean of the 10,000 replicates of the forecast; and
- "RMSEP" is taken as the square root of the unbiased variance of these 10,000 replicates.

The results are evidently very similar to those obtained by the delta method in Table 5-4. The forecasts are slightly different, which can be accounted for by sampling error arising from the limited number of replicates.

6. Model Validation

Chapter summary. Model validation consists of detailed checking that a fitted model is compatible with, and accounts for all features of, the data. There are many diagnostic plots available for this purpose. The present chapter discusses and illustrates a number of these.

Illustration is carried out first in the abstract, and then in relation to a simulated data set, and finally in relation to the actual data set given in Chapter 1 and used in numerical examples throughout this volume. In the case of simulated data, the target model is known, and so its effect of specific model features on some of the diagnostic plots can be clearly illustrated.

6.1. Introduction

Model validation is the process of examining whether the fitted model—both the distributional assumptions and the fitted parameter effects—are acceptable and adequate descriptions of the data being modeled. It is a critical part of building any model—if the assumptions underlying the model are found to be flawed, this then casts doubt on any inferences from that model.

Typically there are three aspects to a model validation:

- Analysis of the distributional assumptions;
- Analysis of the goodness-of-fit of the model; and
- Analysis of the model's predictive performance on data beyond those used in the model estimation.

Of these the third is not usually possible for claims reserving models based on simple triangles (i.e., other than individual claim models, also known as micro-models or granular models), since all the data would normally be used to build the model. Thus, out-of-sample testing is not discussed further here.

In principle, the model validation would begin by validating the choice of distribution and the link function. Of these, the link function is usually determined by the model structure as being that transformation that produces a linear predictor. For example, a multiplicative model implies a log link while an additive model uses an identity link. In terms of model validation, a link function is acceptable if the model passes the other validation tests without requiring an excessive number of interaction terms. Once the link and the distribution have been validated, the user can move onto examine the goodness-of-fit of the model.

One's view of the error distribution is provided by the observed residuals, which depend in turn on the fitted model. No view of the distribution can become available until some model, at least a rudimentary one, has been fitted to the data. Thus the respective forms of the error distribution and model are inter-dependent, and cannot simply be selected sequentially.

In our experience, the estimated mean of the distribution is relatively insensitive to the choice of distribution, and similar findings are reported by Lai and Shih (2003), though, of course, the same is not true of the variance. Thus, our approach to model fitting and validation is generally to select a reasonable set of distribution assumptions using common sense arguments, fit the model and test for goodness of fit, before validating the model distribution assumptions carefully. In more detail, a step-by-step description of this process is as follows:

- Select the appropriate link function (e.g., a multiplicative model implies a log link);
- Select a reasonable distribution—e.g., ODP for a cross-classified model;
- Fit the main effects in the model and any obvious interactions (see Section 7.6);
- Check the residual diagnostics for any gross violations of the distributional assumptions and make changes if necessary;
- Continue with the model fitting using goodness-of-fit tests (primarily comparisons of actual and model fitted values) until a satisfactory goodness-of-fit of cell means is obtained. This may involve the use of interactions in the model;
- Review the distributional diagnostics in detail and make any adjustments required to yield satisfactory results. After any changes, re-check the goodness-of-fit and make changes if necessary. Repeat until a satisfactory model is obtained.

The assessment of the goodness-of-fit and the distributional assumptions is covered in detail below. In practice, the tools used in this assessment are usually graphical, and definitions and examples of all the various graphical tools used are provided.

Following that, some examples of the graphs are given in cases of poor fit and good fit. To facilitate this discussion, simulated data sets are used so that the true underlying model is known with certainty. Finally, model validation will be carried out for the cross-classified model using the example data set.

6.2. Summary of Assumptions and Tests

Before commencing the definition and use of the various model diagnostics later in this chapter, we have gathered together the list of model assumptions and corresponding diagnostics that will be discussed below. This is intended as a reference list that modelers may use to check the fit of their model.

Distributional Assumptions

- The link structure is appropriate:
    - Expectations regarding the modeled quantity will largely determine the choice of link—e.g., a multiplicative model structure requires the use of a log link. It is validated if the model passes the other diagnostics tests without requiring an undue number of interactions.

- The distribution choice is appropriate:
  - Probability-Probability (P-P) plot;
  - Residual plots by accident, development and calendar year periods;
  - Histograms or kernel density plots of the residuals.

Goodness-of-Fit

- The model fits well by accident, development and calendar periods:
  - Plots by accident, development and calendar periods of actual and expected (i.e., the expected value according to the fitted model) in some form, e.g.:
    - actual vs. expected;
    - log(actual) vs. log(expected);
    - Actual/expected;
  - Plots of residuals, also by accident, development and calendar periods;
- All significant interactions have been identified:
  - A triangular (e.g., 2-d) heat map of actual/expected;
  - Actual and expected plots for specific parts of the experience.

6.3. Diagnostic Graphs

All diagnostics graphs involve the comparison of actual and expected quantities, where "expected" is an abbreviation for "expected value according to the fitted model".

The most well-known comparison is that based on residuals but other comparisons such as the quotient of the actual and expected values or plots of actual and expected values are also useful. In more detail, the functions of actual and expected values used are:

- Pearson residuals—both raw and standardized. Refer to Section 2.2.4 for their definition;
- Deviance residuals—both raw and standardized. Refer to Section 2.2.4 for their definition;
- Actual values including sums of actual values across rows ($\sum^{\mathcal{R}(k)}$), columns ($\sum^{\mathcal{C}(j)}$) and diagonals ($\sum^{\mathcal{P}(\mathrm{k+j-1})}$). Depending on the scale of the comparison, the logs of these quantities may be more useful;
- Expected values including sums of expected values across rows, columns and diagonals (denoted by $\sum^{\hat{\mathcal{R}}(k)}$, $\sum^{\hat{\mathcal{C}}(j)}$ and $\sum^{\hat{\mathcal{P}}(\mathrm{k+j-1})}$ respectively). Again, the logs of these quantities may be useful for many reserving problems;
- Actual/expected values in each cell of the triangle—for example $Y_{kj}/\hat{Y}_{kj}$; and
- Actual/expected marginal values by row, column and diagonal. For example, the marginal actual/expected comparison for accident period k is $\sum^{\mathcal{R}(k)}/\sum^{\hat{\mathcal{R}}(k)}$.

Following from the discussion of Pearson and deviance residuals in Section 2.2.4, only deviance residuals will be used in this chapter due to their greater degree of normality when the underlying distribution (Poisson in this case) is not normal. All comments below which discuss normality and homo- and heteroscedasticity of residuals refer to standardized deviance residuals.

Based on these quantities, a number of diagnostic graphs are available to the user to carry out model validation. These graphs are discussed below in Sections 6.3.1 to 6.3.7.

Note that in all of the examples in these sections, the plots are drawn using a correctly specified model of simulated data so that the graphs indicate a well-fitting model.

6.3.1. Scatterplot

A scatterplot of residuals is a simple graph plotting residuals against a relevant variable such as the expected value, accident period, development period or calendar period. Figure 6-1 gives an example of a scatterplot where standardized deviance residuals are plotted against development period.

Departures from a random, homoscedastic plot of deviance residuals suggests problems with the model. A trend in the residuals indicates possible goodness-of-fit issues while heteroscedasticity (e.g., fanning of residuals) often indicates that the dispersion assumptions are inappropriate. As noted above, the example here is taken from a correctly specified model leading to homoscedastic residuals.

6.3.2. Spread Plot

This plot shows some summary statistics of the residuals plotted against a variable of interest (e.g., development period, expected value) to provide the modeler with information on the spread and distribution of the statistics. Specifically, the 25th and 75th percentiles are plotted along with the standard deviation of the residuals. The spread plot is particularly useful for detecting heteroscedasticity of deviance residuals as heteroscedasticity is indicated by widening or narrowing of the inter-quartile range and by significant changes in the standard deviation.

The spread plot corresponding to Figure 6-1 is shown in Figure 6-2 below. Looking past the volatility (particularly in the higher development periods), the interquartile range is reasonably consistent while the standard deviation fluctuates around unity.

Figure 6-1. Scatterplot of Standardized Residuals

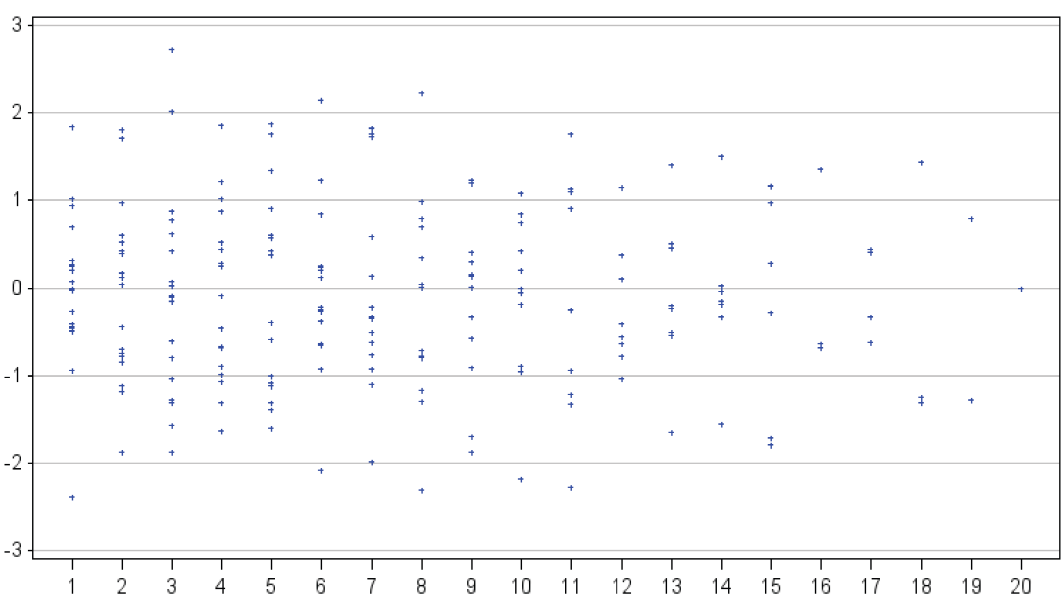
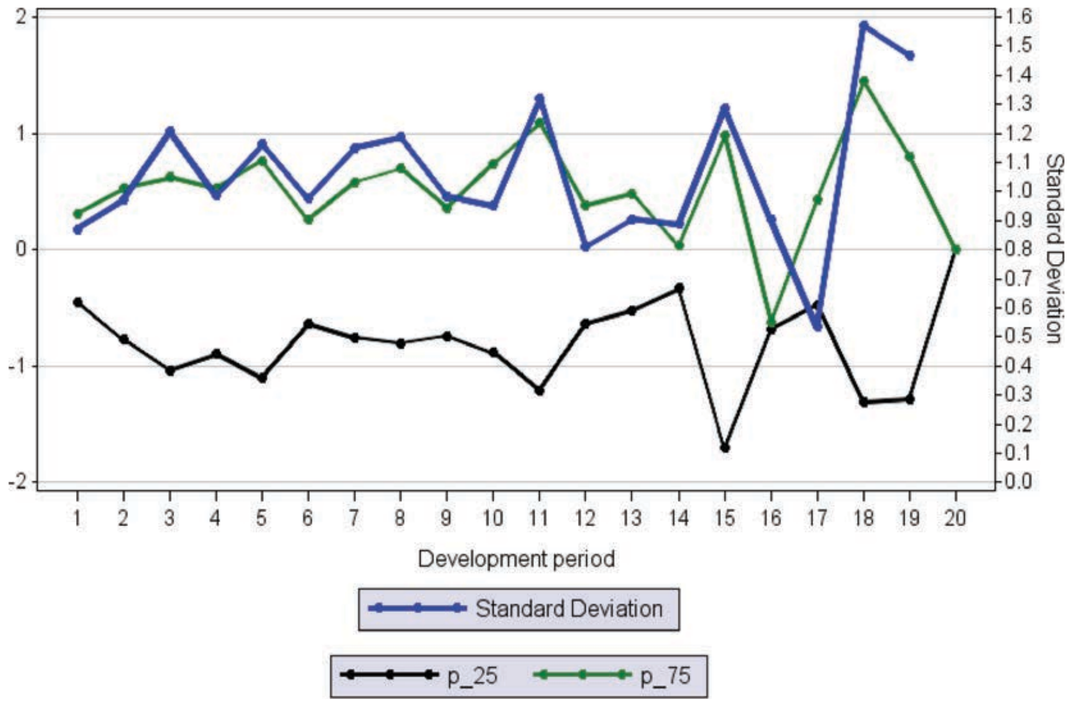


Figure 6-2. Spread Plot



Note that in spread plots, the green and black lines plot the 25th and 75th percentiles while the blue line is the standard deviation of the residuals. If standardized residuals are used, as is the case here, then the standard deviation of these residuals should vary randomly about unity and any systematic departures from this may indicate a problem with the model assumptions.

6.3.3. Actual and Expected Comparison Plots

Actual and expected comparison plots display the actual and expected totals (e.g., by accident, development or calendar period). For example, such a plot by row or accident period shows the actual series ($\Sigma^{\mathcal{R}(k)}$) and the expected series ($\Sigma^{\hat{\mathcal{R}}(k)}$) plotted for $1 \le k \le K$. Areas of poor fit correspond to consistent differences in the actual and expected values. Figure 6-3 is an example of an acceptable graph where the expected values are close to the actual values.

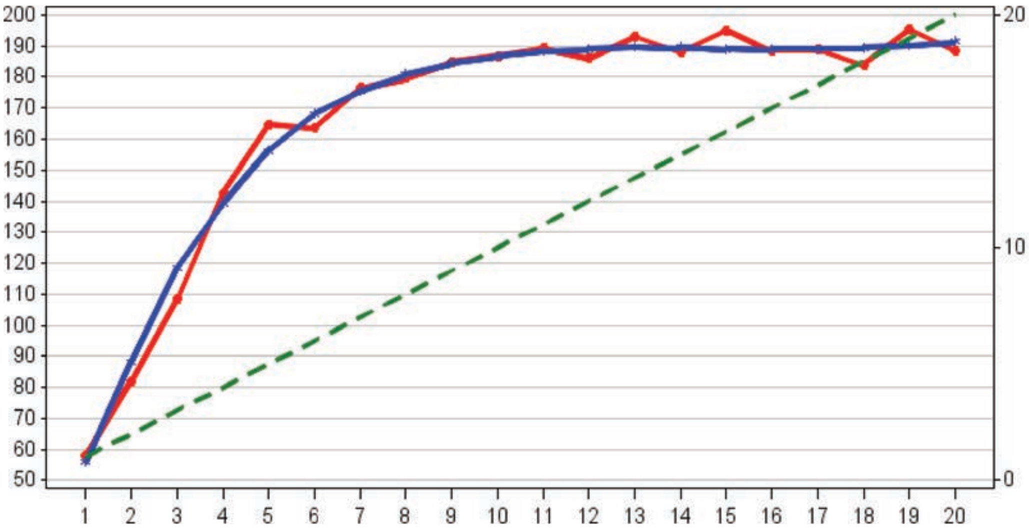
Depending on the scale of the data, it may be more helpful to log the quantities, i.e., log(actual) vs log(expected).

6.3.4. Actual and Expected Ratio Plots

These plots are similar to those in 6.3.3 except that they plot the actual/expected ratio rather than individual actual and expected lines. Systematic deviations away from 100% indicate regions of poor fit.

Figure 6-4 is the ratio plot equivalent to the comparison plot shown in Figure 6-3. Following some volatility in early calendar periods (when there is little data), the ratios fluctuate randomly around 100% indicating an adequate fit.

Figure 6-3. Actual and Expected Comparison Plot by Diagonal (calendar period)



Note that in actual and expected comparison plots, the red line shows the actual totals, the blue line the expected totals while the green dotted line (right-hand scale) shows the cumulative number of data points at each level.

6.3.5. Actual and Expected Ratio 2-D Heat Map

This diagnostic is particularly useful in the loss reserving context where it can be used to look at the goodness-of-fit across a data triangle (or other 2–dimensional array). Specifically, it calculates the actual/expected ratio in each cell of the triangle and applies a formatting conditional on the deviation of the ratio from 100%. In the example in

Figure 6-4. Actual and Expected Ratio Plot by Diagonal (calendar period)

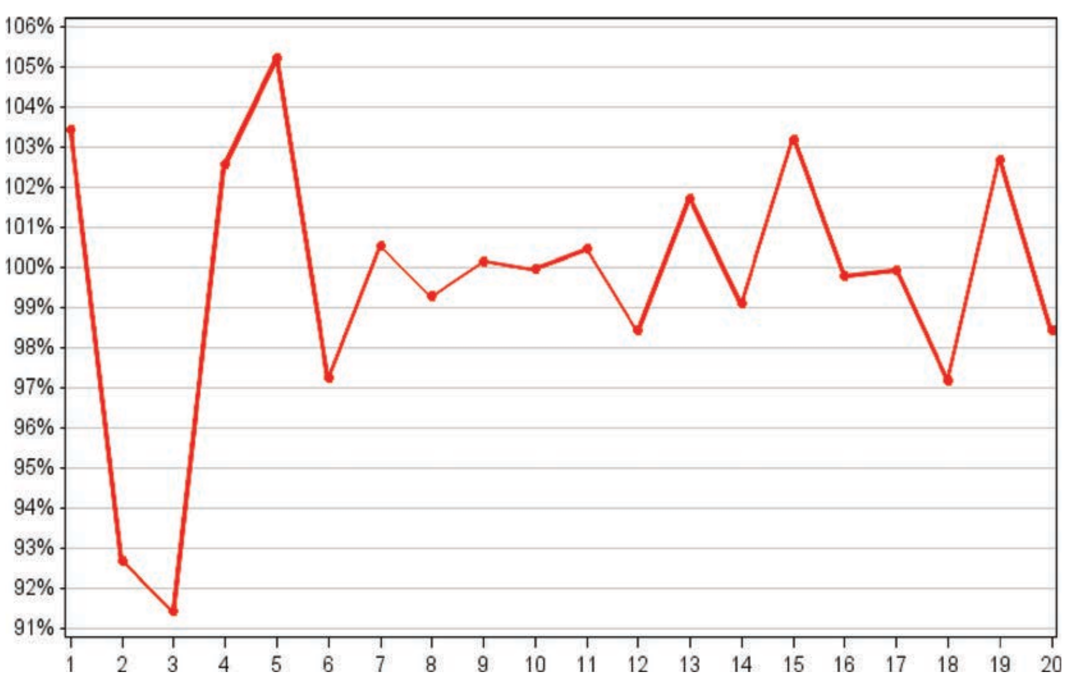


Figure 6-5. Actual and Expected Ratio Heat Map

103%	86%	89%	105%	110%	98%	110%	100%	102%	76%	100%	[illegible]	98%	[illegible]	105%	95%	102%	93%	101%	100%
107%	95%	99%	101%	95%	101%	103%	102%	101%	104%	114%	93%	78%	92%	107%	105%	108%	101%	99%	
91%	106%	108%	97%	99%	100%	107%	93%	101%	103%	96%	93%	[illegible]	90%	91%	95%	90%	106%		
99%	95%	103%	104%	[illegible]	88%	97%	103%	89%	103%	99%	98%	110%	102%	93%	99%	101%			
[illegible]	95%	87%	105%	92%	100%	101%	104%	103%	104%	101%	97%	[illegible]	87%	123%	105%				
94%	103%	[illegible]	98%	[illegible]	106%	93%	107%	95%	[illegible]	88%	101%	102%	109%	[illegible]					
104%	89%	110%	92%	97%	101%	93%	105%	105%	104%	105%	105%	89%	107%						
97%	106%	104%	92%	95%	101%	101%	109%	96%	97%	104%	102%	97%							
101%	111%	103%	101%	104%	99%	104%	95%	95%	87%	100%	97%								
100%	102%	96%	99%	96%	105%	109%	91%	101%	106%	96%									
97%	95%	108%	104%	96%	103%	100%	92%	99%	102%										
88%	104%	88%	100%	103%	96%	95%	111%	110%											
105%	111%	105%	95%	104%	98%	96%	92%												
98%	103%	[illegible]	97%	92%	103%	96%													
101%	89%	103%	105%	101%	98%														
87%	95%	100%	109%	98%															
124%	97%	99%	95%																
94%	107%	97%																	
98%	101%																		
100%																			

Figure 6-5 pink values indicate ratios larger than 100% and blue ratios less than 100%. The more intense the color, the greater the deviation from 100%.

The distribution of colors should be random across the triangle. Clumps of one color indicate areas of poor fit. For example, if the model includes terms for accident and development period effects, then a clumping of colors may indicate the need for further model terms such as interactions between accident and development periods or terms involving calendar periods.

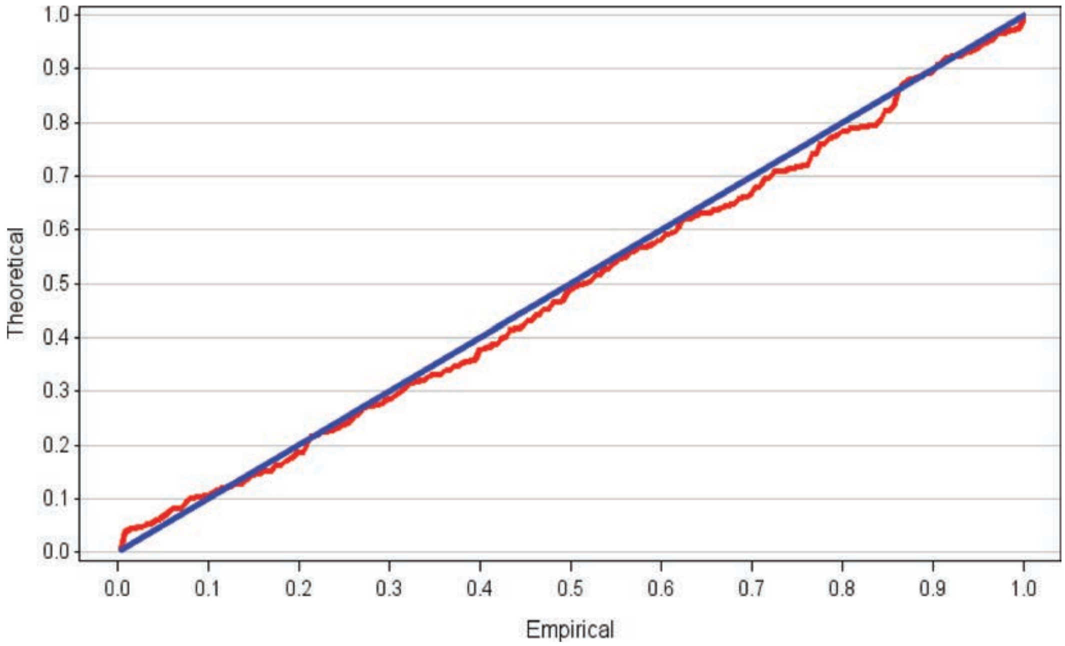
The heat map in Figure 6-5 suggests that the model is not missing interaction or calendar period terms since the blue and pink colors are randomly distributed.

6.3.6. Probability-Probability Plot

A Probability-Probability plot (also known as a **"P-P"** plot or a percent-percent plot) is a graphical method for comparing two probability distributions. A P-P plot plots two cumulative distribution functions ("cdfs") against each other. Given an input u, the plotted points are $(F(u), G(u))$ where F and G represent the cdfs of two probability distributions. Thus, a P-P plot is a parametric graph, whose range is the unit square $[0,1] \times [0,1]$. Each pair of numbers represents the probability of being $\leq u$ under the distributions F and G respectively.

In a GLM application, one distribution will correspond to the selected error distribution (e.g., ODP as discussed in this monograph), referred to as the "theoretical" distribution while the other will correspond to the modelled data (the "empirical" distribution). If the model fits the data well, then the empirical and theoretical distributions should be similar and the resulting P-P plots should be an approximately straight line

Figure 6-6. Probability-Probability Plot



Note that in P-P plots, the blue line is the plot that would be obtained if the actual distribution exactly matched the assumed distribution. The red line is the plot of the assumed theoretical quantiles against the empirical quantiles.

of the form y = x (see Figure 6-6). Pronounced or persistent deviations from a straight line indicate problems with the distributional assumptions.

For the model discussed in this monograph, each observation Y_{kj} is assumed subject to an ODP with mean $\hat{Y}_{kj}$ and variance $\phi_{kj}\hat{Y}_{kj}$. The value of the cdf of this "theoretical" distribution is computed at Y_{kj}. Call it u_{kj}. The empirical distribution, $\hat{u}_{kj}$, may be obtained by sorting by ascending u with $\hat{u}_{kj}$ being the proportion of data points $\leq u_{kj}$. In effect, the empirical readings are simply n equally spaced points in [0,1] where n is the number of observations in the data set.

A related and perhaps better known plot is the Quantile-Quantile ("Q-Q") plot, which plots the quantiles of two distributions against each other. In more detail, the inverse function of a cumulative probability function is the quantile function, i.e., given a cdf F, its quantile function is F^{-1}. Thus, given two cdfs F and G, with associated quantile functions F^{-1} and G^{-1}, a Q-Q plot draws the q^{th} quantile of F against the q^{th} quantile of G for a range of values of q. Thus, the Q-Q plot is a parametric curve indexed over [0,1] with values in the real plane R^2.

The Q-Q plot requires that all observations appearing within it be drawn from the same distribution. This will not usually be the case for the raw observations modeled by a GLM, where the mean may vary from one observation to another. However, a Q-Q plot may be applied to the standardized deviance residuals, which are asymptotically N(0,1). In this case the ordered standardized deviance residuals are plotted against the quantiles of the standard normal distribution. Augustin, Sauleau and Wood (2012) provide some further discussion on the use of Q-Q plots as GLM diagnostics.

6.3.7. Histogram of Residuals

Finally, a simple histogram of standardized deviance residuals is a further useful check on the distributional assumptions—if the model is appropriate, then these residuals should be approximately standard normal, as in Figure 6-7 where magnitude of standardized residuals is represented on the horizontal axis and frequency of their occurrence on the vertical.

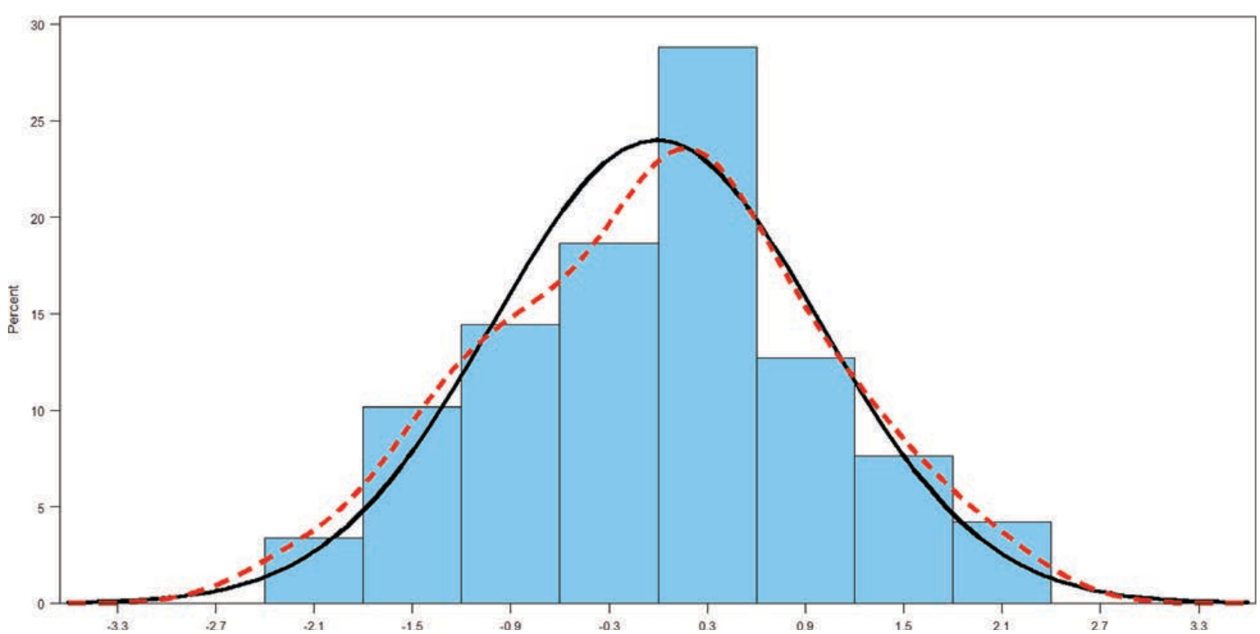
6.4. Simulated Data Set and Fitted Models

Three simulated data sets were generated to illustrate the use of the various model diagnostics in model validation. They are described in Table 6-1. Note that the accident and development period effects used to simulate the data are specified from the formulae given in the table below.

In summary, all three simulated data sets are Poisson distributed. Simulated data set 1 has accident and development period effects only and a constant scale so may be correctly described by a cross-classified model. The second data set is similar to the first except that its scale parameter varies by development period. Thus, a cross-classified model with suitably selected weights is appropriate. Finally the third data set has development effects that vary according to accident period. Thus the cross-classified model cannot adequately model this dataset since it will not capture the interaction between accident and development effects.

A number of different models were fitted, all GLMs of the form $Y_{kj} \sim ODP(\mu_{kj}, \phi_{kj})$. The models differ in the specifics of the definitions of μ_{kj} and ϕ_{kj}, which are given in Table 6-2, together with the data sets to which they were applied.

Figure 6-7. Histogram



Note: the solid line overlay is a normal distribution, fitted using the method of moments, while the dotted line is a kernel density estimator, which may be helpful for small data sets such as those that typically result from reserve estimation using aggregate triangle data.

Table 6-1. Description of Simulated Data

Characteristic	Simulated Data 1	Simulated Data 2	Simulated Data 3
Accident periods	20	20	20
Development periods	20	20	20
Development period effect j=1 to 20	$\beta_j = exp(-0.3[j-1] + 1.5ln[j])$	$\beta_j = \exp(-0.3[j-1] + 1.5ln[j])$	$\beta_j = exp(-0.3[j-1] + 1.5ln[j])$ for k=1 to 10 and $\beta_j = exp(-0.5[j-1] + 2ln[j])$ for k=11 to 20
Accident period effect, k=1 to 20	$\alpha_k = exp(0.05k + 4)$	$\alpha_k = exp(0.05k + 4)$	$\alpha_k = exp(0.05k + 4)$
Scale parameter	1	$min(8, j+1)^2$	1
Distribution	Poisson	Over-dispersed Poisson	Poisson

6.5. Analysis of the Goodness-of-Fit

This aspect of model validation examines the data to ensure that all significant drivers of the target value have been identified. In claims reserving, this corresponds to reviewing the diagnostics by accident, development and calendar period to see if there are any un-modeled trends in the data.

In other words, the model is examined for the quality of fit to the data of its cell expected values. Dispersion and distributional questions will be considered in Section 6.6.

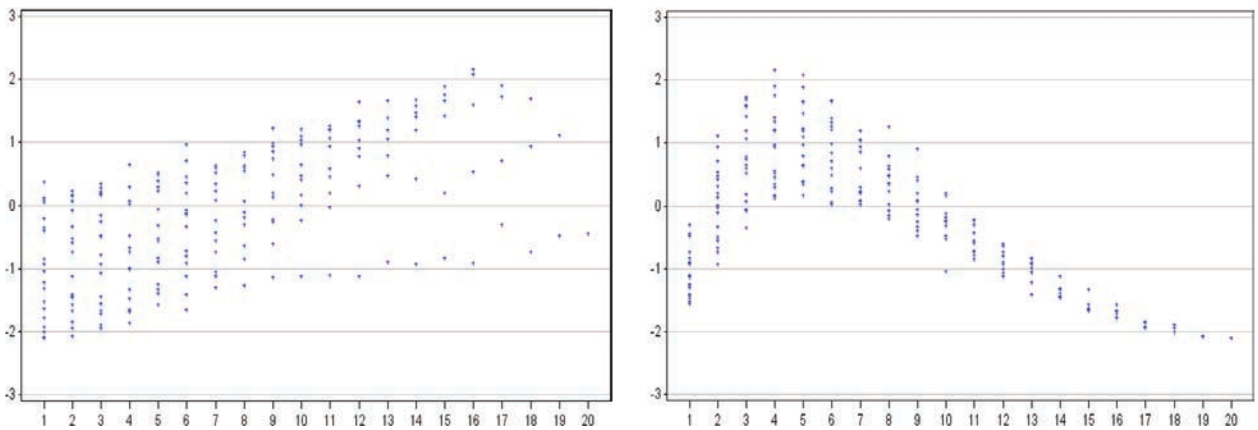
Traditionally this would be carried out by examining the residuals (refer back to Section 2.2.4 for the definition and discussion of Pearson and deviance residuals) for evidence of non-randomness. To illustrate this, the Mean model is fitted to simulated data 1. This model fits a single average to all data points, thereby ignoring the accident

Table 6-2. Models Fitted to Simulated Data

Model Name	Model Description	Simulated Data Set 1	2	3
Mean	$\mu_{kj} = exp(\mu)$ $\phi_{kj} = 1$	Y		
Development	$\mu_{kj} = exp(ln\ \beta_j)$ $\phi_{kj} = 1$	Y		
Full	$\mu_{kj} = exp(ln\ \alpha_k + ln\ \beta_j)$ $\phi_{kj} = 1$	Y	Y	Y
Full weights	$\mu_{kj} = exp(ln\ \alpha_k + ln\ \beta_j)$ $\phi_{kj} = min(8, j+1)^2$		Y	

Figure 6-8. Standardized Deviance Residuals (Mean model)

Accident period | Development period



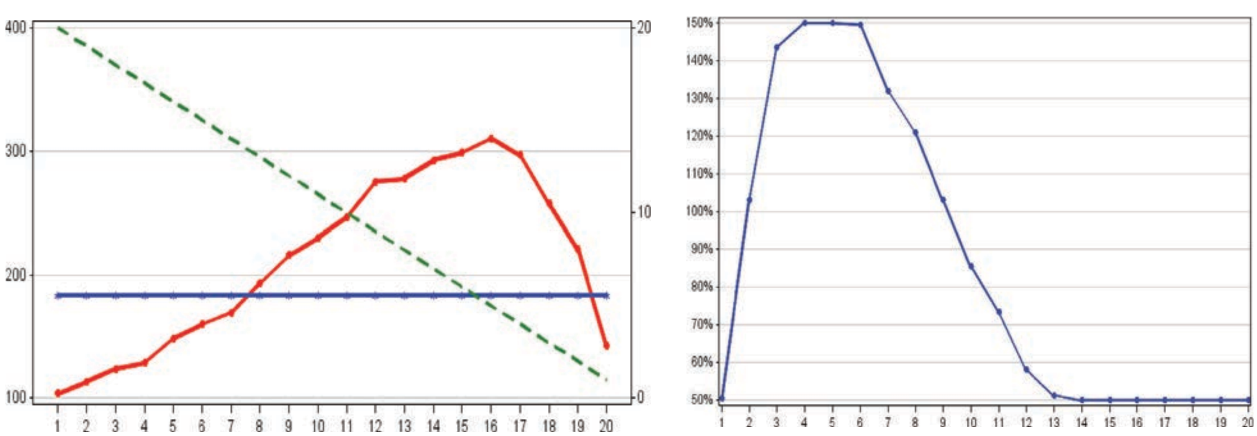
and development period trends that exist in the data. Figure 6-8 shows scatterplots of the deviance residuals by accident and development periods and exhibits clear trends and departures from randomness.

Alternatively actual and expected comparison (Section 6.3.3) or actual and expected ratio (Section 6.3.4) plots may be helpful in providing a clearer view of the goodness-of-fit (or lack thereof).

The trends seen in Figure 6-8 may be clearly seen in the actual and expected plots in Figure 6-9. In general actual and expected plots may often be an easier way of assessing the goodness-of-fit of the data than residual plots. However, residuals plots should not be ignored for this purpose; in particular residual plots are very useful

Figure 6-9. Actual and Expected (Mean model)

Actual and expected comparison plot by accident period | Actual and expected ratio plot by development period



Note: Left hand graph: the red line is the actual line, while the blue line represents the expected values. The green dotted line represents (right-hand scale) the number of data points underlying each plotted point.
Right hand graph: The actual/expected ratios have been truncated to a minimum value of 50% and a maximum value of 150%.

for identifying outliers that may need to be removed prior to fitting a model (refer to Section 2.2.5).

Suppose now that a model with development period effects has been fitted—the Development model. Residual and actual and expected plots by development period are shown in Figure 6-10. Note that, in the right-hand graph, both actual [red] and expected values [blue] have been plotted but they coincide so that only one (the expected line) is actually visible to the reader.

The residual plot no longer appears to contain trends, indicating that the model has captured the development period trends. This is confirmed by the actual and expected comparison plot, where the actual and expected totals are identical.

Note, however, that the ML equations for the ODP models are marginal sum estimation equations. Consequently, the actual and expected marginal totals associated with each model parameter are identical. The Development model contains development period (or column) parameters, and so actual and expected marginal totals by development year are identical (refer to Section 3.2 for further discussion of this point).

Thus, the actual and expected comparison and ratio plots provide no information in this case other than that development period trends have been captured in the slavish manner pre-ordained by marginal sum estimation.

On the other hand, the residual scatterplot does provide some information on the goodness-of-fit; in this case there is a suggestion of heteroscedasticity.

Figure 6-11 shows comparison plots of actual and expected for accident and calendar periods for the Development model. It is clear that the goodness-of-fit is still inadequate. The same plots are shown in Figure 6-12 but in this case for the fully specified cross-classified model, i.e., the Full model. The accident period actuals and expected overlay exactly due to marginal sum estimation in the presence of both accident and development period parameters in the model. The calendar period comparison is very close, suggesting that the model does not contain any calendar period effects.

Figure 6-10. Development Period Diagnostics (Development model)

Scatterplot of residuals

Actual and expected comparison plot by development period

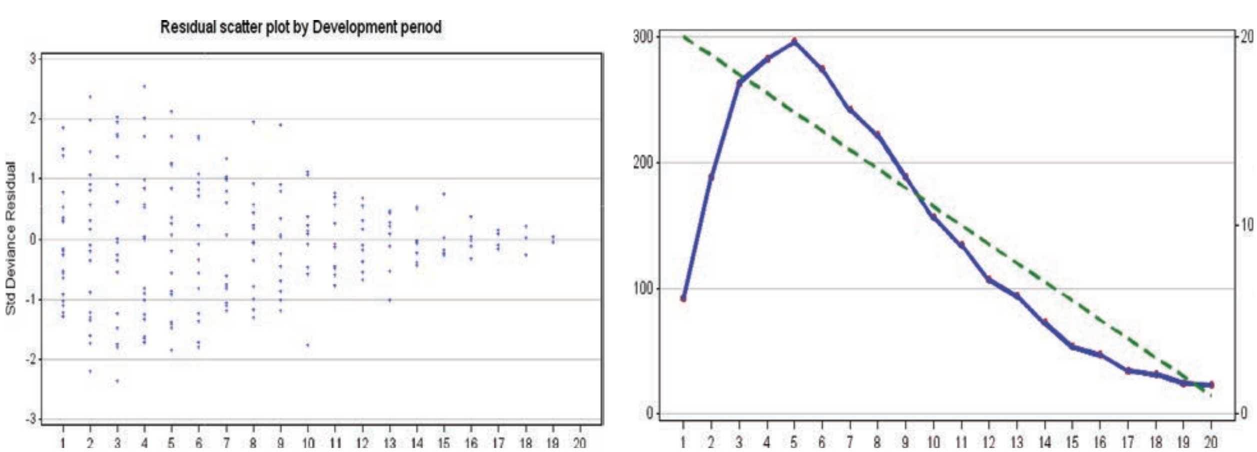
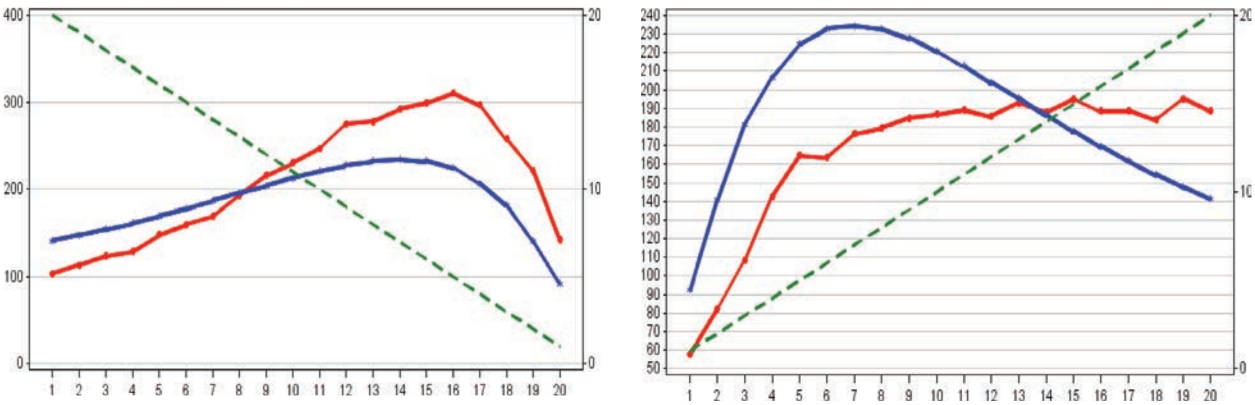


Figure 6-11. Accident and Calendar Period Actual and Expected Comparison Plots (Development model)

Accident period Calendar period



Note: the red lines are the actual lines while the blue lines represent the expected values. The green dotted lines represent (right-hand scale) the number of data points underlying each plotted point.

Other plots which may be useful include the residuals plotted against linear predictor and expected values. These plots are also used later when verifying distribution assumptions.

6.5.1. Identifying Interactions

So far the examples considered have been for a model where there are no interactions. Consider now a case where the development period factors β_j in the cross-classified model change significantly at a point in the past as they do in simulated data 3, and consider the diagnostics under the Full model where one set of development period factors is fitted for all accident periods.

Both the accident period and development period actual and expected comparison (and ratio) graphs are not useful since the actual and expected totals are identical

Figure 6-12. Actual vs Expected for Full CC Model

Accident period Calendar period

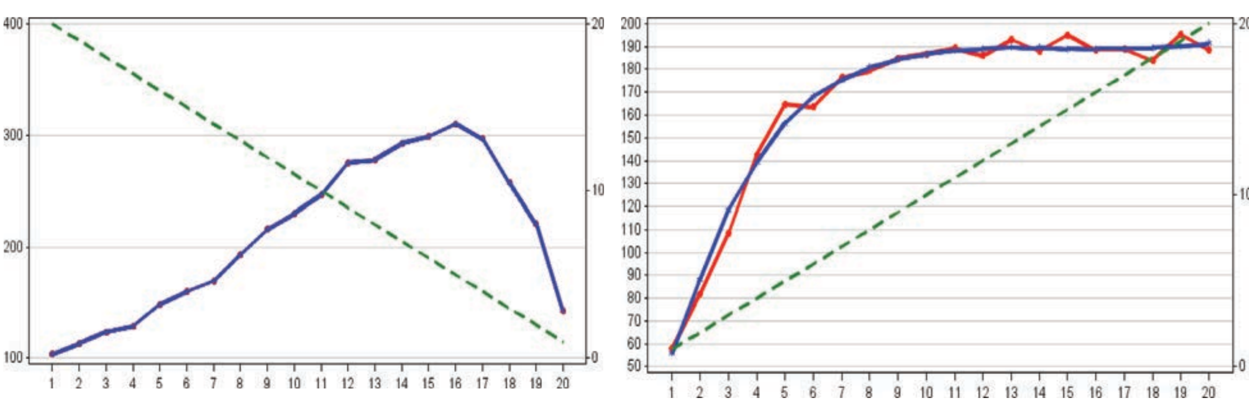
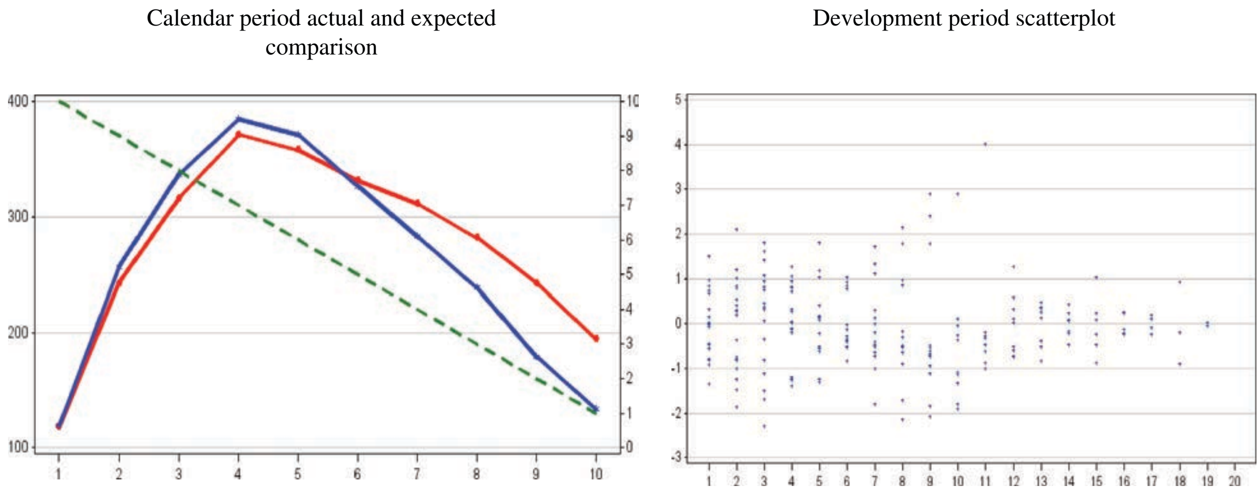


Figure 6-13. Diagnostic Plots—Model with Missing Interaction



Note: the red lines represent the actual observations, while the blue lines represent the expected values. The green dotted lines represent (right-hand scale) the number of data points underlying each plotted point.

due to the use of marginal sum estimation. The calendar period actual and expected comparison plot does suggest areas of poor fit (see Figure 6-13), and some of the residual plots exhibit non-randomness such as that in the development period scatterplot also shown in Figure 6-13.

Since accident and development period effects have been fitted in the model, the missing effect may be either a calendar period effect or an interaction between accident and development period (or both). To determine its nature, a heat map of actual and expected ratios may be helpful.

The heat map is shown in Figure 6-14. The distribution of actual/expected ratios is non-random with clusters of ratios greater than 100% and less than 100%. The lines have been added by judgment to separate out areas that show greater concentrations of ratios greater or less than 100%. Since the clusters appear to be located for specific accident and development period groups rather than along entire diagonals, this suggests that the missing effects are interactions between accident and development periods and not calendar period effects.

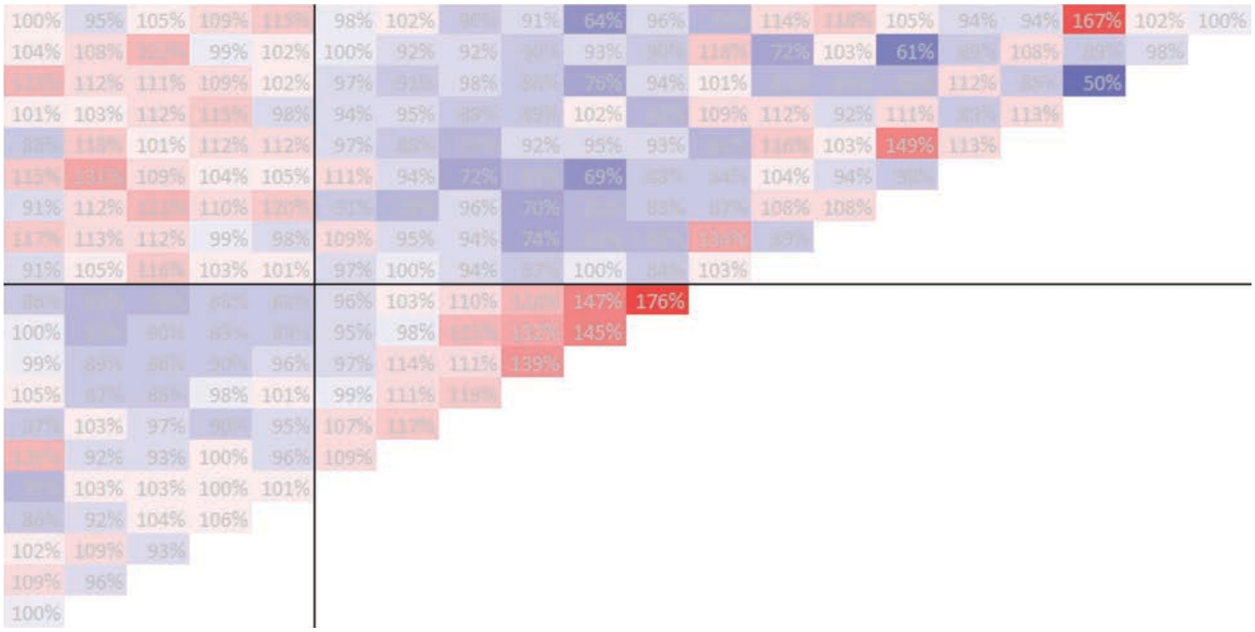
6.6. Analysis of the Distribution Assumptions

The goodness-of-fit tests may be viewed as checking whether the model's cell means provide a good fit to the cell observations. However, in addition to the cell means, it is also important to check whether the model distribution is a good approximation to the data. This is particularly true if the model is intended to be used to assess the variability of the loss reserve estimate.

The main distributional assumptions are:

- The form of the distribution of the data;
- The scale parameter of that distribution; and
- The choice of link.

Figure 6-14. Actual/Expected Heat Map—Full Model for Simulated Data 3



The main tools in checking the distributional assumptions are:

- Plots of residuals; and
- Probability-Probability (P-P) plots.

As discussed in Section 6.1, the recommended approach to model validation was first to fit a simple model and check for any gross violations of the distributional assumptions. At this stage, problems such as a moderate level of heteroscedasticity could be ignored since they may result from poor estimation of the cell means. Providing the residual plots do not indicate a serious problem, the modeler may then continue to fit the model. Once the cell means fit well (based on the goodness of fit tests), the distributional assumptions may be re-examined in fine detail and adjusted as required.

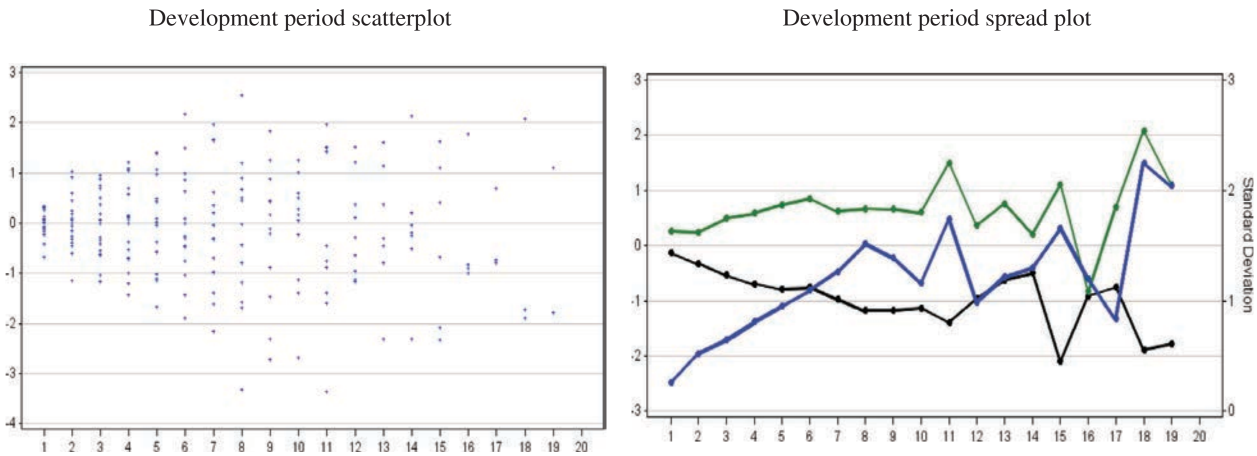
Distribution diagnostics are illustrated for simulated data 2 under the Full model. For simulated data 2, the full model correctly specifies the form of the cell mean but does not correctly specify the variability/scale since it assumes a constant scale parameter rather than a scale that varies by development period. Thus, the diagnostics should show evidence of incorrect dispersion assumptions.

Recall that in the spread plot, the green and black lines represent the interquartile range while the blue line is the standard deviation of the residuals at each development period.

The residuals in Figure 6-15 are clearly heteroscedastic with a fanning out of residuals observable for development periods 1-8, as expected based on the assumptions for the scale parameter (refer to Table 6-1). From the spread plot it is seen clearly that the standard deviation of the residuals increases over the same range of development quarters.

The P-P plot is shown in Figure 6-16. Some deviations from the Poisson distribution may be seen.

Figure 6-15. Diagnostic Plots—Full Model, Scale Parameter Assumed Constant



The model was refitted using the correct formulation for the scale parameter. Strictly speaking, it is the weights, rather than the scale parameter, that require correction. Recall from (2-21) that the scale or dispersion parameter may be written as $\phi_i = \phi/w_{kj}$. In this case, $\phi = 1$ and the weights vary only by development period j and are specified by $w_{kj} = \min(8, j + 1)^2$ (as per the data specification in Table 6-1).

The same plots as in the preceding two figures are shown below in Figure 6-17 after the model refit. The improvement is apparent.

As well as adjustments to the dispersion by means of weights, the modeler should generally consider whether the use of a different distribution, e.g., Gamma rather than Poisson, is more appropriate for the data under consideration.

Figure 6-16. P-P Plot—Full Model, Scale Parameter Assumed Constant

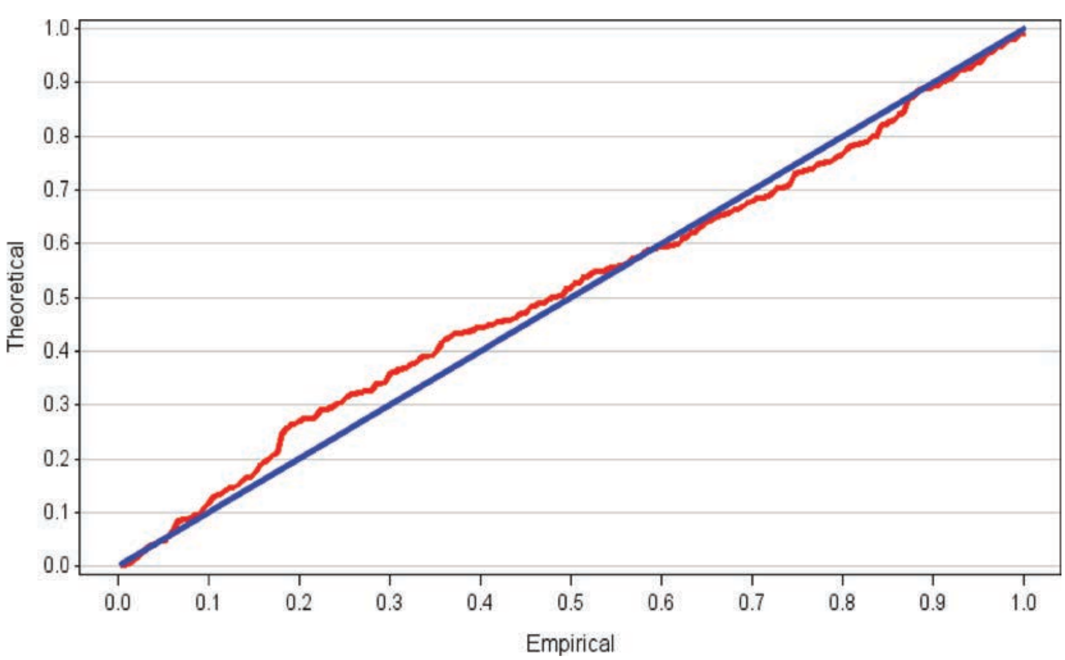


Figure 6-17. Diagnostic Plots for Correctly Specified Variable Scale Model

Finally, there are no particular tests for the choice of the link function. Rather, the link is usually determined by the model structure (e.g., an additive model implies an identity link while a multiplicative model implies a log link), which in turn is often determined *a priori* by the nature of the data being modeled. Generally speaking, if the link function implies a structure that makes sense for the particular data, and if the diagnostics plots are acceptable without requiring an unreasonable number of interactions, then the link function may be considered appropriate.

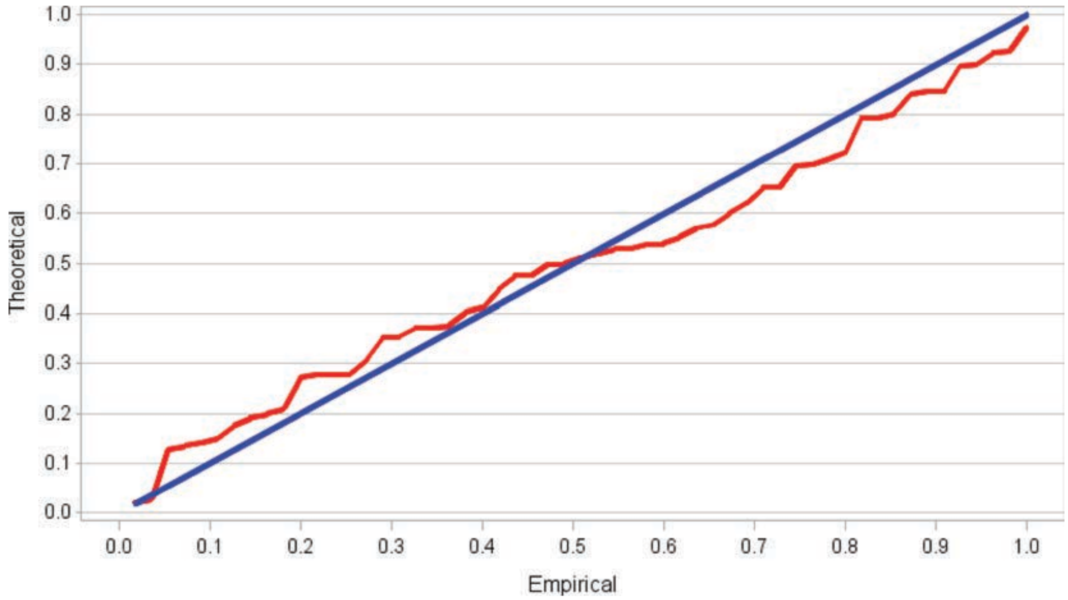
6.7. Model Validation for Real Data

The examples discussed to date have used simulated data so that the true underlying model is known. In practice, this is not the case, so the modeler will need to select the best model using judgment. In the following sections, diagnostics plots for the cross-classified model (Section 3.3.2) based on the data in Table 1-1 are shown together with some commentary.

6.7.1. Initial Check of Distribution Assumptions

As a model of main effects only, the cross-classified model may be easily fitted. Once this is done, the first step in model validation is to check that the distributional assumptions are not grossly violated.

Figure 6-18. P-P Plot for the Cross-Classified Model



The P-P plot for the cross-classified model is shown in Figure 6-18. While there are systematic departures from the straight line, suggesting that the distributional assumptions could be improved, the distortion is not at a level that renders the Poisson log link distributional assumptions unusable as an initial set of assumptions for building a model.

The residual plots should also be checked first for any major problems with the distributional assumptions and second for indications of regions of poor fit.

Figure 6-19 shows the scatter and spread plots by development year for the cross-classified model. As above, the residuals do not suggest a major problem with the distributional assumptions.

However, the spread of the residuals in development years 1 and 2 is greater than in other years, which may suggest a less than optimal fit to the means of the development year 1 and 2 data or that the Poisson assumptions may be inadequate (e.g., perhaps the scale parameter varies by development period).

Figure 6-19. Residuals by Development Year for the Cross-Classified Model

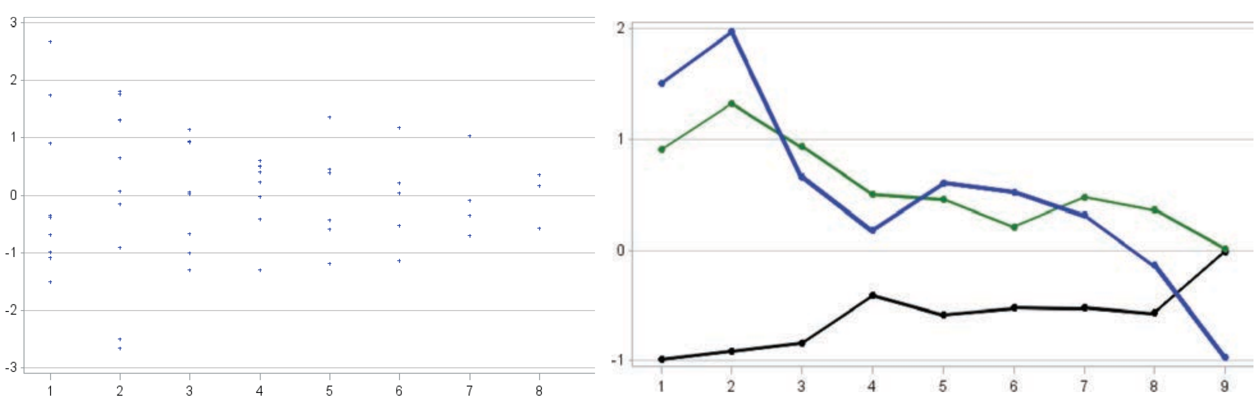
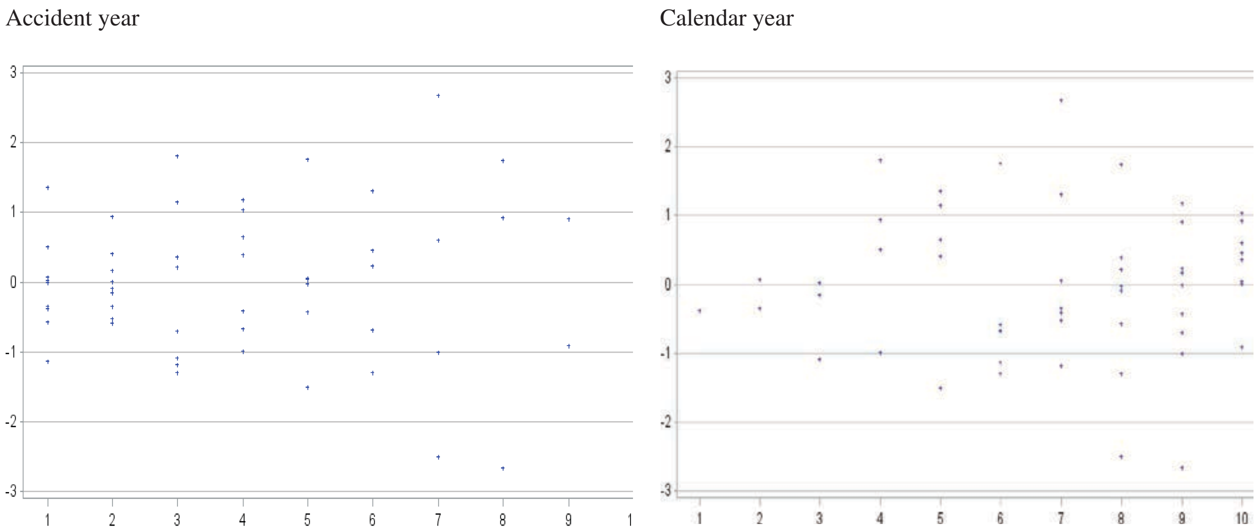


Figure 6-20. Residuals by Accident and Calendar Year for the Cross-Classified Model



The residuals by accident and calendar year are shown in Figure 6-20. The residuals by calendar year, in particular, suggest a problem with the model fitting that should be investigated further.

Thus the conclusions from this stage of the model validation process are that there are areas of poor fit that need further investigation and possible modeling. For the time being the distributional assumptions (Poisson model, constant scale, log link) may continue to be used, but they will need reviewing once the fit of the model has been improved.

6.7.2. Goodness-of-Fit

The next step in the modeling process is to use the various goodness-of-fit diagnostic tools to identify the regions of poor fit better and determine whether these should be modeled.

Since the cross-classified model contains a parameter for each accident and development year, the marginal totals will be identical under ML estimation (Section 3.2). Therefore actual and expected plots by accident and development years will be unhelpful. The comparison plot of actual and expected by calendar year is shown in Figure 6-21 below. This appears satisfactory, even though the residuals by calendar year are problematic (Figure 6-20 above).

This suggests that the poor fit may result from some interactions, so the triangular heat map diagnostic may be useful and is shown in Figure 6-22. This indicates the presence of some missing interactions between accident year and development years 1 and 2 (see the highlighted regions in the plot below).

Even in the absence of evidence of poor fit from the various one-way residual and goodness-of-fit diagnostics, the accident/development 2-d heat map should always be checked in reserving models.

In summary, the fit of the cross-classified model is reasonably good, but there is evidence of some interactions between accident and development years. Chapter 7

Figure 6-21. Actual and Expected Comparison Plot by Calendar Year

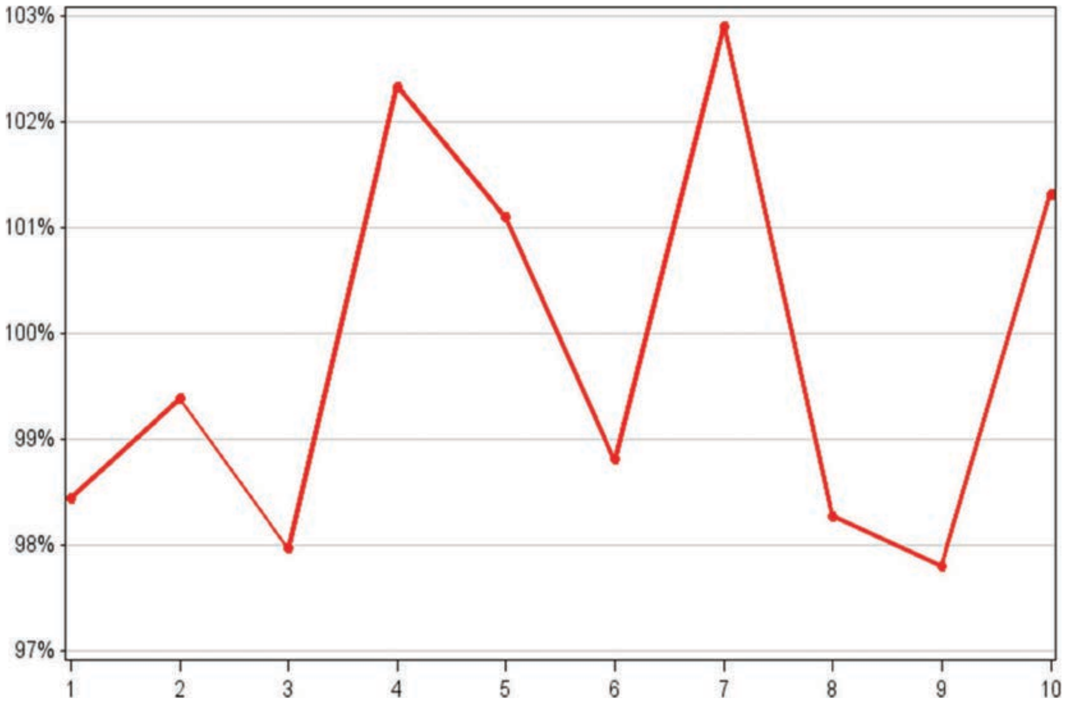


Figure 6-22. 2-D Heat Map for Cross-Classified Model

	Development									
Accident	0	1	2	3	4	5	6	7	8	9
1	98%	100%	100%	104%	113%	87%	96%	92%	100%	100%
2	99%	99%	106%	103%	95%	95%	99%	102%	100%	
3	96%	108%	107%	91%	90%	102%	92%	104%		
4	97%	103%	96%	97%	103%	111%	111%			
5	95%	107%	100%	100%	97%	100%				
6	98%	105%	93%	101%	104%					
7	109%	91%	95%	104%						
8	106%	90%	105%							
9	103%	97%								
10	100%									

deals with extensions to the cross-classified model, including the use of interactions, and the reader is referred there for further discussion concerning their use for this particular loss reserving problem.

Once interactions are included in the model (e.g., as per the discussions in Chapter 7), the modeler should then return to the tests of distributional assumptions and ensure that these are now satisfactory, making adjustments if required.

7. Model Extensions

Chapter summary. It has often been remarked in the literature that the conventional chain ladder involves an excessive number of parameters, with a separate parameter for each accident year and for each development year. The GLM formulations of Chapter 3 follow the same parametric structure, and therefore carry the same excess of parameters.

This chapter is concerned with approaches to parameter reduction, achieved largely by means of generalized additive models. A GAM is obtained by the replacement of each of a number of categorical variates in a GLM with a parametric form that is economical in its parameters. Prime candidates for this sort of parameterization are accident year and development year trends, which are represented by categorical variates in the chain ladder.

The chain ladder assumes a multiplicative structure in the sense that the mean associated with any cell is equal to the product of a row factor and a column factor. Sometimes this model structure will not be supported by the data. The concept of calendar period effects and of interactions, required to correct the structure, is explored.

A parametric form in relation to development year also enables models to be extrapolated beyond the range of development years encompassed by the data. A smooth parametric form will ensure that the model progresses smoothly over development years, both inside and outside the bounds of the data.

Finally, models other than the chain ladder are briefly discussed. These include exposure-based models of claim numbers and payments, models that comprise of a number of sub-models and individual claim models. The chapter concludes with a brief reference to Bayesian models.

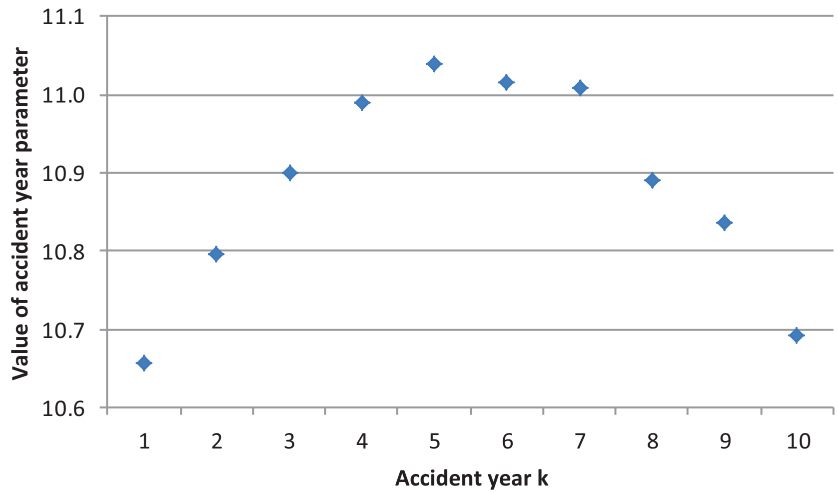
7.1. Chain Ladder Model Revisited

Consider the accident year parameter estimates $ln\ \hat{\alpha}_k$ appearing in Table 5-1. Figure 7-1 plots them against accident year k.

There are 10 parameters plotted. However, they assume a strongly parabolic appearance, raising the question as to whether the 10 values might be adequately represented by means of a smaller number of parameters.

Consider Figure 7-2 in this context. The dotted curves here describe a confidence envelope of ±2 standard errors about the parameter estimates, where the standard errors are also obtained from Table 5-1. The solid line represents the ordinary least squares fit of a quadratic to the parameter estimates.

Figure 7-1. Plot of Accident Year Parameter Estimates



The quadratic curve, which is described by only 3 parameters, appears to track the parameter estimates well and does indeed lie well within the confidence envelope.

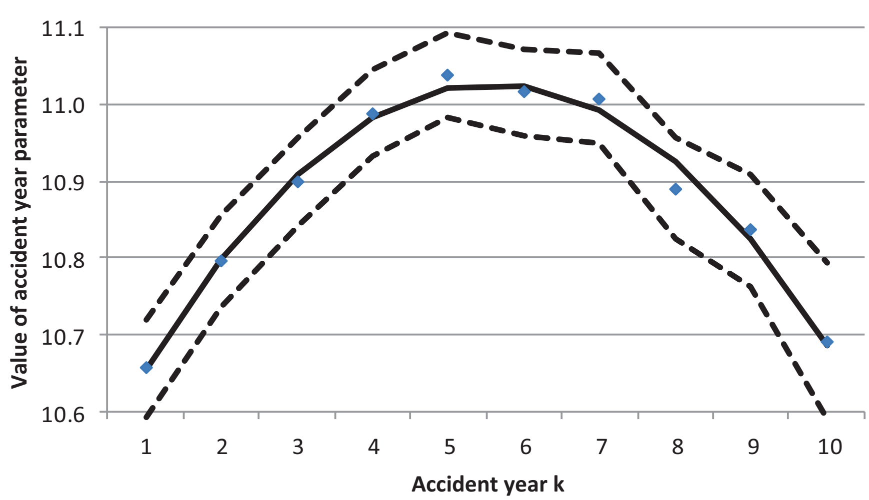
As discussed in Section 4.2.2, an excessive number of model parameters degrades a model's predictive power. A question arises therefore as to whether greater predictive power is obtained when the conventional chain ladder model is replaced by an alternative version in which the 10 accident year effects are represented by a quadratic form.

Curve fitting of this sort might have a physical motivation, or might simply amount to abstract fitting (as in the present case). In either case, one must usually be resigned to the loss of some goodness-of-fit. However, the ultimate justification for such curve fitting is reduction of prediction error as a result of reduced parameterization.

Mathematically, the use of the suggested quadratic form amounts to replacement of (3-18) in the ODP cross-classified model of Section 3.3.2 (i.e., $E[Y_{kj}] = \mu_{kj} = exp(ln\ \alpha_k + ln\ \beta_j)$) by the following:

$$E\left[Y_{kj}\right] = \mu_{kj} = exp\left(a_0 + a_1 k + a_2 k^2 + ln\ \beta_j\right) \tag{7-1}$$

Figure 7-2. Quadratic Fit to Accident Year Parameter Estimates



where there are now 12 free parameters a_0, a_1, a_2, $ln\ \beta_2, \ldots, ln\ \beta_{10}$. Recall that $ln\ \beta_1$ was arbitrarily set to zero in Section 3.3.2 (see Table 3-5) due to parameter redundancy.

It is somewhat convenient to abbreviate this model a little further, thus:

$$E\left[Y_{kj}\right] = \mu_{kj} = exp\left(a_1 k + a_2 k^2 + ln\ \beta_j\right) \tag{7-2}$$

where the degree of freedom lost by deletion of the parameter a_0 is compensated by restoration of $ln\ \beta_1$ as a free parameter. Model (7-2) contains the same number (12) of parameters as (7-1) but those parameters are now a_1, a_2, $ln\ \beta_1, \ldots, ln\ \beta_{10}$.

7.2. Generalized Additive Models

The model (7-2) is an example of a **generalized additive model ("GAM")**. A GAM is a special case of a GLM. Recall the definition of a GLM in Section 2.2.1, and in particular condition (2) of that definition:

$$h\left(\mu_i\right) = x_i^T \beta \tag{7-3}$$

with $x_i^T = (x_{i1}, x_{i2}, \ldots, x_{ip})$, the vector of predictors associated with the i-th observation Y_i.

Now suppose that one or more of the predictors takes the form

$$x_{ij} = u_j\left(z_i\right) \tag{7-4}$$

where u_j is a real-valued function, and z_i is a vector of further covariates: $z_i^T = (z_{i1}, z_{i2}, \ldots, z_{iq})$ which may include components of x_i. The u_j might be basis functions of the type introduced in Section 2.2.2.

When the GLM is defined subject to (7-4), it is a GAM. The model defined by (7-2) provides an example. In the present case,

$$x_i^T = \left(J_{i1}, J_{i2}, \ldots, J_{i,10}, u_1\left(z_i\right), u_2\left(z_i\right)\right) \tag{7-5}$$

where J_{ij} is a 0-1 indicator that takes the value unity if the i-th record relates to development year j and zero otherwise (compare with the design matrix X set out in Section 3.3.2);

$$z_i = \left(k_i\right) \tag{7-6}$$

a 1-vector in which k_i denotes the value of k associated with the i-th record; and

$$u_m\left(k\right) = k^m,\ m = 1, 2 \tag{7-7}$$

The following sections will examine a few applications of GAMs to the data triangle set out in Table 1-1.

7.3. Accident Year Trend

This model has been fitted to the data triangle set out in Table 1-1, and the resulting estimates appear in Table 7-1 under the heading "Simplified model". Those under the heading "Chain ladder" reproduce the estimates from Table 5-1 for comparison. It is evident that the simplification of the model has caused very little difference to the estimated development pattern.

The quadratic representation of the accident year effect (see (7-2)) is $10.471 + 0.2001k - 0.0179k^2$.

The simplified model has been applied to the forecast of outstanding losses, and the associated forecast error estimated by means of a parametric bootstrap. The procedure is parallel to that set out in Section 5.4.2, and its results appear in Table 7-2.

Table 7-2 may be compared with Table 5-5, which contains exactly the same information for the chain ladder model. The comparison indicates that the model simplification has affected the forecast of outstanding losses very little (0.4%), but has resulted in a reduction of 8.4% in estimated forecast error. In short, the reduction in parameterization of the model has resulted in improved forecast efficiency.

Note that, in some lines of business, an exposure measure may be used as an alternative means of capturing accident period trends. This is discussed below in Section 7.8.

7.4. Development Pattern

Consider the development year parameter estimates $ln\ \hat{\beta}_j$ appearing in Table 5-1. Figure 7-3 plots them against development year j.

There are 10 parameters plotted. However, it appears that they might be adequately represented by a linear spline with a knot at $j = 7.5$, again by means of a smaller number of parameters.

Table 7-1. Parameter Estimates for Simplified Model

	$ln\ \hat{\beta}_j$	
j	Chain Ladder	Simplified Model
1	0.000	0.000
2	–0.205	–0.206
3	–0.747	–0.750
4	–1.017	–1.015
5	–1.452	–1.452
6	–1.833	–1.830
7	–2.140	–2.142
8	–2.348	–2.353
9	–2.513	–2.514
10	–2.664	–2.661

Table 7-2. Parametric Bootstrap Estimates of Simplified Model's Forecast Error

	Outstanding Losses		
Accident Year	Forecast	RMSEP	CVP
	$000	$000	%
1989	3,467	885	25.5
1990	8,334	1,295	15.5
1991	14,594	1,659	11.4
1992	22,416	2,000	8.9
1993	32,340	2,312	7.1
1994	45,263	2,614	5.8
1995	62,410	3,076	4.9
1996	79,922	3,658	4.6
1997	104,895	4,844	4.6
Total	373,641	13,086	3.5

Consider Figure 7-4 in this context. The dotted curves here describe a confidence envelope of ±2 standard errors about the parameter estimates, where the standard errors are obtained from Table 5-1. The solid line represents the ordinary least squares fit of the following linear spline to the parameter estimates:

$$b(j) = b_1(j-1) + b_2 \max(0, j-7.5) \tag{7-8}$$

The spline, which is described by only 2 parameters, appears to track the parameter estimates well and does indeed lie well within the confidence envelope with the exception

Figure 7-3. Plot of Development Year Parameter Estimates

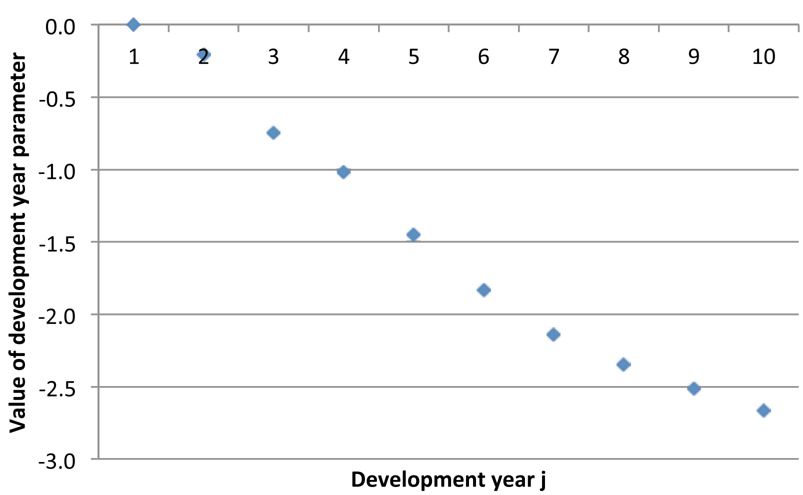
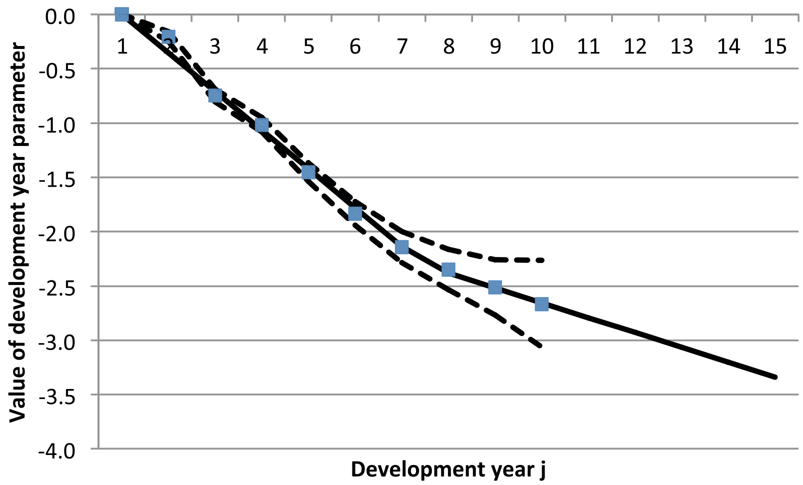


Figure 7-4. Fit of Linear Spline to Development Year Parameter Estimates



of the case $j = 2$. This suggests a model, with the accident year simplification of Section 7.3 incorporated, of the form (7-3) with

$$x_i^T = \left(1, k_i, k_i^2, j_i - 1, \max\left(0, j_i - 7.5\right), J_{i2}\right) \tag{7-9}$$

where j_i denotes the value of j associated with the i-th record. Note the inclusion of the unit regressor, which allows for a constant term in the regression.

Thus the final model takes the form

$$x_i^T\beta = a_0 + a_1 k_i + a_2 k_i^2 + b_1\left(j_i - 1\right) + b_2 \max\left(0, j_i - 7.5\right) + cJ_{i2} \tag{7-10}$$

This model has been fitted to the data triangle set out in Table 1-1, and the resulting estimates appear in Table 7-3.

Table 7-3. Parameter Estimates for Model with Both Accident and Development Year Simplifications

Parameter	Estimate
Accident year parameters	
a_0	10.469
a_1	0.200
a_2	–0.018
Development year parameters	
b_1	–0.358
b_2	0.236
c	0.155

Table 7-4. Parametric Bootstrap Estimates of Forecast Error for Model with Both Accident and Development Year Simplifications

	Outstanding Losses		
Accident Year	Forecast	RMSEP	CVP
	$000	$000	%
1989	3,542	776	21.9
1990	8,410	1,295	15.4
1991	14,490	1,678	11.6
1992	22,201	1,963	8.8
1993	32,572	2,303	7.1
1994	45,660	2,658	5.8
1995	61,592	3,088	5.0
1996	79,975	3,679	4.6
1997	104,959	4,977	4.7
Total	373,403	13,248	3.5

This model has been applied to the forecast of outstanding losses, and the associated forecast error estimated by means of a parametric bootstrap. The procedure is parallel to that set out in Sections 5.4.2 and 7.3, and its results appear in Table 7-4.

The bootstrap estimates of prediction error in Table 7-4 are comparable with those in Table 7-2 for the model that contained only the accident year simplification. It is seen that the development year simplification of the model has caused:

- once again, virtually no change in the forecast of outstanding losses; and
- just a slight increase in the associated CVP (3.50% to 3.55%).

Whether one chooses this model over the one developed in Section 7.3 is largely a matter of taste. The model of the present section reduces the number of model parameters from 12 (19 originally for the chain ladder) to 6, but without any improvement (and, technically, a slight deterioration) in forecast quality. However, it does express the development pattern in parametric form, leading to a smooth tail as well as forming a basis for tail extension, so it may be preferred on this basis. Tail smoothing and extension are discussed further in Section 7.7.

7.5. Calendar Year Trend

The models discussed up to this point have considered accident and development period effects only, or alternatively, the rows and columns of triangles laid out in the manner of Table 1-1. There is a third direction in this triangle—the diagonal or, equivalently, the calendar period—that should be considered.

In practice, calendar year trends are common in insurance data for a number of reasons. Some examples are given below:

- Many lines of business have a clear relationship with economic inflation. For example, changes in wage inflation will impact lines of business such as workers compensation or auto third party bodily injury claims as much of the cost of these claims consists of either income replacement or damages, reflecting pre-injury earnings in either case;
- Award precedents set by court decisions or other environmental change will often apply from a specific point in time, regardless of when the claim occurred;
- Changes in claims management departments such as expansion or contraction of staff numbers may impact the rate at which all claims are closed, which leads to a calendar effect on the insurance data.

A common method for dealing with economic inflation is to adjust the data so that all payments are in the same dollar values, e.g., the dollar values of the valuation date. In this case, the model forecasts will then be in the dollar values of the valuation date, so will need to be adjusted for future economic inflation. This has the advantage of producing forecasts with explicit economic assumptions, rather than an implicit assumption that the rate of economic inflation will be similar to that of the past, as is the case for the chain ladder. This may be useful for scenario tests, or if future rates are expected to be different to past rates, at least in the short term. Furthermore, for a company with multiple lines of business, carrying out a valuation in constant dollar values means that the consistent rates of future economic inflation may be applied across all LOBs. This is helpful both for scenario testing and for estimating variability of reserves since it introduces some correlation (that relating to economic variation) across the different LOBs.

Calendar period changes (both positive and negative) net of changes due to economic inflation are often referred to as superimposed inflation (**"SI"**), terminology introduced by Benktander (1979) and discussed in various parts of Taylor (2000). Typically SI is variable over time. For example, payments might increase at rates beyond economic inflation for a number of years, before measures are put in place to curtail the increase or even reduce claim size. This can lead to nil or even negative SI, which may last for some time, before other factors act to increase claim size once more.

Unmodeled calendar period effects can lead to distortions in the claim size models which would show up in the calendar period and triangular heat map diagnostics discussed in Chapter 6. If the diagnostics suggest calendar period effects, then as a first step, the modeler may wish to consider whether there is a natural economic inflation series for this line of business and, if so, adjust the past claim amounts to the valuation date. If unmodeled effects are still apparent after this step (or if there is no natural series to use), then the modeler should consider including calendar period effects in the model.

Adding calendar period effects to a model such as the cross-classified model must be done with due care. Accident, development and calendar period terms are not independent covariates—knowledge of two of these determines the third. Thus, for the cross-classified model, replacing (3-18) with

$$\mu_{kj} = exp\left(ln\,\alpha_k + ln\,\beta_j + ln\,\gamma_{k+j-1}\right) \tag{7-11}$$

is inappropriate since the collinearity of the accident, development and calendar terms (γ_{k+j}) means that there is no unique solution to the model, and any solutions returned by GLM software will be unstable.

Instead the modeler should impose a simple structure on the calendar period effects, based on examination of the model diagnostics. For example, if SI appears to progress at a constant rate over the first *h* diagonals and to be flat thereafter, then (3-18) could be replaced by

$$\mu_{kj} = exp\left(ln\,\alpha_k + ln\,\beta_j + min\left(h, k + j - 1\right)\varphi\right) \tag{7-12}$$

In practice, selection of an appropriate function should be based on model diagnostics, business knowledge and pragmatism; any calendar period trend will need to be extrapolated into the future for forecasting purposes, so the modeled trend must take this into account.

Recall that, although the Mack model formulation of the chain ladder may appear to be a development year only model, in fact the most recent diagonal of payments in the Mack model functions as accident period effects (see Section 3.2), so the same cautionary note about the addition of calendar period effects applies equally to Mack as to the cross-classified model.

7.6. Interactions

Consider model (7-10). It contains some terms that depend on accident year and others that depend on development year. This means, for example, that the relation between different development years is independent of accident year. In chain ladder, parlance, age-to-age factors are constant across accident years.

Similarly, the relation between different accident years is independent of development year. In these circumstances, the individual components of the linear response are called **main effects**.

In some cases, however, the data may indicate that some development year effects depend on accident year. Consider, for example, Figure 7-5, which displays a heat map for model (7-10).

Features of this map are:

- for development year 1, a distinct area of blue in the earlier accident years;
- for development year 2, a distinct area of pink in the earlier accident years;
- for development year 3, a possible progression from pink to blue with increasing accident year;
- for development year 4, a preponderance of pink over the whole set of accident years.

In effect, it appears that the payment pattern has altered. Traditional actuarial methods typically deal with this by calculating chain ladder factors based on recent diagonals

Figure 7-5. Heat Map for Model with Both Accident and Development Year Simplifications

Accident year	Development year 1	2	3	4	5	6	7	8	9	10
1988	99%	101%	98%	111%	112%	84%	97%	96%	100%	97%
1989	99%	99%	102%	109%	93%	90%	99%	106%	99%	
1990	95%	107%	102%	96%	88%	97%	92%	107%		
1991	97%	103%	94%	104%	102%	107%	113%			
1992	97%	108%	99%	108%	97%	98%				
1993	97%	104%	89%	106%	101%					
1994	110%	92%	93%	112%						
1995	102%	87%	99%							
1996	105%	98%								
1997	101%									

only, e.g., the most recent 3 or 5 diagonals, etc. Essentially this corresponds to one model for older diagonals (even though the chain ladder factors may not be calculated) which is then modified for more recent experience and for projection.

The approach taken by the GLM is similar in principle in that the model is adapted to better fit the changed experience. The above features suggest testing the following additional terms in the model's linear response, listed in the order of the above dot points to which they relate:

$$d_1 J_{i1} K_{i,1-6} + d_2 J_{i2} K_{i,1-6} + d_3 J_{i3} k + d_4 J_{i4} \tag{7-13}$$

where the variate $K_{i,1-6}$ is a 0-1 indicator that takes the value unity if the i-th record relates to an accident year in the range 1 to 6, and zero otherwise (compare with the definition of J_{ij} in Section 7.2).

When these terms are added to (7-10), the complete model becomes (with a slight re-labelling and re-ordering of parameters for logicality):

$$x_i^T \beta = a_0 + a_1 k_i + a_2 k_i^2 + b_1 (j_i - 1) + b_2 \max(0, j_i - 7.5)$$

$$+ c_1 J_{i2} + c_2 J_{i4} + d_1 J_{i1} K_{i,1-6} + d_2 J_{i2} K_{i,1-6} + d_3 J_{i3} k \tag{7-14}$$

When this model is fitted to the data, the parameter estimates are as in Table 7-5. All parameters are significant at levels well below 5%.

The number of parameters has grown to 10, so there is a need to ensure that the additional model terms add to the predictive efficiency of the model.

A comparison of the CVP with that in Table 7-4 shows a substantial reduction of 17% (see Table 7-6). The CVP is now 23% below that of the conventional chain ladder model (see Table 5-5).

The information criteria AIC and BIC were introduced in Section 4.3, while the related measure, GCV, was introduced in Section 4.4. The progression of their values through the sequence of models developed in the present chapter is set out in Table 7-7. The corresponding progression of CVPs is also shown for comparison.

Table 7-5. Parameter Estimates for Model with Interactions

Parameter	Estimate
Accident year parameters	
a_0	10.4900
a_1	0.2066
a_2	–0.0183
Development year parameters	
b_1	–0.3685
b_2	0.2720
c_1	0.0375
c_2	0.0528
Interaction parameters	
d_1	–0.0671
d_2	0.1273
d_3	–0.0113

Table 7-6. Parametric Bootstrap Estimates of Forecast Error for Model with Interactions

	Outstanding Losses		
Accident Year	Forecast	RMSEP	CVP
	$000	$000	%
1989	3,630	569	15.7
1990	8,557	935	10.9
1991	14,563	1,203	8.3
1992	22,193	1,418	6.4
1993	32,505	1,677	5.2
1994	45,771	2,018	4.4
1995	62,998	2,459	3.9
1996	79,601	3,079	3.9
1997	101,742	4,094	4.0
Total	371,559	10,907	2.9

Table 7-7. AIC, BIC and GCV for Various Models

Model	AIC	BIC	GCV	CVP
				%
Conventional chain ladder (ODP cross-classified form)	–509,392	–509,354	6,685,428	3.8
Accident year simplification only	–509,400	–509,376	5,075,351	3.5
Both accident and development year simplifications:				
without interactions	–509,397	–509,385	4,311,874	3.5
with interactions	–509,441	–509,421	1,733,202	2.9

Notes:
- AIC and BIC are defined in Section 4.3. The log likelihood used in their calculation is $\sum_{i=1}^{n} w_i[y_i \log \hat{y}_i - \hat{y}_i]/\phi$ where $w_i = 1$ for all observations and the scale parameter is held constant at the value from the interactions model. The scale parameter is held constant to prevent changes in the scale from distorting the measurement of changed model fit.
- GCV is defined in Section 4.4.
- The values of AIC, BIC and GCV may differ depending on the statistical package. For AIC and BIC, this is because different packages may or may not include an additive constant (depending on the input data only) in the log likelihood expression. Thus the relativities of the scores, rather than their absolute values, are relevant. Additionally, the modeler should satisfy themselves that the measures are calculated appropriately in their package of choice.

The information criteria and GCV were introduced in Sections 4.3 and 4.4 as indicators of model predictive error. All three quantities show an improvement when accident year simplification is introduced and considerable improvement at the introduction of interactions, in line with CVP. On the other hand, the message is more mixed at the introduction of development year simplifications—AIC increases somewhat, BIC and GCV fall somewhat—while CVP remains almost unchanged. This reflects different levels of penalty placed on numbers of parameters—BIC and GCV penalize number of parameters more and therefore the trade-off between worse model predictive accuracy and fewer parameters is acceptable to these measures and not to AIC with its weaker penalty.

Empirical experience indicates that this sort of perverse behavior is not uncommon. In fact, while the information criteria are reasonable indicators of CVP behavior in the case of incremental changes to a model (such as the addition of interactions), they are frequently suspect in the case of wholesale changes (such as the shift from a categorical to a parametric representation). GCV, on the other hand, aligns better with CVP behavior for this particular data set.

Homoscedasticity

The concepts of homoscedasticity and heteroscedasticity were introduced in Sections 2.2.4 and 2.2.5, and the need for ensuring the former before the acceptance of a model discussed in Section 2.2.5.

The above model including interactions is examined for homoscedasticity in Figure 7-6, which plots deviance residuals against accident year, and Figure 7-7, which plots them against development year. Reasonable homoscedasticity appears to have

Figure 7-6. Residual Plot by Accident Year for Model Including Interactions

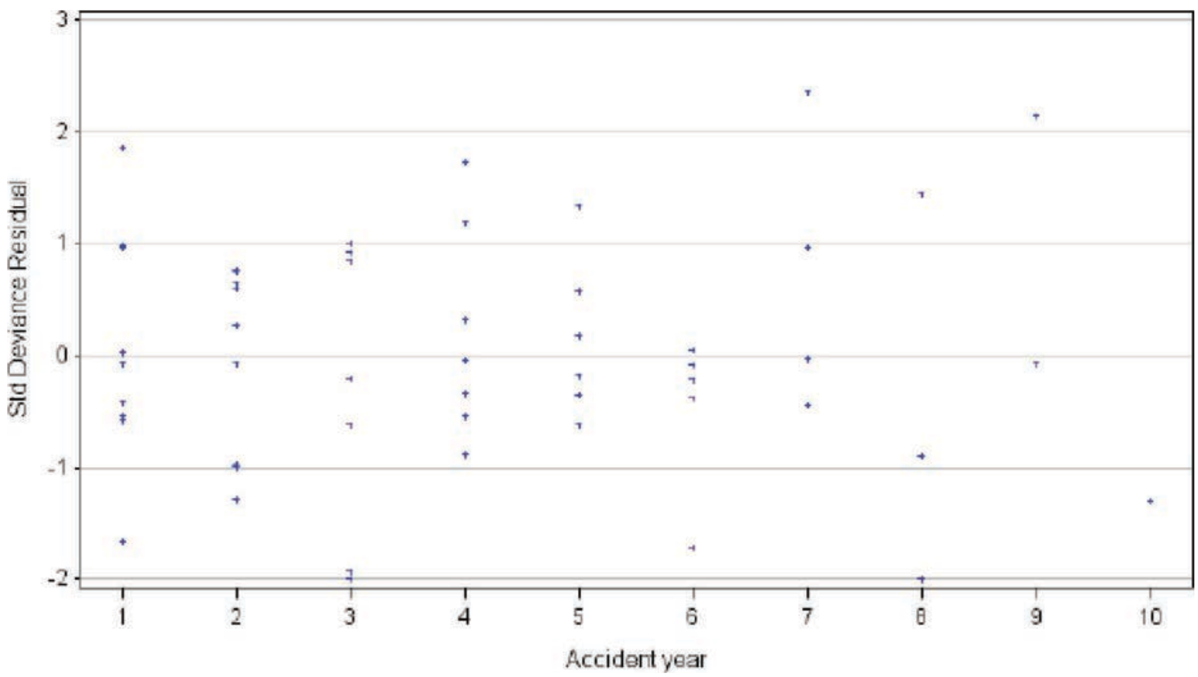
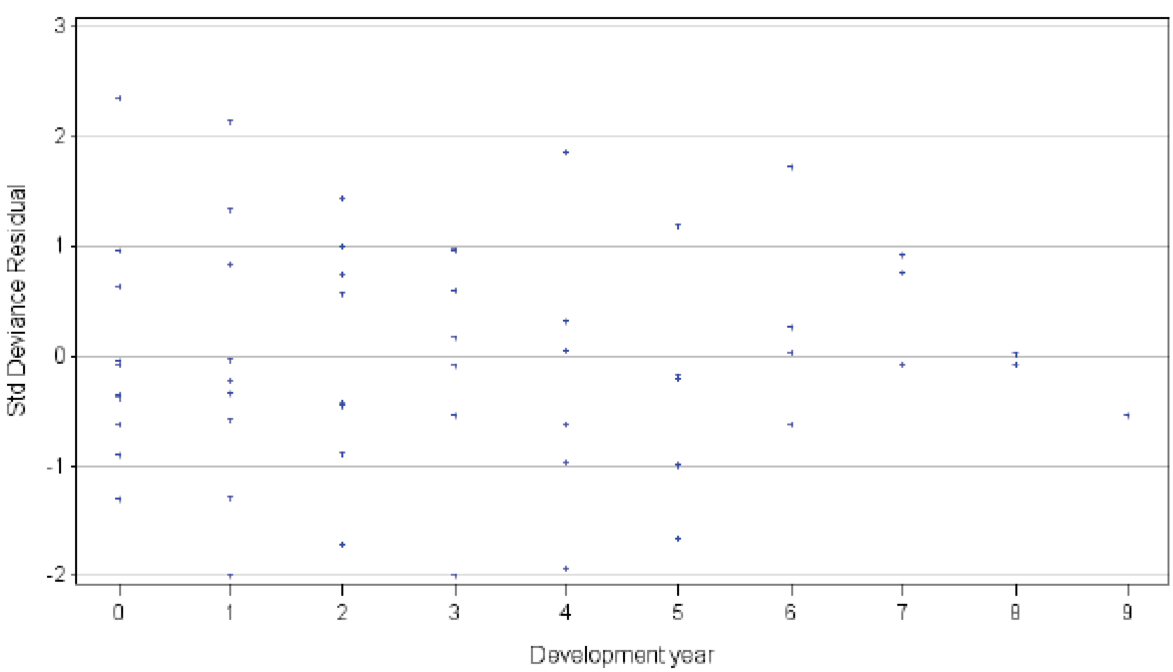


Figure 7-7. Residual Plot by Development Year for Model Including Interactions



been achieved, though perhaps with a slight hint of tapering variance as development year advances beyond about 6. This matter is not pursued further here.

7.7. Tail Smoothing and Extension

7.7.1. Tail Extension

Note that the range of development year has been extended to $j = 15$ in Figure 7-4. The figure indicates that the parametric form fitted to development year may be extended beyond the range of the data, providing a means of **tail extension**.

By (7-14), the linear response $x_i^T\beta$ is linear in $j \geq 8$ for fixed k, with gradient $b_1 + b_2 = -0.097$ (by Table 7-5). According to this model, the linear response decreases by 0.122 from each development year to the next in the tail.

The link function in this example is exponential:

$$E\left[Y_{kj}\right] = \mu_{kj} = exp\left(x_i^T\beta\right) \tag{7-15}$$

(see (7-2)), which implies that $E[Y_{kj}]$ decreases by a factor of $\exp(-0.097) = 0.908$ from each development year to the next in the tail; the tail is pure exponential.

It is necessary to recognize this form of tail extension for what it is, namely an extrapolation beyond the range of the data. In other words there is no direct evidence for the behavior imputed to the tail beyond development year 10, and one must accept the risks of this imputation.

On the other hand, the linear behavior of the linear predictor over the range $j = 8$, 9, 10 gives reasonable cause to believe that the linearity is likely to persist for the next few values of j. The extrapolation becomes steadily more speculative as one progresses to higher development years.

7.7.2. Tail Smoothing

One aspect of the chain ladder that is often problematic is irregularity in the progression of estimated age-to-age factors for the higher development years. As j approaches J in the case of a triangular data set ($J = K$), the number of observations contributing to the estimate $\hat{f}_j$ decreases, until at $j = J - 1$ the estimator (1-8) depends on only the two observations $X_{1,J-1}$ and $X_{1,J}$.

It is evident that parameter estimation on the basis of such a small sample is liable to lead to an estimate with a large standard error. A more reliable estimate might be obtained by the fit of a parametric form (such as (7-8)) to the higher development years j.

As it happens, this was unnecessary in the present example. The development year effects delivered by the unmodified chain ladder (see Table 3-1 or Figure 7-3) were quite smooth. However, other numerical examples would not have yielded such a fortunate result, and a device for smoothing the age-to-age factors for the higher development years would have been beneficial.

An example of this can be found in Table 3.1 of Taylor (2000), where the estimated higher age-to-age factors are as set in Table 7-8.

Table 7-8. An Example of Non-Smooth Age-to-Age Factors

Development Year	Estimated Age-to-Age Factor
10	1.028
11	1.014
12	1.009
13	1.008
14	1.009
15	1.001
16	1.002
17	1.001

7.8. Exposure-Based Methods

It is sometimes the case that there exists a time series $\{e_k\}$ by accident period to which the claims experience of accident period k is expected to be at least roughly proportional. For example, the average number of motor vehicle damage claims in a year would, all else being equal, be expected to be proportional to the number of vehicles insured in that year.

If such a time series can be identified then it may be used to improve the model through the additional (known) time series data. This time series is generally referred to as the exposure, and may be incorporated into the model by (in the case of the cross-classified ODP model) substituting e_k for α_k in (3-18):

$$\mu_{kj} = exp\left(ln\, e_k + ln\, \beta_j\right) \tag{7-16}$$

Note that a crucial difference between (3-18) and (7-16) is that $\{e_k\}$ is a known series whereas $\{\alpha_k\}$ is a series of parameters and must be estimated. In statistical parlance, ln e_k is an offset in the GLM.

Further, it may be shown (Frees and Derrig, 2014, Section 18.3.2) that the inclusion of an exposure offset in a log link model (such as the cross-classified model) results in the remainder of the model terms producing an estimate per unit of the exposure. For example, in a model of ultimate motor vehicle damage claim numbers, with number of vehicles as an offset, the model produces an estimate of claim frequency per vehicle.

As noted in Frees and Derrig (2014), there may be accident period effects in addition to the offset. Thus, (3-18) could be replaced by:

$$\mu_{kj} = exp\left(ln\, e_k + ln\, \alpha_k + ln\, \beta_j\right) \tag{7-17}$$

Simplifications to accident and development period effects as discussed in Sections 7.3 to 7.7 above apply as before, the only difference being that they would now operate per unit of exposure.

It is natural to think of exposure-based models for estimation of the ultimate numbers of claims reported in each accident period (i.e., including IBNR). However, such models are also useful for claim payments. Here, time series based on numbers of claims may be incorporated as an exposure measure to inform the payments model. Within Australian general insurance practice, such models are often used. These models include:

- Payments per claim incurred model (**"PPCI"**): $\{e_k\}$ = ultimate number of claims in accident period k. This model structure is conceptually very similar to the chain ladder model discussed in this monograph, except that the modeled payments are standardized for different numbers of claims incurred in each year. For lines of business with volatile numbers of claims, but similar average payments per claim, this model may be helpful.
- Payments per claim finalized model (**"PPCF"**): Here the time series varies by both accident and development period and is $\{e_{kj}\}$ = number of claims closed in accident period k and development period j. This model is suited to those lines of business where claims tend to settle as lump sums with closure and payment happening in the same cell of the triangle. In this case, the payments would be expected to relate to the number of claims closed in that cell. For example it may be useful for auto bodily injury claims or other liability claims.
- Payments per active claim (**"PPAC"**): As with the PPCF model, the time series varies by both accident and development period. In this case, $\{e_{kj}\}$ = number of active claims during accident period k and development period j. This model is suited to those lines of business where claims have ongoing payments for a number of years. An example would be weekly compensation payments from Workers' Compensation insurance.

Further discussion of the PPCI and PPCF models is given in Taylor (2000) and, in a GLM context, in Frees and Derrig (2014), and the interested reader is directed there. The PPAC model, which may also be referred to as the Payments per Claim Handled (**"PPCH"**) model is discussed in Sawkins (1979) and in Taylor (1986).

Both references given above for the PPCF model discuss the concept of **operational time**, where development period in a model is replaced by the proportion of claims that have finalized to date. This is a useful tool in situations where the rates of claims closure are not constant over time, perhaps due to changes in claims departments or in the wider environment. Operational time may easily be incorporated into a GLM as outlined in Frees and Derrig (2014).

7.9. Beyond a Single Triangle

The exposure measure for a model of ultimate claim numbers is usually a known quantity such as number of vehicles, policy years or wages (e.g., for workers compensation claims). However, the exposure-based payments models rely on counts of claim numbers (ultimate, closed, active) which are not fully known in advance. For example, numbers of claims in recent years may need to be adjusted for IBNR (Incurred but Not Reported) claims. Numbers of claims closed and active claims may be known in the past, but future numbers will require estimation.

Consequently, implementations of the PPCI model may involve two separate models:

1. A model of the ultimate number of claims so that IBNR numbers may be estimated;
2. An exposure-based model of the average payments per claim incurred.

Similarly, implementation of the PPCF model may require three separate models:

1. A model of the ultimate number of claims so that IBNR numbers may be estimated;
2. A model of the rate of closure of claims to allow the timing of future claims finalizations to be estimated;
3. An exposure-based model of the average payments per claim closed.

The prediction error of the PPCF model is the compound of the prediction errors of its component sub-models, and similarly for other models that consist of a number of sub-models. The decision on whether to use models such as these must take into account the additional prediction error introduced by each component and whether this is more than offset by the model's improved representation of the data. Detailed examination of this point may be found in Taylor and Xu (2016), while Taylor (2000) and Frees and Derrig (2014) provide more detail on implementing these models.

The double chain ladder (Martínez Miranda, Nielsen and Verrall, 2012) is another multi-model approach to the estimation of claims reserves. As the name suggests, two chain ladder models are used, one for reported claim numbers and the other for claim payments.

7.9.1. Bootstrapping a Compound Model

Bootstrapping a model such as the PPCI or PPCF is a straightforward extension of the bootstrap for a single triangle model. For each sub-model, n bootstraps are carried out. In the case of the average payments sub-model of the PPCF outlined above, the results of bootstrap b of this model are combined with the bootstrapped ultimate claim numbers from the b-th bootstrap of sub-model 1 and the claim closure pattern that results from the b-th bootstrap of sub-model 2. Further discussion of multiple bootstraps such as these is given in Taylor and Xu (2016).

Note that this process does not allow for correlations between the models apart from those that result from the forecasted value. For example, an increase in claim notifications might cause the finalization rate to slow down due to claims managers having greater numbers of claims to manage. Such an impact will not be captured in the bootstrap process outlined above. However, this type of change is arguably an aspect of model error (Section 4.5), and should be included in the allowance for that error. Scenario testing may also assist in estimating the impacts of such change.

7.10. Individual Models

Up to this point, the models discussed have assumed that the data are available in the form of triangles, such as that in Table 1-1. However, the data actually held by an insurance company will typically be in unit record form, with a considerable amount of information associated with each claim such as claimant information (date of birth, information relevant to the particular policy type such as age, employment, earnings, etc.), claim information (peril, date of accident, notification, finalization, etc.) and transactional details on payments to date. The use of highly summarized triangles, rather than the individual data detail dates back to general insurance practice before the availability of modern computing power, and the need to summarize claims experience into a convenient form for analysis.

This restriction no longer applies, and it is possible to consider the claims experience at an individual claims level. Portfolios may contain thousands or even hundreds of thousands of claims, each associated with a possibly large number of explanatory variables. Contrast this with the small number of observations in a yearly triangle, which is likely to eliminate a considerable amount of useful predictive information. In more technical language, the triangle may not be a sufficient statistic for the mass of detail available.

Currently, there are typically two broad classes of model used in reserving and related problems:

- **Aggregate** or **macro-models:** models applied to aggregate data summarized in arrays of triangular, or some other, shape, such as those described above—the chain ladder, cross-classified model, PPCI, PPCF, etc. The aggregated data is typically aggregated over accident and development periods; and
- **Individual claim, micro-models** or **granular models:** as the name suggests these are applied to individual claim data or to data summarized at a granular level.

The use of individual claim rather than aggregate models may lead to more efficient models.

The application of GLMs to individual claims data proceeds in much the same way as to summarized triangular data. For example, a model with accident and development period effects such as (3-18) can be fitted to individual data. The difference lies in the design matrix, X, where each row corresponds to an individual observation rather than to a triangle cell as it does in Section 3.3.2. Fitting trends by accident, development and calendar periods and model validation proceeds in much the same way as before, the difference being that there are many more data points to inform the modeling process.

Merely fitting the same GLM to individual claim data as was fitted to the aggregate data (triangle) may not produce a markedly different model. However, the use of individual claims data opens up the possibility of using a number of claimant and claim related data as explanatory variables to refine estimates of average claim size. Taylor, McGuire and Sullivan (2008) classify explanatory variables as follows:

- Static variables: constant over the life of a claim (e.g., gender, pre-injury earnings);
- Dynamic variables: these may change over the life of a claim. Dynamic variables may be further categorized as:
  - Time variables: these relate to the passage of time and are therefore future values are known with certainty (e.g., development period, calendar period);
  - Unpredictable variables: future changes in these values are not predictable with certainty (e.g., time until a claim closes, spells off work).

It is evident that, if any unpredictable variables are included in a model, then any forecast of that model will require forecasts of these variables. As noted in Section 7.9, which discusses the same concept for aggregate data models, any decision on the inclusion of an unpredictable variable in a model must offset the increase to the prediction error from use of this variable due to its stochastic nature against the resulting decrease in prediction error due to more accurate modeling.

Consequently, individual reserving models tend to lie on a spectrum from those models with time variables only to models with all types of predictors including unpredictable variables.

Taylor and McGuire (2004) discuss an individual claims reserving model that lies towards the simpler end of the spectrum. This is a model of the average size of auto bodily injury claims, which depends on the time variable accident period and functions of the unpredictable variable, development time until closure of a claim.

McGuire (2007) describes an update to this model where the use of claim severity is found to greatly increase the predictive power of the model. Micro-models are also discussed in detail by Pigeon, Antonio and Denuit (2013) and Antonio and Plat (2014).

At the other end of the spectrum lies the class of individual claims models referred to as Stochastic Case Estimate (**"SCE"**). These are intended to provide estimates of ultimate costs of individual claims that are alternatives to the physical or manual case estimates assigned by claims experts. As such, a model with high discriminatory power is to be preferred and in general, this is achieved by considering a large number of predictors. Further details on the construction of SCE models may be found in Taylor and Campbell (2002), Brookes and Prevett (2004) (which both relate to Australian workers' compensation insurance) and Taylor, McGuire and Sullivan (2008) which applies an SCE to US medical malpractice. The latter paper also includes some discussion of applying a bootstrap to such models.

7.11. Bayesian Models

Although **Bayesian models** and related methods such as **Markov Chain Monte Carlo** (**"MCMC"**) are beyond the scope of this monograph, it is noted that they are increasingly used for stochastic reserving models.

Each GLM considered to this point of the present monograph is non-Bayesian in that its parameters are treated as fixed, though unknown, quantities. It can be transformed into a Bayesian model by representing each unknown parameter as a random quantity deriving from a particular statistical distribution. Put in an alternative manner, a Bayesian model for a particular quantity seeks to estimate the *posterior* distribution of that quantity based on *prior* distributions for the model parameters and the *likelihood* based on observed data.

In many ways, the Bayesian paradigm seems a natural fit to insurance-type problems. The prior distributions of the parameters may be used to codify expert knowledge or *a priori* expectations, and combine this in an objective manner with emerging experience. The similarities with credibility theory are apparent.

For many years, Bayesian analysis was limited for computational reasons; users were forced to restrict themselves largely to combinations of prior distributions and likelihoods that led to closed form analytic solutions (conjugate priors). That changed with the advent of MCMC methods into the wider statistical community, which enabled simulation of full distributions from any posterior distribution. For insurance problems, MCMC enables the modeler to combine a priori knowledge with emerging experience to produce a full distribution of the stochastic reserves.

There have been many papers in the actuarial literature discussing Bayesian models and MCMC, of which a small sample is referenced here. Verrall (2000, 2004), England and Verrall (2002, 2006), England and Verrall (2006), Wüthrich (2007), England, Verrall and Wüthrich (2012) and Taylor and Xu (2016) present various Bayesian models, most of them Bayesian versions of the chain ladder. Scollnik (2001 and 2002), Ntzoufras and Dellaportas (2002), Meyers and Shi (2011), amongst others, describe the implementation of MCMC for insurance data.

All modeling approaches discussed up to this point consist of specification of a particular model, possibly Bayesian but always with a fixed number of parameters, and then estimation of those parameters. More recently, **reversible jump MCMC** (**"RJMCMC"**) methodology has been introduced as a framework containing a complete family of models with differing numbers of parameters. The calibration step then consists of selection of a specific model from the family, as well as estimation of its parameters. A strength of RJMCMC is that it enables the modeler to consider a number of different models simultaneously. For example. Ntzoufras, Katsis and Karlis (2005) use RJMCMC to fit and choose between different models for claims count data, while Verrall and Wüthrich (2012) and Verrall, Hössjer and Björkwall (2012) consider the smoothing of the development period curve in a Bayesian ODP model, allowing RJMCMC to choose the cut-off development period at which parametric functions are used rather than the individual development period parameters.

8. Conclusion

This monograph commenced with the application of the conventional chain ladder algorithm to a data set (Section 1.5). The application was non-stochastic, as is so often the case in practice.

Certain stochastic models were then identified as producing precisely the same forecast as the conventional algorithm (Section 3.3). The stochastic view regards the quantum of outstanding losses as a random variate, and the forecast as an estimate of the mean value of that variate. The stochastic models enable the estimation of the entire distribution of outstanding losses.

The "chain ladder algorithm", as defined here, is absolutely rigid, with no scope for variation according to any eccentricities in the data to which it is applied. In practice, actuaries typically make a number of adjustments to it, such as calibration of the model on the basis of data of only recent years, or limiting in some way the influence of outlying observations.

It was shown (Section 3.4) that some of these adjustments could be formulated within the stochastic models. In consequence, the stochastic model could be made to parallel those used in practice while retaining its ability to estimate the entire distribution of outstanding losses.

Finally, Chapter 7 examined variations of the model that could not be made within the conventional chain ladder framework, but only within the formal stochastic model formulation. These variations explored the much discussed matter of whether or not the conventional chain ladder is over-parameterized, with the degradation of predictive power that comes with over-parameterization.

These model variations took two forms. First, the manner in which accident year was represented as influencing expected paid losses in individual cells of the claim triangle was changed from a separate factor for each accident year to a parametric function of accident year. For example, it was found possible to represent the effects of the 10 separate accident years by a function of only 3 parameters, rather than the 10 parameters required by the conventional chain ladder. The parameterization of development year was similarly reduced.

The second form of model variation introduced was the introduction of interactions. The conventional chain ladder assumes that age-to-age factors are independent of accident year. Frequently, this assumption is violated by data triangles encountered in practice. Violations may be highly localized, affecting only a handful of cells, or they may consist of longer term systematic changes, such as trending age-to-age factors.

In any event, if model interactions are warranted but ignored in the modeling (such as inevitably occurs in the application of the conventional chain ladder), then estimates of accident and development year effects will be distorted.

These changes produce two beneficial results. First, they improve the goodness-of-fit of the model. Second, they reduce the associated prediction error. The end result observed in Table 7-6 was a 17% reduction in prediction error solely by virtue of inclusion of the interactions.

The final prediction error was 23% less than that associated with the conventional chain ladder. It is emphasized that all of these modifications of the conventional chain ladder model are achievable within a GLM framework but not by the conventional approaches that depend essentially on row and column sums or averages.

The chapter concluded by giving an overview of models beyond the chain ladder, discussing exposure-based models (both as a single model, or a model consisting of a number of sub-models in cases where claim numbers form the exposure) and micro- (or granular or individual claim) models which include Stochastic Case Estimate models. A brief introduction to Bayesian models was also provided for the reader's interest.

In summary then, it has been shown that the chain ladder, together with some common variations of it, can be expressed in GLM form. Then it has been further shown that the GLM structure may be extended to a more statistically efficient model in ways that are not achievable without the GLM (or perhaps some other model of a similar level of sophistication).

In the process one has progressed from a heuristic algorithm to a fully stochastic model with diagnostics that are adequate to determine whether that model is a reasonable representation of the data. Further, since the model is fully stochastic, it is capable of producing the full stochastic properties of its forecasts, including prediction error, quantiles, etc.

That is, the GLM is capable of anything of which the conventional chain ladder is capable, but the GLM is capable of many things of which the conventional chain ladder is not.

References

Antonio, K., and R. Plat. 2014. "Micro-Level Stochastic Loss Reserving for General Insurance." *Scandinavian Actuarial Journal* 2014 (7): 649–669.

Augustin, N., E.-A. Sauleau, and S. Wood. 2012. "On Quantile Quantile Plots for Generalized Linear Models." *Computational Statistics and Data Analysis* 56 (8): 2404–2409.

Benktander, G. 1979. "Inflation and Insurance." In Kahane, Y. 1979. *New Frontiers in Insurance: Theory and Practice.* Erhard Center for Higher Studies and Research in Insurance, The Graduate School of Management, Tel Aviv University, Israel.

Bornhuetter, R. L., and R. E. Ferguson. 1972. "The Actuary and IBNR," *Proceedings of the Casualty Actuarial Society* 59: 181–195.

Brookes, R., and M. Prevett. 2004. "Statistical Case Estimation—An Overview of the NSW WorkCover Model." Institute of Actuaries of Australian Xth Accident Compensation Seminar. http://actuaries.asn.au/Library/Statistical_Brookes_Prevett_PPT.pdf.

Buchwalder, M., H. Bühlmann, M. Merz, and M. V. Wüthrich. 2006. "The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited)." *ASTIN Bulletin* 36: 521–542.

Cox, D. R., and D. V. Hinckley. 1974. *Theoretical Statistics.* London: Chapman and Hall.

Efron, B. 1979. "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics* 7: 1–26.

England, P. D., and R. J. Verrall. 2002. "Stochastic Claims Reserving in General Insurance." *British Actuarial Journal* 8 (iii): 443–518.

England, P. D., and R. J. Verrall. 2006. "Predictive Distributions of Outstanding Liabilities in General Insurance." *Annals of Actuarial Science* 1: 221–270.

England, P. D., R. J. Verrall, and M. V. Wüthrich. 2012. "Bayesian Over-dispersed Poisson Model and the Bornhuetter and Ferguson Claims Reserving Method." *Annals of Actuarial Science* 6 (2): 258–281.

Frees, E. W., and R. A. Derrig (eds.). 2014. "Predictive Modeling Applications in Actuarial Science" Volume 1, *Predictive Modeling.* New York: Cambridge University Press.

Hachemeister, C. A., and J. N. Stanard. 1975. "IBNR Claims Count Estimation with Static Lag Functions." Spring Meeting of the Casualty Actuarial Society.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference and Prediction.* New York: Springer.

Jorgensen, B., and M. C. Paes de Souza. 1994. "Fitting Tweedie's Compound Poisson Model to Insurance Claims Data." *Scandinavian Actuarial Journal* 1994 (1): 69–93.

Kahane, Y. 1979. *New Frontiers in Insurance: Theory and Practice.* Erhard Center for Higher Studies and Research in Insurance, The Graduate School of Management, Tel Aviv University, Israel.

Kendall, M. G., and A. Stuart. 1977. *The Advanced Theory of Statistics,* vol. 1. 4th ed. New York: Macmillan.

Lai, T. L., and M.-C. Shih. 2003. "Nonparametric Estimation in Nonlinear Mixed Effects Models." *Biometrika* 90 (1): 1–13.

Lindsey, J. K. 1997. *Applying Generalized Linear Models.* New York: Springer-Verlag.

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

Mack, T., and G. Venter. 2000. "A Comparison of Stochastic Models that Reproduce Chain Ladder Reserve Estimates." *Insurance: Mathematics and Economics* 26 (1): 101–107.

Martínez Miranda, M. D., H. P. Nielsen, and R. Verrall. 2012. "Double Chain Ladder." *ASTIN Bulletin* 42(1): 59–76.

McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models.* 2nd ed. Boca Raton: Chapman and Hall.

McGuire, G. 2007. "Individual Claim Modelling of CTP Data." Institute of Actuaries of Australia XIth Accident Compensation Seminar, Melbourne, Australia. http://actuaries.asn.au/Library/6.a_ACS07_paper_McGuire_Individual%20claim%20modellingof%20CTP%20data.pdf.

Meyers, G. G., and P. Shi. 2011. "The Retrospective Testing of Stochastic Loss Reserve Models." *Casualty Actuarial Society E-Forum,* Summer 2011. http://www.casact.org/pubs/forum/11sumforum/Meyers-Shi.pdf.

Nelder, J. A., and R. W. M. Wedderburn. 1972. "Generalised Linear Models." *Journal of the Royal Statistical Society,* Series A, 135: 370–384.

Ntzoufras, I., and P. Dellaportas. 2002. "Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty" (with discussion). *North American Actuarial Journal* 6: 113–128.

Ntzoufras, I., A. Katsis, and D. Karlis. 2005. "Bayesian Assessment of the Distribution of Insurance Claim Counts Using Reversible Jump MCMC." *North American Actuarial Journal* 9: 90–108.

O'Dowd, C., A. Smith, and P. Hardy. 2005. "A Framework for Estimating Uncertainty in Insurance Claims Cost." XVth General Insurance Seminar, 16–19 October 2005. Institute of Actuaries of Australia. http://www.actuaries.asn.au/Library/gipaper_odowd-smith-hardy0510.pdf.

Peters, G. W., P. V. Shevchenko, and M. V. Wüthrich. 2009. "Model Uncertainty in Claims Reserving within Tweedie's Compound Poisson Models." *ASTIN Bulletin* 39: 1–33.

Pierce, D. A., and D. W. Schafer. 1986. "Residuals in Generalised Linear Models." *Journal of the American Statistical Association* 81: 977–986.

Pigeon, M., K. Antonio, and M. Denuit. 2013. "Individual Loss Reserving with the Multivariate Skew Normal Model." http://ssrn.com/abstract=1996455.

Renshaw, A. E., and R. J. Verrall. 1998. "A Stochastic Model Underlying the Chain-Ladder Technique." *British Actuarial Journal* 4 (iv): 903–923.

Sawkins, R. W. 1979. "Methods of Analysing Claim Payments in General Insurance." *Transactions of the Institute of Actuaries of Australia,* 435–519.

Scollnik, D. P. M. 2001. "Actuarial Modeling with MCMC and BUGS." *North American Actuarial Journal* 5(2): 96–125.

Scollnik, D. P. M. 2002. "Implementation of Four Models for Outstanding Liabilities in WinBUGS: A Discussion of a Paper by Ntzoufras and Dellaportas." *North American Actuarial Journal* 6: 128–136.

Schmidt, K. D., and A. Wünsche. 1998. "Chain Ladder, Marginal Sum and Maximum Likelihood Estimation." *Blätter der Versicherungsmathematiker* 23: 267–277.

Shibata, R. 1997. "Bootstrap Estimate of Kullback-Leibler Information for Model Selection." *Statistica Sinica* 7: 375–394.

Straub, E. 1988. *Non-Life Insurance Mathematics.* Berlin: Springer-Verlag.

Taylor, G. C. 1986. *Claims Reserving in Non-Life Insurance.* New York: Elsevier.

Taylor, G. 2000. *Loss Reserving: An Actuarial Perspective.* Dordrecht, Netherlands: Kluwer Academic Publishers.

Taylor, G. 2009. "The Chain Ladder and Tweedie Distributed Claims Data." *Variance* 3: 96–104.

Taylor, G. 2011. "Maximum Likelihood and Estimation Efficiency of the Chain Ladder." *ASTIN Bulletin* 41: 131–155.

Taylor, G. 2015. "Bayesian Chain Ladder Models." *ASTIN Bulletin* 45(1): 75–99.

Taylor, G., and M. Campbell. 2002. "Statistical Case Estimation." Research paper No 104 of the Centre for Actuarial Studies, University of Melbourne. http://fbe.unimelb.edu.au/_data/assets/pdf_file/0005/806396/104.pdf.

Taylor, G., and G. McGuire. 2004. "Loss Reserving with GLMs: A Case Study." Casualty Actuarial Society 2004 Discussion Paper Program, 327–392.

Taylor, G., G. McGuire, and J. Sullivan. 2008. "Individual Claim Loss Reserving Conditioned by Case Estimates," *Annals of Actuarial Science* 3(1&2): 215–256.

Taylor, G., and J. Xu. 2016. "An Empirical Investigation of the Value of Finalisation Count Information to Loss Reserving," Variance (in press).

Tweedie, M. C. K. 1984. "An Index which Distinguishes between Some Important Exponential Families," In *Statistics: Applications and New Directions, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference,* J. K. Ghosh and J. Roy (Eds.), Indian Statistical Institute, 1984: 579–604.

Verrall, R. J. 2000. "An Investigation into Stochastic Claims Reserving Models and the Chain-Ladder Technique." *Insurance: Mathematics and Economics* 26(1): 91–99.

Verrall, R. J. 2004. "A Bayesian Generalised Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving." *North American Actuarial Journal* 8: 67–89.

Verrall, R. J., O. Hössjer, and S. Björkwall. 2012. "Modelling Claims Run-Off with Reversible Jump Markov Chain Monte Carlo Methods." *ASTIN Bulletin* 42: 35–58.

Verrall, R. J., and M. Wüthrich. 2012. "Reversible Jump Markov Chain Monte Carlo Method for Parameter Reduction in Claims Reserving." *North American Actuarial Journal* 16 (2): 240–259.

Wüthrich, M. V. 2007. "Using a Bayesian Approach for Claims Reserving." *Variance* 1: 292–301.

Author information

Greg Taylor

School of Risk and Actuarial Studies, UNSW School of Business, UNSW Australia, NSW 2052, Australia
gregory.taylor@unsw.edu.au

Gráinne McGuire

Taylor Fry Consulting Actuaries, 55 Clarence Street, Sydney NSW 2000, Australia
grainne.mcguire@taylorfry.com.au

Stochastic Loss Reserving Using Generalized Linear Models

Errata

Page 7, last line of second paragraph: should read "$Y_{k1}(= X_{k1})$".

Page 7, last dot point: replace with the following:

Cape Cod forecast: $B_k = P_k \sum_{i=1}^{K} P_i \omega_i \left[\left(X_{i,K-i+1} + \hat{R}_i\right)/P_i\right] / \sum_{i=1}^{K} P_i \omega_i$ with $\omega_i = 1/\hat{f}_{K-i+1} \dots \hat{f}_{J-1}$.

Page 9, Table 2-1. In the "Inverse Gaussian" row, under the heading $b(\theta)$, the entry $-(-2\theta)^{-½}$ should be $-(-2\theta)^{½}$.

Page 9, sentence immediately following Table 2-1. Add "where n and v are additional parameters providing alternative representations of ϕ".

Page 9, equation (2-5). The factor $\alpha(\phi)$ should be $a(\phi)$.

Page 10, equations (2-12) and (2-13). These are incorrect, and should be deleted. Equation (2-9) holds for $p \neq 1{,}2$, and (2-10) holds for $p \neq 1$. However, in these cases, the form of variance function implies the following:

For $p = 1, b(\theta) = e^{\theta}, \mu = b'(\theta) = e^{\theta}$.

For $p = 2, b(\theta) = -ln(-\theta), \mu = b'(\theta) = -1/\theta$.

Page 11, Table 2-2. In the "Gamma" row, under the heading $b(\theta)$, the entry $ln(-\theta)$ should be $-ln(-\theta)$.

Equation (2-15): Replace by $exp\ c(y, \varphi) = \varphi^{-y/\varphi}[(y/\varphi)!]^{-1}$.

Equation (2-16): Replace by $\pi(y; \mu, \phi) = \frac{(\mu/\phi)^{y/\phi} exp\,(-\mu/\phi)}{(y/\phi)!}$.

Page 29, equation (3-12) require correction in sympathy with the correction to (2-16): replace the term $ln\left(f_{j-1} - 1\right)$ by $ln\left(\frac{f_{j-1}-1}{\phi_{j-1}/X_{k,j-1}}\right)$.

Page 30, 3 lines after equation (3-14): Definition of β should be

$$\beta = (f_1 - 1, f_2 - 1, \dots, f_9 - 1)^T.$$

Page 49, equation (5-21): Replace by

$$\varepsilon_{proc}^{*} = \tilde{Y}_{proc} - \hat{Y}.$$

ABOUT THE SERIES:

CAS monographs are authoritative, peer-reviewed, in-depth works focusing on important topics within property and casualty actuarial practice. For more information on the CAS Monograph Series, visit the CAS website at www.casact.org.

Expertise. Insight. Solutions.

www.casact.org

ESTIMATING THE PREMIUM ASSET ON RETROSPECTIVELY RATED POLICIES

MICHAEL T. S. TENG AND MIRIAM E. PERKINS

Abstract

This paper presents a method for estimating the premium asset on retrospectively rated policies, using the functional relationship between the losses and the retrospective premium. This relationship is examined using the historical premium and loss development data and the retro rating parameters sold in the underlying policy. The cumulative ratio of premium development to loss development, when applied to the expected future loss emergence, gives the expected future premium development on the retro rated policies. The sum of all future premium development is the premium asset.

ACKNOWLEDGEMENT

The authors wish to thank Eric Brosius and Roy Morell for reviewing the technical soundness of the method presented in this paper, and Eugene Connell and members of the CAS Editorial Committee for providing editorial assistance.

1. INTRODUCTION

On retrospectively rated policies, premium that the insurer expects to collect based on the expected ultimate loss experience, less the premium that the insurer has already booked, is called the premium asset. Many insurers call this the Earned But Not Reported premium (EBNR). The admitted portion of the premium asset appears on the balance sheet as the "Asset for Accrued Retrospective Premiums."

In recent years, retro rated policies have become popular for several reasons.

1. A retro rated policy returns premium to the insured for good loss experience. This feature is attractive for a customer who anticipates favorable loss experience through loss control and loss management. By offering retro rated policies, the insurer may be able to attract these good customers.

2. A growing number of commercial insurance buyers are taking advantage of the cash flow feature in a retro rated policy. A retro rated policy allows the insured to pay premium as losses are reported or paid, depending on the contract, rather than paying all premiums up front. This allows the insured to hold on to cash longer.

3. Inflation, rate regulations, uncertainty in claims compensability, increasing utilization of the insurance benefits, and growing attorney involvement have made the cost of insurance much harder to predict today than in the past. Since the premium for a retro rated policy varies directly with the insured's actual loss experience, writing retro policies allows an insurer to shift a large portion of the actual risk to the insured. This makes the insurer more willing to write insurance.

As a result of the growth of retro rated policies, estimating the premium asset for them is a growing need for many commercial lines insurers. This asset frequently exceeds 10% of surplus. Despite the growing importance of the premium asset, there have been few articles written on this subject. Berry [1] and Fitzgibbon [2] have presented methods of calculating the "retro reserve," defined as the difference between the *premium deviation to date* and the *ultimate premium deviation*.[1] The retro reserve is the negative equivalent of the premium asset referred

[1]The ultimate premium deviation is the amount by which the ultimate premium for a retro rated policy is expected to differ from the standard premium (manual premium adjusted for experience rating). The premium deviation to date is the amount by which the currently booked premium differs from the standard premium.

to in this paper. Their approach is to analyze the historical relationship between the loss ratio and the premium deviation using statistical techniques, and then apply such a relationship to the projected loss ratio to calculate a projected ultimate premium deviation. This ultimate premium deviation is then reduced by the premium deviation to date to produce the retro reserve. Berry uses a second approach, which is to estimate ultimate premium using the historical premium emergence pattern, and then subtract current premium to get the retro reserve.

While the statistical methods presented in [1] and [2] may be theoretically sound, they lack intuitive appeal, particularly as they relate to how a retro rating formula actually works. On a retro rated policy, premium is calculated as a function of loss. This function is composed of retro rating parameters such as the loss conversion factor, tax multiplier, retro minimum, and retro maximum; they define how much premium an insurer can collect given a certain amount of loss. Therefore, the premium asset on a retro rated policy should be established as a function of reported losses and the reserve for loss development, where this function is defined by the retro rating parameters.

This paper will present, through an example, a method of calculating the premium asset as a function of current losses, expected future loss emergence, and the retro rating parameters. Specifically, the method looks at how premiums develop as losses develop. The relationship can be expressed as the ratio of premium development to loss development, referred to here as the PDLD ratio. There are two methods of calculating the PDLD ratio—from historical premium and loss development data, and from the retro rating parameters. The latter approach will be developed first, and will be followed by the calculation of the PDLD ratios from historical data. Once the relationship between premium and loss is determined, it can be applied to the expected future loss development to get the expected future premium development. The sum of all future premium development is the premium asset.

This method applies only to retro rated polices (or similar loss sensitive rating plans), and not to prospectively rated policies. There may be a premium asset on prospectively rated polices due to changes in exposure, but this topic will not be discussed here. This method is intended to be applied to an aggregate book of business, or large segment of a book of business, rather than at the individual policy level.

2. THE FORMULA APPROACH TO CALCULATING PDLD RATIOS

The first step is to derive the formula for a PDLD ratio. This starts with the first retro adjustment. On a retro rated policy, the premium calculation is based on a retro formula. A commonly used formula is

$$\mathrm{P}_n = [\mathrm{BP} + (\mathrm{CL}_n \times \mathrm{LCF})] \times \mathrm{TM}, \tag{2.1}$$

where

P_n = Premium at the n^{th} retro adjustment,

BP = Basic premium,

CL_n = Capped loss at the n^{th} adjustment[2],

LCF = Loss conversion factor, and

TM = Tax multiplier.

For example, P_1 denotes the premium computed for the first retro adjustment; P_2 denotes the premium computed for the second retro adjustment. Note that BP, LCF, and TM typically stay the same throughout all retro adjustments. For a more thorough discussion of the retro rating formula, see Gillam and Snader [3].

Using formula (2.1) and denoting L_1 as the amount of loss developed for the first retro adjustment, the first PDLD ratio

[2]Losses that contribute to additional premium: these are total losses subject to a minimum and a maximum amount corresponding to the plan minimum and maximum premiums. Individual claims may also be capped by a per accident limitation, which limits the adverse impact of any single large claim on the premium calculation.

can be stated as follows:

$$\begin{aligned} P_1/L_1 &= [\mathrm{BP} + (\mathrm{CL}_1 \times \mathrm{LCF})] \times \mathrm{TM}/L_1 \\ &= [(\mathrm{BP}/L_1) \times \mathrm{TM}] + [(\mathrm{CL}_1/L_1) \times \mathrm{LCF} \times \mathrm{TM}]. \quad (2.2) \end{aligned}$$

The first term of this formula is $(\mathrm{BP}/L_1) \times \mathrm{TM}$. This is basic premium divided by the loss emerged for the first retro adjustment times the retro tax multiplier. One can approximate this as

$$\mathrm{BP} \times \mathrm{TM}/(\mathrm{SP} \times \mathrm{ELR} \times \%\mathrm{Loss}_1), \qquad (2.3)$$

where

SP = Standard premium,[3]

ELR = Expected loss ratio

= Expected ultimate loss divided by standard premium, and

$\%\mathrm{Loss}_1$ = Expected percentage of loss emerged for the first adjustment.

Formula 2.3 is equivalent to $(\mathrm{BP}/\mathrm{SP}) \times \mathrm{TM}/(\mathrm{ELR} \times \%\mathrm{Loss}_1)$, which is the basic premium factor in a retro rating formula times the tax multiplier, divided by the expected loss ratio emerged for the first retro adjustment. The expected loss ratio for the first retro adjustment would depend on the ultimate expected loss ratio and the percentage of losses emerged at the first adjustment. Typically, losses emerged as of 18 months are used to compute the first retro adjustment.

In Formula 2.2, the term CL_1/L_1 is the ratio of capped losses to uncapped losses. This ratio is referred to as the *loss capping ratio*. Capped losses are losses that contribute to an additional

[3]Manual premium adjusted for experience rating.

premium. Any change in loss, where total loss exceeds the minimum and is below the maximum, will result in additional premium. Conceptually one can view the difference between the capped loss (CL) and the uncapped loss (L) as the portion of loss outside the boundaries of the retro maximum and minimum. On plans that cap the losses with a per accident loss limit, the capped loss would also exclude the losses exceeding this limit, since they do not contribute to additional premium. The loss capping ratio usually decreases as the data becomes more mature. This is because an increasing portion of the loss development occurs outside of loss limitations. The loss capping ratio can be derived by comparing the capped and the uncapped loss development, if such data are available; often they are not. In this paper, the loss capping ratio is derived using a loss ratio distribution. Because the explanation of this method is somewhat detailed, it is presented after the example of the PDLD ratio calculation, in Section 5.

If the loss data used is already capped (i.e., L_n equals CL_n for all n), then the loss capping ratio will be one. Otherwise, this ratio will have to be estimated. The example assumes that the loss capping ratio is 0.85 for losses developed through the first retro adjustment. This means that 15 percent of the losses developed through the first retro adjustment are eliminated by the net effect of the retro maximums, minimums, and per accident limitations.

To show how Formula 2.2 can be used to estimate the PDLD ratio, the example assumes the following retro rating parameters:

$$\text{Basic premium factor} = 0.20$$

$$\text{Expected loss ratio} = 0.70$$

$$\text{Loss conversion factor} = 1.20$$

$$\text{Tax multiplier} = 1.03$$

$$\%\text{Loss}_1 = 78.4\%.$$

These retro rating parameters may be computed as the average of the sold retro parameters. Substituting these values into Formula 2.2, one gets a PDLD ratio for the first retro adjustment of

$$[0.20 \times 1.03/(0.70 \times 78.4\%)] + (0.85 \times 1.20 \times 1.03) = 1.426.$$

The PDLD ratio for the second retro adjustment period refers to the *incremental premiums* developed between the first and the second retro adjustments, divided by the *incremental losses* developed between these two adjustments. Typically, successive retro adjustments occur at one year intervals. One can view the PDLD ratio for the second retro adjustment period as the ratio of the *change in premium* divided by the *change in loss*. Algebraically, this equals

$$\begin{aligned}(\mathrm{P}_2 - \mathrm{P}_1)/(\mathrm{L}_2 - \mathrm{L}_1) \\ &= (\mathrm{CL}_2 - \mathrm{CL}_1) \times \mathrm{LCF} \times \mathrm{TM}/(\mathrm{L}_2 - \mathrm{L}_1) \\ &= [(\mathrm{CL}_2 - \mathrm{CL}_1)/(\mathrm{L}_2 - \mathrm{L}_1)] \times \mathrm{LCF} \times \mathrm{TM}. \qquad (2.4)\end{aligned}$$

This example assumes an incremental loss capping ratio of 0.58 for the second retro adjustment period. Substituting this loss capping ratio and the retro rating parameters into Formula 2.4, one gets a PDLD ratio of $0.58 \times 1.20 \times 1.03 = 0.717$. The PDLD ratios for the third and subsequent retro adjustments are calculated in a similar manner.

The advantage of using the retro formula to estimate the PDLD ratio is that it responds to changes in the retro rating parameters that are sold, whereas the PDLD ratios derived from the historical data may not be indicative of the future PDLD ratios. If the retro rating parameters change significantly over time, one should give more weight to the PDLD ratios derived by formula than those derived from the historical data. A summary of the formula PDLD ratios is shown in Exhibit 4, Part 2.

When possible one should retrospectively test the PDLD ratios derived by formula against actual emergence in the subsequent retro adjustment periods to determine if any bias exists. A possible source of bias is the use of average parameters for the LCF, tax multiplier, maximum, minimum, and per accident limitation. One should study the appropriateness of the selections and adjust them as necessary. Such a study could lead to better parameter selections and more accurate premium estimates.

3. THE EMPIRICAL APPROACH TO CALCULATING PDLD RATIOS

The use of empirical data is another way to calculate the PDLD ratios. Two types of data are needed for the empirical approach: booked premium development and reported loss development.[4] For the example presented in this paper, premium booked by policy effective quarter by valuation quarter is displayed in Exhibit 6 and reported loss data is shown in Exhibit 7. The calculation of the PDLD ratios is shown in Exhibit 4. The PDLD ratio after the sixth retro adjustment is selected at zero, which assumes that there are no further retro adjustments.[5]

Data should be segregated into homogeneous groups by size of account and by the type of rating plan sold. When appropriate, other criteria should be used in grouping the data. Policies are grouped based on the calendar quarter in which they became effective. These groups will be referred to as policy effective quarters. The first policy effective quarter of 1994 will be

---

[4]Booked premium on a retro rated policy is the premium computed using the retro rating formula and the most recent loss valuation. Reported loss is the amount of loss that has been reported to the insurer. It does not include future loss development for unreported claims, for such losses are often not entered into the premium calculation.

[5]The NCCI and ISO retrospective rating manuals prescribe a maximum premium adjustment period of 3 to 4 years. The actual maximum adjustment period varies from one retro policy to another. A maximum premium adjustment period of six years is common among major commerical line retro policies. However, due to increasing uncertainty of loss costs and growing usage of cash flow financing of premiums, retro policies will probably be written with longer premium adjustment periods in the future.

denoted as 1994.1, the second quarter will be denoted as 1994.2, and so on.

The first retro premium computation is usually based on losses developed through 18 months. However, it takes time to do the retro calculation and to record adjusted premiums. This paper assumes that due to time lags in processing and recording, premiums are recorded 3 to 9 months following the recording of losses. Therefore, it is assumed that premiums booked through 27 months are the result of the first retro adjustment. Since retro adjustments are usually done in annual intervals, premiums recorded through 39 months would be the result of the second retro adjustment, using losses evaluated at 30 months. Premiums recorded through 51 months would be the result of the third retro adjustment, using losses evaluated at 42 months, and so on. In practice, the actual length of the retro adjustment period and the premium booking lag may vary from one insurer to another.

The PDLD ratio for the first retro adjustment equals premiums booked through 27 months divided by losses reported through 18 months. At the first retro adjustment period, the PDLD ratio indicated by an overall average of the historical data is 1.460 (see Exhibit 4, Part 1). However, there is an upward trend in the responsiveness of premium to loss over the latest several policy quarters and these PDLD ratios are higher than the historical average. Such a trend could be the result of more liberal retro rating parameters (higher maximum, minimum, or per accident limitation), but this is probably not the case here since the PDLD ratio calculated by formula is 1.426 and it reflects the plan parameters currently being sold. A more likely explanation for the trend is an improvement in loss experience, either due to chance or to known changes in the system such as workers compensation reform. A larger portion of the loss is within the boundaries of the retro maximum and the per accident limitation, resulting in more additional premium per dollar of loss. The formula approach will not reflect a change in loss experience unless the formula is revised. (This revision is discussed in Section 5.) In recognition of these changing conditions, a PDLD ratio of 1.750 was selected for the first adjustment.

The PDLD ratio for the second retro adjustment period is the *incremental premiums* developed between the first and the second retro adjustments divided by the *incremental losses* developed between these two adjustments. It is assumed that losses developed through 30 months are used to calculate the premiums for the second retro adjustment and that the resulting premiums are booked at the 39 month valuation. The selected PDLD ratio from historical data is 0.700, which is close to the formula ratio of 0.717. The PDLD ratios from the two methods also compare closely at the third adjustment.

The historical PDLD ratios may fluctuate significantly after the first retro adjustment period. This is because the premium and loss development on a few policies can be a large component of the total incremental development on policy quarter data. Historical PDLD ratios for an individual policy quarter could even be negative in spite of upward aggregate loss development—this could happen when there is upward development in high loss layers (resulting in no additional premium) and downward development (and return premium) on layers that are still within loss limitations. Where the historical PDLD ratios fluctuate significantly, one should use an average of as many historical data points as possible. In situations like this, the PDLD ratios derived by formula may provide a better indication of the relationship between premium and loss.

In the example, the historical and formula PDLD ratios begin to diverge after the third retro adjustment period. Several factors could be contributing to this. First, since the historical ratios are lower than the formula ratios, worse than expected loss experience during the mid-1980s may have caused a larger portion of the loss to be outside the boundaries of the retro maximum and the per accident limitation than the formula approach would predict. This is the opposite situation from the one described at the first retro adjustment period above. Second, average retrospective rating parameters may be changing over time. In the case of shifting parameters over time, a single selected PDLD ratio may not be the best estimate of development for all exposure periods. As with loss development analysis, the actuary must decide how best to develop each period to "square the triangle." For the fourth through sixth adjustment periods, the PDLD ratios were selected between those indicated by the two methods.

4. CUMULATIVE PDLD RATIOS

The ultimate goal of this method is to estimate the premium asset, which is the sum of all future premium adjustments based on the expected future loss emergence. As shown before, the relationship between premium and loss can be expressed by the PDLD ratios. However, the PDLD ratios are incremental factors. To estimate how much premium can be expected based on all future loss development, one needs to calculate the cumulative PDLD ratios, or the CPDLD ratios.

A CPDLD ratio is the average of the PDLD ratios in all subsequent retro adjustment periods, weighted by the percentage of losses to emerge in each period. For instance, the CPDLD ratio at the second retro adjustment is the average of the PDLD ratios for the second and subsequent retro adjustment periods, weighted by the percentage of losses emerged in each period. The CPDLD ratio at the third adjustment is the average of the PDLD ratios for the third and subsequent retro adjustment periods, weighted by the percentage of losses emerged in each period. The loss emergence pattern is shown at the bottom of Exhibit 7.

Using the loss emergence pattern derived from the loss development data in Exhibit 7 and the selected PDLD ratios from Exhibit 4, one can calculate the CPDLD ratios. For example, the first CPDLD ratio equals 1.492, which is computed as follows:

$$\frac{(1.750 \times 78.4\% + 0.700 \times 9.3\% + 0.550 \times 4.4\% + 0.450 \times 2.9\% + 0.400 \times 3.0\% + 0.350 \times 1.6\%)}{(78.4\% + 9.3\% + 4.4\% + 2.9\% + 3.0\% + 1.6\% + 0.4\%)}.$$

The second CPDLD ratio is 0.556, which is computed as follows:

$$\frac{(0.700 \times 9.3\% + 0.550 \times 4.4\% + 0.450 \times 2.9\% + 0.400 \times 3.0\% + 0.350 \times 1.6\%)}{(9.3\% + 4.4\% + 2.9\% + 3.0\% + 1.6\% + 0.4\%)}.$$

The calculation of the remaining CPDLD ratios is shown in Exhibit 3.

The CPDLD ratio tells how much premium an insurer can expect to collect for a dollar of loss that has yet to emerge. For instance, the first CPDLD ratio is 1.492, which means that each dollar of loss emerged provides the insurer one dollar and 49 cents of premium. The second CPDLD ratio is 0.556, which means that after the first retro adjustment, each additional dollar of loss provides the insurer 56 cents of premium.

The relationship of premium development to loss development is usually greater than unity at the first retro adjustment. This is because the basic premium is included in the first retro premium computation, and because only a small portion of loss is limited by the retro maximum or per accident limitation at this early maturity. The application of the loss conversion factor and the tax multiplier results in more than a dollar of premium per dollar of loss. As time goes on, however, a decreasing portion of incremental loss development results in additional premium. Incremental premium, equal to the loss capping ratio times LCF and TM, will generally be less than loss and hence the CPDLD ratios should be less than 1.0 at the later adjustments.

Having calculated the CPDLD ratios, the next step is to multiply these ratios by the expected future loss emergence to get the expected future premiums. Adding future premiums to the booked premiums gives ultimate premiums. For example, at 12/31/94, policy effective quarters 1993.1 through 1994.4 have not yet had the first retro adjustment (they are all less than 27 months old). The expected loss amount for these policy effective quarters, as computed in Exhibit 2, is $280,844,000 ($196,767,000 from 1993, plus $84,077,000 from 1994). Since the marginal premium per dollar of loss is $1.492, this means $280,844,000 $\times$ 1.492 or $419,019,000 of future premium is expected. Since there was no prior retro adjustment, the expected ultimate premium for these policy effective quarters is $419,019,000.

At 12/31/94, policy quarters 1992.1 through 1992.4 have had one retro adjustment (they are older than 27 months but not yet 39 months old). For these policy periods, the expected amount of loss yet to emerge is $50,747,000 (see Exhibit 2). Exhibit 3 shows that for each dollar of loss emerged after the first retro adjustment, the insurer can expect $0.556 of premium. This means the insurer can expect to collect $50,747,000 $\times$ 0.556 or $28,216,000 in additional premium. Adding this to the $328,778,000 of premium booked from the first retro adjustment (the premium for 1992.1 through 1992.4 evaluated as of 27 months), gives an expected ultimate premium of $356,993,000. Exhibit 1 shows the calculation of the ultimate premium for each policy period.

The final step is to subtract premium booked as of 12/31/94 from the estimated ultimate premium to get the premium asset as of 12/31/94. The sum of the premium assets for all policy periods as calculated in Exhibit 1 is $43 million.

Note that the premiums booked as of 12/31/94 (Column (7) of Exhibit 1) are close to but not equal to the premiums booked from the prior retro adjustments (Column (5) of Exhibit 1). This may be due to differences in the timing of retro adjustments, minor premium adjustments, or interim premium booking that occurs between the regularly scheduled retro adjustments.

5. LOSS CAPPING RATIO

We now return to the subject of the loss capping ratio. The loss capping ratio, CL/L, is the ratio of capped loss development to uncapped loss development. This term is essential to the calculation of the PDLD ratio, which expresses the relationship between premium development and loss development on a retro rated policy. Capped loss development includes the effect of the retro maximum and minimum, and the per accident loss limit. It is often difficult to obtain capped loss development data, especially as it pertains to losses eliminated by the retro maximum and minimum. Hence, it may be necessary to use a Table M[6] approach to estimate the impact of the retro plan maximum and minimum on loss development. If a per accident limit is purchased, the treatment of the losses eliminated by the limit is similar to that for losses eliminated by retro maximum and minimum.

The loss capping ratio can be solved for using the relationship

$$\mathrm{CLR} = \mathrm{LR}(1 - \chi - \mathrm{LER}),$$

where

χ = Table M net insurance charge
= Table M charge at max – Table M savings at min,

LER = Percent of losses eliminated due to the per accident limitation,

CLR = capped loss ratio
= capped loss divided by standard premium, and

LR = uncapped loss ratio
= uncapped loss divided by standard premium.

[6] Also called the Table of Insurance Charges. Table M is used to calculate the insurance charge associated with a retro plan's maximum and minimum. Gillam and Snader [3] give a detailed description of this table.

The loss capping ratio is then:

$$\mathrm{CLR/LR} = (1 - \chi - \mathrm{LER}). \tag{5.1}$$

To calculate the loss capping ratio, one needs the net insurance charge at each retro adjustment period. The insurance charge is typically determined from the values of the retro rating parameters sold under the plan and the presumed loss ratio distribution underlying Table M. However, the percentage of losses actually affected by the retro maximum or minimum will differ from expected due to the random nature of insurance losses and the fact that losses are not at their ultimate valuation. Therefore, the charge and savings computed at each retro adjustment period should be a function of the actual loss ratio as opposed to the expected ultimate loss ratio under the plan.

If it is assumed that the loss ratio probability distribution function has the same shape throughout all development stages, then at each retro adjustment one may enter Table M by defining two entry ratios:

Entry ratio at the max = (loss ratio at max/actual loss ratio), and

Entry ratio at the min = (loss ratio at min/actual loss ratio).

Loss ratios at the retro maximum and minimum should be estimated from the sold retro rating parameters. The loss ratio at maximum is the standard premium loss ratio at which the net retro premium reaches the maximum premium; for this example, we assume it is 1.200. Similarly, the loss ratio at minimum is the standard premium loss ratio at which the net retro premium reaches the minimum premium; for this example, we assume it is 0.100.

The actual loss ratio may be computed by dividing the actual loss at each retro adjustment period by the standard premium. Alternatively, it can be estimated as the expected loss ratio (expected ultimate loss divided by standard premium) times the expected percentage of losses emerged at each retro adjustment. For instance, if the expected loss ratio is 0.700 and 78.4% of losses emerge by the first retro adjustment, one can estimate the actual loss ratio at the first retro adjustment to be $0.700 \times 78.4\%$, or 0.549.

If actual loss experience differs from the expected experience underlying Table M, one should multiply the estimate of the actual loss ratio by a factor representing the relationship between actual and expected losses. For example, if the original expected loss ratio was 0.700 but actual loss experience produces an average loss ratio of 0.800, multiply 0.549 by a factor of 0.800/0.700. Such an adjustment factor is needed to calculate the correct entry ratios for Table M.

The two entry ratios for the first retro adjustment can be computed as:

$$\text{Entry ratio at the max} = (1.200/0.549) = 2.19, \text{ and}$$

$$\text{Entry ratio at the min} = (0.100/0.549) = 0.18.$$

Table M also requires one to estimate the average size of the accounts insured by the retro rated policies. For this example, the average size is assumed to be \$750,000 in standard premium. This may be estimated from the sold policy information. The use of the average policy size is another potential source of bias between the PDLD ratios calculated using the formula method and the PDLD ratios that actually emerge. One way to reduce this bias is by grouping the data according to policy size. The net insurance charge for a \$750,000 account at 2.19 and 0.18 entry ratios is calculated to be 0.109. This is shown in Exhibit 5.

In the event that a per accident loss limit is sold, losses eliminated by such limit divided by total losses should also be considered in the calculation of the loss capping ratio. Furthermore, the Table M insurance charge should be adjusted to reflect the per accident loss limit. One method of making such an adjustment is presented by Robbin [4]. In this example we assume that 4.2% of losses are eliminated by the per accident limitation as of the first retro adjustment. Thus, the loss capping ratio at the first retro adjustment is one minus 0.109 (the net insurance charge) minus 0.042 (the per-accident loss elimination ratio), or 85%. Loss capping ratios for the second and subsequent retro adjustment periods are calculated in Exhibit 5.

By using Table M to calculate the loss capping ratios, one major assumption is that the loss ratio probability distribution function underlying Table M is appropriate for all retro adjustment periods. This may not be true. The procedure can be refined by using a loss ratio distribution that is more appropriate for each retro adjustment period. Such distributions may be calculated from empirical data at the proper evaluation dates, and be used to replace or modify the Table M distribution, depending on the credibility of the empirical data.

Thus far the loss capping ratios calculated are those developed as of each retro adjustment. Since the PDLD ratios are incremental, one needs to calculate the incremental loss capping ratios, using the loss capping ratios developed through each retro adjustment. This is done by algebraic manipulation. For example, the incremental loss capping ratio for the second retro adjustment period is $[(\mathrm{CL}_2 - \mathrm{CL}_1)/(\mathrm{L}_2 - \mathrm{L}_1)]$ which may be stated as

$$\frac{[(\mathrm{CL}_2/\mathrm{L}_2) \times (\mathrm{ELR} \times \%\mathrm{Loss}_2) - (\mathrm{CL}_1/\mathrm{L}_1) \times (\mathrm{ELR} \times \%\mathrm{Loss}_1)]}{[(\mathrm{ELR} \times \%\mathrm{Loss}_2) - (\mathrm{ELR} \times \%\mathrm{Loss}_1)]}. \tag{5.2}$$

Note L_n is the amount of losses emerged as of the nth retro adjustment, and $\mathrm{CL}_n/\mathrm{L}_n$ is the loss capping ratio developed as of the nth retro adjustment. The ELR is the expected loss ratio, and $\%\mathrm{Loss}_n$ is the expected percentage of losses emerged as of the nth retro adjustment. The incremental loss capping ratios are calculated in Exhibit 5.

6. FURTHER ISSUES

The method described in this paper can be used to calculate the premium asset for all types of loss-sensitive rating plans, as long as the rating formula reflects what is being sold to the insured. Further issues to think about are:

1. The definition of loss may include allocated loss adjustment expense (ALAE). Frequently, retro rated policies are written with ALAE included in the definition of loss. This allows the insurer to pass on to the insured not only losses, but attorney expenses as well. The loss data used in computing the PDLD ratios should be consistent with that used in the rating plan.

2. Changes in the mix of business may change the PDLD ratio. Changes in the mix of business by state, industry group, or even geographical region can alter the average rating parameters sold and the underlying claim frequency and claim severity. This will in turn affect how sensitive the premium is to loss.

3. Collectibility of premium should be considered. When the premium asset is secured, there is little question as to its collectibility. If a portion of the premium asset is not secured, then a provision should be made to anticipate bad debt.

REFERENCES

[1] Berry, Charles H., "A Method of Setting Retro Reserves," *PCAS* LXVII, 1980, pp. 226–238.

[2] Fitzgibbon, W. J. Jr., "Reserving for Retrospective Returns," *PCAS* LII, 1965, pp. 203–214.

[3] Gillam, W. R. and R. H. Snader, "Fundamentals of Individual Risk Rating," National Council on Compensation Insurance, 1992.

[4] Robbin, Dr. I., "Overlap Revisited—The 'Insurance Charge Reflecting Loss Limitation' Procedure," 1990 CAS Discussion Paper Program, Vol. II, pp. 809–855.

EXHIBIT 1

CALCULATION OF FUTURE PREMIUM EMERGENCE AND PREMIUM ASSET

(dollars in thousands)

Policy Periods (1)	Expected Future Loss Emergence (2)	CPDLD Ratios (3)	Expected Future Premium (2)x(3) (4)	Premiums Booked from Prior Adjustment (5)	Estimated Total Premium (4)+(5) (6)	Premium Booked as of 12/94 (7)	Premium Asset (6)-(7) (8)
1987.1 to 1987.4	-262	0.000	0	494,927	494,927	494,927	0
1988.1 to 1988.4	3,282	0.285	935	467,388	468,324	467,796	528
1989.1 to 1989.4	9,146	0.354	3,238	460,660	463,897	460,716	3,181
1990.1 to 1990.4	21,347	0.390	8,325	453,525	461,850	452,520	9,331
1991.1 to 1991.4	25,397	0.447	11,352	336,654	348,007	337,966	10,041
1992.1 to 1992.4	50,747	0.556	28,216	328,778	356,993	330,216	26,777
1993.1 to 1994.4	280,844	1.492	419,019	0	419,019	425,590	-6,570
					3,013,018	2,969,730	43,288

Notes:
(2) From Exhibit 2, Column (7a).
(3) From Exhibit 3, Column (7).
(5) From Exhibit 4.
(7) From the latest diagonal of Exhibit 6.

EXHIBIT 2

LOSS PROJECTIONS

(dollars in thousands)

Policy Eff. Quarter (1)	Losses Reported as of 12/94 (2)	Loss Develop. Factors (3)	Percent Earned as of 12/94 (4)	Ultimate Losses (2)x(3)x(4) (5)	Annual Total (5a)	Losses Reported at Prior Retro Adjust. (6)	Annual Total (6a)	Expected Loss Emer-gence (5)-(6) (7)	Annual Total (7a)
1987.1	102,064	1.000	100%	102,064		102,059		5	
1987.2	65,339	1.000	100%	65,339		65,264		75	
1987.3	155,738	1.000	100%	155,738		155,950		-212	
1987.4	94,067	1.000	100%	94,067	417,208	94,197	417,470	-130	-262
1988.1	88,908	1.001	100%	89,040		87,781		1,259	
1988.2	57,763	1.002	100%	57,880		58,054		-174	
1988.3	152,121	1.004	100%	152,683		151,031		1,652	
1988.4	96,809	1.010	100%	97,734	397,337	97,189	394,055	545	3,282
1989.1	81,384	1.010	100%	82,203		80,475		1,728	
1989.2	55,898	1.012	100%	56,582		55,541		1,041	
1989.3	131,539	1.020	100%	134,190		130,944		3,246	
1989.4	91,423	1.030	100%	94,189	367,164	91,058	358,018	3,131	9,146
1990.1	89,715	1.032	100%	92,599		87,639		4,960	
1990.2	56,032	1.039	100%	58,239		54,473		3,766	
1990.3	118,268	1.052	100%	124,399		117,202		7,197	
1990.4	73,037	1.063	100%	77,660	352,897	72,236	331,550	5,424	21,347
1991.1	60,399	1.067	100%	64,424		57,620		6,804	
1991.2	34,136	1.068	100%	36,459		33,064		3,395	
1991.3	60,696	1.085	100%	65,883		58,493		7,390	

EXHIBIT 2

PART 2

Policy Eff. Quarter (1)	Losses Reported as of 12/94 (2)	Loss Develop. Factors (3)	Percent Earned as of 12/94 (4)	Ultimate Losses (2)x(3)x(4) (5)	Annual Total (5a)	Losses Reported at Prior Retro Adjust. (6)	Annual Total (6a)	Expected Loss Emergence (5)-(6) (7)	Annual Total (7a)
1991.4	61,068	1.100	100%	67,187	233,953	59,380	208,556	7,807	25,397
1992.1	53,455	1.107	100%	59,152		49,161		9,991	
1992.2	37,393	1.114	100%	41,652		33,060		8,592	
1992.3	62,118	1.140	100%	70,830		53,069		17,761	
1992.4	50,766	1.166	100%	59,188	230,822	44,785	180,075	14,403	50,747
1993.1	39,519	1.184	100%	46,799		0		46,799	
1993.2	30,286	1.215	100%	36,793		0		36,793	
1993.3	51,005	1.276	100%	65,093		0		65,093	
1993.4	34,516	1.393	100%	48,082	196,767	0	0	48,082	196,767
1994.1	21,189	1.661	90%	31,672		0		31,672	
1994.2	11,381	2.416	70%	19,245		0		19,245	
1994.3	14,339	4.397	45%	28,372		0		28,372	
1994.4	1,862	12.856	20%	4,788	84,077	0	0	4,788	84,077
TOTAL	2,118,031			2,247,065	2,196,148	1,889,725	1,889,725	357,340	306,423

Notes:

(2) Figures on the latest diagonal of the loss data in Exhibit 7.

(3) Derived from loss development data in Exhibit 7.

(4) These earning ratios reflect the fact that policies written in the latest four quarters are not fully earned.

(6) These represent losses recorded as of prior retro adjustments (Exhibit 7). Policy effective quarters 1993.1 through 1994.4 would not have had any retro adjustments as of 12/31/94; therefore, the losses recorded are 0. Policy effective quarters 1992.1 through 1992.4 would have had one retro adjustment; therefore, losses evaluated at 18 months were entered into this column.

EXHIBIT 3

CPDLD RATIO CALCULATION

Retro Adjustment Periods (1)	Selected PDLD Ratios (2)	% Loss Emerged (3)	PDLD Ratio x Loss Emg (2)x(3) (4)	Upward Cumulative of Col. (4) (5)	Upward Cumulative of Col. (3) (6)	CPDLD Ratios (5)/(6) (7)
First	1.750	78.4%	1.371	1.492	100.0%	1.492
Second	0.700	9.3%	0.065	0.120	21.6%	0.556
Third	0.550	4.4%	0.024	0.055	12.3%	0.447
Fourth	0.450	2.9%	0.013	0.031	7.9%	0.390
Fifth	0.400	3.0%	0.012	0.017	4.9%	0.354
Sixth	0.350	1.6%	0.006	0.006	2.0%	0.285
Subsequent	0.000	0.4%	0.000	0.000	0.4%	0.000

Notes:
(2) From Exhibit 4.
(3) From Exhibit 7.

EXHIBIT 4

PART 1

PDLD RATIO CALCULATION

(dollars in thousands)

Policy Eff. Quarter	First Retro Adjustment Loss 0-18	First Retro Adjustment Prem 0-27	First Retro Adjustment PDLD Ratio	Second Retro Adjustment Loss 19-30	Second Retro Adjustment Prem 28-39	Second Retro Adjustment PDLD Ratio	Third Retro Adjustment Loss 31-42	Third Retro Adjustment Prem 40-51	Third Retro Adjustment PDLD Ratio
1983.3	42,461	52,436	1.235	5,515	4,012	0.727	4,533	2,351	0.519
1983.4	20,151	26,222	1.301	2,738	2,722	0.994	1,480	576	0.389
1984.1	23,076	29,189	1.265	2,142	1,927	0.900	2,076	1,086	0.523
1984.2	19,243	23,422	1.217	1,032	1,904	1.844	507	740	1.461
1984.3	54,927	69,310	1.262	8,900	6,371	0.716	3,804	3,432	0.902
1984.4	33,393	43,305	1.297	4,308	3,189	0.740	2,819	1,274	0.452
1985.1	46,100	59,203	1.284	3,384	3,349	0.990	2,312	1,347	0.583
1985.2	27,696	38,717	1.398	2,679	2,120	0.791	2,675	1,687	0.631
1985.3	96,041	133,094	1.386	9,717	7,926	0.816	6,465	3,054	0.472
1985.4	49,481	66,351	1.341	7,193	4,063	0.565	4,268	2,560	0.600
1986.1	63,095	87,173	1.382	5,865	4,249	0.724	4,045	2,298	0.568
1986.2	42,163	57,654	1.367	3,904	2,283	0.585	3,882	1,981	0.510
1986.3	115,105	160,838	1.397	12,006	10,917	0.909	12,037	7,932	0.659
1986.4	58,712	84,641	1.442	6,627	3,536	0.534	3,737	3,579	0.958
1987.1	77,373	103,693	1.340	7,879	8,776	1.114	4,795	2,987	0.623
1987.2	49,770	68,397	1.374	4,867	3,467	0.712	4,029	993	0.246
1987.3	120,053	171,434	1.428	15,117	9,858	0.652	8,909	4,189	0.470
1987.4	73,502	101,483	1.381	7,479	5,701	0.762	5,101	2,290	0.449
1988.1	71,999	98,806	1.372	6,083	4,745	0.780	4,138	1,006	0.243
1988.2	45,861	63,885	1.393	5,253	2,688	0.512	3,392	853	0.252
1988.3	115,461	161,154	1.428	13,462	6,642	0.652	7,128	2,854	0.470
1988.4	79,063	109,253	1.382	7,723	3,974	0.515	5,082	2,604	0.512

EXHIBIT 4

PART 1—PAGE 2

1989.1	71,471	99,777	1.396	3,744	913	0.244	4,462	2,131	0.478
1989.2	49,486	67,553	1.365	2,976	2,923	0.982	2,834	1,906	0.672
1989.3	108,330	153,443	1.416	12,886	6,525	0.506	6,733	5,732	0.851
1989.4	72,082	104,838	1.454	10,903	8,693	0.797	7,288	4,224	0.580
1990.1	76,452	107,468	1.406	8,989	7,055	0.785	2,199	319	0.145
1990.2	46,393	70,127	1.512	6,513	5,695	0.874	1,566	1,159	0.740
1990.3	102,035	158,027	1.549	13,486	6,788	0.503	1,682	901	0.536
1990.4	57,548	91,918	1.597	8,522	2,840	0.333	6,166	1,227	0.199
1991.1	54,037	81,901	1.516	3,583	2,595	0.724			
1991.2	30,240	54,045	1.787	2,824	666	0.236			
1991.3	55,325	94,797	1.713	3,168	2,552	0.805			
1991.4	54,302	97,650	1.798	5,078	2,448	0.482			
1992.1	49,161	82,057	1.669						
1992.2	33,060	59,279	1.793						
1992.3	53,069	99,074	1.867						
1992.4	44,785	88,367	1.973						
Selection Based on Historical Averages									
Average All			1.460			0.730			0.556
Weighted Average All			1.455			0.680			0.532
Selected			1.750			0.700			0.550
Selection Based on Retro Formula									
LCF			1.20			1.20			1.20
TM			1.03			1.03			1.03
Loss Capping Ratio			85%			58%			45%
Implied PDLD Ratio			1.426 *			0.717			0.556
Final Selection			1.750			0.700			0.550

* Also assumes a basic premium factor of 0.2, an expected loss ratio of 0.7, and an expected loss emergence of 78.4% at first adjustment.

EXHIBIT 4

PART 2

PDLD RATIO CALCULATION

(dollars in thousands)

Policy Eff. Quarter	Fourth Retro Adjustment Loss 43-54	Prem 52-63	PDLD Ratio	Fifth Retro Adjustment Loss 55-66	Prem 64-75	PDLD Ratio	Sixth Retro Adjustment Loss 67-78	Prem 76-87	PDLD Ratio
1983.3	1,925	763	0.397	2,057	712	0.346	1,170	75	0.064
1983.4	1,078	662	0.615	64	56	0.867	525	186	0.355
1984.1	1,139	883	0.776	827	526	0.636	1,123	-103	-0.092
1984.2	1,137	573	0.504	906	593	0.655	165	15	0.088
1984.3	2,949	1,159	0.393	2,619	635	0.243	2,475	137	0.055
1984.4	1,424	206	0.145	1,378	46	0.033	1,329	86	0.065
1985.1	1,538	267	0.173	2,265	120	0.053	528	615	1.165
1985.2	2,026	773	0.381	1,730	189	0.109	1,072	210	0.196
1985.3	6,525	2,670	0.409	6,604	2,611	0.395	3,566	155	0.043
1985.4	3,049	1,196	0.392	2,194	1,091	0.497	2,533	958	0.378
1986.1	1,700	1,243	0.731	3,519	874	0.248	1,477	621	0.421
1986.2	2,480	63	0.025	1,476	888	0.601	1,969	194	0.099
1986.3	5,380	2,703	0.502	8,623	1,693	0.196	4,364	1,601	0.367
1986.4	3,316	561	0.169	3,032	728	0.240	1,907	84	0.044
1987.1	5,508	1,796	0.326	4,720	1,522	0.322	1,784	69	0.039

EXHIBIT 4

PART 2—PAGE 2

1987.2	2,521	207	0.082	2,970	869	0.293	1,107	416	0.375
1987.3	7,089	2,571	0.363	3,589	2,532	0.705	1,191	-320	-0.268
1987.4	4,456	1,199	0.269	3,277	572	0.175	381	226	0.593
1988.1	3,267	1,498	0.458	2,294	82	0.036			
1988.2	2,461	894	0.363	1,086	102	0.094			
1988.3	6,284	3,014	0.363	8,696	108	0.012			
1988.4	4,351	2,528	0.581	970	698	0.720			
1989.1	798	339	0.425						
1989.2	245	147	0.601						
1989.3	2,996	1,043	0.348						
1989.4	785	472	0.601						
Selection Based on Historical Averages									
Average All			0.400			0.340			0.222
Weighted Average All			0.385			0.266			0.182
Selected			0.400			0.300			0.200
Selection Based on Retro Formula									
LCF			1.20			1.20			1.20
TM			1.03			1.03			1.03
Loss Capping Ratio			40%			40%			40%
Implied PDLD Ratio			0.494			0.494			0.494
Final Selection			0.450			0.400			0.350

EXHIBIT 5

PART 1

LOSS CAPPING RATIO CALCULATION

(with per accident limitation)

Retro Adjustment (1)	Ultimate Standard Premium Loss Ratio (2)	Percent of Total Losses Emerged (3)	Emerged Loss Ratio (2)x(3) (4)	Loss Ratio at Retro Maximum (5)	Loss Ratio at Retro Minimum (6)	Entry Ratio at Retro Maximum (5)/(4) (7)	Entry Ratio at Retro Minimum (6)/(4) (8)
First	0.700	78.4%	0.549	1.200	0.100	2.19	0.18
Second	0.700	87.7%	0.614	1.200	0.100	1.95	0.16
Third	0.700	92.1%	0.645	1.200	0.100	1.86	0.16
Fourth	0.700	95.1%	0.665	1.200	0.100	1.80	0.15
Fifth	0.700	98.0%	0.686	1.200	0.100	1.75	0.15
Sixth	0.700	99.6%	0.697	1.200	0.100	1.72	0.14
Subsequent	0.700	100.0%	0.700	1.200	0.100	1.71	0.14

EXHIBIT 5

PART 2

LOSS CAPPING RATIO CALCULATION

(with per accident limitation)

Retro Adjustment (1)	Insurance Charge at Retro Maximum (9)	Insurance Saving at Retro Minimum (10)	% of Losses Eliminated by Retro Max/Min (9)-(10) (11)	Loss Elimination Ratio from Per Accident Limitation (12)	Cumulative Loss Capping Ratios 1.0-(11)-(12) (13)	Incremental Loss Capping Ratios (14)	Selected Incremental Loss Capping Ratios (15)
First	0.113	0.004	10.9%	4.2%	84.9%	84.9%	85.0%
Second	0.133	0.003	13.0%	5.0%	82.0%	58.0%	58.0%
Third	0.142	0.003	13.9%	5.9%	80.2%	45.0%	45.0%
Fourth	0.148	0.003	14.5%	6.5%	79.0%	39.6%	40.0%
Fifth	0.154	0.003	15.1%	7.1%	77.8%	38.7%	40.0%
Sixth	0.156	0.002	15.4%	7.4%	77.2%	41.7%	40.0%
Subsequent	0.158	0.002	15.6%	7.5%	76.9%	3.3%	0.0%

Notes:

(2) By judgment.

(3) Based on loss development pattern. See Exhibit 7.

(5),(6) Based on the retro rating values on the policies sold.

(9),(10) From NCCI Table of Insurance Charges, assuming $750,000 standard premium at the entry ratios listed in Columns (7) and (8), with losses used for loss group estimation adjusted for the per accident limitation.

(12) From a study of the percentage of losses eliminated due to per accident limitation.

(14) = [(13)x(4) - (Prior 13)x(Prior 4)] / [(4) - (Prior 4)].

(15) By judgment.

EXHIBIT 6

PART 1

BOOKED PREMIUM

(dollars in thousands)

POL EFF	EVALUATED AT (MONTHS)													
QUARTER	3	6	9	12	15	18	21	24	27	30	33	36	39	42
1983.3	18,087	23,481	33,550	40,867	45,075	49,911	50,174	50,629	52,436	52,388	52,428	53,735	56,448	56,442
1983.4	7,545	10,684	15,697	19,696	22,407	24,608	25,438	25,367	26,222	26,135	26,109	26,653	28,944	28,886
1984.1	7,930	13,516	19,135	24,101	25,408	28,018	28,777	28,947	29,189	28,967	28,960	29,563	31,116	32,031
1984.2	6,277	10,386	14,770	18,422	20,366	22,441	22,258	23,261	23,422	23,608	23,714	24,239	25,326	25,319
1984.3	20,221	31,438	45,913	56,490	61,851	66,697	67,910	68,094	69,310	69,648	69,804	73,363	75,681	75,727
1984.4	9,581	21,089	29,224	38,056	41,299	43,344	43,489	43,897	43,305	43,851	43,794	44,982	46,495	46,645
1985.1	15,110	28,734	40,623	49,872	52,336	57,881	58,742	56,075	59,203	59,164	59,814	60,692	62,552	62,780
1985.2	9,345	19,304	26,189	33,115	37,010	37,650	39,114	40,131	38,717	38,855	39,247	39,753	40,836	40,780
1985.3	43,187	69,712	99,219	120,931	128,832	138,488	138,401	132,641	133,094	134,019	134,241	138,147	141,020	141,198
1985.4	18,627	36,106	51,122	61,567	63,169	67,119	68,046	67,651	66,351	67,117	67,131	68,357	70,414	70,703
1986.1	27,390	46,053	62,065	76,616	79,958	83,746	88,382	83,949	87,173	87,630	87,438	88,964	91,422	91,780
1986.2	15,906	28,997	39,082	48,402	53,764	56,713	59,748	58,482	57,654	57,171	57,156	57,801	59,937	59,994
1986.3	75,944	99,936	128,593	152,088	157,264	167,176	168,696	160,320	160,838	161,506	161,702	167,381	171,755	173,178
1986.4	34,837	47,808	61,795	75,233	79,268	83,931	85,404	82,732	84,641	84,662	83,954	85,334	88,177	89,027
1987.1	43,330	57,756	74,709	90,700	96,064	103,569	105,732	100,641	103,693	105,325	105,992	107,900	112,468	113,333
1987.2	21,776	37,152	46,687	57,190	62,681	67,875	69,209	68,077	68,397	67,799	68,083	70,573	71,863	71,866
1987.3	81,929	97,806	130,510	154,795	165,210	174,935	172,841	172,374	171,434	170,357	170,034	176,167	181,292	181,375
1987.4	40,213	56,338	71,591	88,365	91,969	98,370	99,258	99,738	101,483	101,762	101,940	104,615	107,184	107.014

EXHIBIT 6

PART 1—PAGE 2

1988.1	42,723	54,489	70,630	85,547	90,156	96,939	97,210	97,349	98,806	99,398	99,426	101,516	103,551	103,824
1988.2	22,257	34,740	43,526	53,014	58,682	62,781	62,671	62,232	63,885	63,771	63,912	66,090	66,573	66,593
1988.3	72,106	88,814	119,264	144,686	150,957	155,317	157,645	157,669	161,154	159,364	159,056	163,070	167,796	167,503
1988.4	39,617	54,419	70,154	90,176	96,199	100,143	101,400	103,318	109,253	109,553	109,928	111,601	113,227	113,259
1989.1	39,587	52,001	70,440	86,851	90,443	92,953	95,370	95,048	99,777	99,118	98,874	99,321	100,691	100,996
1989.2	22,202	34,226	45,230	58,091	61,548	63,860	65,452	65,312	67,553	67,514	67,325	68,442	70,476	70,691
1989.3	65,381	87,768	116,546	143,674	150,281	154,254	155,500	153,310	153,443	152,289	151,898	153,635	159,968	162,505
1989.4	39,328	59,317	79,875	101,224	107,600	109,565	108,541	107,785	104,838	105,096	106,006	106,285	113,531	113,513
1990.1	41,480	58,906	82,540	101,782	105,099	109,581	107,738	108,319	107,468	108,595	108,481	112,426	114,523	114,013
1990.2	24,980	44,112	57,612	69,544	72,525	73,996	74,014	72,759	70,127	71,153	71,796	73,669	75,822	75,729
1990.3	63,128	90,937	124,828	150,974	157,555	164,785	161,333	157,804	158,027	154,559	155,939	161,526	164,815	164,057
1990.4	39,431	56,685	75,859	96,782	102,985	103,811	105,995	102,283	91,918	90,293	89,057	92,789	94,759	94,621
1991.1	33,478	44,023	65,178	77,912	80,929	82,418	85,231	83,748	81,901	81,593	81,741	83,321	84,497	84,444
1991.2	14,280	28,297	40,717	52,368	56,505	57,790	58,229	54,008	54,045	53,647	53,094	54,363	54,711	54,940
1991.3	45,699	57,962	76,470	90,684	94,509	99,243	96,431	97,296	94,797	95,483	95,037	95,884	97,348	97,351
1991.4	40,472	55,839	71,234	96,004	103,650	110,764	108,317	107,345	97,650	97,275	97,453	98,816	100,098	
1992.1	33,004	48,424	63,007	77,569	80,923	83,160	82,743	79,426	82,057	81,207	82,305	83,654		
1992.2	15,320	28,687	42,953	55,414	59,215	59,989	61,432	62,709	59,279	59,531	59,450			
1992.3	43,587	62,778	79,724	96,319	102,225	102,001	100,433	101,661	99,074	98,745				
1992.4	39,809	55,591	64,921	84,643	91,176	95,711	95,421	96,272	88,367					
1993.1	23,915	33,188	43,949	53,188	56,028	57,200	57,259	57,304						
1993.2	18,976	30,311	39,234	46,519	49,357	50,776	50,887							
1993.3	45,269	55,794	70,411	84,974	87,886	89,838								
1993.4	28,913	40,173	51,901	64,487	69,025									
1994.1	24,902	33,572	41,305	49,737										
1994.2	20,642	28,772	34,494											
1994.3	42,916	55,331												
1994.4	18,975													

EXHIBIT 6

PART 2

BOOKED PREMIUM

(dollars in thousands)

POL EFF	EVALUATED AT (MONTHS)														
QUARTER	45	48	51	54	57	60	63	66	69	72	75	78	81	84	87
1983.3	56,698	57,448	58,799	58,904	58,875	59,306	59,562	59,301	59,335	59,492	60,275	60,309	60,321	60,221	60,350
1983.4	28,923	29,122	29,520	29,490	29,463	29,492	30,182	30,048	30,048	30,204	30,238	30,172	30,164	30,041	30,424
1984.1	31,993	31,896	32,202	32,473	32,489	32,513	33,085	32,997	33,046	33,223	33,612	33,612	33,612	33,735	33,509
1984.2	25,433	25,717	26,066	26,095	26,330	26,365	26,639	26,622	26,944	26,961	27,232	27,241	27,070	27,086	27,247
1984.3	76,163	77,812	79,113	79,140	79,358	79,730	80,272	79,940	80,017	80,328	80,907	80,813	80,850	80,934	81,044
1984.4	46,535	47,080	47,768	47,815	47,798	47,833	47,974	47,844	47,424	47,527	48,020	47,977	47,971	47,873	48,106
1985.1	62,880	63,239	63,899	64,075	64,072	63,911	64,166	63,818	63,785	63,794	64,286	64,360	64,348	64,502	64,901
1985.2	41,156	41,412	42,523	42,446	42,635	42,814	43,296	42,989	42,975	43,080	43,485	43,517	43,371	43,682	43,695
1985.3	141,185	143,140	144,073	144,137	144,236	145,032	146,743	146,395	146,591	147,980	149,354	149,163	149,109	149,224	149,509
1985.4	70,606	71,420	72,974	72,741	73,185	73,342	74,170	73,882	73,688	74,169	75,261	75,226	74,893	75,060	76,219
1986.1	91,767	92,532	93,720	94,000	93,988	94,319	94,963	94,514	94,611	94,869	95,837	95,762	95,762	96,108	96,458
1986.2	60,463	61,202	61,919	61,909	61,941	61,664	61,982	61,909	62,439	62,630	62,869	63,097	63,109	63,093	63,064
1986.3	173,315	174,277	179,686	179,875	180,311	181,404	182,390	181,594	181,531	181,942	184,082	184,903	184,896	185,286	185,683
1986.4	89,240	90,156	91,756	91,817	91,734	91,698	92,317	92,091	91,988	92,023	93,045	93,071	93,055	92,899	93,129
1987.1	113,294	114,273	115,455	115,468	115,768	116,457	117,251	116,837	116,912	117,887	118,773	118,924	118,864	118,536	118,843
1987.2	71,824	72,598	72,856	73,264	73,165	72,941	73,063	72,576	72,622	73,138	73,933	74,079	74,081	74,175	74,348

EXHIBIT 6

PART 2—PAGE 2

1987.3	181,456	183,852	185,481	185,695	185,960	185,921	188,052	188,062	187,991	188,412	190,584	190,639	190,735	190,555	190,265
1987.4	106,769	108,691	109,474	109,590	109,674	109,553	110,673	110,406	110,472	110,766	111,245	111,243	111,236	111,279	111,472
1988.1	103,719	103,937	104,557	104,769	104,822	104,662	106,055	105,656	105,466	105,683	106,137	106,218	106,230	106,292	
1988.2	66,453	66,542	67,426	67,585	67,579	68,302	68,320	68,055	67,964	67,923	68,422	68,448	68,462		
1988.3	167,536	168,059	170,650	172,206	172,539	172,583	173,664	172,956	172,995	172,960	173,772	173,984			
1988.4	113,847	114,158	115,831	115,967	115,909	116,218	118,360	117,566	117,541	117,189	119,058				
1989.1	101,165	102,126	102,822	102,919	102,750	102,955	103,161	103,367	103,574	103,781					
1989.2	70,687	72,016	72,382	72,362	72,385	72,457	72,529	72,602	72,675						
1989.3	162,799	163,947	165,700	166,040	165,811	165,390	166,743	166,034							
1989.4	113,494	115,806	117,754	117,872	117,990	118,108	118,226								
1990.1	114,032	114,738	114,842	113,920	113,959	113,443									
1990.2	75,670	76,523	76,981	77,086	76,983										
1990.3	164,447	164,886	165,716	166,107											
1990.4	94,655	94,805	95,986												
1991.1	84,616	85,643													
1991.2	54,874														

EXHIBIT 7

PART 1

REPORTED LOSSES

(dollars in thousands)

POL EFF QUARTER	EVALUATED AT (MONTHS) 3	6	9	12	15	18	21	24	27	30	33	36	39	42
1983.3	5,121	15,662	24,950	36,667	41,044	42,461	44,191	45,528	46,321	47,976	48,898	49,439	50,413	52,508
1983.4	1,336	5,853	10,153	15,218	18,928	20,151	21,293	21,821	22,244	22,889	23,270	23,466	23,802	24,369
1984.1	2,746	6,798	11,408	17,227	20,523	23,076	23,954	24,351	24,730	25,218	25,765	26,065	26,897	27,294
1984.2	1,393	5,284	8,929	14,264	18,651	19,243	20,080	20,759	20,954	20,276	20,143	20,504	20,549	20,782
1984.3	6,618	17,632	31,538	46,077	51,318	54,927	58,153	59,814	61,636	63,827	65,053	65,301	66,273	67,631
1984.4	2,417	9,115	17,939	24,030	30,204	33,393	35,277	36,237	36,887	37,701	38,636	39,640	39,583	40,520
1985.1	3,847	13,981	22,898	34,132	41,523	46,100	46,809	47,718	48,853	49,485	50,490	50,804	50,748	51,796
1985.2	2,164	6,559	12,772	20,433	26,052	27,696	28,720	29,155	29,654	30,375	31,281	31,597	32,095	33,051
1985.3	11,514	34,201	57,070	84,782	92,911	96,041	98,225	101,501	103,660	105,758	106,336	108,721	110,196	112,223
1985.4	4,252	14,692	28,032	38,282	46,101	49,481	52,737	54,006	55,152	56,674	58,126	59,382	59,953	60,942
1986.1	6,670	20,522	32,059	46,939	56,233	63,095	63,724	65,044	65,887	68,960	69,613	70,511	71,368	73,005
1986.2	3,531	10,917	19,770	31,772	40,050	42,163	43,055	43,374	44,117	46,067	46,648	47,551	47,994	49,949
1986.3	14,331	38,985	63,615	99,612	109,338	115,105	118,102	121,840	124,300	127,112	129,821	131,477	133,952	139,149
1986.4	4,768	17,534	32,808	45,213	54,351	58,712	61,941	62,548	63,559	65,340	66,158	66,925	67,544	69,077
1987.1	8,142	23,354	38,500	56,764	67,709	77,373	79,521	81,580	83,222	85,252	86,910	87,978	88,660	90,047
1987.2	4,329	13,671	25,907	37,740	49,263	49,770	50,940	52,428	53,018	54,637	56,396	56,994	57,460	58,666
1987.3	13,373	36,138	63,600	101,166	113,655	120,053	124,721	128,107	130,850	135,171	137,613	139,866	141,475	144,080
1987.4	6,190	20,923	41,460	56,846	67,724	73,502	76,331	76,489	78,388	80,981	82,914	84,044	84,277	86,082
1988.1	6,916	20,545	34,772	50,554	63,347	71,999	73,728	74,396	75,834	78,082	79,412	80,223	80,531	82,220
1988.2	4,087	11,179	21,183	34,005	44,314	45,861	46,854	47,968	49,482	51,114	52,248	53,089	53,407	54,507
1988.3	12,952	35,571	63,806	98,684	110,159	115,461	120,036	123,446	124,724	128,922	131,384	133,661	134,948	136,050
1988.4	5,451	22,014	40,250	57,358	71,190	79,063	85,346	85,263	86,493	86,786	88,053	88,746	89,616	91,868

EXHIBIT 7

PART 1—PAGE 2

1989.1	7,869	21,725	34,204	50,512	62,907	71,471	71,946	72,275	72,442	75,215	76,405	77,279	77,522	79,677
1989.2	3,615	12,348	23,914	36,257	46,573	49,486	50,566	50,507	51,567	52,462	54,081	54,467	54,940	55,296
1989.3	11,397	33,593	61,655	92,142	102,949	108,330	112,488	116,584	119,203	121,216	124,649	125,375	126,285	127,949
1989.4	4,842	19,715	39,063	52,236	64,224	72,082	76,852	78,864	80,320	82,985	87,691	88,199	88,955	90,273
1990.1	9,511	25,559	39,469	56,952	68,215	76,452	78,668	79,502	81,249	85,440	86,905	86,250	86,565	87,639
1990.2	3,836	12,606	22,351	33,833	42,925	46,393	48,984	50,954	52,012	52,906	53,893	53,431	53,847	54,473
1990.3	11,677	32,321	57,374	83,863	94,633	102,035	106,144	109,426	112,652	115,521	116,449	116,762	116,920	117,202
1990.4	5,196	14,997	30,880	42,939	52,459	57,548	62,316	63,183	64,537	66,070	70,479	70,789	71,707	72,236
1991.1	7,133	19,024	29,513	39,914	47,837	54,037	56,184	56,093	56,697	57,620	59,265	59,537	59,577	60,025
1991.2	2,129	6,883	13,419	20,833	28,953	30,240	31,317	32,484	32,779	33,064	33,768	33,584	33,779	33,771
1991.3	5,053	14,983	30,591	47,533	53,959	55,325	55,647	57,417	57,666	58,493	59,106	59,916	59,814	60,696
1991.4	3,797	14,259	28,024	40,545	48,147	54,302	58,775	58,620	58,270	59,380	59,800	60,963	61,068	
1992.1	6,135	16,850	27,327	36,669	43,512	49,161	51,595	52,060	51,888	52,340	53,333	53,455		
1992.2	3,052	8,282	16,036	24,417	31,064	33,060	35,210	35,543	36,320	36,680	37,393			
1992.3	4,619	15,383	31,122	45,310	50,764	53,069	55,737	59,254	60,047	62,118				
1992.4	3,596	12,562	23,922	33,822	41,006	44,785	47,814	49,581	50,766					
1993.1	3,786	12,318	18,973	26,870	31,880	34,297	36,340	39,519						
1993.2	2,091	7,172	12,834	20,247	24,915	27,847	30,286							
1993.3	5,349	14,393	25,756	40,138	46,561	51,005								
1993.4	2,881	10,585	20,030	27,906	34,516									
1994.1	6,241	9,031	14,755	21,189										
1994.2	1,357	6,287	11,381											
1994.3	5,083	14,339												
1994.4	1,862													
Wt'd Avg 16	2.924	1.820	1.454	1.192	1.092	1.050	1.026	1.016	1.022	1.024	1.007	1.006	1.014	1.016
Selected	2.924	1.820	1.454	1.192	1.092	1.050	1.026	1.016	1.022	1.024	1.007	1.006	1.014	1.016
Cumulative	12.856	4.397	2.416	1.661	1.393	1.276	1.215	1.184	1.166	1.140	1.114	1.107	1.100	1.085
% Emerged	7.8%	22.7%	41.4%	60.2%	71.8%	78.4%	82.3%	84.4%	85.8%	87.7%	89.8%	90.4%	90.9%	92.1%

EXHIBIT 7

PART 2

REPORTED LOSSES

(dollars in thousands)

POL EFF	EVALUATED AT (MONTHS)														
QUARTER	45	48	51	54	57	60	63	66	69	72	75	78	81	84	87
1983.3	53,010	52,952	52,987	54,433	54,844	54,908	55,464	56,489	56,740	56,907	56,968	57,660	58,013	58,319	58,893
1983.4	24,713	24,840	25,085	25,447	25,598	25,913	25,430	25,511	25,527	25,784	25,905	26,036	26,257	26,439	26,396
1984.1	27,873	28,168	28,054	28,433	28,783	28,888	28,887	29,261	29,942	29,959	29,857	30,384	30,332	30,414	30,496
1984.2	21,110	21,056	21,408	21,919	22,129	22,130	22,260	22,825	22,942	22,721	22,697	22,990	23,351	23,331	23,639
1984.3	68,177	68,656	69,571	70,581	70,598	70,630	70,699	73,199	73,429	73,673	74,127	75,675	76,027	76,207	76,291
1984.4	41,687	41,786	41,658	41,944	42,229	42,602	42,760	43,323	43,870	44,107	44,278	44,652	45,046	45,296	45,746
1985.1	52,028	52,557	52,925	53,334	54,047	54,796	55,246	55,600	55,619	55,744	55,635	56,128	56,031	56,762	56,822
1985.2	33,713	33,797	34,161	35,077	35,629	35,888	35,816	36,806	37,422	37,734	37,721	37,878	37,977	38,073	38,029
1985.3	113,596	114,633	115,792	118,749	120,458	121,027	121,826	125,352	126,865	127,252	127,387	128,918	129,448	129,815	130,355
1985.4	61,543	62,769	63,048	63,992	64,759	65,053	65,140	66,186	67,419	68,036	68,140	68,719	69,068	69,417	69,467
1986.1	73,711	73,961	73,350	74,704	76,207	76,344	77,005	78,223	79,734	80,044	80,055	79,700	79,828	80,048	80,598
1986.2	50,459	50,798	51,492	52,429	52,185	52,937	53,475	53,905	54,533	54,659	54,858	55,874	55,290	55,033	55,170
1986.3	141,803	141,276	142,446	144,528	148,673	149,820	150,424	153,151	154,266	155,533	156,039	157,514	158,046	157,870	158,282
1986.4	70,913	71,454	71,585	72,392	73,624	73,956	74,188	75,425	76,073	76,103	76,481	77,332	77,791	77,708	77,451
1987.1	91,273	91,974	92,426	95,555	96,936	97,662	98,212	100,275	100,547	100,919	101,009	102,059	102,416	102,215	102,064
1987.2	59,490	59,750	60,308	61,187	62,269	62,746	62,840	64,157	64,589	64,652	64,690	65,264	65,408	65,176	65,339
1987.3	146,880	148,007	148,826	151,169	152,187	152,686	153,406	154,758	155,466	155,732	156,009	155,950	156,064	155,915	155,738
1987.4	88,294	89,146	89,288	90,538	92,262	92,860	92,825	93,815	93,787	93,977	93,987	94,197	94,251	94,124	94,067

EXHIBIT 7

PART 2—PAGE 2

1988.1	83,299	83,605	84,078	85,487	85,962	86,393	86,795	87,781	89,697	89,780	89,457	89,473	89,042	88,908	
1988.2	55,421	55,937	56,157	56,968	57,732	57,913	57,606	58,054	58,287	58,285	57,871	57,691	57,763		
1988.3	139,179	139,464	140,598	142,334	145,198	150,268	150,526	151,031	151,733	151,487	151,590	152,121			
1988.4	92,734	93,316	94,645	96,219	97,757	97,788	97,436	97,189	97,467	97,076	96,809				
1989.1	80,263	79,879	79,483	80,475	81,137	81,113	80,897	81,253	81,546	81,384					
1989.2	55,591	55,467	55,322	55,541	55,780	55,751	55,663	55,652	55,898						
1989.3	129,957	129,598	129,944	130,944	131,516	131,650	131,233	131,539							
1989.4	91,621	91,261	91,296	91,058	91,020	91,296	91,423								
1990.1	91,080	90,996	90,437	89,844	89,689	89,715									
1990.2	55,529	55,502	55,535	55,556	56,032										
1990.3	117,673	117,528	117,746	118,268											
1990.4	73,643	73,101	73,037												
1991.1	60,440	60,399													
1991.2	34,136														
Wt'd Avg 16	1.001	1.003	1.011	1.012	1.007	1.002	1.010	1.008	1.002	1.000	1.006	1.002	1.001	1.001	
Selected	1.001	1.003	1.011	1.012	1.007	1.002	1.010	1.008	1.002	1.000	1.006	1.002	1.001	1.001	Tail
Cumulative	1.068	1.067	1.063	1.052	1.039	1.032	1.030	1.020	1.012	1.010	1.010	1.004	1.002	1.001	1.000
% Emerged	93.6%	93.8%	94.0%	95.1%	96.2%	96.9%	97.1%	98.0%	98.8%	99.0%	99.1%	99.6%	99.8%	99.9%	100.0%

DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXXIII

ESTIMATING THE PREMIUM ASSET ON RETROSPECTIVELY RATED POLICIES

MIRIAM PERKINS AND MICHAEL T. S. TENG

DISCUSSION BY SHOLOM FELDBLUM

1. INTRODUCTION

Perkins and Teng have provided us with a new and remarkably intuitive procedure for estimating the accrued retrospective premium asset: the PDLD (premium development to loss development) approach. This reserve is often significant—amounting to half a billion dollars or more for some of the major workers compensation carriers—and it has been difficult to accurately estimate with traditional procedures. The paper by Perkins and Teng should greatly enhance our actuarial repertoire.

Specifically, the PDLD method has several distinct advantages over other procedures:

1. It is modeled directly on the retrospective rating formula, so it is easily explained to underwriters and claims personnel who are familiar with retrospectively rated policies.
2. Its emphasis on the premium sensitivity in the retrospective rating formula parallels the risk-based capital loss-sensitive contract offset in the underwriting risk charges and the new loss-sensitive contract Part 7 of Schedule P. For regulators familiar with the risk-based capital formula and with the statutory accounting requirements, this loss reserving approach is a natural complement to the statutory procedures.

3. The procedure may prove particularly useful when changes in the retrospective rating plan parameters distort the indications of other methods.

There are few existing methods for estimating the accrued retrospective premium asset, and the indications are often highly uncertain. The PDLD approach will enable actuaries to estimate this asset more accurately.

This discussion has two parts.

1. The complexity of the reserve estimation procedures for the accrued retrospective premium asset often hides the rationale of these methods from the average reader. The first part of this discussion uses graphical representations of Fitzgibbon's method and of the PDLD method to show the rationale behind each method and to explain the advantages of the latter method.[1] We then show how to combine the better parts of the two methods to improve the PDLD procedure.
2. The second part of this discussion highlights the implications of the Perkins and Teng procedure for the calculation of the loss-sensitive contract offset to the underwriting risk charges in the risk-based capital formula and for the use of Schedule P, Part 7, to estimate premium sensitivity.[2]

2. THE PDLD PROCEDURE

This section addresses two issues:

1. How does the PDLD procedure differ intuitively from Fitzgibbon's procedure, and in what ways is it better?

---

[1]See Fitzgibbon [6], F. J. Hope [8], Unthoff [11], Berry [2], and Morell [10]. The term "Fitzgibbon's method" in the text includes the enhancements provided by Berry and Morell.

[2]The term "premium sensitivity" stems from the term "loss-sensitive contracts." This paper uses the term "premium responsiveness" to refer to the same phenomenon.

2. What aspects of Fitzgibbon's procedure can be added to the PDLD procedure to enhance its accuracy?

Let us begin our inquiry with a more fundamental question. Why not estimate the accrued retrospective premium asset the same way that we estimate loss reserves? That is, why not use a chain-ladder development procedure on historical triangles of either collected premium or billed premium? This would be the premium analogue to a chain-ladder development procedure using either paid losses or reported losses.

Indeed, Schedule P already does this. Part 6 of Schedule P shows historical triangles of exposure year earned premiums by line of business (for all types of contracts), and Part 7 of Schedule P shows historical triangles of policy year earned premium on loss-sensitive contracts (all lines of business combined). Why go through the complexities of Fitzgibbon's method or the PDLD method when a straightforward chain ladder development method would suffice?

The underlying rationale of Fitzgibbon's method and the PDLD method is that

a. estimates of ultimate incurred losses can be obtained sooner than estimates of retrospective premiums can be obtained, and

b. retrospective premiums depend on incurred losses.

In workers compensation, for instance, a good estimate of ultimate incurred losses is generally available soon after the expiration of the policy, since claims emerge rapidly and development on known claims is relatively stable. The first retrospective adjustment, however, occurs about six months after the expiration of the policy. The retrospective premium may not be billed and collected for an additional three months after the adjustment is done.

Using Fitzgibbon's method or the PDLD method, an initial estimate of the accrued retrospective premium asset can be produced soon after the policy expires, using the known loss information and the relationships between incurred losses and retrospective premium. Similarly, the accrued retrospective premium asset estimate can be updated each quarter, as new loss data becomes available. If a chain-ladder premium development procedure is used, however, the initial estimate cannot be produced until at least nine months after the policy expiration, and it can be updated only annually thereafter.

The reserve estimation procedures in both Fitzgibbon's method and the PDLD method are based upon the retrospective rating formula. They differ in the details, not the concept, although the details can be crucial for reserve estimation. Using graphs to clarify the methods, the two approaches will be compared and contrasted using the following steps:

- how premium is determined in the retrospective rating formula;

- how Fitzgibbon, followed by Berry, converts the premium determination procedure to a reserve estimation procedure;

- what problems arise in the reserve estimation procedure, and how Berry resorts to a second reserve estimation procedure to resolve them;

- how the PDLD procedure modifies the original Fitzgibbon procedure to solve the aforementioned problems, without having to resort to a second reserve estimation procedure; and

- how part of Fitzgibbon's procedure can be used to enhance the PDLD procedure, giving users the best of both worlds.

Retrospective Premium Determination

Fitzgibbon's method and the PDLD method both seek to replicate the premium determination procedure in the retrospective rating formula. Of course, a single reserving formula cannot perfectly replicate hundreds of slightly different rating plans. Nevertheless, the more successfully the reserving procedure can replicate the rating procedure, the more accurate will be the reserve estimates. So let us begin with the premium determination formula.

The retrospective premium is composed of two parts:

1. Part of the premium covers the incurred losses, as well as any expenses associated with these losses, such as loss adjustment expenses. However, not all losses enter the retrospective rating formula. There is a loss limit, which means that individual losses exceeding a certain amount—such as \$250,000—do not affect the retrospective premium adjustments. In addition, state premium taxes, as well as other state assessments (such as involuntary market loads) are levied on the premiums, whether they are standard premiums or retrospective premium adjustments.

   The retrospective rating plan expresses this part of the premium as

   $$\begin{aligned}&(\text{loss conversion factor}) \times (\text{incurred losses})\\ &\qquad \times (\text{tax multiplier}),\end{aligned}$$

   where the loss conversion factor (LCF) covers primarily loss adjustment expenses.

2. The other part of the premium covers company expenses and the insurance charge. Company expenses are all expenses that are not a direct function of losses, such as underwriting expenses and acquisition expenses.

The insurance charge results from the maximum and minimum limitations on the retrospective premium. Having a maximum premium, of course, is the whole purpose of insurance. The insured needs protection against the unanticipated large losses that it cannot prudently retain. But the insurer must collect premium to cover these large losses. So the insurance charge is the difference between

a. the expected loss (to the insurer) caused by the maximum premium and

b. the expected gain (to the insurer) caused by the minimum premium.

The expected loss is the average additional amount of premium that the insurer would have collected had there been no maximum premium limitation. The expected gain is the average amount of premium that it would not have collected had there been no minimum premium limitation.

This charge must also cover any premiums lost because of the loss limits, which cap the individual loss values entering the retrospective rating plan.[3]

As before, a provision must be added for state premium taxes and other state assessments. This part of the premium may be expressed as

$$[(\text{expense provision}) + (\text{insurance charge})$$

$$+ (\text{excess loss charge})] \times (\text{tax multiplier}).$$

[3]The computation of the insurance charge is the standard Table M and Table L calculation. For the "formula" approach in the PDLD method, which can be used with Fitzgibbon's method as well, the reserving actuary may have to recompute certain Table M or Table L charges.

For simplicity, the first three components are combined into the basic premium, so the expression above can be restated as

$$(\text{basic premium}) \times (\text{tax multiplier}).$$

Thus, the formula for the retrospective premium is

$$\text{Retrospective premium} = (\text{tax multiplier})$$
$$\times [(\text{basic premium}) + ((\text{loss conversion factor})$$
$$\times (\text{limited incurred losses}))].$$

The Reserving Formula

The formula above is the rationale for Fitzgibbon's reserve formula. Premium is assumed to be a linear function of the incurred losses, or

$$\text{Retrospective premium} = C + B \times \text{Losses}.$$

The pricing formula becomes the reserving formula. For application to an entire book of business, Fitzgibbon and Berry make two modifications to this basic equation:

1. They use ratios to standard premium. That is, they write

   $$\text{Retrospective premium} \div \text{Standard Premium}$$
   $$= K + B \times \text{Standard Loss Ratio},$$

   where $K = C \div$ Standard Premium.

2. They examine the retrospective adjustment. In other words, they subtract unity from both sides of the equation above, to get

   $$\text{Retro Adjustment} = A + B \times \text{Standard Loss Ratio},$$

   where $A = K - 1$.

The Historical Regression

Fitzgibbon and Berry estimate the parameters *A* and *B* from a historical regression, using standard loss ratios and retrospective adjustments from mature policy years. But the attentive reader might observe that the two parameters in Fitzgibbon's formula depend on the parameters in the retrospective rating formula. So why do they use a regression analysis on past experience? Why don't they just walk over to the pricing actuary in the next office and ask what parameters are used in the retrospective rating plan?

Actuarial reserves are typically estimated on an aggregate basis, for all states, all insureds, all policy years. The parameters, however, vary from year to year, from state to state, and from plan to plan. For instance:

- A small insured may purchase a plan with a low maximum premium and therefore a large insurance charge, whereas a large insured may prefer a plan with a high maximum premium and a low insurance charge. Also, larger insureds may be offered plans with lower expense provisions, since their underwriting and acquisition expenses as a percentage of standard premium are lower than for smaller insureds.
- Premium taxes differ from state to state. In addition, some retrospective rating plans include involuntary market expense loads as a part of the tax multiplier, and the involuntary market loads vary widely among jurisdictions.
- The basic premium may vary from year to year. It may be low when interest rates are high and the insurer expects to earn its required profit margin from investment income. It may be higher when interest rates are low, or if the insurer uses a cash flow plan, such as a paid loss retro, so little investment income is retained by the insurer.

In theory, the reserving actuary could collect the hundreds of needed plan combinations and match these with the appropriate experience and calculate the reserve. Or the actuary, to save a few months of work, might determine the average parameters by means of a regression analysis on historical data.

This is what Fitzgibbon and Berry have done. The regression analysis calculates the average retrospective rating plan parameters from past experience. In fact, this method is probably more accurate than might be achieved by collecting all the parameters actually used in each state and each policy year for each insured. Most companies allow their underwriters and agents substantial flexibility in rating workers compensation contracts. The pricing actuary may recommend a basic premium charge of 30% of standard premium, but the underwriter or salesperson may reduce the basic premium charge to 25% of standard premium. The pricing actuary's recommended parameters may not match the plan parameters that are actually used in practice. The reserving actuary needs to know the premiums that are actually charged, not the pricing actuary's indicated premiums. So the reserving actuary turns to the regression analysis, not to the pricing actuary's rate book.[4]

---

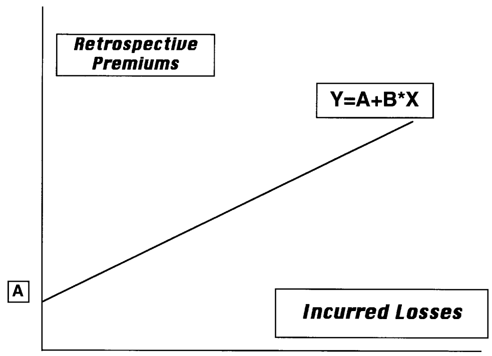
[4]How is it then, that Perkins and Teng manage to estimate PDLD ratios from the retrospective rating plan parameters in their formula approach? Moreover, they need to estimate more numbers than Fitzgibbon and Berry need to estimate, so how are they able to do this when Fitzgibbon and Berry found it unmanageable?

The answer is that the Perkins and Teng paper presents the method only. In practice, estimating the PDLD ratios from the retrospective rating plan parameters is exceedingly difficult, particularly if the company writes business in different states and for different types of insureds, if the company has changed its plan parameters over time, or if the company allows its underwriters and agents discretion in modifying the plan parameters to attract potentially good risks. Perhaps Ms. Perkins or Mr. Teng can elaborate on the relative ease or difficulty of estimating the PDLD ratios in various scenarios.

As pointed out by Robert Finger, the regression approach is not without its difficulties as well. Rating plan factors and aggregate loss ratios change over time, so a regression performed on historical data may not be equally applicable to current policies. Moreover, the observed values are actually the result of many changes at the individual plan level. The premium on individual plans is not a simple function of total incurred losses. For instance, premium may decrease on an adjustment when incurred losses increase, since there may be positive development on a claim that was already limited and negative development on claims that were below the per accident limit. See also Morell [10], which discusses this same issue.

FIGURE 1

FITZGIBBON'S METHOD



Graphical Representations

To see the difference between Fitzgibbon's method and the PDLD method, let us look at these procedures graphically. Fitzgibbon's method represents the relationship between the net earned premium[5] on the retrospectively rated book of business (as a percentage of standard premium) and the total incurred losses on this book of business (again, as a percentage of standard premium) as a straight line, as shown in Figure 1.[6] Algebraically, the straight line is $Y = A + B^*X$, where A is the constant factor and B is the slope factor.

One interpretation of this graph is as follows: if there are no incurred losses on this book of business, then the ratio of net premium to standard premium equals A. The constant factor A represents the basic premium percentage in the retrospective

---

[5]Net earned premium is earned premium after retrospective adjustments; see Feldblum [3].

[6]The figures on both axes of this graph are shown as ratios to standard earned premium. Alternatively, one could show both sets of figures as absolute dollar amounts. Berry uses ratios, though he shows the vertical axis as ratios of retrospective premium *returns* to standard premium. The three methods are equivalent.

rating formula.[7] As losses are incurred, and the loss ratio to standard premium increases, we move to the right and up along the straight line, and the net premium as a percentage of the standard premium increases. For each dollar of additional loss, the net retrospective premium increases by B dollars.

The slope factor B is the premium responsiveness for this book of business. The slope is not exactly unity, for several reasons. First, some losses exceed the loss limit, or they cause the retrospective premium to reach the maximum premium, even before the first adjustment, thereby reducing the slope of the line segment. Second, in some plans the minimum premium exceeds the basic premium. Third, a loss conversion factor and a tax multiplier are applied to the incurred losses in the retrospective rating formula, thereby changing the slope of the line segment. The combined effect depends on the "swing" of the plan. For plans with narrow swing, generally sold to small accounts, the slope would be less than unity. For plans with wide swing, generally sold to large accounts, the slope might be greater than unity.[8]

Projections versus Reality

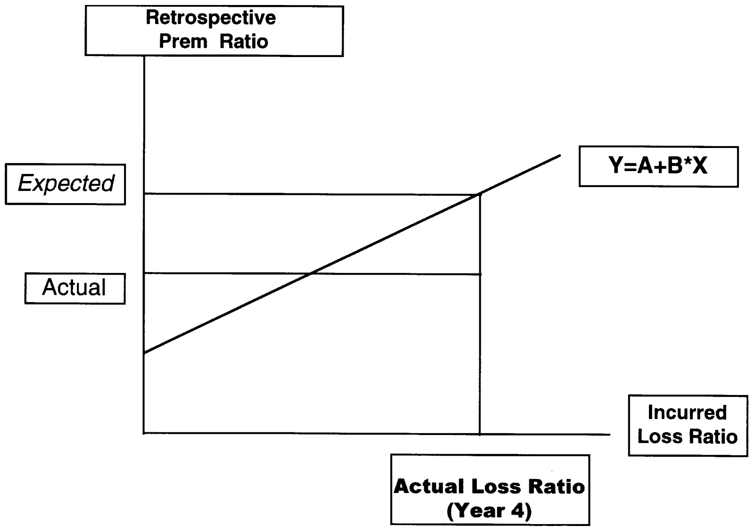
The problem with this method, as Berry points out, is that it does not consider the emerging experience on the book of business itself. This emerging experience may differ from that expected from the graph for several reasons. First, the A and B factors are only estimated from the regression; they are not known with certainty. Moreover, they may vary from year to year. Second, the pattern of losses among the individual policies

[7] Since the A factor is fitted by a regression on the aggregate book of business, it would not necessarily equal the basic factor on any particular plan.

[8] Fitzgibbon and Berry might say that this is not an exact interpretation of their regression line. Their regression line relates the *ultimate* loss ratio to the retrospective premium percentage. Their graph is not necessarily intended to represent the *movement* from no losses at policy inception to ultimate losses many years later. However, the purpose here is to highlight the contrast with the PDLD method, not to explain Fitzgibbon's method itself.

FIGURE 2

ACTUAL VERSUS EXPECTED RESULTS



affects the results. One large loss may have the same effect on the aggregate loss ratio as a dozen small losses. The effect on the net premium may differ because of loss limits and maximum premiums.

Suppose that after four years, the actual experience on this book of business shows less premium responsiveness than had been initially anticipated, as shown in Figure 2. The book of business is relatively mature after four years. The projection produced by this method does not change from year to year (as long as the incurred losses do not change), so it will continue to give an estimate of retrospective premium that is too high.

Berry's solution is to gradually discard this method, and to substitute a method that relies on the actual experience of the book of business (his "DR2" method). Initially, his reserve estimate relies entirely on this method. As time goes on, and more information becomes available from the actual book of business, he assigns progressively less weight to this method and more weight to his "DR2" method.

The Perkins and Teng Solution

Perkins and Teng transform Fitzgibbon's graph to solve this problem. Think of Fitzgibbon's graph in a slightly different fashion: as the movement over time of reported losses, net earned premium, and reported loss ratio. At policy inception, reported losses are \$0, so the reported loss ratio is 0% and the ratio of net premium to standard premium equals A, the constant factor in Fitzgibbon's regression equation, or the Y-intercept in Fitzgibbon's graph.

There are two ways to interpret the chart in Figure 1. Only the first of these reflects the intentions of Fitzgibbon and Berry. The second reflects the PDLD method. The alternative interpretations are:

1. the graph relates the ultimate loss ratio and the ultimate retrospective premium ratio among different books of business or different years of experience, or
2. the graph relates the reported loss ratio and the net earned premium at different points in time for a single book of business.

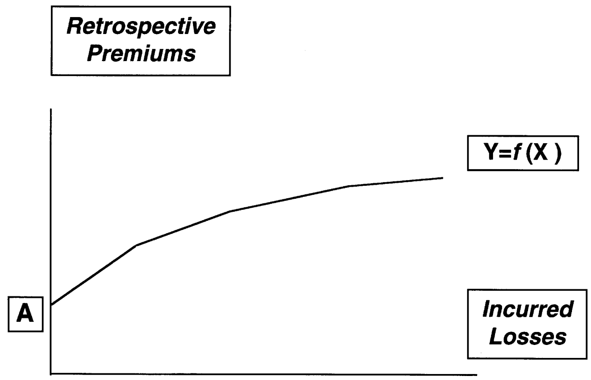
Decreasing Slopes

These two types of graphs seem similar. In truth, they look quite different. The first relationship is drawn by Fitzgibbon and Berry as a straight line. Actually, the curve is concave, as explained below, but a straight line is a close enough approximation for the majority of the curve.[9] The second relationship, however,

[9]It is a poor approximation at high loss ratios and at low loss ratios, though, where the maximum and minimum premium limitations flatten the curve. Fitzgibbon and Berry were aware of the approximation problems at the end points, and adjustments could always be made where necessary.

FIGURE 3

THE PERKINS AND TENG "PDLD" GRAPH



is not a straight line at all. Rather, it is a set of line segments, of steadily decreasing slope as we move to the right, as shown in Figure 3.[10]

The differing slopes of these line segments result from the loss limits and the maximum premiums in the retrospective rating plans. Most reported losses from policy inception until the first retrospective adjustment are rateable losses, which means that they are generally not truncated by the loss limit, and the retrospective premium is generally not capped by the maximum premium. The slope of the line segment is therefore close to unity. That is, for each dollar of reported loss, the insurer receives about a dollar of premium.

During subsequent periods, new reported losses stem from the emergence of IBNR claims and from development on known

[10]We use a series of line segments because retrospective adjustments are done annually, and the PDLD method reflects this by using line segments with different slopes for each adjustment period. In truth, a continuous concave curve better reflects reality, though it would not lead to a feasible reserving method.

claims. In workers compensation, for instance, new reported losses after the first adjustment may arise from the re-evaluation of a lower back sprain from a temporary total injury to a permanent total injury, with a corresponding re-estimation of the incurred loss from $25,000 to $500,000. This loss may be truncated by the loss limit in the retrospective rating formula, and the resulting retrospective premium may also be capped by the maximum premium.

This example is not contrived. On the contrary, it is quite common in workers compensation. Persons unfamiliar with industrial accidents often think of lifetime pension cases as quadriplegics or workers who have lost arms or legs in workplace accidents. Such injuries would be recognized immediately as high-cost, permanent total disabilities. These claims, which are recognized well before the first retrospective adjustment, are the ones that are most likely to be curtailed by the loss limits and maximum premium. This might lead some actuaries to think that the slope of the line segment in our graph should be flattest in the initial period.

In fact, accidents resulting in quadriplegia or the loss of arms or legs are rare. Most lifetime pension cases stem from sprains and strains and similar injuries that seem at first to be only temporary. After several years, when it becomes evident that the injured employee will not be returning to work, the claim is recorded as a permanent total injury and the benefit amount is re-estimated.[11]

We may state this as a general rule:[12]

1. *As a book of business matures, premium responsiveness on loss-sensitive contracts declines.*

---

[11]In the company at which the PDLD method was developed, fewer than 20% of claims that will ultimately be lifetime pension cases are recognized as such by the claims department at the first retrospective adjustment.

[12]As with any general rule, there are exceptions in particular instances.

In other words, as policies mature, a greater percentage of loss development is excluded from retrospective rating by the maximum premium and by the loss limit.

A second factor contributing to the declining slopes of the line segments is the overall increase in the reported loss ratio. It is not just that late-reported losses may be capped by the loss limit. Even a small claim will not increase the retrospective premium if the maximum premium has already been reached. Suppose the retrospective premium equals the maximum premium two years after policy inception. Then small claims reported during the first two years would have a premium responsiveness exceeding unity (because of the loss conversion factor and the tax multiplier), while small claims reported after the first two years would show a premium responsiveness of zero. We can state this second phenomenon as a general rule as well:

2. *At higher loss ratios, premium responsiveness on loss-sensitive contracts declines.*

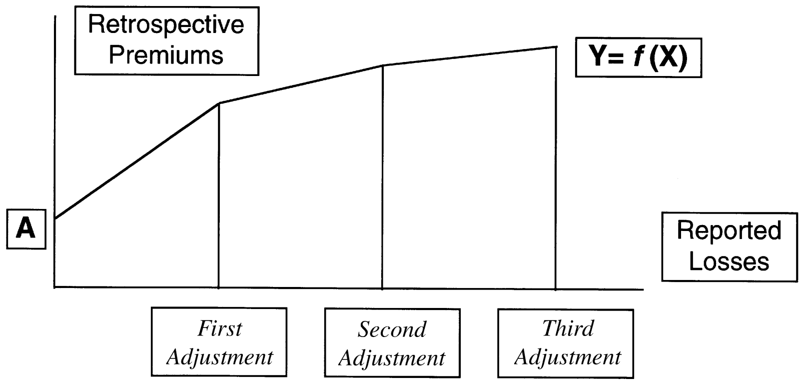
This last phenomenon relates to the overall loss ratio, not to the types of claims reported in any particular period. At higher overall loss ratios, more policyholders have reached their maximum premiums, so premium responsiveness is lower. Thus, it applies not only to the PDLD method, but to Fitzgibbon's method as well. That is, Fitzgibbon's graph is not really a straight line. In theory, it is a curve that is concave downwards, with steadily decreasing slope as the loss ratio increases.

Let us return to the PDLD method. At policy inception, the projected premium responsiveness graph is shown in Figure 4. Each line segment represents one period. The first line segment is from policy inception to the first retrospective adjustment, at about 21 months.[13] Subsequent periods are each one year long.

[13]The billing of retrospective premium generally lags the incurral of additional losses by about three months (on average) for an individual policy and by about nine months (on average) for a policy year. See below in the text for a full explanation of the lag times and effects that these may have on the observed premium responsiveness.

FIGURE 4

THE PDLD SEGMENTED GRAPH



The horizontal axis represents reported losses. For clarity, the graph is not drawn to scale. That is, the change in reported losses from policy inception to the first retrospective adjustment may be 50 percentage points or more in workers compensation, whereas the change in reported losses between adjustments at late maturities may be only a few percentage points. However, the graph shows all the line segments of equal length, so that the difference in their slopes can be seen clearly.

Actual versus Expected Experience

At the first adjustment, actual experience may differ in two ways from the experience that would be expected from the theoretical graph.

1. Actual reported losses may differ from the projected reported losses. For instance, at policy inception, the projected reported loss ratio to standard earned premium at 21 months may have been 55%. The actual reported loss ratio to standard earned premium at 21 months may be 50%.

2. The relationship between reported losses and retrospective premium may differ from that projected at policy inception. For instance, suppose that the Y-intercept in the graph is 20% and the slope of the first line segment is 1.100. Then for an actual reported loss ratio of 50% at the first retrospective adjustment, the ratio of net premium to standard premium is expected to be $20\% + 1.100 * 50\% = 75\%$. Suppose, however, that the actual ratio of net premium to standard premium at the first retrospective adjustment is only 72%.

These effects are shown in Figure 5 (not drawn precisely to scale).

- The projected experience at policy inception was for a reported loss ratio of 55% and a retrospective premium ratio of 80.5% [$= 20\% + 1.100 * 55\%$].
- For a reported loss ratio of 50% at the first retrospective adjustment, the graph projects a retrospective premium ratio of 75%.
- Actual experience at the first retrospective adjustment shows a reported loss ratio of 50% and a retrospective premium ratio of 72%.

The Perkins and Teng Assumptions

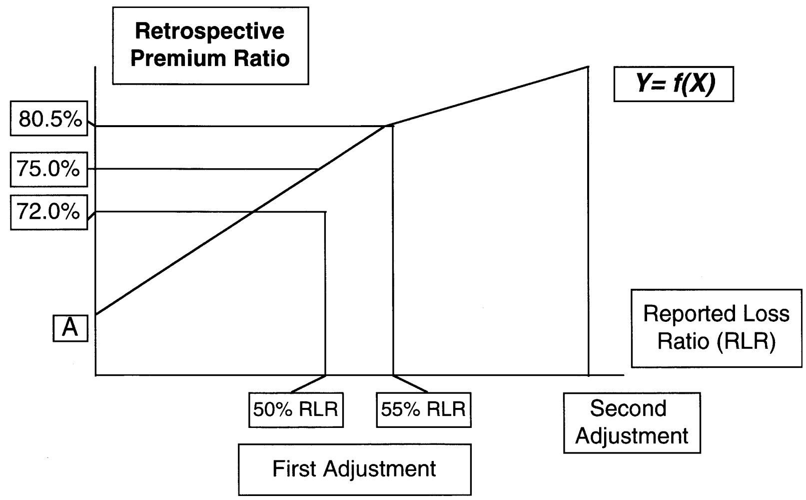
Two assumptions underlie the PDLD method. These are:

A. The premium responsiveness during subsequent adjustments is independent of the premium responsiveness during preceding adjustments.

B. The slope of the line segment depends on the time period, not on the beginning loss ratio or the beginning retrospective premium ratio.

FIGURE 5

PDLD METHOD: ACTUAL VERSUS EXPECTED RESULTS



We illustrate this for the first two line segments in Figure 5. Suppose the slope of the second line segment is 0.800. Think of the second line segment as an infinite number of parallel lines, all with slope of 0.800. At policy inception, we expected the second line segment to start at the point (55%, 80.5%) and to continue onwards with a slope of 0.800. As it turns out, the second line segment begins at the point (50%, 72%), but it still continues onwards with a slope of 0.800.

Compare the illustration with the two assumptions. We had expected a 75% retrospective premium ratio with a 50% reported loss ratio, but we actually get a 72% retro premium ratio. In other words, the slope of the first line segment is lower than we had originally expected. Nevertheless, we do not change our expectations for the slope of the second line segment. This is Assumption A.

The second assumption relates to when we change from the first line segment to the second line segment. From the appearance of the graph in Figure 5, one might think that we change when the reported loss ratio reaches 55%. That is not the meaning of the graph. Rather, we change at the first adjustment, regardless of the reported loss ratio at that time.

The manner in which the PDLD method solves Berry's problem should now be clear. Fitzgibbon's graph relates the *ultimate* loss ratio to the ultimate retrospective premium ratio. If actual experience differs from expected experience along the way, there is no way to get back on track. The PDLD method relates the *reported* loss ratio to the retrospective premium ratio. If actual experience differs from expected experience along the way, the next line segment begins at a starting point that corresponds to the actual experience.

The PDLD method quantifies the accrued retrospective premium asset in two steps.

1. Project the future loss development in each adjustment period.

2. Estimate the future premium revenue by the product of the future loss development in each period and the slope of the line segment in that period. The sum of these products is the accrued retrospective premium asset.

The PDLD method can be thought of as follows. The line segments represent a mountain being climbed, from the 0% reported loss ratio at policy inception to the ultimate loss ratio when all losses are settled. At each retrospective adjustment, the remaining part of the climb is shifted, both horizontally and vertically, but the shape of the climb is not changed (that is, the slopes of each line segment remain fixed).[14]

An Enhancement

In Figure 5, the first line segment begins at a point on the Y-axis representing the amount of retrospective premium when the reported loss ratio is 0%; that is, the Y-intercept is positive. This is the proper way to estimate the accrued retrospective premium asset. Perkins and Teng, however, have the first line segment passing through the origin; that is, the Y-intercept is 0. As a result, Perkins and Teng get a slope for the first line segment of 1.750. In fact, empirical data in their Exhibit 4, Sheet 1 for the most recent four quarters shows an average slope of 1.825.

Perkins and Teng's numbers combine two separate items: the basic premium ratio and the slope of the first line segment (when drawn properly). By failing to distinguish between these two elements, the method becomes less intuitive: how does one explain a slope of 1.825 or 1.750?

Similarly, the combination of these two elements leads to confusing interpretations. For instance, when discussing the cumulative premium development to loss development ratios (CPDLD), Perkins and Teng write:

> The CPDLD ratio tells how much premium an insurer can expect to collect for a dollar of loss that has yet to emerge. For instance, the first CPDLD ratio is 1.492, which means that each dollar of loss emerged provides the insurer one dollar and 49 cents of premium. The second CPDLD ratio is 0.556, which means that after the first retro adjustment, each additional dollar of loss provides the insurer 56 cents of premium.

[14]Actually, although the slopes of each line segment remain fixed, the length of the line segments may be changed. At each retrospective adjustment, Perkins and Teng re-estimate the losses expected to be reported in each subsequent period. These revisions, however, are generally minor.

The interpretation of the second CPDLD ratio is correct. The interpretation of the first CPDLD ratio, however, is mistaken. The first CPDLD ratio relates to all the expected losses from policy inception, at least according to the procedure in the Perkins and Teng paper.

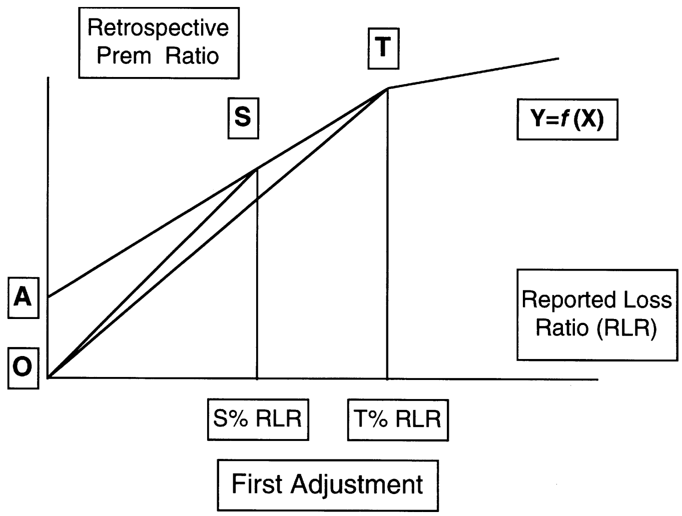
How should we interpret the 1.492 CPDLD ratio from policy inception to the first retrospective adjustment? Consider a relatively wide-swing retrospective rating plan: that is, a plan with high loss limits and maximum premiums. The amount of expected premium for each dollar of loss equals the loss conversion factor times the tax multiplier, minus a small amount for the non-rateable losses. This product may be about 1.200. The remainder of the first CPDLD ratio which Perkins and Teng calculate is the basic premium charge divided by the expected loss ratio (as a function of standard premium). For a basic premium charge of 25% and a standard loss ratio of 85%, this calculation gives $0.25 \div 0.85 = 0.294$. Adding 1.200 to 0.294 gives 1.494, which is about equal to the empirical figure which Perkins and Teng compute. In other words, when the basic premium charge is disentangled from the slopes of the line segments, the Perkins and Teng procedure corresponds intuitively with the actual retrospective rating formula.[15]

The failure to separate these two issues makes it harder for the actuary to analyze changes in the figures over time. For instance, what causes the steady rise in the slope of the first line segment from an average of 1.254 in policy year 1963 to an average of 1.825 in policy year 1992 (see Exhibit 4, Sheet 1 in the original paper)? Is it caused by a change in the average basic

[15]For a plan with significant loss limits or maximum premiums, the intuitive is analogous. The lower the loss limits, or the lower the maximum premium, the weaker will be the premium responsiveness, but the basic premium charge will be greater, because the insurance charge will be larger. These two effects will offset each other, since the insurance charge is calculated as the expected losses arising from the loss limits and maximum premiums.

FIGURE 6

LOSS REPORTING PATTERNS IN THE PDLD GRAPH



premium ratio, or is it caused by a change in premium responsiveness during the first period? These two factors are shown separately in the graphs drawn in this discussion, but they are not easily distinguished in the way that Perkins and Teng show their procedure.

This change could also be caused by a lengthening of the loss reporting pattern. This is an equally likely cause, and a graphical representation of it is illuminating.

In Figure 6, the basic premium ratio and the slope of the first line segment are not changed, but the percentage of losses expected to be reported before the first adjustment is decreased. That is, the expected ultimate loss ratio remains the same, but the expected reported loss ratio at the first adjustment decreases from T to S. The first line segment is therefore shorter, though it has the same slope. In the PDLD procedure, however, the slope of the first line segment appears to increase. That is, the slope from 0 to S is greater than the slope from 0 to T.[16]

Fortunately, it is simple to adjust the PDLD method to show the basic premium ratio separately from the true slope of the first line segment. One need only estimate the average basic premium charge as a ratio to the standard loss ratio, and then subtract this figure from the first CPDLD.

3. LOSS-SENSITIVE CONTRACTS AND UNDERWRITING RISK

Insurance serves several important economic functions, such as the transfer of the risk of financial loss from the consumer to the insurance company. Because of the unlimited nature of workers compensation benefits, a single severe workplace injury might financially impair a small employer. The transfer of this risk from the employer to the insurance company is a societal benefit of workers compensation insurance.

A societal downside to insurance is moral hazard. If there were no workers compensation insurance, then employers would take great pains to keep their workplaces as safe as possible, since they would shoulder any cost of workplace accidents. Insurance has two effects on employers' safety efforts. On the one hand, the loss engineering staffs of most workers compensation carriers can identify potential workplace hazards and improve employers' safety procedures. On the other hand, some employers become less concerned with employee safety, since they no longer bear all the costs.

An increase in moral hazard hurts both employees and employers. It hurts employees since workplace accidents may in-

[16]The effect is even more pronounced in the Perkins and Teng graph, which is drawn as a concave curve instead of a series of line segments.

crease. It hurts employers in numerous ways: there are training costs for new employees, work flows are interrupted, and workers compensation premiums increase to cover the higher loss costs.

Retrospectively rated contracts are an attempt to achieve the benefits of insurance while reducing the drawbacks. Employers are protected from the risk of large losses that might otherwise bankrupt the firm. But they still bear the cost of most other injuries, so moral hazard is kept low.

Insurance involves the transfer of risk from the consumer to the insurer. In retrospectively rated contracts, some of this risk is transferred back to the consumer. The NAIC has developed the loss-sensitive contract offset to the underwriting risk charges in the risk-based capital formula in order to reflect the fact that the risk on retrospectively rated contracts differs from the risk on prospectively rated contracts. Previous actuarial studies had not addressed this question, and the American Academy of Actuaries Task Force on Risk-Based Capital had little actuarial or statistical data to give to the NAIC.

The PDLD procedure, however, provides a direct answer. In fact, the Perkins and Teng paper sheds light on the potential limitations of both the risk-based capital loss-sensitive contract offset and the loss-sensitive contract exhibits in Part 7 of Schedule P.

Underwriting Risk

The insurance contract transfers the risk of random loss occurrences from the consumer to the insurance company. This risk is primarily process risk. For instance, suppose the consumer is an employer concerned with industrial accidents. The employer may estimate that there is a one in one hundred chance of a severe accident in his workplace this year. The primary risk that this employer faces is *not* that he has misestimated the probability—

that it is truly one in ninety, not one in one hundred. Nor is it the risk that the cost of such accidents may change, say from an average of $20,000 per accident to $25,000 per accident. Rather, the primary risk is that an accident will indeed occur this year in his workplace.

The risk to the insurance company is different. It is primarily parameter risk, not process risk. If the book of business is large enough, process risk effectively disappears. However, the risk that the probability of an accident is truly one in ninety, or the risk that the average cost of these accidents is truly $25,000, are serious concerns for the insurer. A relatively small error in the estimation of these parameters may wipe out the expected profits of the insurer.

Loss-sensitive contracts mitigate this risk for the insurance company. The insured is still protected against random large losses by the loss limit in the retrospective rating plan and by the maximum premium. Meanwhile, the insurance company is protected against the accumulation of more losses than expected, or a rise in the average cost per claim, by the responsiveness of retrospective premiums to incurred losses.[17]

Underwriting risk has two facets. Premium risk (or "written premium risk," in the NAIC risk-based capital terminology) is the risk that future premiums will prove inadequate to cover the future losses and expenses. This risk takes a variety of forms. For instance, there is a market risk that the competitive pressures of an underwriting cycle downturn will force premium rates below adequate levels. There is a regulatory/political risk that needed premium increases will not be approved or that new types of claims will be deemed compensable by the courts.

Reserving risk is the risk that the reserves held for accidents that have already occurred may prove inadequate. Once again,

[17]For a full discussion of the effects of loss-sensitive contracts on workers compensation reserving risk, see Hodes, Feldblum and Blumsohn [7].

this risk takes a variety of forms. For instance, there is the economic risk that a recession will cause injured employees to remain on disability for longer periods, since there may be no jobs to return to (workers compensation). Or there may be judicial risk, that courts or juries may grant higher awards to claimants (general liability).

Loss-Sensitive Contracts and Underwriting Risk

Loss-sensitive contracts reduce the risks to the insurer, since if losses are higher than expected, additional premiums are collected from the insureds. When the NAIC instituted its risk-based capital formula, which quantified the capital needed to guard against written premium risk and reserving risk, several large commercial lines insurers argued that a capital requirement that is appropriate for prospectively rated business is too high for retrospectively rated business, since the retrospective rating formula itself protects against unexpectedly high losses.

But how effective are these contracts in mitigating risk? In other words, how responsive are the premiums to unexpected losses?

If there were no loss limits or maximum premiums in the retrospective rating plans, the premium responsiveness would equal the product of the loss conversion factor and the tax multiplier. We term this 100% responsiveness, since the loss conversion factor generally covers loss-related expenses and the tax multiplier pays for premium taxes (and other state assessments) that depend upon the losses incurred or the premium collected. In other words, with 100% responsiveness, the insurer would get \$1.00 in extra premium for each \$1.00 in additional losses and loss-related expenses.

If there were no loss limits or maximum premiums in the retrospective rating plans, then the insurer would not be exposed to underwriting risk. If underwriting results are worse than expected, or if reserves develop adversely, the insurer would collect the full loss from the insured through retrospective premium adjustments. There remain some other risks, such as the credit risk that the insured will not be able to pay the retrospective premiums when they come due, but these risks are usually far smaller than the underwriting risk.

In practice, of course, there are loss limits and maximum premiums. Premium responsiveness is less than 100%. So the NAIC instituted a 30% loss-sensitive contract offset on primary insurance policies and a 15% loss-sensitive contract offset on reinsurance treaties. The loss-sensitive contract offset of 30% means that if the risk-based capital underwriting risk charge for a block of prospectively rated business is $\$X$, then the corresponding charge for the same book of business written on loss-sensitive contracts is $\$X * (1 - 30\%)$.[18]

In other words, the primary insurance loss-sensitive contract offset assumes (conservatively) that the premium responsiveness is only 30%. That is to say, for each $1.00 in additional losses and loss-related expenses, $0.30 of additional premium (on average) is collected.

The 30% figure was not based on definitive data because credible industry data on premium responsiveness was not available. The consulting firm Tillinghast/Towers Perrin conducted an industry-wide survey of 16 large writers of retrospectively rated contracts, and calculated an average premium responsiveness of 65%. The survey asked insurance companies how responsive they thought their loss-sensitive contracts were to unexpected loss emergence or unexpected loss development. The 65% was a rough average of the company estimates. Adjusting this figure downward for conservatism and for the potential

[18]For a complete description of the loss-sensitive contract offset in the risk-based capital formula, see Feldblum [5].

credit risk led to the 30% offset factor in the risk-based capital formula.[19]

In order to obtain industry data to more accurately estimate the loss-sensitive contract offset factor, the NAIC added Part 7 to Schedule P. The exhibits in this section of Schedule P are designed to allow the estimation of premium responsiveness on loss-sensitive contracts. These exhibits are a considerable advance over the information available previously, but they are far less useful than the information provided by reserving studies using the PDLD method.

In the future, insurance companies will seek to better quantify the effects of loss-sensitive contracts on underwriting risk, and state regulators will attempt more accurate estimations of the appropriate offset factor for these contracts. The study by Perkins and Teng highlights several areas that must be carefully considered.

Time Frames

The Schedule P Part 7 exhibits are the NAIC's attempt to quantify premium responsiveness, using the same method as Perkins and Teng, but with annual reporting of premiums and losses. The Perkins and Teng paper shows that the Schedule P results will be distorted in several ways, possibly to the extent that premium responsiveness will not be shown at all. Some of the problems can be corrected (in theory, at least) by means of the procedures in the Perkins and Teng paper; other distortions may be more difficult to remove.

---

[19]The rationale given by the Tillinghast study and adopted by the NAIC for the lower (15%) offset factor used for reinsurance treaties reflects the different types of loss-sensitive contracts generally used by primary companies and by reinsurers. The primary company retrospective rating plan adjusts the premiums billed for adverse loss experience. Some of these plans have extremely wide swings, in that the final premium may be as much as 100% more than the standard premium. Reinsurers generally use sliding scale commissions, in that the reinsurance commission remitted to the ceding company depends upon the loss experience on the book of business. Since the commission rate

The intended use of the Schedule P Part 7 exhibits is not explained in the Annual Statement Instructions, and few actuaries understand how these exhibits purport to quantify premium responsiveness. Let us first clarify the intention of this part of Schedule P with an illustration. We will then explain the problems with the statutory exhibits by a comparison with the Perkins and Teng paper.

The risk-based capital reserving risk charge is based on the loss reserves—both case and IBNR reserves—that are shown by the company's Schedule P, Part 2, minus Schedule P, Part 3. The reserving risk charge quantifies the capital needed to protect against the risk that these reserves may develop adversely in a worst-case scenario. The loss-sensitive contract offset factor reduces this capital requirement to reflect the additional premium that the insurer expects to receive in this worst-case scenario.

The dollar amount of adverse development of the loss reserve equals the dollar amount of adverse development of the incurred losses in Schedule P, Part 2. Part 7 of Schedule P displays incurred losses on loss-sensitive contracts and the corresponding adverse or favorable premium development relative to the adverse or favorable loss development.

An Illustration

An example should clarify this. Suppose we are given the extracts from Schedule P, Part 7A, Sections 2 through 5 shown in Table 1 (figures are in thousands of dollars). The actual exhibits contain more cells, but these extracts suffice to illustrate the quantification techniques. We wish to determine premium responsiveness from 24 to 36 months and from 36 to 48 months.

The sections of Schedule P, Part 7A, contain the following historical triangles, by policy year and valuation date, of experi-

---

is bounded below by 0%, and in many treaties it is bounded below by an even higher amount, the swing of the typical reinsurance treaty is much narrower than that of many primary retrospective rating plans.

TABLE 1

SCHEDULE P, PART 7A, SECTIONS 2, 3, 4, AND 5, SELECTED ENTRIES ($000)

Section 2	1994	1995	1996	1997
1994	1,000	2,200	2,400	2,500
1995		1,100	2,500	2,650
1996			1,200	3,000
1997				1,500

Section 3	1994	1995	1996	1997
1994	350	550	300	200
1995		400	600	450
1996			450	650
1997				500

Section 4	1994	1995	1996	1997
1994	1,500	3,150	3,300	3.350
1995		1,650	3,600	3,700
1996			1,800	4,200
1997				2,000

Section 5	1994	1995	1996	1997
1994	0	200	150	110
1995		0	210	155
1996			0	220
1997				0

ence on loss-sensitive contracts:[20]

- Section 2: Incurred losses and ALAE on loss-sensitive contracts
- Section 3: IBNR plus bulk loss and ALAE reserves on loss-sensitive contracts
- Section 4: Earned premium on loss-sensitive contracts

[20]For a full description of Schedule P, Part 7, see Feldblum [4].

- Section 5: Accrued retrospective premium reserves on loss-sensitive contracts.

This illustration is contrived. It is designed to show how Part 7 of Schedule P was intended to be used. We then examine how the Perkins and Teng paper explains the problems with this use of the Part 7 exhibits.

These exhibits are policy year exhibits, not accident year losses (as in Parts 2, 3, and 4 of Schedule P) or exposure year premiums (as in Part 6 of Schedule P). In Section 2 of Part 7, the incurred losses as of 24 months are about twice the incurred losses as of 12 months. This makes sense: the policy year 1994 incurred losses as of 12 months are those losses on policies written in 1994 that have occurred by December 31, 1994. These are about half of the policy year 1994 losses. By December 31, 1995, all of the policy year 1994 losses have occurred (though they have not necessarily all been reported by this time), so the 24 month figure is about twice as great as the 12 month figure.

The same is true for Section 4, showing the policy year earned premiums. By the end of the policy year, all the premiums have been written (though not necessarily collected), but only about half of these premiums have been earned.

This example assumes that the initial written premium for this block of business is the estimated ultimate net premium. Initially, there is no retrospective premium reserve. At the first retrospective adjustment, some premiums are returned to policyholders, since not all losses have yet been recorded, even though the insurer knows that there will probably be development on the reported losses. The accrued retrospective premium asset becomes positive after the first adjustment. For companies that charge initial premiums below the estimated ultimate net premium (for competitive reasons), the accrued retrospective premium asset will be positive from policy inception.

Quantifying Premium Responsiveness

Consider first the premium responsiveness from 24 to 36 months. Only policy years 1994 and 1995 in our illustration are mature enough to measure this.[21] For policy year 1994, losses develop from $2.20 million to $2.40 million from 24 months to 36 months, for a change of $0.20 million. Premiums develop from $3.15 million to $3.30 million during the same time period, for a change of $0.15 million. The premium responsiveness is $0.15 million ÷ $0.20 million, or 75%.

For policy year 1995, losses develop from $2.50 million to $2.65 million from 24 months to 36 months, for a change of $0.15 million. Premiums develop from $3.60 million to $3.70 million during the same time period, for a change of $0.10 million. The premium responsiveness is $0.10 million ÷ $0.15 million, or 67%.

As the estimated premium responsiveness from 24 months to 36 months, we might take the average of these two numbers. Alternatively, we might give more weight to the 1995 policy year, particularly if the rating plan parameters had changed in 1995.

For the premium responsiveness from 36 months to 48 months, only policy year 1994 is sufficiently mature to provide the needed figures. Losses develop from $2.40 million to $2.50 million from 36 months to 48 months, for a change of $0.10 million. Premiums develop from $3.30 million to $3.35 million during the same time period, for a change of $0.05 million. The premium responsiveness is $0.05 million ÷ $0.10 million, or 50%.

This is consistent with the Perkins and Teng paper. As reserves mature, premium responsiveness diminishes, since more losses are censored by the loss limit and more premiums are capped

[21] In an actual Schedule P, all earlier policy years would also show this relationship.

by the maximum premium. In addition, at later maturities, some retrospective rating plans are closed.

This example was designed to illustrate the intended use of the Schedule P exhibits; it would rarely be encountered in practice. The incurred losses here develop smoothly upward, and the premiums follow them equally smoothly. An adequately reserved company should show flat incurred losses along development periods, and similarly flat earned premiums. After all, these incurred losses include IBNR and bulk reserves, and the earned premiums include the accrued retrospective premium asset. The changes in incurred losses from period to period would be sometimes small and sometimes large, sometimes positive and sometimes negative, resulting primarily from random loss fluctuations. The changes in earned premiums from period to period would be equally variable, resulting again from random loss fluctuations as well as from censoring by the loss limits and capping by the premium maximums.[22]

We have two series of variable figures with means of zero, since favorable and adverse development are equally likely (in theory, at least). The ratios of these series will be even more variable, sometimes very high, sometimes very low, sometimes positive, and sometimes negative. These ratios may not tell us much about premium responsiveness.

Reported Losses and Billed Premium

As the Perkins and Teng paper shows, premium responsiveness does not deal with the relationship of changes in total earned premium to changes in total incurred losses. Rather, it deals with the relationship of changes in billed premium to changes in re-

[22]The date of recognition of additional losses or additional accrued retrospective premium reserves would add to the variability in the two series of changes, one of incurred losses and one of earned premiums. That is, the reserving actuary may recognize the potential increase in ultimate losses in one year, but he or she may not book the corresponding increase in the accrued retrospective premium reserves until some later time.

ported losses. Accordingly, Schedule P, Part 7 allows that analysis to be performed as well.

Section 2 of Part 7 shows incurred losses, and Section 3 shows IBNR and bulk reserves. The difference between Sections 2 and 3 represents reported losses.[23] Similarly, Section 4 shows total earned premiums, and Section 5 shows the net reserve for premium adjustments and accrued retrospective premiums. The difference between Sections 4 and 5 represents billed premium.

Let us repeat the premium responsiveness calculations using the simulated Schedule P, Part 7 exhibits provided above. For the premium responsiveness from 24 months to 36 months, we have data from policy years 1994 and 1995. For policy year 1994, reported losses develop from ($2.2 million–$0.55 million) at 24 months to ($2.4 million–$0.3 million) at 36 months, for a change of $0.45 million. Billed premium develops from ($3.15 million–$0.2 million) at 24 months to ($3.3 million–0.15 million) at 36 months, for a change of $0.20 million. Premium responsiveness from 24 months to 36 months is $0.20 million ÷ $0.45 million = 44.4%.

For policy year 1995, reported losses develop from ($2.50 million–$0.60 million) at 24 months to ($2.65 million–0.45 million) at 36 months, for a change of $0.30 million. Billed premium develops from ($3.6 million–$0.21 million) at 24 months to ($3.70 million–$0.155 million) at 36 months, for a change of $0.155 million. Premium responsiveness from 24 months to 36 months is $0.155 million ÷ $0.30 million = 51.7%.

Anticipated Emergence versus Unanticipated Development

These figures do indeed reflect reality, but is this reality related to the risk-based capital loss-sensitive contract offset factor?

---

[23] This is the same as the calculation of accident year reported losses as Part 2 of Schedule P minus Part 4 of Schedule P.

The risk-based capital reserving risk charge seeks to quantify the amount of capital needed to guard against unanticipated adverse development of loss reserves. For instance, if in a worst-case (but still reasonable) scenario, the company's reserves would develop adversely by \$15 million, then the company should hold \$15 million of capital to ensure its solvency.

The figures calculated in the preceding section measure the responsiveness of retrospective premiums to the emergence of anticipated losses. They do not tell us how responsive the retrospective premiums would be to the emergence of unanticipated losses.

An example should clarify this. Suppose we are examining the premium responsiveness from 24 months to 36 months on a workers compensation retrospectively rated plan with an average swing. Suppose that at 24 months the reported losses are \$100 million, and the anticipated reported losses at 36 months are \$120 million. The expected ultimate losses are \$150 million.

From our hypothetical experience, we find a premium responsiveness for this period of 50%. That is to say, when reported losses increase by \$20 million, the billed premium increases by \$10 million. What are the implications for large and unanticipated adverse loss development, as envisioned in the risk-based capital worst-case scenario? For example, if the ultimate losses are re-estimated at \$180 million at 36 months instead of \$150 million, will the accrued retrospective premium asset increase by an additional \$15 million, or 50% of the additional losses of \$30 million?

Consider the real-world characteristics of the numerical example given above. The development of reported losses from \$100 million to \$120 million from 24 months to 36 months may be decomposed into several parts. One part is the lengthening of some temporary cases for another few months, or an increase in some medical benefits. This development is rateable, so premium responsiveness is high. Another part is the reclassification of some temporary total cases, such as lower back sprains, into lifetime pension cases, when it becomes clear that the injured employee will not be returning to work. Only some of this development is rateable, and the rest is truncated by the loss limits or the maximum premiums.

Large and unanticipated adverse loss development has a heavy proportion of this nonrateable element. The re-estimation of the ultimate losses from $150 million to $180 million may result from the re-classification of several back sprains as severe and permanent disabilities, or from a judicial or legislative decision that certain disease claims, or psychiatric claims, are compensable. These claims are generally large and they are paid over a long period of time. A large part of these claims may not be rateable.

The Perkins and Teng paper discusses these issues. As noted above in this discussion, the premium responsiveness depends on the maturity of the losses as well as on the average loss ratio in the block of business. The emergence of anticipated losses differs from the unanticipated adverse development of the expected losses in that:

- the anticipated losses are generally paid sooner than the unanticipated losses, and
- the anticipated losses generally represent a lower loss ratio than do the unanticipated losses.

Since the anticipated losses are generally paid sooner, they are accompanied by a stronger premium responsiveness. Since the anticipated losses are generally in a lower loss ratio environment, they are associated with a stronger premium responsiveness. In sum, the figures derived from the historical triangles in Schedule P, Part 7 may not be relevant to the scenarios with which risk-based capital is concerned.

Reserving Risk Offset versus Premium Risk Offset

The NAIC risk-based capital formula uses the same loss-sensitive contract offsets for reserving risk as for written premium risk: 30% for primary insurance contracts and 15% for reinsurance contracts. As the Perkins and Teng paper shows, the offset should be much higher for written premium risk than for reserving risk.[24]

For the written premium risk loss-sensitive contract offset, one must examine the first CPDLD factor in a Perkins and Teng study. However, one must separate the basic premium charge from the premium responsiveness to losses, or the offset factor will be overstated; see the discussion above for further explanation of this. Moreover, one must remove the effects of the loss conversion factor and the tax multiplier, which would also overstate the appropriate offset factor.

For the reserving risk loss-sensitive contract offset, one must examine the CPDLD factors at each maturity. One would then weight these CPDLD factors by the distribution of reserves at each maturity. As is true for the written premium risk loss-sensitive contract offset, one must remove the effects of the loss conversion factor and the tax multiplier.

The difference between premium responsiveness to the emergence of anticipated losses and premium responsiveness to unanticipated adverse loss development (or unanticipated adverse un-

[24]The appropriate figures depend on the types of plans sold by the insurance company. The indicated range of figures is wide, and the type of analysis used by Perkins and Teng must be applied to each company's book of business. For instance, for a workers compensation carrier selling wide-swing plans to large national accounts, the appropriate figures may be between 80% and 85% for the written premium risk loss-sensitive contract offset and between 60% and 65% for the reserving risk loss-sensitive contract offset. For a company selling narrow swing plans to small risks, the offsets are much smaller, extending down as far as the figures used in the NAIC risk-based capital formula. For a full analysis of premium sensitivity on plans sold to small accounts, see Bender [1] and Mahler [9].

derwriting results) can be significant. In the Perkins and Teng framework, the CPDLD's should be based on a book of business with an overall loss ratio equal to the worst-case year loss ratio in the NAIC risk-based capital scenario. Empirical data for such CPDLD's are not readily accessible. Approximations by curve-fitting techniques to the CPDLD's that are empirically available may have to be substituted.

Premium Billing Lags

Another section of the Perkins and Teng paper brings to light an equally significant problem with the Schedule P exhibits. When quantifying premium responsiveness, it is important to use corresponding premiums and losses. Premium billing occurs about 3 months after the retrospective adjustment. This implies that the premium billing lags the average loss occurrence by 3 to 15 months.

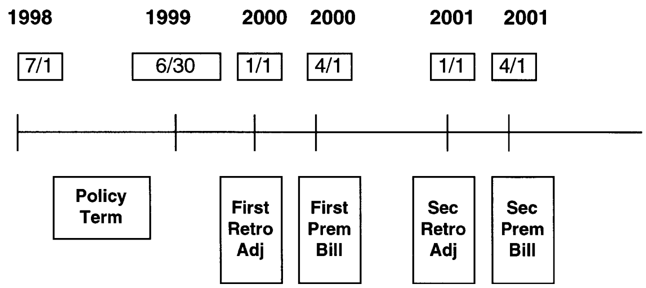
An example should clarify these figures. Suppose a policy is effective from July 1, 1998 through June 30, 1999. Retrospective adjustments are done six months after the policy's expiration and every 12 months subsequently. For this policy, the retrospective adjustments will be done on each January 1, starting with January 1, 2000. The resulting retrospective premium adjustment will be billed or returned to the policyholder on each April 1.

Each retrospective premium adjustment is driven by losses that are reported between 15 months and 3 months prior to the premium billing date. For this policy, losses reported between January 1 and December 31 affect the premium adjustment that will be billed on April 1. The schematic in Figure 7 shows this graphically.

The average lag between loss reporting and premium billing is 9 months. This is the lag used by Perkins and Teng. If one does not use any lag, as was the intention of the designers of Schedule P, Part 7, the results will be distorted. To see this most

FIGURE 7

PREMIUM AND LOSS DATES FOR RETROSPECTIVELY RATED POLICIES



clearly, suppose that:

- the retrospective premium billing is done on July 1,
- all losses occur on July 1,
- there is 100% premium responsiveness, and
- the annual incurred losses alternate between \$1,000 and \$0.

The Schedule P, Part 7, premium responsiveness test would show the following:

Year	1	2	3	4	5	6
Change in incurred losses	\$1000	\$0	\$1000	\$0	\$1000	\$0
Change in billed premium	—	\$1000	\$0	\$1000	\$0	\$1000

The premium billing shows up a year after the loss occurs. In this example, there is 100% premium responsiveness, but Schedule P, Part 7, shows a -100% premium responsiveness.[25]

[25] If X denotes the change in incurred losses, and Y is the change in billed premium, than 100% premium responsiveness is represented as $Y = 100\% * X$. This policy's experience shows a line of $Y = \$1000 - 100\% * X$. In the actual calculations of premium respon-

In practice, simplistic examinations of premium responsiveness may yield regression coefficients which are negative or seemingly random. The reserving actuary may conclude that the data are incorrect, when the true problem is an improper matching of premiums and losses.

The Perkins and Teng paper shows a possible solution to our problem. Ideally, one should use quarterly data, with a 9-month lag between premium billing dates and loss reporting dates. Few insurers have this data, and the costs of obtaining such data far outweigh any benefits from these exhibits. As a practical alternative, one should use a 12-month lag in the quantification of premium responsiveness. A 12-month lag is not ideal, but it is better than no lag at all. Moreover, this requires no change in the exhibit completion process: the same exhibits may be used, but the quantification procedure would be modified.

4. CONCLUSION

Miriam Perkins and Michael Teng have put together an excellent paper, based on eight years of carefully examining the accrued retrospective premium reserves in workers compensation, general liability, and commercial auto for one of the country's largest writers of retrospectively rated policies. They methodically analyzed how premium responsiveness changes by reserve maturity and by aggregate loss ratio, and they systematically tested the lags between loss reporting and premium billing in the company's book of business.

The Perkins and Teng procedure is important not just for reserve projections but also for risk analysis. Our profession has much to gain as other actuaries learn the techniques presented by Perkins and Teng and use them to quantify the risk and rewards of loss-sensitive contracts.

---

siveness, of course, one does not use successive adjustments for a single policy or block of policies, but successive calendar years for the same adjustment for successive blocks of policies. The underlying concepts are the same, though the schematic becomes more complex.

REFERENCES

[1] Bender, Robert K., "Aggregate Retrospective Premium Ratio as a Function of the Aggregate Incurred Loss Ratio," *PCAS* LXXXI, 1994, pp. 36–74.

[2] Berry, Charles H., "A Method for Setting Retro Reserves," *PCAS* LXVII, 1980, pp. 226–238.

[3] Feldblum, Sholom, "Workers' Compensation Ratemaking," CAS Examination 5 Study Note, 1993.

[4] Feldblum, Sholom, "Completing and Using Schedule P," *Regulation and the Casualty Actuary*, NAIC, 1997.

[5] Feldblum, Sholom, "NAIC Property-Casualty Risk-Based Capital Requirements," *PCAS* LXXXIII, 1996, pp. 297–435.

[6] Fitzgibbon, Walter J., Jr., "Reserving for Retrospective Returns," *PCAS* LII, 1965, pp. 203–214.

[7] Hodes, Douglas M., Sholom Feldblum, and Gary Blumsohn, "Workers Compensation Reserve Uncertainty," Casualty Actuarial Society *Forum*, Summer 1996, pp. 61–149.

[8] Hope, F. J., Discussion of "Reserving for Retrospective Returns," *PCAS* LIII, 1966, pp. 185–187.

[9] Mahler, Howard C., Discussion of "Aggregate Retrospective Premium Ratio as a Function of the Aggregate Incurred Loss Ratio," *PCAS* LXXXI, 1994, pp. 75–90.

[10] Morrell, Roy K., Discussion of "A Method for Setting Retro Reserves," *PCAS* LXVIII, 1981, pp. 107–110.

[11] Uhthoff, D. R., Discussion of "Reserving for Retrospective Returns," *PCAS* LIII, 1996, pp. 187–189.

TESTING THE ASSUMPTIONS OF AGE-TO-AGE FACTORS

GARY G. VENTER

Abstract

The use of age-to-age factors applied to cumulative losses has been shown to produce least-squares optimal reserve estimates when certain assumptions are met. Tests of these assumptions are introduced, most of which derive from regression diagnostic methods. Failures of various tests lead to specific alternative methods of loss development.

ACKNOWLEDGEMENT

I would like to thank the Committee on Review of Papers for comments leading to numerous expository improvements over previous drafts. Any obscurity that remains is of course my own.

INTRODUCTION

In his paper "Measuring the Variability of Chain Ladder Reserve Estimates" Thomas Mack presented the assumptions needed for least-squares optimality to be achieved by the typical age-to-age factor method of loss development (often called "chain ladder"). Mack also introduced several tests of those assumptions. His results are summarized below, and then other tests of the assumptions are introduced. Also addressed is what to do when the assumptions fail. Most of the assumptions, if they fail in a particular way, imply least-squares optimality for some alternative method.

The organization of the paper is to first show Mack's three assumptions and their result, then to introduce six testable implications of those assumptions, and finally to go through the testing of each implication in detail.

PRELIMINARIES

Losses for accident year w evaluated at the end of that year will be denoted as being as of age 0, and the first accident year in the triangle is year 0. The notation below will be used to specify the models. Losses could be either paid or incurred. Only development that fills out the triangle is considered. Loss development beyond the observed data is often significant but is not addressed here. Thus age ∞ will denote the oldest possible age in the data triangle.

Notation

$c(w,d)$:	cumulative loss from accident year w as of age d
$c(w,\infty)$:	total loss from accident year w when end of triangle reached
$q(w,d)$:	incremental loss for accident year w from $d-1$ to d
$f(d)$:	factor applied to $c(w,d)$ to estimate $q(w,d+1)$
$F(d)$:	factor applied to $c(w,d)$ to estimate $c(w,\infty)$

Assumptions

Mack showed that some specific assumptions on the process of loss generation are needed for the chain ladder method to be optimal. Thus if actuaries find themselves in disagreement with one or another of these assumptions, they should look for some other method of development that is more in harmony with their intuition about the loss generation process. Reserving methods more consistent with other loss generation processes will be discussed below. Mack's three original assumptions are slightly restated here to emphasize the task as one of predicting future incremental losses. Note that the losses $c(w,d)$ have an evaluation date of $w+d$.

1. $\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)c(w,d)$.

   In words, the expected value of the incremental losses to emerge in the next period is proportional to the total losses emerged to date, by accident year. Note that in Mack's definition of the chain ladder, $f(d)$ does not depend on w, so the factor for a given age is constant across accident years. Note also that this formula is a linear relationship with no constant term. As opposed to other models discussed below, the factor applies directly to the cumulative data, not to an estimated parameter, like ultimate losses. For instance, the Bornhuetter-Ferguson method assumes that the expected incremental losses are proportional to the ultimate for the accident year, not the emerged to date.

2. Unless $v = w$, $c(w,d)$ and $c(v,g)$ are independent for all v, w, d and g.

   This would be violated, for instance, if there were a strong diagonal, when all years' reserves were revised upwards. In this case, instead of just using the chain ladder method, most actuaries would recommend eliminating these diagonals or adjusting them. Some model-based methods for formally recognizing diagonal effects are discussed below.

3. $\mathrm{Var}[q(w,d+1) \mid \text{data to } w+d] = a[d,c(w,d)]$.

   That is, the variance of the next increment observation is a function of the age and the cumulative losses to date. Note that $a(\cdot,\cdot)$ can be any function but does not vary by accident year. An assumption on the variance of the next incremental losses is needed to find a least-squares optimal method of estimating the development factors. Different assumptions, e.g., different functions $a(\cdot,\cdot)$ will lead to optimality for different methods of estimating the factor f. The form of $a(\cdot,\cdot)$ can be tested by trying different forms, estimating the f's, and seeing if the variance formula holds. There will almost always be some function $a(\cdot,\cdot)$ that reasonably accords with the observations, so the issue with this assumption is not its validity but its implications for the estimation procedure.

Results (Mack)

In essence what Mack showed is that under the above assumptions the chain ladder method gives the minimum variance unbiased linear estimator of future emergence. This gives a good justification for using the chain ladder in that case, but the assumptions need to be tested. Mack assumed that $a[d,c(w,d)] = k(d)c(w,d)$, that is, he assumed that the variance is proportional to the previous cumulative loss, with possibly a different proportionality factor for each age. In this case, the minimum variance unbiased estimator of $c(w,\infty)$ from the triangle of data to date $w+d$ is $F(d)c(w,d)$, where the age-to-ultimate factor $F(d) = [1+f(d)][1+f(d+1)]\cdots$, and $f(d)$ is calculated as:

$$f(d) = \sum_w q(w,d+1) \bigg/ \sum_w c(w,d),$$

where the sum is over the w's mutually available in both columns (assuming accident years are on separate rows and ages are in separate columns). Actuaries often use a modified chain ladder that uses only the last n diagonals. This will be one of the alternative methods to test if Mack's assumptions fail. Using only part of the data when all the assumptions hold will reduce the accuracy of the estimation, however.

Extension

In general, the minimum variance unbiased $f(d)$ is found by minimizing

$$\sum_w [f(d)c(w,d) - q(w,d+1)]^2 k(d)/a[d,c(w,d)].$$

This is the usual weighted least-squares result, where the weights are inversely proportional to the variance of the quantity being estimated. Because only proportionality, not equality, to the variance is required, $k(d)$ can be any convenient function of d—usually chosen to simplify the minimization.

For example, suppose $a[d,c(w,d)] = k(d)c(w,d)^2$. Then the $f(d)$ produced by the weighted least-squares procedure is the average of the individual accident year d to $d+1$ ratios, $q(w,d+1)/c(w,d)$. For $a[d,c(w,d)] = k(d)$, each $f(d)$ regression above is then just standard unweighted least squares, so $f(d)$ is the regression coefficient $\sum_w c(w,d)q(w,d+1)/\sum_w c(w,d)^2$. (See Murphy [8].) In all these cases, $f(d)$ is fit by a weighted regression, and so regression diagnostics can be used to evaluate the estimation. In the tests below just standard least-squares will be used, but in application the variance assumption should be reviewed.

Discussion

Without going into Mack's derivation, the optimality of the chain ladder method is fairly intuitive from the assumptions. In particular, the first assumption is that the expected emergence in the next period is proportional to the losses emerged to date. If that were so, then a development factor applied to the emerged to date would seem highly appropriate. Testing this assumption will be critical to exploring the optimality of the chain ladder. For instance, if the emergence were found to be a constant plus a percent of emergence to date, then a different method would be indicated—namely, a factor plus constant development method. On the other hand, if the next incremental emergence were proportional to ultimate rather than to emerged to date, a Bornhuetter-Ferguson type approach would be more appropriate.

To test this assumption against its alternatives, the development method that leads from each alternative needs to be fit, and then a goodness-of-fit measure applied. This is similar to trying a lot of methods and seeing which one you like best, but it is different in two respects: (1) each method tested derives from an alternative assumption on the process of loss emergence; (2) there is a specific goodness-of-fit test applied. Thus the fitting is a test of the emergence patterns that the losses are subject to, and not just a test of estimation methods.

TESTABLE IMPLICATIONS OF ASSUMPTIONS

Verifying a hypothesis involves finding as many testable implications of that hypothesis as possible, and verifying that the tests are passed. In fact a hypothesis can never be fully verified, as there could always be some other test you haven't thought of. Thus the process of verification is sometimes conceived as being really a process of attempted falsification, with the current tentatively-accepted hypothesis being the strongest (i.e., most easily testable) one not yet falsified. (See Popper [9].) The assumptions (1)–(3) are not directly testable, but they have testable implications. Thus they can be falsified if any of the implications are found not to hold, which would mean that the optimality of the chain ladder method could not be shown for the data in question. Holding up under all of these tests would increase the actuary's confidence in the hypothesis, still recognizing that no hypothesis can ever be fully verified. Some of the testable implications are:

1. Significance of factor $f(d)$.
2. Superiority of factor assumption to alternative emergence patterns such as:
   (a) linear with constant: $E[q(w,d+1) \mid \text{data to } w+d] = f(d)c(w,d) + g(d)$;
   (b) factor times parameter: $E[q(w,d+1) \mid \text{data to } w+d] = f(d)h(w)$;
   (c) including calendar year effect: $E[q(w,d+1) \mid \text{data to } w+d] = f(d)h(w)g(w+d)$.

Note that in these examples the notation has changed slightly so that $f(d)$ is a factor used to estimate $q(w, d+1)$, but not necessarily applied to $c(w,d)$. These alternative emergence models can be tested by goodness of fit, controlling for number of parameters.

3. Linearity of model: look at residuals as a function of $c(w,d)$.
4. Stability of factor: look at residuals as a function of time.
5. No correlation among columns.
6. No particularly high or low diagonals.

The remainder of this paper consists of tests of these implications.

TESTING LOSS EMERGENCE—IMPLICATIONS 1 & 2

The first four of these implications are tests of assumption (1). Standard diagnostic tests for weighted least-squares regression can be used as measures.

Implication 1: Significance of Factors

Regression analysis produces estimates for the standard deviation of each parameter estimated. Usually the absolute value of a factor is required to be at least twice its standard deviation for the factor to be regarded as significantly different from zero. This is a test failed by many development triangles, which means that the chain ladder method is not optimal for those triangles.

The requirement that the factor be twice the standard deviation is not a strict statistical test, but more like a level of comfort. For the normal distribution this requirement provides that there is only a probability of about 4.5% of getting a factor of this absolute value or greater when the true factor is zero. Many analysts are comfortable with a factor with absolute value 1.65 times its standard deviation, which could happen about 10% of the time by chance alone. For heavier-tailed distributions, the same ratio of factor to standard deviation will usually be more likely to occur by chance. Thus, if a factor were to be considered not significant for the normal distribution, it would probably be even less significant for other distributions. This approach could be made into a formal statistical test by finding the distribution that the factors follow. The normal distribution is often satisfactory, but it is not unusual to see some degree of positive skewness, which would suggest the lognormal. Some of the alternative models discussed below are easier to estimate in log form, so that is not an unhappy finding.

It may be tempting to do the regression of cumulative on previous cumulative and test the significance of that factor in order to justify the use of the chain ladder. However it is only the incrementals that are being predicted, so this would have to be carefully interpreted. In a cumulative-to-cumulative regression, the significance of the difference of the factor from unity is what needs to be tested. This can be done by comparing that difference to the standard deviation of the factor, which is equivalent to testing the significance of the factor in the incremental-to-cumulative regression. Some alternative methods to try when this assumption fails are discussed below.

Implication 2: Superiority to Alternative Emergence Patterns

If alternative emergence patterns give a better explanation of the data triangle observed to date, then assumption (1) of the chain ladder model is also suspect. In these cases development based on the best-fitting emergence pattern would be a natural option to consider. The sum of the squared errors (SSE) would be a way to compare models (the lower the better) but this should be adjusted to take into account the number of parameters used. Unfortunately it appears that there is no generally accepted method to make this adjustment. One possible adjustment is to compare fits by using the SSE divided by $(n-p)^2$, where n is the number of observations and p is the number of parameters. More parameters give an advantage in fitting but a disadvantage in prediction, so such a penalty in adjusting the residuals may be appropriate. A more popular adjustment in recent years is to base goodness of fit on the Akaike Information Criterion, or AIC (see Lütkepohl [5]). For a fixed set of observations, multiplying the SSE by $e^{2p/n}$ can approximate the effect of the AIC. The AIC has been criticized as being too permissive of over-parameterization for large data sets, and the Bayesian Information Criterion, or BIC, has been suggested as an alternative. Multiplying the SSE by $n^{p/n}$ would rank models the same as the BIC. As a comparison, if you have 45 observations, the improvement in SSE needed to justify adding a 5th parameter to a 4 parameter model is about 5%, $4\frac{1}{2}$%, and almost 9%, respectively, for these three adjustments. In the model testing below the sum of squared residuals divided by $(n-p)^2$ will be the test statistic, but in general the AIC and BIC should be regarded as good alternatives.

Note again that this is not just a test of development methods but is also a test to see what hypothetical loss generation process is most consistent with the data in the triangle.

The chain ladder has one parameter for each age, which is less than for the other emergence patterns listed in implication 2. This gives it an initial advantage, but if the other parameters improve the fit enough, they overcome this advantage. In testing the various patterns below, parameters will be fit by minimizing the sum of squared residuals. In some cases this will require an iterative procedure.

Alternative Emergence Pattern 1: Linear with Constant

The first alternative mentioned is just to add a constant term to the model. This is often significant in the age 0 to age 1 stage, especially for highly variable and slowly reporting lines, such as excess reinsurance. In fact, in the experience of myself and other actuaries who have reported informally, the constant term has often been found to be more statistically significant than the factor itself. If the constant is significant and the factor is not, a different development process is indicated. For instance in some triangles earning of additional exposure could influence the 0-to-1 development. It is important in such cases to normalize the triangle as much as possible, e.g., by adjusting for differences among accident years in exposure and cost levels (trend). With these adjustments a purely additive rather than a purely multiplicative method could be more appropriate.

Again, the emergence assumption underlying the linear with constant method is:

$$\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)c(w,d) + g(d).$$

If the constant is statistically significant, this emergence pattern is more strongly supported than that underlying the chain ladder.

Alternative Emergence Pattern 2: Factor Times Parameter

The chain ladder model expresses the next period's loss emergence as a factor times losses emerged so far. An important alternative, suggested by Bornhuetter and Ferguson (BF) in 1972, is to forecast the future emergence as a factor times estimated ultimate losses. While BF use some external measure of ultimate losses in this process, others have tried to use the data triangle itself to estimate the ultimate (e.g., see Verrall [13]). In this paper, models that estimate emerging losses as a percent of ultimate will be called parameterized BF models, even if they differ from the original BF method in how they estimate the ultimate losses.

The emergence pattern assumed by the parameterized BF model is:

$$\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)h(w).$$

That is, the next period expected emerged loss is a lag factor $f(d)$ times an accident year parameter $h(w)$. The latter could be interpreted as expected ultimate for the year, or at least proportional to that. This model thus has a parameter for each accident year as well as for each age (one less actually, as you can assume the $f(d)$'s sum to one—which makes $h(w)$ an estimate of ultimate losses; thus multiplying all the $f(d)$'s, $d > 0$, by a constant and dividing all the h's by the same constant will not change the forecasts). For reserving purposes there is even one fewer parameter, as the age 0 losses are already in the data triangle, so $f(0)$ is not needed. Thus, for a complete triangle with n accident years the BF has $2n-2$ parameters, or twice the number as the chain ladder. This will result in a penalty to goodness of fit, so the BF has to produce much lower fit errors than the chain ladder to give a better test statistic.

Testing the parameterized BF emergence pattern against that of the chain ladder cannot be done just by looking at the statistical significance of the parameters, as it could with the linear plus constant method, as one is not a special case of the other. This testing is the role of the test statistic, the sum of squared residuals divided by the square of the degrees of freedom. If this statistic is better for the BF model, that is evidence that the emergence pattern of the BF is more applicable to the triangle being studied. That would suggest that loss emergence for that book can be more accurately represented as fluctuating around a proportion of ultimate losses rather than a percentage of previously emerged losses.

Stanard [10] assumed a loss generation scheme that resulted in the expected loss emergence for each period being proportional to the ultimate losses for the period. This now can be seen to be the BF emergence pattern. Then by generating actual loss emergence stochastically, he tested some loss development methods. The chain ladder method gave substantially larger estimation errors for ultimate losses than his other methods, which were basically different versions of BF estimation. This illustrates how far off reserves can be when one reserving technique is applied to losses that have an emergence process different from the one underlying the technique.

A simulation in accord with the chain ladder emergence assumption would generate losses at age j by multiplying the simulated emerged losses at age $j-1$ by a factor and then adding a random component. In this manner the random components influence the expected emergence at all future ages. This may seem an unlikely way for losses to emerge, but it is for the triangles that follow this emergence pattern that the chain ladder will be optimal. The fact that Stanard used the simulation method consistent with the BF emergence pattern, and this was not challenged by the reviewer, John Robertson, suggests that actuaries may be more comfortable with the BF emergence assumptions than with those of the chain ladder. Or perhaps it just means that no one would be likely to think of simulating losses by the chain ladder method.

An important special case of the parameterized BF was developed by some Swiss and American reinsurance actuaries at a meeting in Cape Cod, and is sometimes called the Cape Cod method (CC). It is given by setting $h(w)$ to just a single h for all accident years. CC seems to have one more parameter than the chain ladder, namely h. However, any change in h can be offset by inverse changes in all the f's. CC thus has the same number of parameters as the chain ladder, and so its fit measure is not as heavily penalized as that of BF. However a single h requires a relatively stable level of loss exposure across accident years. Again it would be necessary to adjust for known exposure and price level differences among accident years, if using this method. The chain ladder and BF can handle changes in level from year to year as long as the development pattern remains consistent.

The BF model often has too many parameters. The last few accident years especially are left to find their own levels based on sparse information. Reducing the number of parameters, and thus using more of the information in the triangle, can often yield better predictions, especially in predicting the last few years. It could be that losses follow the BF emergence pattern, but this is disguised in the test statistic due to too many parameters. Thus, testing for the alternate emergence pattern should also include testing reduced parameter BF models.

The full BF not only assumes that losses emerge as a percentage of ultimate, but also that the accident years are all at different mean levels and that each age has a different percentage of ultimate losses. It could be, however, that several years in a row, or all of them, have the same mean level. If the mean changes, there could be a gradual transition from one level to another over a few years. This could be modeled as a linear progression of accident year parameters, rather than separate parameters for each year. A similar process could govern loss emergence. For instance, the 9th through 15th periods could all have the same expected percentage development. Finding these relationships and incorporating them in the fitting process will help determine what emergence process is generating the development.

The CC model can be considered a reduced parameter BF model. The CC has a single ultimate value for all accident years, while the BF has a separate value for each year. There are numerous other ways to reduce the number of parameters in BF models. Simply using a trend line through the BF ultimate loss parameters would use just two accident year parameters in total instead of one for each year. Another method might be to group years using apparent jumps in loss levels and fit an h parameter separately to each group. Within such groupings it is also possible to let each accident year's h parameter vary somewhat from the group average, e.g., via credibility, or to let it evolve over time, e.g., by exponential smoothing.

Alternative Emergence Patterns Example

Table 1 shows incremental incurred losses by age for some excess casualty reinsurance. As an initial test, the statistical sig-

TABLE 1

INCREMENTAL INCURRED LOSSES

	Age									
Year	0	1	2	3	4	5	6	7	8	9
0	5,012	3,257	2,638	898	1,734	2,642	1,828	599	54	172
1	106	4,179	1,111	5,270	3,116	1,817	−103	673	535	
2	3,410	5,582	4,881	2,268	2,594	3,479	649	603		
3	5,655	5,900	4,211	5,500	2,159	2,658	984			
4	1,092	8,473	6,271	6,333	3,786	225				
5	1,513	4,932	5,257	1,233	2,917					
6	557	3,463	6,926	1,368						
7	1,351	5,596	6,165							
8	3,133	2,262								
9	2,063									

TABLE 2

STATISTICAL SIGNIFICANCE OF FACTORS

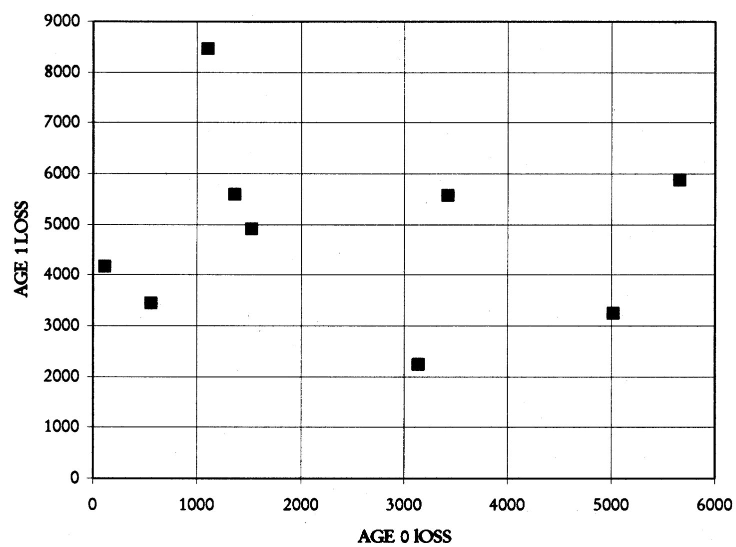
	0 to 1	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8
a	5,113	4,311	1,687	2,061	4,064	620	777	3,724
Std. Dev. *a*	1,066	2,440	3,543	1,165	2,242	2,301	145	0.000
b	−0.109	0.049	0.131	0.041	−0.100	0.011	−0.008	−0.197
Std. Dev. *b*	0.349	0.309	0.283	0.071	0.114	0.112	0.008	0.000

nificance of the factors was tested by regression of incremental losses against the previous cumulative losses. In the regression the constant is denoted by a and the factor by b. This provides a test of implication 1—significance of the factor, and also one test of implication 2—alternative emergence patterns. In this case the alternative emergence patterns tested are factor plus constant and constant with no factor. Here they are being tested by looking at whether or not the factors and the constants are significantly different from zero, rather than by any goodness-of-fit measure.

Table 2 shows the estimated parameters and their standard deviations. As can be seen, the constants are usually statistically

FIGURE 1

AGE 1 VS. AGE 0 LOSSES



significant (parameter nearly double its standard deviation, or more), but the factors never are. The chain ladder assumes the incremental losses are proportional to the previous cumulative, which implies that the factor is significant and the constant is not. The lack of significance of the factors and the significance of many of the constants both suggest that the losses to emerge at any age $d+1$ are not proportional to the cumulative losses through age d. The assumptions underlying the chain ladder model are thus not supported by this data. A constant amount emerging for each age usually appears to be a reasonable estimator, however.

Figure 1 illustrates this. A factor by itself would be a straight line through the origin with slope equal to the development factor, whereas a constant would give a horizontal line at the height of the constant. As an alternative, the parameterized BF model was fit to the triangle. As this is a non-linear model, fitting is a little more involved. A statistical package that includes non-linear regression could ease the estimation. A method of fitting the parameters without such a package will be discussed, followed by an analysis of the resulting fit.

To do the fitting, an iterative method can be used to minimize the sum of the squared residuals, where the (w,d) residual is $[q(w,d) - f(d)h(w)]$. Weighted least squares could also be used if the variances of the residuals are not constant over the triangle. For instance, the variances could be proportional to $f(d)^p h(w)^q$ for some values of p and q, usually 0, 1, or 2, in which case the regression weights would be $1/f(d)^p h(w)^q$.

A starting point for the f's or the h's is needed to begin the iteration. While almost any reasonable values could be used, such as all f's equal to $1/n$, convergence will be faster with values likely to be in the ballpark of the final factors. A natural starting point thus might be the implied $f(d)$'s from the chain ladder method. For ages greater than 0, these are the incremental age-to-age factors divided by the cumulative-to-ultimate factors. To get a starting value for age 0, subtract the sum of the other factors from unity. Starting with these values for $f(d)$, regressions were performed to find the $h(w)$'s that minimize the sum of squared residuals (one regression for each w). These give the best h's for that initial set of f's. The standard linear regression formula for these h's simplifies to:

$$h(w) = \sum_d f(d)q(w,d) \Big/ \sum_d f(d)^2.$$

Even though that gives the best h's for those f's, another regression is needed to find the best f's for those h's. For this step the usual regression formula gives:

$$f(d) = \sum_w h(w)q(w,d) \Big/ \sum_w h(w)^2.$$

TABLE 3

BF PARAMETERS

Age d	0	1	2	3	4	5	6	7	8	9
$f(d)$ 1st	0.106	0.231	0.209	0.155	0.117	0.083	0.038	0.032	0.018	0.011
$f(d)$ ult.	0.162	0.197	0.204	0.147	0.115	0.082	0.037	0.030	0.015	0.009
Year w	0	1	2	3	4	5	6	7	8	9
$h(w)$ 1st	17,401	15,729	23,942	26,365	30,390	19,813	18,592	24,154	14,639	12,733
$h(w)$ ult.	15,982	16,501	23,562	27,269	31,587	20,081	19,032	25,155	13,219	19,413

Now the h regression can be repeated with the new f's, etc. This process continues until convergence occurs, i.e., until the f's and h's no longer change with subsequent iterations. It may be possible that this procedure would converge to a local rather than the global minimum, which can be tested by using other starting values.

Ten iterations were used in this case, but substantial convergence occurred earlier. The first round of f's and h's and those at convergence are in Table 3. Note that the h's are not the final estimates of the ultimate losses, but are used with the estimated factors to estimate future emergence. In this case, in fact, $h(0)$ is less than the emerged to date. As the h's are unique only up to a constant of proportionality, which can be absorbed by the f's, it may improve presentations to set $h(0)$ to the estimated ultimate losses for year 0.

Standard regression assumes each observation q has the same variance, which is to say the variance is proportional to $f(d)^p h(w)^q$, with $p = q = 0$. If $p = q = 1$ the weighted regression formulas become:

$$h(w)^2 = \sum_d [q(w,d)^2/f(d)] \Big/ \sum_d f(d) \qquad \text{and}$$

$$f(d)^2 = \sum_w [q(w,d)^2/h(w)] \Big/ \sum_w h(w).$$

TABLE 4

DEVELOPMENT FACTORS

	Incremental								
Prior	0 to 1	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8	8 to 9
	1.22	0.57	0.26	0.16	0.10	0.04	0.03	0.02	0.01
	Ultimate								
	0 to 9	1 to 9	2 to 9	3 to 9	4 to 9	5 to 9	6 to 9	7 to 9	8 to 9
	6.17	2.78	1.77	1.41	1.21	1.10	1.06	1.03	1.01
	Incremental/Ultimate								
0.162	0.197	0.204	0.147	0.115	0.082	0.037	0.030	0.015	0.009

For comparison, the development factors from the chain ladder are shown in Table 4. The incremental factors are the ratios of incremental to previous cumulative. The ultimate ratios are cumulative to ultimate. Below them are the ratios of these ratios, which represent the portion of ultimate losses to emerge in each period. The zeroth period shown is unity less the sum of the other ratios. These factors were the initial iteration for the $f(d)$s shown above.

Having now estimated the BF parameters, how can they be used to test what the emergence pattern of the losses is?

A comparison of this fit to that from the chain ladder can be made by looking at how well each method predicts the incremental losses for each age after the initial one. The SSE adjusted for number of parameters will be used as the comparison measure, where the parameter adjustment will be made by dividing the SSE by the square of the difference between the number of observations and the number of parameters, as discussed earlier. Here there are 45 observations, as only the predicted points count as observations. The adjusted SSE was 81,169 for the BF, and 157,902 for the chain ladder. This shows that the emergence pattern for the BF (emergence proportional to ultimate) is much more consistent with this data than is the chain ladder emergence pattern (emergence proportional to previous emerged).

TABLE 5

FACTORS IN CC METHOD

Age d	0	1	2	3	4	5	6	7	8	9
$f(d)$	0.109	0.220	0.213	0.148	0.124	0.098	0.038	0.028	0.013	0.008

The CC method was also tried for this data. The iteration proceeded similarly to that for the BF, but only a single h parameter was fit for all accident years. Now:

$$h = \sum_{w,d} f(d)q(w,d) \Big/ \sum_{w,d} f(d)^2.$$

This formula for h is the same as the formula for $h(w)$ except the sum is taken over all w. The estimated h is 22,001, and the final factors f are shown in Table 5. The adjusted SSE for this fit is 75,409. Since the CC is a special case of the BF, the unadjusted SSE is necessarily worse than that of the BF method (in this case 59M vs. 98M), but with fewer parameters in the CC, the adjustment makes them similar. These are close enough that which is better depends on the adjustment chosen for extra parameters. The BIC also favors the CC, but the AIC is better for the BF. As is often the case, the statistics can inform decision-making but not determine the decision.

Intermediate special cases could be fit similarly. If, for instance, a single factor were sought to apply to just two accident years, the sum would be taken over those years to estimate that factor, etc.

This is a case where the BF has too many parameters for prediction purposes. More parameters fit the data better but use up information. The penalty in the fit measure adjusts for this problem, and the penalty used finds the CC to be a somewhat better model. Thus the data is consistent with random emergence around an expected value that is constant over the accident years.

TABLE 6

TERMS IN ADDITIVE CHAIN LADDER

Age d	1	2	3	4	5	6	7	8	9
$g(d)$	4,849.3	4,682.5	3,267.1	2,717.7	2,164.2	839.5	625.0	294.5	172.0

Again, the CC method would probably work even better for loss ratio triangles than for loss triangles, as then a single target ultimate value makes more sense. Adjusting loss ratios for trend and rate level could increase this homogeneity.

In addition, an additive development was tried, as suggested by the fact that the constant terms were significant in the original chain ladder, even though the factors were not. The development terms are shown in Table 6. These are just the average loss emerged at each age. The adjusted sum of squared residuals is 75,409. This is much better than the chain ladder, which might be expected, as the constant terms were significant in the original significance-test regressions while the factors were not. The additive factors in Table 6 differ from those in Table 2 because there is no multiplicative factor in Table 6.

Is it a coincidence that the additive chain ladder gives the same fit accuracy as the CC? Not really, in that they both estimate each age's loss levels with a single value. Let $g(d)$ denote the additive development amount for age d. As the notation suggests, this does not vary by accident year. The CC method fits an overall h and a factor $f(d)$ for each age such that the estimated emergence for age d is $f(d)h$. Here too the predicted development varies by age but is a constant for each accident year. If you have estimated the CC parameters you can just define $g(d) = f(d)h$. Alternatively, if the additive method has been fit, no matter what h is estimated, the f's can be defined as $f(d)h = g(d)$. As long as the parameters are fit by least-squares they have to come out the same: if one came out lower, you could have used the equations in the two previous sentences to get this same lower value for

TABLE 7

BF-CC PARAMETERS

Age d	0	1	2	3	4	5	6	7	8	9
$f(d)$	$*$	0.230	0.230	0.160	0.123	0.086	0.040	0.040	0.017	0.017
Year w	0	1	2	3	4	5	6	7	8	9
$h(w)$	14,829	14,829	20,962	25,895	30,828	20,000	20,000	20,000	20,000	20,000

the other. The two models have the same age and accident year relationships and so will always come out the same when fit by least-squares. They are defined differently, however, and so other methods of estimating the parameters may come up with separate estimates, as in Stanard [10]. In the remainder of this paper, the models will be used interchangeably.

Finally, an intermediate BF-CC pattern was fit as an example of the possible approaches of this type. In this case ages 1 and 2 are assumed to have the same factor, as are ages 6 and 7 and ages 8 and 9. This reduces the number of f parameters from 9 to 6. The number of accident year parameters was also reduced: years 0 and 1 have a single parameter, as do years 5 through 9. Year 2 has its own parameter, as does year 4, but year 3 is the average of those two. Thus there are 4 accident year parameters, and so 10 parameters in total. Any one of these can be set arbitrarily, with the remainder adjusted by a factor, so there are really just 9. The selections were based on consideration of which parameters were likely not to be significantly different from each other.

The estimated factors are shown in Table 7. The factor to be set arbitrarily was the accident year factor for the last 5 years, which was set to 20,000. The other factors were estimated by the same iterative regression procedure as for the BF, but the factor constraints change the simplified regression formula. The adjusted sum of squared residuals is 52,360, which makes it the best approach tried. This further supports the idea that claims emerge as a percent of ultimate for this data. It also indicates that the various accident years and ages are not all at different levels. The actual and fitted values from this, the chain ladder, and CC are in Exhibit 1. The fitted values in Exhibit 1 were calculated as follows. For the chain ladder, the factors from Table 4 were applied to the cumulative losses implied from Table 1. For the CC the fitted values are just the terms in Table 6. For the BF-CC they are the products of the appropriate f and h factors from Table 7. The parameters for all the models to this point are summarized in Exhibit 2.

Alternative Emergence Patterns-Summary

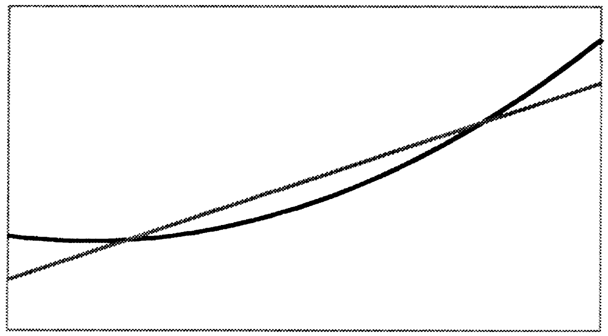
The chain ladder assumes that future emergence for an accident year will be proportional to losses emerged to date. The BF methods take expected emergence in each period to be a percentage of ultimate losses. This could be interpreted as regarding the emerged to date to have a random component that will not influence future development. If this is the actual emergence pattern, the chain ladder method will apply factors to the random component, and thus increase the estimation error.

The CC and additive chain ladder methods assume in effect that years showing low losses or high losses to date will have the same expected future dollar development. Thus a bad loss year may differ from a good one in just a couple of emergence periods, and have quite comparable loss emergence in all other periods. The chain ladder and the most general form of the BF, on the other hand, assume that a bad year will have higher emergence than a good year in most periods.

The BF and chain ladder emergence patterns are not the only ones that make sense. Some others will be reviewed when discussing diagonal effects below.

Which emergence pattern holds for a given triangle is an empirical issue. Fitting parameters to the various methods and looking at the significance of the parameters and the adjusted sum of squared residuals can test this.

FIGURE 2



RESIDUAL ANALYSIS—TESTING IMPLICATIONS 3 & 4

So far the first two of the six testable implications of the chain ladder assumptions have been addressed. Looking at the residuals from the fitting process can test the next two implications.

Implication 3: Test of Linearity—Residuals as Function of Previous

Figure 2 shows a straight line fit to a curve. The residuals can be seen to be first positive, then negative then all positive. This pattern of residuals is indicative of a non-linear process with a linear fit. The chain ladder model assumes the incremental losses at each age are a linear function of the previous cumulative losses.

A scatter plot of the incremental against the previous cumulative, as in Figure 3, can be used to check linearity; looking for this characteristic non-linear pattern (i.e., strings of positive and negative residuals) in the residuals plotted against the previous cumulative is equivalent. This can be tested for each age to see if a non-linear process may be indicated. Finding this would suggest that emergence is a non-linear function of losses to date. In

FIGURE 3

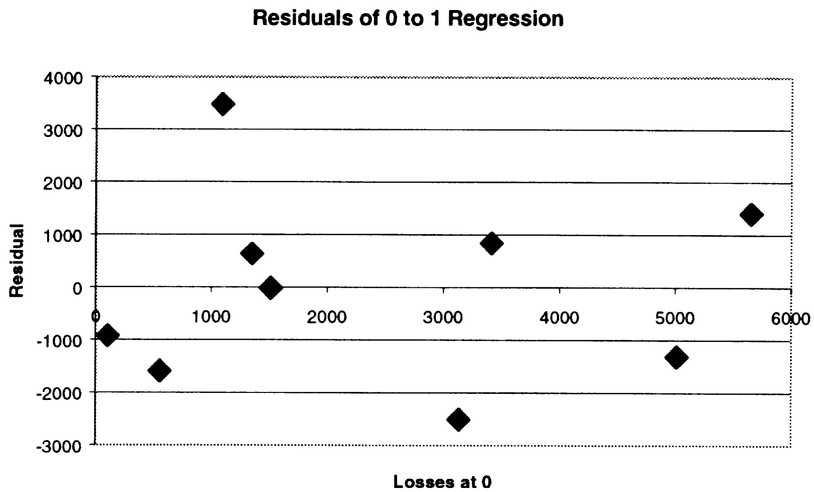


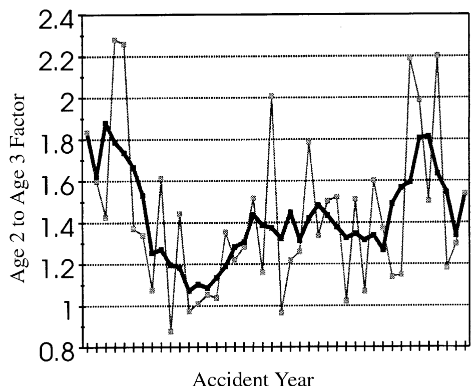
Figure 3 there are no apparent strings of consecutive positive or negative residuals, so non-linearity is not indicated.

Implication 4: Test of Stability—Residuals Over Time

If a similar pattern of sequences of high and low residuals is found when plotted against time, instability of the factors may be indicated. If the factors appear to be stable over time, all the accident years available should be used to calculate the development factors, in order to reduce the effects of random fluctuations. When the development process is unstable, the assumptions for optimality of the chain ladder are no longer satisfied. A response to unstable factors over time might be to use a weighted average of the available factors, with more weight going to the more recent years, e.g., just use the last 5 diagonals. A weighted average should be used when there is a good reason for it, e.g., when residual analysis shows that the factors are changing, but otherwise it will increase estimation errors by over-emphasizing some observations and under-emphasizing others.

FIGURE 4

2ND TO 3RD FIVE-TERM MOVING AVERAGE



Another approach to unstable development would be to adjust the triangle for measurable instability. For instance, Berquist and Sherman [1] suggest testing for instability by looking for changes in the settlement rate of claims. They measured this by looking at the changes in the percentage of claims closed by age. If instability is found, the triangle is adjusted to the latest pattern. The adjusted triangle, however, should still be tested for stability of development factors by residual analysis and as illustrated below.

Figure 4 shows the 2nd to 3rd factor by accident year from a large development triangle (data in Exhibit 3) along with its five-term moving average. The moving average is the more stable of the two lines, and is sometimes in practice called "the average of the last five diagonals." There is apparent movement of the factor over time as well as a good deal of random fluctuation. There is a period of time in which the moving average is as low as 1.1 and other times it is as high as 1.8. This is the kind of variability that would suggest using the average of recent diagonals instead of the entire triangle when estimating factors. This is not suggested due to the large fluctuations in factors, but rather because of the changes over time in the level around which the factors are fluctuating. A lot of variability around a fixed level would in fact suggest using all the data.

It is not clear from the data what is causing the moving average factors to drift over time. Faced with data like this, the average of all the data would not normally be used. Grouping accident years or taking weighted averages would be useful alternatives.

The state-space model in the Verall and Zehnwirth references provides a formal statistical treatment of the types of instability in a data triangle. This model can be used to help analyze whether to use all the data, or to adopt some form of weighted average that de-emphasizes older data. It is based on comparing the degree of instability of observations around the current mean to the degree of instability in the mean itself over time. While this is the main statistical model available to determine weights to apply to the various accident years of data, a detailed discussion is beyond the scope of this paper.

INDEPENDENCE—TESTING IMPLICATIONS 5 & 6

Implications 5 and 6 have to do with independence within the triangle. Mack's second assumption above is that, except for observations in the same accident year, the columns of incremental losses need to be independent. He developed a correlation test and a high-low diagonal test to check for dependencies. The data may have already been adjusted for known changes in the case reserving process. For instance, Berquist and Sherman recommend looking at the difference between paid and incurred case severity trends to determine if there has been a change in case reserve adequacy, and if there has, adjusting the data accordingly. Even after such adjustments, however, correlations may exist within the triangle.

TABLE 8

SAMPLE CORRELATION = $-1.35/(146.37 \times 0.20)^{1/2} = -.25$

Year	X = 0 to 1	Y = 1 to 2	$(X - \mathrm{E}[X])^2$	$(Y - \mathrm{E}[Y])^2$	$(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])$
1	0.65	0.32	54.27	0.14	2.78
2	39.42	0.26	986.46	0.19	−13.71
3	1.64	0.54	40.70	0.02	0.98
4	1.04	0.36	48.63	0.11	2.31
5	7.76	0.66	0.07	0.00	0.01
6	3.26	0.82	22.63	0.01	−0.57
7	6.22	1.72	3.24	1.05	−1.85
8	4.14	0.89	15.01	0.04	−0.74
Average	8.02	0.70	146.37	0.20	−1.35

Implication 5: Correlation of Development Factors

Mack developed a correlation test for adjacent columns of a development factor triangle. If a year of high emergence tends to follow one with low emergence, then the development method should take this into account. Another correlation test would be to calculate the sample correlation coefficients for all pairs of columns in the triangle, and then see how many of these are statistically significant, say at the 10% level. The sample correlation for two columns is just the sample covariance divided by the product of the sample standard deviations for the first n elements of both columns, where n is the length of the shorter column. The sample correlation calculation in Table 8 shows that for the triangle in Table 1 above, the correlation of the first two development factors is -25%.

Letting r denote the sample correlation coefficient, define $T = r[(n-2)/(1-r^2)]^{1/2}$. A significance test for the correlation coefficient can be made by considering T to be t-distributed with $n-2$ degrees of freedom. If T is greater than the t-statistic for 0.9 at $n-2$ degrees of freedom, for instance, then r can be considered significant at the 10% level. (See Miller and Wichern [7, p. 214].)

In this example, $T = -0.63$, which is not significant even at the 10% level. This level of significance means that 10% of the pairs of columns could show up as significant just by random happenstance. A single correlation at this level would thus not be a strong indicator of correlation within the triangle. If several columns are correlated at the 10% level, however, there may be a correlation problem.

To test this further, if m is the number of pairs of columns in the triangle, the number that display significant correlation could be considered a binomial variate in m and 0.1, which has standard deviation $0.3m^{1/2}$. Thus more than $0.1m + m^{1/2}$ significant correlations (mean plus 3.33 standard deviations) would strongly suggest there is actual correlation within the triangle. Here the 10% level and 3.33 standard deviations were chosen for illustration. A single correlation that is significant at the 0.1% level would also be indicative of a correlation problem, for example.

If there is such correlation, the product of development factors is not unbiased, but the relationship $\mathrm{E}[XY] = (\mathrm{E}[X])(\mathrm{E}[Y]) + \mathrm{Cov}(X,Y)$ could be used to correct the product, where here X and Y are development factors.

Implication 6: Significantly High or Low Diagonals

Mack's high-low diagonal test counts the number of high and low factors on each diagonal, and tests whether or not that is likely to be due to chance. Here another high-low test is proposed: use regression to see if any diagonal dummy variables are significant. This test also provides alternatives in case the pure chain ladder is rejected. An actuary will often have information about changes in company operations that may have created a diagonal effect. If so, this information could lead to choices of modeling methods—e.g., whether to assume the effect is permanent or temporary. The diagonal dummies can be used to measure the effect in any case, but knowledge of company operations will help determine how to use this effect. This is particularly so if the effect occurs in the last few diagonals.

A diagonal in the loss development triangle is defined by $w + d$ = constant. Suppose for some given data triangle, the diagonal $w + d = 7$ has been estimated to be 10% higher than normal. Then an adjusted BF estimate of a cell might be:

$$q(w,d) = 1.1f(d)h(w) \qquad \text{if} \quad w + d = 7, \qquad \text{and}$$

$$q(w,d) = f(d)h(w) \qquad \text{otherwise.}$$

This is an example of a multiplicative diagonal effect. Additive diagonal effects can also be estimated, using regression with diagonal dummies.

	Age			
Year	0	1	2	3
1	2	5	4	
3	8	9		
7	10			
7				

Incr. Ages 1–3	Cum. Age 0	Cum. Age 1	Cum. Age 2	Dummy 1	Dummy 2
2	1	0	0	0	0
8	3	0	0	1	0
10	7	0	0	0	1
5	0	3	0	1	0
9	0	11	0	0	1
4	0	0	8	0	1

The small sample triangle of incremental losses here will be used as an example of how to set up diagonal dummies in a chain ladder model. The goal is to get a matrix of data in the form needed to do a multiple regression. First the triangle (except the first column) is strung out into a column vector. This is the dependent variable, and forms the first column of the matrix above. Then columns for the independent variables are added. The second column is the cumulative losses at age 0 corresponding to the loss entries that are at age 1, and zero for the other loss entries. The regression coefficient for this column would be the 0 to 1 cumulative-to-incremental factor. The next two columns are cumulative losses at age 1 and age 2 corresponding to the age 2 and age 3 data in the first column. The last two columns are the diagonal dummies. They pick out the elements of the last two diagonals. The coefficients for these columns would be additive adjustments for those diagonals, if significant.

This method of testing for diagonal effects is applicable to many of the emergence models. In fact, if diagonal effects are found to be significant in chain ladder models, they probably are needed in the BF models of the same data. Thus tests of the chain ladder vs. BF should be done with the diagonal elements included. Some examples are given in the Appendix. Another popular modeling approach is to consider diagonal effects to be a measure of inflation (e.g., see Taylor [11]). In a payment triangle this would be a natural interpretation, but a similar phenomenon could occur in an incurred triangle. In this case the latest diagonal effects might be projected ahead as estimates of future inflation. An understanding of the aspects of company operations that drive the diagonal effects would help address these issues.

This approach incorporates diagonal effects right into the emergence model. For instance, an emergence model might be:

$$\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)g(w+d).$$

Here $g(w+d)$ is a diagonal effect, but every diagonal has such a factor. The usual interpretation is that g measures the cumulative claims inflation applicable to that diagonal since the first accident year. It would even be possible to add accident year effects $h(w)$ as well, e.g.,

$$\mathrm{E}[q(w,d+1) \mid \text{data to } w+d] = f(d)h(w)g(w+d).$$

There are clearly too many parameters here, but a lot of them might reasonably be set equal. For instance, the inflation might be the same for several years, or several accident years might be at the same level. Note that since g is cumulative inflation, a constant inflation level could be achieved by setting $g(w+d) = (1+j)^{w+d}$. Then j is the only inflation parameter to be estimated.

The age and accident year parameters might also be able to be written as trends rather than individual factors. If $f(d) = (1+i)^d$ and $h(w) = h \times (1+k)^w$, then the model reduces to four parameters h, i, j, and k. However it would be more usual to need more parameters than this, possibly written as changing trends. That is, i, j, and k might be constant for some periods, then change for others. Note that if they are constant for all periods, the estimator $h(1+i)^d(1+j)^{w+d}(1+k)^w$ is $h(1+i+j+ij)^d(1+k+j+jk)^w$, which eliminates the parameter j, as i becomes $i+j+ij$ and k becomes $k+j+jk$.

It might be better to start without the accident year trend and keep the calendar year trend, especially if the triangle has been normalized for accident year changes. The model for the (w,d) cell would then be $h(1+i)^d(i+j)^{w+d}$, which has just three parameters.

As with the BF model, the parameters of models with diagonal trends can be estimated iteratively. With reasonable starting values, fix two of the three sets of parameters, and fit the third by least squares, and rotate until convergence is reached. Alternatively, a non-linear search procedure could be utilized. As an example of the simplest of these approaches, modeling $\mathrm{E}[q(w,d+1) \mid \text{data to } w+d]$ as just $6{,}756(0.7785)^d$ gives an adjusted sum of squares of 57,527 for the reinsurance triangle above. This is not the best fitting model, but it is better than some and has only two parameters $h = 6{,}756$ and $i = -0.2215$.

Calendar year trend accounts for inflation in the time between loss occurrence and loss settlement, which many actuaries believe has an impact on ultimate losses. Whether it is influencing a given loss triangle can be investigated by testing for diagonal effects.

CONCLUSION

The first test that will quickly indicate the general type of emergence pattern faced is the test of significance of the cumulative-to-incremental factors at each age. This is equivalent to testing if the cumulative-to-cumulative factors are significantly different from unity. When this test fails, the future emergence is not proportional to past emergence. It may be a constant amount, or it may be proportional to ultimate losses, as in the BF pattern.

When this test is passed, the addition of an additive component may give an even better fit. Even when the test is failed, including an additive term may make the factor significant. In either case the BF emergence pattern may still produce a better fit. Reduced parameter BF models could also give better performance, as they will be less responsive to random variation. If an additive component is significant, then converting the triangle to on-level loss ratios may improve the forecasts.

Tests of stability and for diagonal effects may lead to further improvements in the model. However, if the emergence is stable, excluding data by using only the last *n* diagonals will lead to higher estimation errors on average.

An actuary might advise: "If the chain ladder doesn't work, try Bornhuetter-Ferguson." This is a reasonable conclusion, with the interpretation of "doesn't work" to mean "fails the assumptions of least-squares optimality," and "try" to mean "test the underlying assumptions of."

REFERENCES

[1] Berquist, James R., and Richard E. Sherman, "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," *PCAS* LXIV, 1977, pp. 123–184.

[2] Bornhuetter, Ronald L., and R. E. Ferguson, "The Actuary and IBNR," *PCAS* LIX, 1972, pp. 181–195.

[3] Gerber and Jones, "Credibility Formulas with Geometric Weights," *Society of Actuaries Transactions*, 1975.

[4] de Jong, P., and B. Zehnwirth, "Claims Reserving, State-space Models and the Kalman Filter," *Journal of the Institute of Actuaries*, 1983, pp. 157–181.

[5] Lütkepohl, *Introduction to Multiple Time Series Analysis*, Springer-Verlag, 1993.

[6] Mack, Thomas, "Measuring the Variability of Chain Ladder Reserve Estimates," Casualty Actuarial Society *Forum* 1, Spring 1994, pp. 101–182.

[7] Miller, Robert, and Dean Wichern, *Intermediate Business Statistics*, Holt, Rinehart and Winston, 1977.

[8] Murphy, Daniel M., "Unbiased Loss Development Factors," *PCAS* LXXXI, 1994, pp. 154–222.

[9] Popper, Karl R., *Conjectures and Refutations*, Routledge, 1969.

[10] Stanard, James N., "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques," *PCAS* LXXII, 1985, pp. 124–148.

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

[12] Verrall, Richard, "A State Space Representation of the Chain Ladder Linear Model," *Journal of the Institute of Actuaries* 116, Part III, December 1989, pp. 589–610.

[13] Zehnwirth, Ben, "A Linear Filtering Theory Approach to Recursive Credibility Estimation," *ASTIN Bulletin* 15, April 1985, pp. 19–36.

[14] Zehnwirth, Ben, "Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals, and Risk Based Capital," Casualty Actuarial Society *Forum* 2, Spring 1994, pp. 447–606.

EXHIBIT 1

COMPARATIVE FITS

Chain Ladder									
	1	2	3	4	5	6	7	8	9
Actual	3,257	2,638	898	1,734	2,642	1,828	599	54	172
Fit	6,101	4,705	2,846	1,912	1,350	656	580	296	172
% Error	87%	78%	217%	10%	−49%	−64%	−3%	448%	0%
Actual	4,179	1,111	5,270	3,116	1,817	−103	673	535	
Fit	129	2,438	1,408	1,728	1,374	632	499	257	
% Error	−97%	119%	−73%	−45%	−24%	−714%	−26%	−52%	
Actual	5,582	4,881	2,268	2,594	3,479	649	603		
Fit	4,151	5,116	3,619	2,614	1,868	900	736		
% Error	−26%	5%	60%	1%	−46%	39%	22%		
Actual	5,900	4,211	5,500	2,159	2,658	984			
Fit	6,883	6,574	4,113	3,444	2,336	1,057			
% Error	17%	56%	−25%	60%	−12%	7%			
Actual	8,473	6,271	6,333	3,786	225				
Fit	1,329	5,442	4,131	3,591	2,588				
% Error	−84%	−13%	−35%	−5%	1,050%				
Actual	4,932	5,257	1,233	2,917					
Fit	1,842	3,667	3,053	2,095					
% Error	−63%	−30%	148%	−28%					
Actual	3,463	6,926	1,368						
Fit	678	2,287	2,856						
% Error	−80%	−67%	109%						
Actual	5,596	6,165							
Fit	1,644	3,953							
% Error	−71%	−36%							
Actual	2,262								
Fit	3,814								
% Error	69%								
CC									
	1	2	3	4	5	6	7	8	9
Actual	3,257	2,638	898	1,734	2,642	1,828	599	54	172
Fit	4,364	3,746	2,287	1,631	1,082	336	188	59	17
% Error	34%	42%	155%	−6%	−59%	−82%	−69%	9%	−90%
Actual	4,179	1,111	5,270	3,116	1,817	−103	673	535	
Fit	4,364	3,746	2,287	1,631	1,082	336	188	59	
% Error	4%	237%	−57%	−48%	−40%	−426%	−72%	−89%	
Actual	5,582	4,881	2,268	2,594	3,479	649	603		
Fit	4,364	3,746	2,287	1,631	1,082	336	188		
% Error	−22%	−23%	1%	−37%	−69%	−48%	−69%		
Actual	5,900	4,211	5,500	2,159	2,658	984			
Fit	4,364	3,746	2,287	1,631	1,082	336			
% Error	−26%	−11%	−58%	−24%	−59%	−66%			
Actual	8,473	6,271	6,333	3,786	225				

EXHIBIT 1

(CONTINUED)

Fit	4,364	3,746	2,287	1,631	1,082				
% Error	−48%	−40%	−64%	−57%	381%				
Actual	4,932	5,257	1,233	2,917					
Fit	4,364	3,746	2,287	1,631					
% Error	−12%	−29%	85%	−44%					
Actual	3,463	6,926	1,368						
Fit	4,364	3,746	2,287						
% Error	26%	−46%	67%						
Actual	5,596	6,165							
Fit	4,364	3,746							
% Error	−22%	−39%							
Actual	2,262								
Fit	4,364								
% Error	93%								
BF-CC									
	1	2	3	4	5	6	7	8	9
Actual	3,257	2,638	898	1,734	2,642	1,828	599	54	172
Fit	3,411	3,411	2,373	1,824	1,275	593	593	252	252
% Error	5%	29%	164%	5%	−52%	−68%	−1%	367%	47%
Actual	4,179	1,111	5,270	3,116	1,817	−103	673	535	
Fit	3,411	3,411	2,373	1,824	1,275	593	593	252	
% Error	−18%	207%	−55%	−41%	−30%	−676%	−12%	−53%	
Actual	5,582	4,881	2,268	2,594	3,479	649	603		
Fit	4,821	4,821	3,354	2,578	1,803	838	838		
% Error	−14%	−1%	48%	−1%	−48%	29%	39%		
Actual	5,900	4,211	5,500	2,159	2,658	984			
Fit	5,956	5,956	4,143	3,185	2,227	1,036			
% Error	1%	41%	−25%	48%	−16%	5%			
Actual	8,473	6,271	6,333	3,786	225				
Fit	7,090	7,090	4,932	3,792	2,651				
% Error	−16%	13%	−22%	0%	1,078%				
Actual	4,932	5,257	1,233	2,917					
Fit	4,600	4,600	3,200	2,460					
% Error	−7%	−12%	160%	−16%					
Actual	3,463	6,926	1,368						
Fit	4,600	4,600	3,200						
% Error	33%	−34%	134%						
Actual	5,596	6,165							
Fit	4,600	4,600							
% Error	−18%	−25%							
Actual	2,262								
Fit	4,600								
% Error	103%								

EXHIBIT 1

(CONTINUED)

Additive with Multiplicative Diagonals and Accident Years

	1	2	3	4	5	6	7	8	9
Actual	3,257	2,638	898	1,734	2,642	1,828	599	54	172
Fit	3,185	3,185	2,148	2,730	1,995	660	660	660	477
% Error	−2%	21%	139%	57%	−24%	−64%	10%	1,122%	177%
Actual	4,179	1,111	5,270	3,116	1,817	−103	673	535	
Fit	3,185	3,185	3,465	2,730	1,995	660	660	477	
% Error	−24%	187%	−34%	−12%	10%	−741%	−2%	−11%	
Actual	5,582	4,881	2,268	2,594	3,479	649	603		
Fit	4,036	6,508	4,390	3,460	2,529	836	604		
% Error	−28%	33%	94%	33%	−27%	29%	0%		
Actual	5,900	4,211	5,500	2,159	2,658	984			
Fit	6,508	6,508	4,390	3,460	2,529	604			
% Error	10%	55%	−20%	60%	−5%	−39%			
Actual	8,473	6,271	6,333	3,786	225				
Fit	5,136	5,136	3,465	2,730	1,442				
% Error	−39%	−18%	−45%	−28%	541%				
Actual	4,932	5,257	1,233	2,917					
Fit	5,136	5,136	3,465	1,972					
% Error	4%	−2%	181%	−32%					
Actual	3,463	6,926	1,368						
Fit	5,136	5,136	2,503						
% Error	48%	−26%	83%						
Actual	5,596	6,165							
Fit	5,136	3,710							
% Error	−8%	−40%							
Actual	2,262								
Fit	3,710								
% Error	64%								

EXHIBIT 2

SUMMARY OF PARAMETERS

	0	1	2	3	4	5	6	7	8	9
BF $f(d)$	0.162	0.197	0.204	0.147	0.115	0.082	0.037	0.030	0.015	0.009
BF $h(w)$	15,982	16,501	23,562	27,269	31,587	20,081	19,032	25,155	13,219	19,413
CC $f(d)$	0.109	0.220	0.213	0.148	0.124	0.098	0.038	0.028	0.013	0.008
Additive Chain	—	4,849.3	4,682.5	3,267.1	2,717.7	2,164.2	839.5	625.0	294.5	172.0
BF-CC $f(d)$	—	0.230	0.230	0.160	0.123	0.086	0.040	0.040	0.017	0.017
BF-CC $h(w)$	14,829	14,829	20,962	25,895	30,828	20,000	20,000	20,000	20,000	20,000

EXHIBIT 3

2ND TO 3RD FACTORS FROM LARGE TRIANGLE

2nd to 3rd→		1.81	1.60	1.41	2.29	2.25	1.38
	1.36	1.07	1.60	0.89	1.42	0.99	1.01
	1.03	1.02	1.35	1.21	1.28	1.51	1.17
	2.00	0.98	1.21	1.24	1.79	1.32	1.48
	1.51	1.01	1.51	1.06	1.60	1.10	1.11
	2.20	2.00	1.50	2.20	1.19	1.28	1.52

APPENDIX

DIAGONAL EFFECTS IN BF MODELS

As an example, a test for diagonal effects in the CC model was made in the reinsurance triangle as follows. The CC is the same as the additive chain ladder, so it can be expressed as a linear model. This can be estimated via a single multiple regression in which the dependent variable is the entire list of incremental losses for ages 1 to 9 and all accident years—45 items in all. That is, the triangle beyond age 0 is strung out into a single vector. Age and diagonal dummy independent variables can be established in a design matrix to pick out the right elements of the parameter vector of age and diagonal terms to estimate each incremental loss cell. For the additive chain ladder, the column dummy variables will be 1 or 0, as opposed to cumulative losses or 0 in the chain ladder example. Then the coefficient of that column will be the additive element for the given age.

The later columns of the design matrix would be diagonal dummies, as in the chain ladder example. By doing a multiple linear regression for the incremental loss column in terms of the age and diagonal dummies, additive terms by age and by diagonal will be estimated. The regression can tell which terms are statistically significant, and the others can be dropped from the specification.

With the reinsurance triangle tested above, the first three diagonals turned out to be lower than the others, as was the last diagonal. Also, the first two ages were not significantly different from each other, nor were the last four. This produced a model with five age parameters and two diagonal parameters—one for the first three diagonals combined, and one for the last diagonal. The parameters are shown in Table 9.

The sum of squared residuals for this model is 49,673.4 when adjusted for seven parameters used. This is considerably better

TABLE 9

TERMS IN ADDITIVE CHAIN LADDER WITH DIAGONAL EFFECTS

Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Diag 1–3	Diag 9
5,569.0	5,569.0	3,739.2	2,881.8	2,361.1	993.3	993.3	993.3	993.3	−2,319.9	−984.7

than the model without diagonal effects. The multiple regression found the diagonals to be statistically significant and adding them to the model improved the fit.

A problem with the diagonal analysis is how to use them in forecasting. One reason for diagonal effects is a change in company practice, particularly in the claims handling process. If the age effects are considered the dominant influence with occasional distortion by diagonal effects, then including diagonal dummy variables will give better estimates for the underlying age terms. Then these, but not the diagonal effects, would be used in forecasting.

Having identified the significant diagonal effects through linear regression, it may be more reasonable to convert them to multiplicative effects through non-linear regression. The model could be of the form:

$$q(w,d) = f(d)g(w+d),$$

where $f(d)$ is the additive age term for age d, and $g(w+d)$ is the factor for the $w+d$th diagonal. Again this can be estimated iteratively by fixing the f's to estimate the g's by linear regression, then fixing those g's to estimate the next iteration of f's, until convergence is reached. The previous model was refit with the diagonals as factors with the result in Table 10. This had a slightly better adjusted sum of squared residuals of 49,034.8.

Diagonal factors can be used in conjunction with accident year factors as in:

$$q(w,d) = f(d)g(w+d)h(w).$$

TABLE 10

ADDITIVE CHAIN LADDER WITH MULTIPLICATIVE DIAGONAL EFFECTS

Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Diag 1–3	Diag 9
5,692.3	5,692.3	3,823.0	2,816.1	2,416.7	672.1	672.1	672.1	672.1	.5598	.6684

TABLE 11

ADDITIVE CHAIN LADDER WITH MULTIPLICATIVE DIAGONAL & AY EFFECTS

Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Diag 1-3	Diag 9	AY 3-4
5,135.6	5,135.6	3,464.7	2,730.1	1,995.4	660.1	660.1	660.1	660.1	.6201	.7225	1.2672

As an example, a factor was added to the above model to represent accident years 3 and 4, and the 4th age term was forced to be the average of the 3rd and 5th. The result is in Table 11.

The adjusted sum of squared residuals came down to 44,700.9, which is considerably better than the previous best-fitting model, and almost twice as good as in the original BF model, which in turn was almost twice as good as the chain ladder. It appears that accident year effects and diagonal effects are significant in this data. The fit is shown as the last section of Exhibit 1. The numerous examples fit to this data were for the sake of illustration. Some models of the types discussed may still fit better than the particular ones shown here.

Errata to
Testing the Assumptions of Age-to-Age Factors
By Venter, G.G. in *PCAS* LXXXV, 1998

Casualty Actuarial Society[1]

Version 1.0, January 31, 2020

This note presents errata to material in Venter's paper on "Testing the Assumptions of Age-to-Age Factors." Items printed in **red** indicate an update, clarification, or change.

1. Errata

The following passage of Venter Factors (page 833) should be amended from:

> Letting r denote the sample correlation coefficient, define $T = r[(n - 2) / (1 - r^2)]^{1/2}$. A significance test for the correlation coefficient can be made by considering T to be t-distributed with n – 2 degrees of freedom. If T is greater than the t-statistic for 0.9 at n – 2 degrees of freedom, for instance, then r can be considered significant at the 10% level. (See Miller and Wichern [7, p. 214].)

to:

> Letting r denote the sample correlation coefficient, define $T = r[(n - 2) / (1 - r^2)]^{1/2}$. Given correlations can be positive or negative, a two-tailed significance test for the correlation coefficient can be made by considering T to be t-distributed with n – 2 degrees of freedom. If |T| is greater than the t-statistic for 0.95 at n – 2 degrees of freedom, for instance, then r can be considered significant at the 10% level. (See Miller and Wichern [7, p. 214].)

[1] This note was prepared by the Exam 7 Syllabus Committee.

© 2020, Casualty Actuarial Society, All Rights Reserved

Obtaining Predictive Distributions for Reserves Which Incorporate Expert Opinion

by R. J. Verrall

ABSTRACT

This paper shows how expert opinion can be inserted into a stochastic framework for loss reserving. The reserving methods used are the chain-ladder and Bornhuetter-Ferguson, and the stochastic framework follows England and Verrall [8]. Although stochastic models have been studied, there are two main obstacles to their more frequent use in practice: ease of implementation and adaptability to user needs. This paper attempts to address these obstacles by utilizing Bayesian methods, and describing in some detail the implementation, using freely available software and programs supplied in the Appendix.

KEYWORDS

Bayesian statistics, Bornhuetter-Ferguson, chain-ladder, claims reserving, expert opinion, risk

1. Introduction

There has been a lot of attention given to stochastic reserving methods in the actuarial literature over recent years. Useful summaries can be found in England and Verrall [8] and Taylor [17]. The reader is strongly recommended to read England and Verrall [8], which contains more details on the basic models, before reading this paper.

There have been many useful things that have resulted from the recent papers on stochastic loss reserving: it is now possible to use a variety of methods to obtain reserve estimates, prediction intervals, predictive distributions, and so on. It is possible to use these methods for assessing the reserving risk, and for modeling a portfolio, line of business, or a whole company in a dynamic financial analysis. In short, the research published in recent years has been very successful in enhancing the understanding of loss reserving methods. This has been done by establishing stochastic approaches to models that are commonly used for loss reserving—for example, the chain-ladder technique, the Hoerl curve, and other parametric and non-parametric models. The stochastic approaches have added further models to the range of possible approaches. To take just one example, England and Verrall [7] showed how a nonparametric approach can be used to define a complete spectrum of models, with the chain-ladder technique at one end and the Hoerl curve at the other end.

In practical terms, it appears that the stochastic approaches that have found most popularity are those that are the simplest to implement. To pick out two examples, both Mack's model ([11]) and the bootstrap ([6] and [5]) are straightforward to implement in a spreadsheet. In contrast, using the full statistical model requires the use of statistical software, with some careful programming. It is not surprising, therefore, that a practitioner requiring prediction intervals as well as reserve estimates, or simply wanting to investigate the use of a stochastic approach, should choose the methods that are simplest to implement.

One aspect of reserving that has not, so far, received a great deal of attention in the literature is the question of intervention in the process by the actuary. In other words, the stochastic models have largely concentrated on providing a framework for the basic, standard methods. When these are used in practice, it is common to apply some expert knowledge or opinion to adjust the results before they are used. Examples of situations when intervention may be desirable is when there has been a change in the payment pattern due to a change in company policy, or where legislatures have enacted benefit limitations that restrict the potential for loss development and require an adjustment to historical development factors.

While it is possible to intervene in some models, the tendency is for this intervention to disrupt the assumptions made in the stochastic framework. For example, it is possible to change one or more of the residuals before applying a bootstrapping procedure, if the observed residuals appear to be out of line with what might be expected. But if this is done, the validity of the stochastic assumptions may be compromised. To take another example, consider the chain-ladder technique. This method involves the estimation of development factors, but it is often the case that these are adjusted before being applied to obtain reserve estimates. If this is done, the estimates from the stochastic model are being abandoned, and it is not clear what effect this might have on the prediction errors. For example, it is possible to calculate estimation errors for any parameter estimated in a stochastic model, but what estimation error should be used for a parameter that is simply inserted? The only way to address this properly is to use the Bayesian approach, and this provides an important motivation for the ideas discussed in this paper.

A second area where expert knowledge is applied is when the Bornhuetter-Ferguson [1] technique is used. This method uses the development factors from the chain-ladder technique, but it does not apply these to the latest cumulative losses to estimate the outstanding losses. Instead, an estimate is first procured separately, using background knowledge about the claims. This is then used with the development factors to obtain reserve estimates. Although not originally formulated using a Bayesian philosophy, the Bornhuetter-Ferguson technique is quite clearly suited to this approach because of the basic idea of what it is trying to do: incorporate expert opinion. Thus, we have a second important motivation for considering the use of Bayesian reserving methods. These are two very important examples of reserving approaches commonly used, which are best modeled using Bayesian methods. Among previous papers to discuss Bayesian loss reserving, we would mention de Alba [4] and Ntzoufras and Dellaportas [13].

One important property of Bayesian methods that makes them suitable for use with a stochastic reserving model is that they allow us to incorporate expert knowledge in a natural way, overcoming any difficulties about the effect on the assumptions made. In this paper, we consider the use of Bayesian models for loss reserving in order to incorporate expert opinion into the prediction of reserves. We concentrate on two areas as mentioned above: the Bornhuetter-Ferguson technique and the insertion of prior knowledge about individual development factors in the chain-ladder technique. The possibility of including expert knowledge is an important property of Bayesian models, but there is another equally important point: the ease with which they can be implemented. This is due to modern developments in Bayesian methodology based on so-called "Markov chain Monte Carlo" (MCMC) methods. It is difficult to emphasize enough the effect these methods have had on Bayesian statistics, but the books by Congdon ([3] and [2]) give some idea of the scope of the applications for which they have been used. The crucial aspect as far as this paper is concerned is that they are based on simulation, and therefore have some similarities with bootstrapping methods that, as was mentioned above, have gained in popularity for loss reserving.

It is also important that easy-to-use software is now available that allows us to implement the Bayesian models for loss reserving. While it is straightforward to define a Bayesian model, it is not always so easy to find the required posterior distributions for the parameters and predictive distributions for future observations. However, this has been made much easier in recent years by the development of MCMC methods, and by the software package winBUGS [16]. This software package is freely available from http://www.mrc-bsu.cam.ac.uk/bugs, and the programs for carrying out the Bayesian analysis for the models described in this paper are contained in the Appendix. Section 6.1 provides instructions on downloading this software. An excellent reference for actuarial applications of MCMC methods using winBUGS is Scollnik [15].

The basic idea behind MCMC methods is to simulate the posterior distribution by breaking the simulation process down into a number of simulations that are as easy to carry out as possible. This overcomes a common problem with Bayesian methods—that it can be difficult to derive the posterior distribution, which may in many cases be multidimensional. Instead of trying to simulate all the parameters at once, MCMC methods use the conditional distribution of each parameter, given all the others. In this way, the simulation is reduced to a univariate distribution, which is much easier to deal with. A Markov chain is formed because each parameter is considered in turn, and it is a simulation-based method: hence the term Markov chain Monte Carlo. For the readers for whom this is the first time they have encountered MCMC methods, it is suggested that they simply accept that they are a neat way to get the posterior distributions for Bayesian models and continue reading the paper. If they like the ideas and would like to find out more, Scollnik [15] gives a much fuller account than is possible here, and the reader is advised to spend time working through some simpler examples with the help of the Scollnik paper.

This paper is set out as follows. In Section 2, we describe the notation and basic methods used, and in Section 3 we summarize the stochastic models used in the context of the chain-ladder technique. Sections 4 and 5 describe the Bayesian models for incorporating prior information into the reserving process. In Section 6 we describe in some detail how to implement the Bayesian models so that the reader can investigate the use of these models, using the programs given in the Appendix. In Section 7 we state some conclusions.

2. Notation and basic methods

To begin with, we define the notation used in this paper, and in doing so we briefly summarize the chain-ladder technique and the Bornhuetter-Ferguson method.

Although the methods can also be applied to other shapes of data, in order that the notation should not get too complicated we make the assumption that the data is in the shape of a triangle. Thus, without loss of generality, we assume that the data consist of a triangle of incremental losses:

$$\begin{array}{llll} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & \cdots & C_{2,n-1} & \\ \vdots & & & \\ C_{n1} & & & \end{array}.$$

This can also be written as $\{C_{ij} : j = 1,\ldots,n-i+1; i = 1,\ldots,n\}$, where n is the number of accident years. C_{ij} is used to denote incremental losses, and D_{ij} is used to denote the cumulative losses, defined by:

$$D_{ij} = \sum_{k=1}^{j} C_{ik}. \tag{2.1}$$

One of the methods considered in this paper is the chain-ladder technique, and the development factors $\{\lambda_j : j = 2,\ldots,n\}$. The usual estimates of the development factors from the standard chain-ladder technique are

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} D_{ij}}{\sum_{i=1}^{n-j+1} D_{i,j-1}}. \tag{2.2}$$

Note that we only consider forecasting losses up to the latest development year (n) so far observed, and no tail factors are applied. It would be possible to extend this to allow a tail factor, using the same methods, but no specific modeling is carried out in this paper of the shape of the run-off beyond the latest development year. Thus, we refer to cumulative losses up to development year n, $D_{in} = \sum_{k=1}^{n} C_{ik}$, as "ultimate losses." For the chain-ladder technique, the estimate of outstanding losses is $D_{i,n-i+1}(\hat{\lambda}_{n-i+2} \cdot \hat{\lambda}_{n-i+3} \cdots \hat{\lambda}_n - 1)$.

The first case we consider is when these development factor estimates are not used for all rows. In other words, we consider the more general case where there is a separate development factor in each row, $\lambda_{i,j}$. The standard chain-ladder model sets $\lambda_{i,j} = \lambda_j$, for $i = 1,2,\ldots,n-j+1$; $j = 2,3,\ldots,n$, but we consider allowing the more general case where development factors can change from row to row. Section 4 describes the Bayesian approach to this, allowing expert knowledge to be used to set prior distributions for these parameters. In this way, we will be able to intervene in the estimation of the development factors, or else simply leave them for the standard chain-ladder model to estimate.

In Section 5 we consider the Bornhuetter-Ferguson method. This method uses the development factors from the chain-ladder technique, but it incorporates knowledge about the "level" of each row by replacing the chain-ladder estimate of outstanding claims, $D_{i,n-i+1}(\hat{\lambda}_{n-i+2}\hat{\lambda}_{n-i+3}\cdots\hat{\lambda}_n - 1)$ by $M_i(1/(\hat{\lambda}_{n-i+2}\hat{\lambda}_{n-i+3}\cdots\hat{\lambda}_n))(\hat{\lambda}_{n-i+2}\cdot\hat{\lambda}_{n-i+3}\cdots\hat{\lambda}_n - 1)$. Here, M_i denotes a value for the ultimate losses for accident year i that is obtained using expert knowledge about the losses (for example, taken from the premium calculation). Thus, $M_i(1/(\hat{\lambda}_{n-i+2}\hat{\lambda}_{n-i+3}\cdots\hat{\lambda}_n))$ replaces the latest cumulative losses for accident year i, to which the usual chain-ladder parameters are applied to obtain the estimate of outstanding losses. From this, it can be seen that the difference between the Bornhuetter-Ferguson method and the chain-ladder technique is that the Bornhuetter-Ferguson technique uses an external estimate of the "level" of each row in the triangle, while the chain-ladder technique uses the data in that row itself. The Bornhuetter-Ferguson method can be formulated using a Bayesian approach, with the information about the external estimates for each row being used to form the prior distributions, as in Section 5.

This section has defined the notation used in the paper, and outlined the basic reserving methods that will be considered using stochastic approaches. In order to do this, a brief introduction to the stochastic models is needed, and this is given in Section 3.

3. Stochastic models for the chain-ladder technique

This section gives a brief summary of stochastic models that are related to the chain-ladder technique. A much fuller account may be found in England and Verrall [8], and in that paper's references and discussion. We consider the chain-ladder technique and note that it is possible to apply Bayesian methods in a similar way to other models.

There are a number of different approaches that can be taken to the chain-ladder technique, with various positivity constraints, all of which give the same reserve estimates as the chain-ladder technique. The connections between the chain-ladder technique and various stochastic models have been explored in a number of previous papers. For example, Mack [11] takes a non-parametric approach and specifies only the first two moments for the cumulative losses. In Mack's model the conditional mean and variance of $D_{ij} \mid D_{i,j-1}, \lambda_j, \sigma_j^2$ are $\lambda_j D_{i,j-1}$ and $\sigma_j^2 D_{i,j-1}$, respectively. Estimates of all the parameters are derived, and the properties of the model are examined. As was stated in the introduction, one of the advantages of this approach is that the parameter estimates and prediction errors can be obtained using a spreadsheet, without having recourse to a statistical package or any complex programming. The consequence of not specifying a distribution for the data is that there is no predictive distribution. Also, there are separate parameters in the variance that must also be estimated, separately from the estimation of the development factors.

As a separate stream of research, generalized linear models have also been considered. Renshaw and Verrall [14] used an approach based on generalized linear models [12] and examined the over-dispersed Poisson model for incremental losses:

$C_{ij} \mid c, \alpha, \beta, \varphi \sim$ independent over-dispersed Poisson, with mean, m_{ij}, where $\log(m_{ij}) = c + \alpha_i + \beta_j$, and $\alpha_1 = \beta_1 = 0$.

The term "over-dispersed" requires some explanation. It is used here in connection with the Poisson distribution, and it means that if $X \sim \text{Poisson}(\mu)$, then $Y = \varphi X$ follows the over-dispersed Poisson distribution with $E(Y) = \varphi\mu$ and $V(Y) = \varphi^2 E(X) = \varphi^2\mu$. φ is usually greater than 1—hence the term "over-dispersed"—but this is not a necessity. It can also be used for other distributions, and we make use of it for the negative binomial distribution. As with the Poisson distribution, the over-dispersed negative binomial distribution is defined such that if $X \sim$ negative binomial then $Y = \varphi X$ follows the over-dispersed negative binomial distribution. Furthermore, a quasi-likelihood approach is taken so that the loss data are not restricted to the positive integers.

It can be seen that this formulation has some similarities with the model of Kremer [9], but it has a number of advantages. It does not necessarily break down if there are negative incremental loss values, it gives the same reserve estimates as the chain-ladder technique, and it has been found to be more stable than the log-normal model of Kremer. For these reasons, we concentrate on it in this paper. There are a number of ways of writing this model, which are useful in different contexts (note that the reserve estimates are unaffected by the way the model is written). In a strict sense, the formulation requires that the data are positive—otherwise it is more difficult to justify and interpret the inferences made from the data. However, in a purely practical context, it is useful to note that the estimation does not break down in the presence of some negative values.

Another way of writing the over-dispersed Poisson model for the chain-ladder technique is as follows:

$C_{ij} \mid x, y, \varphi \sim$ independent over-dispersed Poisson, with mean $x_i y_j$, and $\sum_{k=1}^{n} y_k = 1$.

Here $x = \{x_1, x_2, \ldots, x_n\}$ and $y = \{y_1, y_2, \ldots, y_n\}$ are parameter vectors relating to the rows (accident years) and columns (development years), respectively, of the run-off triangle. The parameter $x_i = E[D_{in}]$, and so represents expected ultimate cumulative losses (up to the latest development year so far observed, n) for the ith accident year. The column parameters, y_j, can be interpreted as the proportions of ultimate losses that emerge in each development year.

Although the over-dispersed Poisson models give the same reserve estimates as the chain-ladder technique (as long as the row and column sums of incremental claims are positive), the connection with the chain-ladder technique is not immediately apparent from this formulation of the model. For this reason, the negative binomial model was developed by Verrall [20], building on the over-dispersed Poisson model. Verrall showed that the same predictive distribution can be obtained from a negative binomial model (also with the inclusion of an over-dispersion parameter). In this recursive approach, the incremental claims have an over-dispersed negative binomial distribution, with mean and variance

$(\lambda_j - 1)D_{i,j-1}$ and $\varphi\lambda_j(\lambda_j - 1)D_{i,j-1}$, respectively.

Again, the reserve estimates are the same as the chain-ladder technique, and the same positivity constraints apply as for the over-dispersed Poisson model. It is clear from this that the column sums must be positive, since a negative sum would result in a development factor less than 1 ($\lambda_j < 1$), causing the variance to be negative. It is important to note that exactly the same predictive distribution can be obtained from either the Poisson or negative binomial models. Verrall [20] also argued that the model could be specified either for incremental or cumulative losses, with no difference in the results. The negative binomial model has the advantage that the form of the mean is exactly the same as that which naturally arises from the chain-ladder technique. In fact, by adding the previous cumulative losses, an equivalent model for $D_{ij} \mid D_{i,j-1}, \lambda_j, \varphi$ has an over-dispersed negative binomial distribution, with mean and variance

$\lambda_j D_{i,j-1}$ and $\varphi\lambda_j(\lambda_j - 1)D_{i,j-1}$, respectively.

Here the connection with the chain-ladder technique is immediately apparent because of the format of the mean.

Another model, which is not considered further in this paper, is closely connected with Mack's model, and deals with the problem of negative incremental claims. This model replaces the negative binomial by a normal distribution, whose mean is unchanged, but whose variance is altered to accommodate the case when $\lambda_j < 1$. Preserving as much of $\lambda_j(\lambda_j - 1)D_{i,j-1}$ as possible, the variance is still proportional to $D_{i,j-1}$, with the constant of proportionality depending on j, but a normal approximation is used for the distribution of incremental claims. Thus, $C_{ij} \mid D_{i,j-1}, \lambda_j, \varphi_j$ is approximately normally distributed, with mean and variance

$$D_{i,j-1}(\lambda_j - 1) \text{ and } \varphi_j D_{i,j-1}, \text{ respectively,}$$

or $D_{ij} \mid D_{i,j-1}, \lambda_j, \varphi_j$ is approximately normally distributed, with mean and variance

$$\lambda_j D_{i,j-1} \text{ and } \varphi_j D_{i,j-1}, \text{ respectively.}$$

As for Mack's model, there is now another set of parameters in the variance that needs to be estimated.

For each of these models, the mean square error of prediction can be obtained, allowing the construction of prediction intervals, for example. Loss reserving is a predictive process: given the data, we try to predict future loss emergence. These models apply to all the data, both observed and future observations. The estimation is based on the observed data, and we require predictive distributions for the future observation.

We use the expected value of the distribution of future losses as the prediction. When considering variability, attention is focused on the root mean squared error of prediction (RMSEP), also known as the prediction error. To explain what this is, we consider, for simplicity, a random variable, y, and a predicted value $\hat{y}$. The mean squared error of prediction (MSEP) is the expected square difference between the actual outcome and the predicted value, $E[(y - \hat{y})^2]$, and can be written as follows:

$$E[(y - \hat{y})^2] = E[((y - E[y]) - (\hat{y} - E[y]))^2]. \tag{3.1}$$

In order to obtain an estimate of this, it is necessary to plug in $\hat{y}$ instead of y in the final expectation. Then the MSEP can be expanded as follows:

$$\begin{aligned} E[(y - \hat{y})^2] \approx\ & E[(y - E[y])^2] \\ & - 2E[(y - E[y])(\hat{y} - E[\hat{y}])] \\ & + E[(\hat{y} - E[\hat{y}])^2]. \end{aligned} \tag{3.2}$$

Assuming future observations are independent of past observations, the middle term is zero, and

$$E[(y - \hat{y})^2] \approx E[(y - E[y])^2] + E[(\hat{y} - E[\hat{y}])^2]. \tag{3.3}$$

In words, this is

prediction variance = process variance
+ estimation variance.

It is important to understand the difference between the prediction error and the standard error. Strictly, the standard error is the square root of the estimation variance. The prediction error is concerned with the variability of a forecast, taking account of uncertainty in parameter estimation and also of the inherent variability in the data being forecast. Further details of this can be found in England and Verrall [8].

Using non-Bayesian methods, these two components—the process variance and the estimation variance—are estimated separately, and Section 7 of England and Verrall [8] goes into detail about this. The direct calculation of these quantities can be a tricky process, and this is one of the reasons for the popularity of the bootstrap. The bootstrap uses a fairly simple simulation approach to obtain simulated estimates of the prediction variance in a spreadsheet. Fortunately, the same advantages apply to the Bayesian methods: the full predictive distribution can be found using simulation methods, and the RMSEP can be obtained directly by calculating its standard deviation. In addition, it is preferable to have the full predictive distribution, rather than just the first two moments, which is another advantage of Bayesian methods.

The purpose of this paper is to show how expert opinion, from sources other than the specific data set under consideration, can be incorporated into the predictive distributions of the reserves. We use the approach of generalized linear models outlined in this section, concentrating on the over-dispersed Poisson and negative binomial models. We begin with considering how it is possible to intervene in the development factors for the chain-ladder technique in Section 4, and then consider the Bornhuetter-Ferguson method in Section 5.

4. Incorporating expert opinion about the development factors

In this section, the approach of Verrall and England [21] is used to show how to specify a Bayesian model that allows the practitioner to intervene in the estimation of the development factors for the chain-ladder technique. There are a number of ways in which this could be used, and we describe some possibilities in this section. It is expected that a practitioner would be able to extend these to cover situations that, although not specifically covered here, would also be useful. The cases considered here are the intervention in a development factor in a particular row, and the choice of how many years of data to use in the estimation. The reasons for intervening in these ways could be that there is information that the settlement pattern has changed, making it inappropriate to use the same development factor for each row.

For the first case, what may happen in practice is that a development factor in a particular row is simply changed. Thus, although the same development parameters (and hence run-off pattern) are usually applied for all accident years, if there is some exogenous information that indicates that this is not appropriate, the practitioner may decide to apply a different development factor (or set of factors) in some, or all, rows.

In the second case, it is common to look at, say, five-year volume weighted averages in calculating the development factors, rather than using all the available data in the triangle. The Bayesian methods make this particularly easy to do and are flexible enough to allow many possibilities.

We use the negative binomial model described in Section 3, with different development factors in each row. This is the model for the data, and we then specify prior distributions for the development factors. In this way, we can choose prior distributions that reproduce the chain-ladder results, or we can intervene and use prior distributions based on external knowledge. The model for incremental claims, $C_{ij} \mid D_{i,j-1}, \lambda_{i,j}, \varphi$, is an over-dispersed negative binomial distribution, with mean and variance

$$(\lambda_{i,j} - 1)D_{i,j-1} \text{ and } \varphi\lambda_{i,j}(\lambda_{i,j} - 1)D_{i,j-1}, \text{ respectively.}$$

We next need to define prior distributions for the development factors, $\lambda_{i,j}$. It is possible to set some of these equal to each other (within each column) in order to revert to the standard chain-ladder model. This is done by setting

$$\lambda_{i,j} = \lambda_j \qquad \text{for} \quad i = 1,2,\ldots,n-j+1; \quad j = 2,3,\ldots,n$$

and defining vague prior distributions for λ_j ($j = 2,3,\ldots,n$). This was the approach taken in Section 8.4 of England and Verrall [8] and is very similar to that taken by de Alba [4]. This can provide a very straightforward method to obtain prediction errors and predictive distributions for the chain-ladder technique.

However, we really want to move away from the basic chain-ladder technique, and construct Bayesian prior distributions that encompass the expert opinion about the development parameters. Suppose, for example, that we have a 10×10 triangle. We consider the two possibilities for incorporating expert knowledge described above.

To illustrate the first case, suppose that there is information that implies that the second development factor (from Column 2 to Column 3) should be given the value 1.5, for rows 8, 9, and 10, and that there is no indication that the other parameters should be treated differently from the standard chain-ladder technique. An appropriate way to treat this would be to specify

$$\lambda_{i,j} = \lambda_j \quad \text{for} \quad i = 1,2,\ldots,n-j+1; \quad j = 2,4,5,\ldots,n$$

$$\lambda_{i,3} = \lambda_3 \quad \text{for} \quad i = 1,2,\ldots,7$$

$$\lambda_{8,3} = \lambda_{9,3} = \lambda_{10,3}.$$

The means and variances of the prior distributions of the parameters are chosen to reflect the expert opinion:

$\lambda_{8,3}$ has a prior distribution with mean 1.5 and variance W, where W is set to reflect the strength of the prior information.

λ_j have prior distributions with large variances.

For the second case, we divide the data into two parts using the prior distributions. To do this, we set

$$\lambda_{i,j} = \lambda_j \quad \text{for} \quad i = n-j-3, n-j-2, n-j-1, n-j, n-j+1$$

$$\lambda_{i,j} = \lambda_j^* \quad \text{for} \quad i = 1,2,\ldots,n-j-4$$

and give both λ_j and λ_j^* prior distributions with large variances so that they are estimated from the data. Adjustments to the specification are made in the later development years, where there are less than five rows. For these columns there is just one development parameter, λ_j.

The specific form of the prior distribution (gamma, log-normal, etc.) is usually chosen so that the numerical procedures in winBUGS work as well as possible.

These models are used as illustrations of the possibilities for incorporating expert knowledge about the development pattern, but it is (of course) possible to specify many other prior distributions. In the Appendix, the winBUGS code is supplied, which can be cut and pasted directly in order to examine these methods. Section 6 contains a number of examples, including the ones described in this section.

Table 1. Parameters, mean and variance of a gamma distribution

α_i	β_i	M_i	M_i/β_i
10000	10	1000	100
1000	1	1000	1000
100	0.1	1000	10000

5. A Bayesian model for the Bornhuetter-Ferguson method

In this section, we show how the Bornhuetter-Ferguson method can be considered in a Bayesian context, using the approach of Verrall [19]. For further background on the Bornhuetter-Ferguson method, see Mack [10].

In Section 3, the over-dispersed Poisson model was defined as follows.

$C_{ij} \mid x, y, \varphi \sim$ independent over-dispersed Poisson, with mean $x_i y_j$, and $\sum_{k=1}^{n} y_k = 1$.

In the Bayesian context, we also require prior distributions for the parameters. The Bornhuetter-Ferguson method assumes that there is expert opinion about the level of each row, and we therefore concentrate first on the specification of prior distributions for these. The most convenient form to use is gamma distributions:

$$x_i \mid \alpha_i, \beta_i \sim \text{ independent } \Gamma(\alpha_i, \beta_i). \quad (5.1)$$

There is a wide range of possible choices for the parameters of these prior distributions, α_i and β_i. It is easiest to consider the mean and variance of the gamma distribution, α_i/β_i and α_i/β_i^2, respectively. These can be written as M_i and M_i/β_i, from which it can be seen that, for a given choice of M_i, the variance can be altered by changing the value of β_i. To consider a simple example, suppose it has been decided that $M_i = 1000$. Table 1 shows how the value of β_i affects the variance of the prior distribution, while M_i is kept constant.

Clearly, choosing a larger value of β_i implies we are more sure about the value of M_i, and choosing a smaller value means we are less sure.

We now consider the effect of using these prior distributions on the model for the data. Recall that, for the chain-ladder technique, the mean of the distribution of incremental claims may be written as $(\lambda_j - 1)D_{i,j-1}$.

Using a similar approach, Verrall [20] and Verrall [19] derive the distribution of C_{ij}, given the past data, after the row parameters have been estimated. In a Bayesian context, this means first deriving the posterior distribution of the row parameters given the data using a standard prior-posterior analysis:

$$f(x_i \mid y; data) \propto f(data \mid x, y) f(x_i \mid \alpha_i, \beta_i). \tag{5.2}$$

Note that, if we are considering C_{ij}, the *data* used here is $C_{i1}, C_{i2}, \ldots, C_{i,j-1}$. Having obtained this distribution, the distribution of the next observation can be found as follows:

$$f(C_{ij} \mid y; data) = \int f(C_{ij} \mid x_i, y) f(x_i \mid y; data) dx_i. \tag{5.3}$$

This result is derived in detail in Verrall [19], where it is shown that it is possible to rewrite it in terms of the usual chain-ladder development factors, λ_j, rather than using the column parameters y_j. For full details of the derivation, the reader is referred to Verrall [19]. For the purposes of this paper, the important point is that the mean of C_{ij} for the Bayesian model is

$$Z_{ij}(\lambda_j - 1)D_{i,j-1} + (1 - Z_{ij})(\lambda_j - 1)M_i \frac{1}{\lambda_j \lambda_{j+1} \ldots \lambda_n},$$

where

$$Z_{ij} = \frac{\sum_{k=1}^{j-1} y_k}{\beta_i \varphi + \sum_{k=1}^{j-1} y_k}.$$

This can be seen to be in the form of a credibility formula, and is a trade-off between the chain-ladder $((\lambda_j - 1)D_{i,j-1})$ and the Bornhuetter-Ferguson $((\lambda_j - 1)M_i(1/(\lambda_j \lambda_{j+1} \ldots \lambda_n)))$. The credibility factor, Z_{ij}, governs the trade-off between the prior mean and the data. We can influence the balance of this trade-off through the choice of β_i. In line with the discussion above, the larger the value of β_i the closer we get to the Bornhuetter-Ferguson method, and the smaller the value of β_i, the closer we get to the chain-ladder technique. In this way, we can use different specifications of the prior distributions for the row parameters in order to use the chain-ladder technique, the Bornhuetter-Ferguson method, or a complete spectrum of methods between these two extremes. If we choose to use prior distributions with large variances, we do not influence the parameter estimates and the result will be the same as (or extremely close to) the chain-ladder technique. If we use very small variances, we are saying that we are very sure what the parameter values should be and the results will be the same as (or very close to) the Bornhuetter-Ferguson method. Thus, we can use these methods within a stochastic framework, and we can also consider using the whole range of models that lie between these two.

We have yet to consider the estimation of the column parameters, other than to point out that the Bornhuetter-Ferguson method, being deterministic, simply plugs in the chain-ladder parameter estimates. We now consider this issue in more detail and define a Bayesian approach to the Bornhuetter-Ferguson method. One option is to simply use plug-in estimates, obtained, for example, from the straightforward chain-ladder technique. This is the approach used in the deterministic application of the Bornhuetter-Ferguson method, but it is not suitable here since we would prefer a stochastic approach. A better option is to define improper prior distributions for the column parameters, and estimate the column parameters *first*, before applying prior distributions for the row parameters and estimating these. This second option allows us to take into account the fact that the column parameters have been estimated when calculating the prediction errors, predictive distribution, etc. It is not required to include any information about the column parameters, and hence we use improper gamma distributions for the column parameters, and derive the posterior distributions of these using a standard Bayesian prior-posterior analysis. The result of this is a distribution that looks similar to the negative binomial model for the chain-ladder technique, but which is recursive in i instead of j:

$C_{ij} \mid C_{1,j}, C_{2,j}, \ldots, C_{i-1,j}, x, \varphi \sim$ over-dispersed negative binomial, with mean $(\gamma_i - 1)\sum_{m=1}^{i-1} C_{m,j}$.

Comparing this to the mean of the chain-ladder model, $(\lambda_j - 1)D_{i,j-1} = (\lambda_j - 1)\sum_{m=1}^{j-1} C_{i,m}$, it can be seen that they are identical in form, with the recursion either being across the rows or down the columns.

In the context of the Bornhuetter-Ferguson method, we now have the stochastic version of this model. The Bornhuetter-Ferguson method inserts values for the expected ultimate claims in each row, x_i, in the form of the values M_i. In the Bayesian context, prior distributions will be defined for the parameters x_i, as discussed above. However, the model has been reparameterized, with a new set of parameters, γ_i. Hence, it is necessary to define the relationship between the new parameters, γ_i, and the original parameters, x_i. This is given in the equations below, which can be used to find values of γ_i from the values of x_i given in the prior distributions. Note that there was an error in the equation given in Verrall [19], and I am grateful to Peter England for pointing this out.

The Bornhuetter-Ferguson technique can be reproduced by using strong prior information for the row parameters, x, and the chain-ladder technique can be reproduced by using improper priors for the row parameters. In other words, the Bornhuetter-Ferguson technique assumes that we are completely sure about the values of the row parameters, and their prior distributions have very small variances, while the chain-ladder technique assumes there is no information and has very large variances.

The preceding equations have now defined a stochastic version of the Bornhuetter-Ferguson technique. Since the column parameters (the development factors) are dealt with first, using improper prior distributions, their estimates will be those implied by the chain-ladder technique. Prior information can be defined in terms of distributions for the parameters x_i, which can then be converted into values for the parameters γ_i, and this is implemented in Section 6.

6. Implementation

This section explains how the Bayesian models can be implemented, using the software package winBUGS [16] which is available from http://www.mrc-bsu.cam.ac.uk/bugs. The programs used in these illustrations are contained in the Appendix.

The data set used in this section is taken from Taylor and Ashe [18], and has also been used in a number of previous papers on stochastic reserv-

$$\begin{aligned}
\gamma_1 &= 1\\
\gamma_2 &= 1 + \frac{x_2\left(1 - \dfrac{1}{\lambda_n}\right)}{C_{1n}}\\
\gamma_i &= 1 + \frac{x_i\left(1 - \dfrac{1}{\sum_{k=n-i+2}^{n}\lambda_k}\right)}{\sum_{m=1}^{i-1} C_{m,n} + \sum_{k=n-i+3}^{n}\left[\left(\prod_{l=n-k+2}^{i-1}\gamma_l\right)\sum_{m=1}^{n-k+1} C_{m,k}\right]} \qquad i = 3,\ldots,n.
\end{aligned} \tag{5.4}$$

Table 2. Data from Taylor and Ashe [18] with the chain-ladder estimates

Accident Year	Development Year 1	2	3	4	5	6	7	8	9	10
1	357,848	766,940	610,542	482,940	527,326	574,398	146,342	139,950	227,229	67,948
2	352,118	884,021	933,894	1,183,289	445,745	320,996	527,804	266,172	425,046	
3	290,507	1,001,799	926,219	1,016,654	750,816	146,923	495,992	280,405		
4	310,608	1,108,250	776,189	1,562,400	272,482	352,053	206,286			
5	443,160	693,190	991,983	769,488	504,851	470,639				
6	396,132	937,085	847,498	805,037	705,960					
7	440,832	847,631	1,131,398	1,063,269						
8	359,480	1,061,648	1,443,370							
9	376,686	986,608								
10	344,014									
Chain-ladder development factors:										
3.4906	1.7473	1.4574	1.1739	1.1038	1.0863	1.0539	1.0766	1.0177		
Chain-ladder reserve estimates:										
2	94,634									
3	469,511									
4	709,638									
5	984,889									
6	1,419,459									
7	2,177,641									
8	3,920,301									
9	4,278,972									
10	4,625,811									
Overall	18,680,856									

ing. The incremental loss data is given in Table 2, together with the chain-ladder results for comparison purposes.

Before looking at the uses of the Bayesian models, we should discuss the nuisance parameter φ. In a full Bayesian analysis, we should also give this a prior distribution and estimate it along with the other parameters. However, for ease of implementation we instead use a plug-in estimate, in line with the approach taken in classical methods (in England and Verrall [8], for example). The value used is that obtained from the straightforward application of the over-dispersed Poisson model, estimating the row and column parameters using maximum likelihood estimation (it is possible to use S-Plus or Excel for this).

6.1. Using the software

Before considering the results from the programs in any detail, we first describe how to set up the software and run one of the programs from scratch. An excellent reference in the context of actuarial modeling is Skollnik [15]. Table 2 shows the standard chain-ladder results, and in this section we will implement the model described in Section 5, but use the assumptions of the chain-ladder technique, rather than the Bornhuetter-Ferguson method. This means that we will use large variances for the prior distributions for the ultimate claims in each row, implying that there is no prior knowledge about them, and hence the results we obtain should be close to the chain-ladder results. Thus, we will first reproduce the results that can also be obtained using non-Bayesian methods (see England and Verrall [8] for more details of the non-Bayesian methods). After going through this example in detail, the remainder of this section will show how the Bayesian models incorporating prior knowledge described in Sections 4 and 5 can be implemented, and illustrate the effect that the choice of prior distributions can have.

The steps necessary for implementing the chain-ladder technique in winBUGS are listed below.

1. Go to the web site, download the latest version of the software and install.
2. Go back to the web site and register, and you will be sent a copy of the key to unlock the software. Follow the instructions in the email for unlocking the software.
3. Once you have a fully functioning version of winBUGS, you can run the programs in the Appendix. Open winBUGS and click on "File" in the top toolbar, and then "New" in the pop-down list. This will open a new window.
4. Copy the program in (i) of the Appendix, including the word "model" at the top and all the data at the bottom, right down to where the next subsection begins at (ii). The last line is 0,0,0,0,0,0,0,0,0)). Paste all of this into the new window in winBUGS.
5. In winBUGS, select "Model" in the toolbar at the top and "Specification" in the pop-down list. This opens a new window called "Specification Tool."
6. Highlight the word "model" at the top of the program, and then click "check model" in the Specification Tool window. If all is well, it will say "model is syntactically correct" in the bottom left corner.
7. Now move down in the window containing the program until you get to #DATA. Highlight the word "list" immediately below that, and click "load data" in the Specification Tool window. It should say "data loaded" in the bottom left corner.
8. Click "compile" in the Specification Tool window. After a few seconds, it should say "model compiled" in the bottom left corner.
9. Now move down in the window containing the program until you get to #INITIAL VALUES. Highlight the word "list" immediately below that, and click "load inits" in the Specification Tool window. It should say "model is initialised" in the bottom left corner.
10. Select "Model" in the toolbar at the top and "Update" in the pop-down list. This opens a new window called "Update Tool." The number of iterations in the simulation process can be changed in this window by changing the figure next to "updates." Just at the moment, 1,000 is sufficient, so click on "update." This runs 1,000 simulations without storing the results. This may take a few minutes: don't be concerned if nothing appears to be happening! When it is complete, a message appears in the bottom left corner saying how long the updates took (for my laptop it was 221 seconds).
11. Select "Inference" in the toolbar at the top and "Samples" in the pop-down list. This opens a new window called "Sample Monitor Tool." We want to look at the row totals and overall total, which have been defined as a vector R and Total in the program. In the Sample Monitor Tool window, click in the box to the right of the word "node" and type R. Then click on "set." Repeat for Total, noting that it is case sensitive.
12. Return to the Update Tool Window and click on Update to perform 1,000 simulations. This should be quicker (6 seconds for my laptop). This time the values of R and Total will be stored.
13. Return to the Sample Monitor Tool window, type * in the box to the right of the word "node," and click "stats." This will give a new window with something like the results below. This completes the steps necessary for fitting the Bayesian model.

Table 3. Results

Node	Mean	sd	MC Error	2.5%	Median	97.5%	Start	Sample
R[2]	92750.0	110600.0	2963.0	779.2	56320.0	412800.0	1001	1000
R[3]	473900.0	223100.0	6424.0	1.52E+5	4.4E+5	1.011E+6	1001	1000
R[4]	7.05E+5	2.58E+5	9085.0	307600.0	674500.0	1.288E+6	1001	1000
R[5]	985800.0	304600.0	8127.0	467600.0	960600.0	1.667E+6	1001	1000
R[6]	1.417E+6	378300.0	13430.0	768500.0	1.399E+6	2.217E+6	1001	1000
R[7]	2.174E+6	5.19E+5	16850.0	1.271E+6	2.132E+6	3.233E+6	1001	1000
R[8]	3.925E+6	776900.0	28100.0	2.585E+6	3.885E+6	5.555E+6	1001	1000
R[9]	4.284E+6	1.066E+6	36840.0	2.464E+6	4.207E+6	6.731E+6	1001	1000
R[10]	4.641E+6	2.002E+6	61630.0	1.73E+6	4.407E+6	9.345E+6	1001	1000
Total	1.87E+7	3.056E+6	101600.0	1.314E+7	1.861E+7	2.554E+7	1001	1000

The columns of Table 3 headed "mean" and "sd" give the predicted reserves and prediction errors, and these values can be compared with the chain-ladder results in Table 2. Since this is a simulation process, the results will depend on the prior distributions, the initial values, and the number of iterations carried out. The prior distributions in the program had reasonably large variances, so the results should be close to the chain-ladder results. More simulations should be used in steps 10 and 12 (we use 10,000 in the illustrations below), and the prior variances could be increased. Using this number of simulations gives the results shown in Table 4.

The results certainly confirm that we can reproduce the chain-ladder results, and produce the prediction errors. It is also possible to obtain other information about the model from winBUGS. For example, it is possible to produce full predictive distributions, using "density" in the Sample Monitor Tool window.

We have now described one implementation of a Bayesian model using winBUGS. In the rest of this section, we consider the Bayesian models described in Sections 4 and 5 in order to consider how expert opinion can be incorporated into the predictive distribution of reserves. In each case, the programs are available in the Appendix, and the results can be reproduced using steps 3 to 13, above. It should be noted that this is a simulation-based program, so the results obtained may not exactly match the results given below. However, there should be no significant differences, with the differences that there are being due to simulation error.

Table 4. Chain-ladder results. the prediction error is equal to the Bayesian standard deviation

	Chain-Ladder Reserve	Bayesian Mean	Bayesian Standard Deviation	Prediction Error (%)
Year 2	94,634	94,440	111,100	118%
Year 3	469,511	471,400	219,400	47%
Year 4	709,638	716,300	263,600	37%
Year 5	984,889	991,600	308,100	31%
Year 6	1,419,459	1,424,000	374,700	26%
Year 7	2,177,641	2,186,000	497,200	23%
Year 8	3,920,301	3,935,000	791,000	20%
Year 9	4,278,972	4,315,000	1,068,000	25%
Year 10	4,625,811	4,671,000	2,013,000	43%
Overall	18,680,856	18,800,000	2,975,000	16%

6.2. Intervention in the chain-ladder technique

We now consider using a prior distribution to intervene in some of the parameters of the chain-ladder model, instead of using prior distributions with large variances that just reproduce the chain-ladder estimates. The implementation is set up in Section (ii) of the Appendix, and the program can be cut and pasted into winBUGS and run following steps 3 onwards, above.

We consider two cases, as discussed in Section 4. For the first case, we assume that there is information that implies that the second develop-

Table 5. Individual development factors

Accident Year	Development Year								
	2	3	4	5	6	7	8	9	10
1	3.143	1.543	1.278	1.238	1.209	1.044	1.04	1.063	1.018
2	3.511	1.755	1.545	1.133	1.084	1.128	1.057	1.086	
3	4.448	1.717	1.458	1.232	1.037	1.12	1.061		
4	4.568	1.547	1.712	1.073	1.087	1.047			
5	2.564	1.873	1.362	1.174	1.138				
6	3.366	1.636	1.369	1.236					
7	2.923	*1.878*	1.439						
8	3.953	*2.016*							
9	3.619								

ment factor (from Column 2 to Column 3) should be given the value 1.5 for rows 7, 8, 9, and 10, and that there is no indication that the other parameters should be treated differently from the standard chain-ladder technique. In order to implement this, the parameter for the second development factor for rows 7–10 is given a prior distribution with mean 1.5. We then look at two different choices for the prior variance for this parameter. Using a large variance means that the parameter is estimated separately from the other rows, but using the data without letting the prior mean influence it too greatly. We then use a standard deviation of 0.1 for the prior distribution, so that the prior mean has a greater influence.

We consider first the estimate of the second development factor. The chain ladder estimate is 1.7473 and the individual development factors for the triangle are shown in Table 5. The rows for the second development factor that are modeled separately are shown in italics. The estimate using the Bayesian models is 1.68 for rows 1–6. When a large variance is used for the prior distribution of the development factor for rows 7–10, the estimate using the Bayesian model is 1.971. With the smaller variance for this prior distribution, the estimate is 1.673, and has been drawn down towards the prior mean of 1.5. This clearly shows how the prior distributions can be used to influence the parameter estimates.

The effect on the reserve estimates is shown in Table 6, which compares the reserves and prediction errors for the two cases outlined above with the results for the chain-ladder model (which could be produced using the program in Section 6.1 on this set of data). The chain-ladder figures are slightly different from those given in Table 4 because this is a simulation method.

It is interesting to note that, in this case, the intervention has not had a marked effect on the prediction errors (in percentage terms). However, the prediction errors themselves have changed considerably, and this indicates that it is important to think of the prediction error as a percentage of the prediction. Other prior distributions could have a greater effect on the percentage prediction error.

The second case we consider is when we use only the most recent data for the estimation of each development factor. For the last three development factors, all the data is used because there is no more than three years for each. For the other development factors, only the three most recent years are used. The estimates of the development factors are shown in Table 7. The estimates of the first development factor are not affected by the change in the model (the small differences could be due to simulation error or the changes elsewhere). For the other development factors, the estimates can be seen to be affected by the model assumptions.

The effect of using only the latest three years in the estimation of the development factors in

Table 6. Reserves and prediction errors for the chain-ladder and Bayesian models

	Chain-Ladder		Large Variance		Small Variance	
	Reserve	Prediction Error	Reserve	Prediction Error	Reserve	Prediction Error
Year 2	97,910	115%	95,920	116%	95,380	117%
Year 3	471,200	46%	475,700	46%	470,500	47%
Year 4	711,100	38%	721,700	37%	714,400	37%
Year 5	989,200	31%	996,800	31%	994,700	31%
Year 6	1,424,000	27%	1,429,000	26%	1,428,000	27%
Year 7	2,187,000	23%	2,196,000	23%	2,185,000	23%
Year 8	3,930,000	20%	3,937,000	20%	3,932,000	20%
Year 9	4,307,000	24%	4,998,000	27%	4,044,000	25%
Year 10	4,674,000	43%	5,337,000	44%	4,496,000	43%
Overall	18,790,000	16%	20,190,000	17%	18,360,000	16%

Table 7. Development factors using three most recent years' data separately

Accident Year	Development Year 2	3	4	5	6	7	8	9	10
1	3.143	1.543	1.278	1.238	1.209	1.044	1.04	1.063	1.018
2	3.511	1.755	1.545	1.133	1.084	1.128	1.057	1.086	
3	4.448	1.717	1.458	1.232	1.037	1.12	1.061		
4	4.568	1.547	1.712	1.073	1.087	1.047			
5	2.564	1.873	1.362	1.174	1.138				
6	3.366	1.636	1.369	1.236					
7	2.923	1.878	1.439						
8	3.953	2.016							
9	3.619								
Earlier rows	3.575	1.688	1.513	1.197	1.139	1.045			
Recent rows	3.579	1.852	1.393	1.155	1.085	1.099	1.054	1.076	1.018
All rows	3.527	1.751	1.46	1.175	1.104	1.087	1.054	1.076	1.018

the forecasting of outstanding claims can be seen in Table 8.

In this case, the effect on the reserves is not particularly great. The prediction errors have increased for most years, although the effect is not great on these either. The importance of the Bayesian method is to actually be able to assess the effect of using different sets of data on the uncertainty of the outcome.

6.3. The Bornhuetter-Ferguson method

In this section, we consider intervention on the level of each row, using the Bornhuetter-Ferguson method. We consider two examples. The first uses small variances for the prior distributions of the row parameters, thus reproducing the Bornhuetter-Ferguson method. The second example uses less strong prior information, and produces results that lie between the Bornhuetter-Ferguson method and the chain-ladder technique. We use the negative binomial model for the data that was described in Section 5, and the winBUGS code for this is given in the Appendix,

Table 8. Reserve estimates using three most recent years' data

	Chain-Ladder		Bayesian Model	
	Reserve	Prediction Error	Reserve	Prediction Error
Year 2	97,910	115%	94,860	115%
Year 3	471,200	46%	469,300	46%
Year 4	711,100	38%	712,900	37%
Year 5	989,200	31%	1,042,000	30%
Year 6	1,424,000	27%	1,393,000	27%
Year 7	2,187,000	23%	2,058,000	24%
Year 8	3,930,000	20%	3,468,000	22%
Year 9	4,307,000	24%	4,230,000	27%
Year 10	4,674,000	43%	4,711,000	47%
Overall	18,790,000	16%	18,180,000	18%

Table 9. Negative binomial model: Bayesian model with precise priors for all rows: mean and prediction error of reserves

	Bayesian Mean Reserve	Bayesian Prediction Error	Bayesian Prediction Error %	Bornhuetter-Ferguson Reserve
Year 2	95,680	111,100	116%	95,788
Year 3	482,500	211,900	44%	480,088
Year 4	736,400	250,100	34%	736,708
Year 5	1,118,000	296,500	27%	1,114,999
Year 6	1,533,000	339,700	22%	1,527,444
Year 7	2,305,000	410,300	18%	2,308,139
Year 8	3,474,000	497,500	14%	3,466,839
Year 9	4,547,000	555,000	12%	4,550,270
Year 10	5,587,000	610,900	11%	5,584,677
Overall	19,880,000	1,854,000	9%	19,864,951

section (i). Section 6.1 used this method with large variances for the prior, thereby reproducing the chain-ladder technique.

First we consider the Bornhuetter-Ferguson method, exactly as it is usually applied. For this, we begin by using prior distributions for the row parameters, which all have standard deviation 1,000 (which is small compared with the means), and whose means are:

x_2	x_3	x_4	x_5
5,500,000	5,500,000	5,500,000	5,500,000
x_6	x_7	x_8	x_9
5,500,000	6,000,000	6,000,000	6,000,000
x_{10}			
6,000,000			

In order to implement this, using the code in the Appendix, it is necessary to change the "DATA" section of the program (just before the "INITIAL VALUES" section). It is explained in the program exactly what changes to make.

The Bornhuetter-Ferguson estimates of outstanding losses, and the results from the Bayesian model are shown in Table 9.

In this case, it can be seen that the results are very close to those of the Bornhuetter-Ferguson technique. Thus, if it is desired to use the Bornhuetter-Ferguson method within this stochastic framework, this is the approach that should be used. The added information available is the prediction errors. Further, it is possible to generate predictive distributions rather than just the mean and prediction error.

The Bornhuetter-Ferguson technique assumes that there is strong prior information about the row parameters, so that the standard deviations of the prior distributions used in this example are small. The other end of the spectrum is constituted by the chain-ladder technique, when large standard deviations are used for the prior distributions. Between these two extremes is a whole range of possible models, which can be specified by using different standard deviations. We now illustrate the results when less strongly informative prior distributions are used for the row parameters. We use the same prior means as above, but this time use a standard deviation of 1,000,000. We are incorporating prior belief about the ultimate losses for each year, but allowing for uncertainty in this information. The associated reserve results are shown in Table 10. Notice that the reserves are between the chain-ladder and Bornhuetter-Ferguson results. Notice also that the precision of the prior has influenced the prediction errors, but to a lesser extent. This provides an extra level of flexibility, allowing for the choice of a range of models in a continuous spectrum between the chain-ladder technique and Bornhuetter-Ferguson.

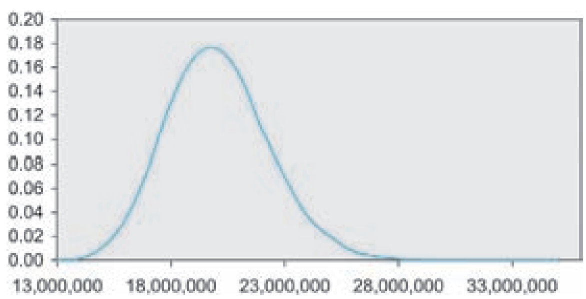
7. Conclusions

This paper has shown how expert opinion, separate from the reserving data, can be incorporated into the prediction intervals for a stochastic model. The advantages of a stochastic approach are that statistics associated with the predictive distribution are also available, rather than just a point estimate. In fact, it is possible to produce the full predictive distribution, rather than just

Table 10. Negative binomial model: Bayesian model with informative priors: mean and prediction error of reserves

	Bayesian Mean Reserve	Bayesian Prediction Error	Bayesian Prediction Error	Bornhuetter-Ferguson Reserve	Chain-Ladder Reserve
Year 2	94,660	111,500	118%	95,788	94,634
Year 3	470,400	218,800	47%	480,088	469,511
Year 4	717,100	265,900	37%	736,708	709,638
Year 5	994,900	308,900	31%	1,114,999	984,889
Year 6	1,431,000	376,800	26%	1,527,444	1,419,459
Year 7	2,198,000	488,900	22%	2,308,139	2,177,641
Year 8	3,839,000	727,200	19%	3,466,839	3,920,301
Year 9	4,417,000	865,500	20%	4,550,270	4,278,972
Year 10	5,390,000	1,080,000	20%	5,584,677	4,625,811
Overall	19,550,000	2,252,000	12%	19,864,951	18,680,856

Figure 1. Distribution of reserve for Bornhuetter-Ferguson estimation



the first two moments. As was emphasized by England and Verrall [8], the full predictive distribution contains a lot more information than just its mean and standard deviation, and it is a great advantage to be able to look at this distribution. As an illustration of this, Figure 1 shows the predictive distribution of outstanding losses for the final example considered above, in Section 6.3, Table 10.

A further possibility for including expert knowledge within a stochastic framework applies when the Bornhuetter-Ferguson technique is used. This is an adaptation of the method used in Sections 5 and 6.3, whereby the reserve is specified rather than the ultimate losses, u_i. The reserve value can be used to infer a value for u_i, from which the stochastic version of the Bornhuetter-Ferguson method can be applied.

We have concentrated on two important situations that we believe are most common when expert opinion is used. However, the same approach could also be taken in other situations and for other modeling methods, such as the Hoerl curve. This would allow us to add tail factors to the models considered in this paper. This paper has been more concerned with the general approach rather than specific reserving methods. However, we acknowledge that methods based on the chain-ladder setup are very commonly used and we hope that, by using this framework, we enable actuaries to appreciate the suggestions made in this paper, and to experiment with the programs supplied.

References

[1] Bornhuetter, R. L. and R. E. Ferguson, "The Actuary and IBNR," *Proceedings of the Casualty Actuarial Society* 59, 1972, pp. 181–195.

[2] Congdon, Peter, *Applied Bayesian Modelling*, Somerset, NJ: Wiley, 2001.

[3] Congdon, Peter, *Bayesian Statistical Modelling*, Somerset, NJ: Wiley, 2001.

[4] de Alba, Enrique, "Bayesian Estimation of Outstanding Claims Reserves," *North American Actuarial Journal* 6:4, 2002, pp. 1–20.

[5] England, P. D., Addendum to "Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving," *Insurance: Mathematics and Economics* 31, 2002, pp. 461–466.

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

[7] England, P. D. and R. J. Verrall, "A Flexible Framework for Stochastic Claims Reserving," *Proceedings of the Casualty Actuarial Society* 88, 2001, pp. 1–38.

[8] England, P. D. and R. J. Verrall, "Stochastic Claims Reserving in General Insurance" (with discussion), *British Actuarial Journal* 8, 2002, pp. 443–554.

[9] Kremer, E., "IBNR Claims and the Two Way Model of ANOVA," *Scandinavian Actuarial Journal* 1, 1982, pp. 47–55.

[10] Mack, T., "Credible Claims Reserves: The Benktander Method," *ASTIN Bulletin* 30, 2000, pp. 333–347.

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

[12] McCullagh, P. and J. Nelder. *Generalized Linear Models* (2nd edition), London: Chapman and Hall, 1989.

[13] Ntzoufras, I. and P. Dellaportas, "Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty," *North American Actuarial Journal* 6:1, 2002, pp. 113–128.

[14] Renshaw, A. E. and R. J. Verrall, "A Stochastic Model Underlying the Chain-Ladder Technique," *British Actuarial Journal* 4, 1998, pp. 903–923.

[15] Scollnik, D. P. M., "Actuarial Modeling with MCMC and BUGS," *North American Actuarial Journal* 5:2, 2001, pp. 96–124.

[16] Spiegelhalter, D. J., A. Thomas, N. G. Best, and W. R. Gilks, *BUGS 0.5: Baysian Inference using Gibbs Sampling*, MRC Biostatics Unit, Cambridge, UK, 1996.

[17] Taylor, G. C., *Loss Reserving: An Actuarial Perspective*, Norwell, MA: Kluwer, 2000.

[18] Taylor, G. C. and F. R. Ashe, "Second Moments of Estimates of Outstanding Claims," *Journal of Econometrics* 23, 1983, pp. 37–61.

[19] Verrall, R. J., "A Bayesian Generalised Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving," *North American Actuarial Journal* 8:3, 2004, pp. 67–89.

[20] Verrall, R. J., "An Investigation Into Stochastic Claims Reserving Models and the Chain-Ladder Technique," *Insurance: Mathematics and Economics* 26, 2000, pp. 91–99.

[21] Verrall, R. J. and P. D. England, "Incorporating Expert Opinion into a Stochastic Model for the Chain-Ladder Technique," *Insurance: Mathematics and Economics* 37, 2005, pp. 355–370.

Appendix

The code for winBUGS is shown below for the models used in Section 6. This is available from the author on request and can be cut and pasted directly into winBUGS. Anything to the right of "#" is ignored, so the code can be changed by adding and removing this at the start of a line.

(i) This section contains the code for the Bornhuetter-Ferguson method in Section 5, which was used for the illustrations in Sections 6.1 and 6.3.

```
model
{
#Model for Data
        for( i in 1 : 45 ) {
        Z[i]<-Y[i]/1000
        pC[i]<-D[i]/1000
#Zeros trick
        zeros[i]<-0
        zeros[i]~dpois(phi[i])
        phi[i]<-(-pC[i]*log(1/(1+g[row[i]]))-Z[i]*log(g[row[i]]/(1+g[row[i]])))/scale
                }
#Cumulate down the columns:
        DD[3]<-DD[1]+Y[46]
        for( i in 1 : 2 ) {DD[4+i]<-DD[4+i-3]+Y[49+i-3]}
        for( i in 1 : 3 ) {DD[7+i]<-DD[7+i-4]+Y[52+i-4]}
        for( i in 1 : 4 ) {DD[11+i]<-DD[11+i-5]+Y[56+i-5]}
```

```
        for( i in 1 : 5 ) {DD[16+i]<-DD[16+i-6]+Y[61+i-6]}
        for( i in 1 : 6 ) {DD[22+i]<-DD[22+i-7]+Y[67+i-7]}
        for( i in 1 : 7 ) {DD[29+i]<-DD[29+i-8]+Y[74+i-8]}
        for( i in 1 : 8 ) {DD[37+i]<-DD[37+i-9]+Y[82+i-9]}

#Needed for the denominator in definition of gammas
        E[3]<-E[1]*gamma[1]
        for( i in 1 : 2 ) {E[4+i]<-E[4+i-3]*gamma[2]}
        for( i in 1 : 3 ) {E[7+i]<-E[7+i-4]*gamma[3]}
        for( i in 1 : 4 ) {E[11+i]<-E[11+i-5]*gamma[4]}
        for( i in 1 : 5 ) {E[16+i]<-E[16+i-6]*gamma[5]}
        for( i in 1 : 6 ) {E[22+i]<-E[22+i-7]*gamma[6]}
        for( i in 1 : 7 ) {E[29+i]<-E[29+i-8]*gamma[7]}
        for( i in 1 : 8 ) {E[37+i]<-E[37+i-9]*gamma[8]}

        EC[1]<-E[1]/1000
        EC[2]<-sum(E[2:3])/1000
        EC[3]<-sum(E[4:6])/1000
        EC[4]<-sum(E[7:10])/1000
        EC[5]<-sum(E[11:15])/1000
        EC[6]<-sum(E[16:21])/1000
        EC[7]<-sum(E[22:28])/1000
        EC[8]<-sum(E[29:36])/1000
        EC[9]<-sum(E[37:45])/1000

#Model for future observations
        for( i in 46 : 90 ) {
                a1[i]<-max(0.01,a[row[i]]*DD[i-45]/(1000*scale))
                b1[i]<-1/(gamma[row[i]]*1000*scale)
                Z[i]~dgamma(a1[i],b1[i])
                Y[i]<-Z[i]
                fit[i]<-Y[i]
                        }
scale<-52.8615
#Convert row parameters to gamma using (5.6)
        for (k in 1:9) {
                gamma[k]<-1+g[k]
                g[k]<-u[k]/EC[k]
                a[k]<-g[k]/gamma[k]
                        }
#Prior distributions for row parameters.
        for (k in 1:9) {
                        u[k]~dgamma(au[k],bu[k])
```

```
                    au[k]<-bu[k]*(ultm[k+1]*(1-1/f[k]))
                    bu[k]<-(ultm[k+1]*(1-1/f[k]))/pow(ultsd[k+1],2)
                        }
#The prior distribution can be changed by changing the data input values for the
#vectors ultm and ultsd

#Row totals and overall reserve
    R[1]<-0
    R[2]<-fit[46]
    R[3]<-sum(fit[47:48])
    R[4]<-sum(fit[49:51])
    R[5]<-sum(fit[52:55])
    R[6]<-sum(fit[56:60])
    R[7]<-sum(fit[61:66])
    R[8]<-sum(fit[67:73])
    R[9]<-sum(fit[74:81])
    R[10]<-sum(fit[82:90])
    Total<-sum(R[2:10])
    }

#DATA
list(
row=c(1,1,1,1,1,1,1,1,1,
2,2,2,2,2,2,2,2,
3,3,3,3,3,3,3,4,4,
4,4,4,4,5,5,5,5,5,
6,6,6,6,7,7,7,8,
8,9,1,2,2,3,3,3,4,4,4,
4,5,5,5,5,5,6,6,6,6,6,6,
7,7,7,7,7,7,7,8,8,8,8,8,
8,8,8,9,9,9,9,9,9,9,9,
9),
Y=c(352118,884021,933894,1183289,445745,320996,527804,266172,425046,
290507,1001799,926219,1016654,750816,146923,495992,280405,
310608,1108250,776189,1562400,272482,352053,206286,
443160,693190,991983,769488,504851,470639,
396132,937085,847498,805037,705960,
440832,847631,1131398,1063269,
359480,1061648,1443370,
376686,986608,
344014,
NA,
NA,NA,
```

```
NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA),

D=c(357848,766940,610542,482940,527326,574398,146342,139950,227229,
709966,1650961,1544436,1666229,973071,895394,674146,406122,
1000473,2652760,2470655,2682883,1723887,1042317,1170138,
1311081,3761010,3246844,4245283,1996369,1394370,
1754241,4454200,4238827,5014771,2501220,
2150373,5391285,5086325,5819808,
2591205,6238916,6217723,
2950685,7300564,
3327371,
NA,
NA,NA,
NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA),

DD=c(67948,
652275,NA,
686527,NA,NA,
1376424,NA,NA,NA,
1865009,NA,NA,NA,NA,
3207180,NA,NA,NA,NA,NA,
6883077,NA,NA,NA,NA,NA,NA,
7661093,NA,NA,NA,NA,NA,NA,NA,
8287172,NA,NA,NA,NA,NA,NA,NA,NA),

E=c(67948,
652275,NA,
686527,NA,NA,
1376424,NA,NA,NA,
1865009,NA,NA,NA,NA,
3207180,NA,NA,NA,NA,NA,
```

```
6883077,NA,NA,NA,NA,NA,NA,
7661093,NA,NA,NA,NA,NA,NA,NA,
8287172,NA,NA,NA,NA,NA,NA,NA,NA),

f=c(1.017724725, 1.095636823, 1.154663551, 1.254275641, 1.384498969,
1.625196481, 2.368582213, 4.138701016, 14.44657687),
ultm=c(NA,5500, 5500, 5500, 5500, 5500, 6000, 6000, 6000, 6000),
ultsd=c(NA,10000,10000,10000,10000,10000,10000,10000,10000,10000))
```

These values for the ultsd will give the chain-ladder results. To obtain the Bornhuetter-Ferguson results, replace the last line with the following line:
ultsd=c(NA,1,1,1,1,1,1,1,1,1))
The other illustration in section 6.3 uses:
ultsd=c(NA,1000,1000,1000,1000,1000,1000,1000,1000,1000))

```
#INITIAL VALUES
list(u=c(5500, 5500, 5500, 5500, 5500, 6000, 6000, 6000, 6000),
Z=c(NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,
NA,NA,
NA,
0,
0,0,
0,0,0,
0,0,0,0,
0,0,0,0,0,
0,0,0,0,0,0,
0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0))
```

(ii) Code for the model in section 4, which was used for the illustrations in section 6.2.

```
model
{
#Model for data:
	for( i in 1 : 45 ) {
		Z[i]<-Y[i]/(scale*1000)
```

```
            pC[i]<-D[i]/(scale*1000)
            C[i]<-Z[i]+pC[i]
        zeros[i]<-0
            zeros[i]~dpois(phi[i])
            phi[i]<-(loggam(Z[i]+1)+loggam(pC[i])-loggam(C[i])-
pC[i]*log(p1[row[i],col[i]])-Z[i]*log(1-p1[row[i],col[i]]))
            }
DD[3]<-DD[2]+Y[47]
        for( i in 1 : 2 ) {DD[4+i]<-DD[4+i-1]+Y[49+i-1]}
        for( i in 1 : 3 ) {DD[7+i]<-DD[7+i-1]+Y[52+i-1]}
        for( i in 1 : 4 ) {DD[11+i]<-DD[11+i-1]+Y[56+i-1]}
        for( i in 1 : 5 ) {DD[16+i]<-DD[16+i-1]+Y[61+i-1]}
        for( i in 1 : 6 ) {DD[22+i]<-DD[22+i-1]+Y[67+i-1]}
        for( i in 1 : 7 ) {DD[29+i]<-DD[29+i-1]+Y[74+i-1]}
        for( i in 1 : 8 ) {DD[37+i]<-DD[37+i-1]+Y[82+i-1]}

#Model for future observations
        for( i in 46 : 90 ) {
            a1[i]<-max(0.01,(1-p1[row[i],col[i]])*DD[i-45]/(1000*scale))
                b1[i]<-p1[row[i],col[i]]/(1000*scale)
                Z[i]~dgamma(a1[i],b1[i])
                Y[i]<-Z[i]
                    }
        scale<-52.8615

#Set up the parameters of the negative binomial model.
        for (k in 1:9) {
            p[k]<-1/lambda[k]
            lambda[k]<-exp(g[k])+1
            g[k]~dnorm(0.5,1.0E-6)
        }
#Choose one of the folllowing (1,2 or 3) and delete the "#" at the start of each line before running.

#1. Vague Priors: Chain-ladder model
#       for (j in 1:9) {
#       for (i in 1:10) {p1[i,j]<-p[j]}
#                   }

#2. Intervention in second development factor.
#       for (i in 1:10) {p1[i,1]<-p[1]}
#       for (i in 1:6) {p1[i,2]<-p[2]}
#       p1[7,2]<-p82
#       p1[8,2]<-p82
```

```
#         p1[9,2]<-p82
#         p1[10,2]<-p82
#         for (j in 3:9) {
#         for (i in 1:10) {p1[i,j]<-p[j]}
#                           }
#         lambda82<-g82+1
#         p82<-1/lambda82
#Use one of the following 2 lines:
#         g82~dgamma(0.005,0.01) #This is a prior with a large variance
#         g82~dgamma(25,50) #This is a prior with a small variance

#3. Using latest 3 years for estimation of development factors.
#         for (j in 1:6) {
#         for (i in 1:(7-j)) {p1[i,j]<-op[j]}
#         for (i in (8-j):10) {p1[i,j]<-p[j]}
#         }
#         for (j in 7:9) {
#         for (i in 1:10) {p1[i,j]<-p[j]}
#                   }
#         for (k in 1:6) {
#                   op[k]<-1/olambda[k]
#                   olambda[k]<-exp(og[k])+1
#                   og[k]~dnorm(0.5,1.0E-6)
#                           }

#Row totals and overall reserve
                    R[1]<-0
                    R[2]<-Y[46]
                    R[3]<-sum(Y[47:48])
                    R[4]<-sum(Y[49:51])
                    R[5]<-sum(Y[52:55])
                    R[6]<-sum(Y[56:60])
                    R[7]<-sum(Y[61:66])
                    R[8]<-sum(Y[67:73])
                    R[9]<-sum(Y[74:81])
                    R[10]<-sum(Y[82:90])
                    Total<-sum(R[2:10])
          }
#DATA
list(
row=c(1,1,1,1,1,1,1,1,1,
2,2,2,2,2,2,2,2,
3,3,3,3,3,3,3,4,4,
```

```
4,4,4,4,5,5,5,5,5,
6,6,6,6,7,7,7,8,
8,9,2,3,3,4,4,
4,5,5,5,5,6,6,6,6,6,
7,7,7,7,7,7,8,8,8,8,
8,8,8,9,9,9,9,9,9,9,
9,10,10,10,10,10,10,10,10,10),
col=c(1,2,3,4,5,6,7,8,9,
1,2,3,4,5,6,7,8,
1,2,3,4,5,6,7,1,2,3,
4,5,6,1,2,3,4,5,1,
2,3,4,1,2,3,1,
2,1,9,8,9,7,8,9,
6,7,8,9,5,6,7,8,9,4,
5,6,7,8,9,3,4,5,6,7,
8,9,2,3,4,5,6,7,8,9,
1,2,3,4,5,6,7,8,9),
Y=c(
766940,610542,482940,527326,574398,146342,139950,227229,67948,
884021,933894,1183289,445745,320996,527804,266172,425046,
1001799,926219,1016654,750816,146923,495992,280405,
1108250,776189,1562400,272482,352053,206286,
693190,991983,769488,504851,470639,
937085,847498,805037,705960,
847631,1131398,1063269,
1061648,1443370,
986608,
NA,
NA,NA,
NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA),
D=c(
357848,1124788,1735330,2218270,2745596,3319994,3466336,3606286,3833515,
352118,1236139,2170033,3353322,3799067,4120063,4647867,4914039,
290507,1292306,2218525,3235179,3985995,4132918,4628910,
310608,1418858,2195047,3757447,4029929,4381982,
443160,1136350,2128333,2897821,3402672,
396132,1333217,2180715,2985752,
```

```
440832,1288463,2419861,
359480,1421128,
376686,
NA,
NA,NA,
NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,NA),
DD=c(5339085,
4909315,NA,
4588268,NA,NA,
3873311,NA,NA,NA,
3691712,NA,NA,NA,NA,
3483130,NA,NA,NA,NA,NA,
2864498,NA,NA,NA,NA,NA,NA,
1363294,NA,NA,NA,NA,NA,NA,NA,
344014,NA,NA,NA,NA,NA,NA,NA,NA))

#INITIAL VALUES
This is what is used for 1.

For 2, replace the first line by
list(g=c(0,0,0,0,0,0,0,0,0), g82=0.5,

For 3, replace the first line by
list(g=c(0,0,0,0,0,0,0,0,0), og=c(0,0,0,0,0,0),

list(g=c(0,0,0,0,0,0,0,0,0),
Z=c(NA,NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,NA,
NA,NA,NA,NA,NA,
NA,NA,NA,NA,
NA,NA,NA,
NA,NA,
NA,
0,
0,0,
```

```
0,0,0,
0,0,0,0,
0,0,0,0,0,
0,0,0,0,0,0,
0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0))
```

Errata

"Obtaining Predictive Distributions for Reserves Which Incorporate Expert Opinion"
by R. J. Verrall, *Variance* vol. 1, no. 1, 2007.

Correction 1.
On page 63, in equation 5.4 for γ_i a summation sign in the numerator should be corrected to be a product sign:

$x_i\left(1-\frac{1}{\sum_{k=n-i+2}^{n}\lambda_k}\right)$ *SHOULD READ:* $x_i\left(1-\frac{1}{\prod_{k=n-i+2}^{n}\lambda_k}\right)$

Correction 2.
On page 63, in equation 5.4 for γ_i the upper limit on the first sum in the denominator should be corrected:

$\sum_{m=1}^{i-1}C_{m,n}$ *SHOULD READ:* $\sum_{m=1}^{i-1}C_{m,n-i+2}$

The corrected equation for γ_i is:

$$\gamma_i = 1+\frac{x_i\left(1-\frac{1}{\prod_{k=n-i+2}^{n}\lambda_k}\right)}{\sum_{m=1}^{i-1}C_{m,n-i+2}+\sum_{k=n-i+3}^{n}\left[\left(\prod_{l=n-k+2}^{i-1}\gamma_l\right)\sum_{m=1}^{n-k+1}C_{m,k}\right]} \qquad i=3,\ldots,n.$$