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.

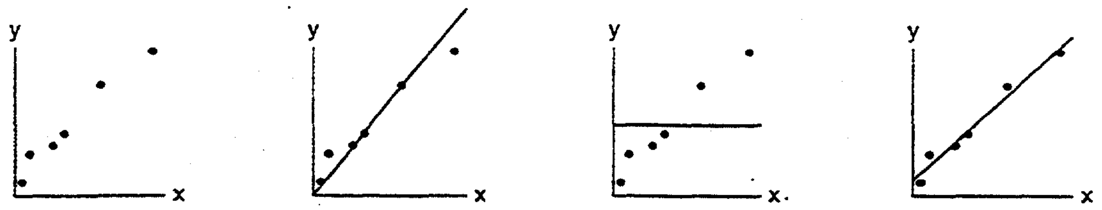
The link ratio method The traditional link ratio method estimates y as $L(x) = cx$, where c is a "selected link ratio". The value of c is chosen after a review of the observed link ratios from previous years—as an average of several years, perhaps, or as a weighted average. The choice is not easy in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



$L(x) = cx$ $L(x) = k$ $L(x) = a + bx$

a. Raw data b. Link ratios c. Budgeted loss d. Least squares

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.

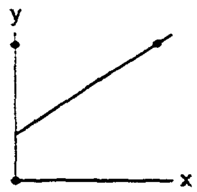


Figure 2: The simple model

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 $b = 0.968$, $a = 6{,}023$, $\bar{x} = 21{,}139$, and $\bar{y} = 26{,}482$. If we set $d = \bar{x}/\bar{y} = 0.798$, 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

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

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.

	Reported Loss Ratio					
AY	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.

	Reported Loss Ratio					
AY	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.

		Ultimate	
AY	EP	Loss Ratio	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.