DISCUSSION OF PAPER PUBLISHED IN VOLUME LXXI

EXTRAPOLATING, SMOOTHING, AND INTERPOLATING DEVELOPMENT FACTORS

RICHARD E. SHERMAN

VOLUME LXXI

DISCUSSION BY STEPHEN P. LOWE AND DAVID F. MOHRMAN

Mr. Sherman's paper presents a potpourri of practical applications involving the fitting of a parametric equation to loss development factor data. The particular equation utilized is called the inverse power curve, the form of which is

$$f(t) = 1 + \frac{a}{(t + c)^b} \tag{1}$$

where a, b and c are parameters to be estimated, and t represents time as it relates to the maturity of the body of claims.

It should be readily seen that the parameter c provides a linear transformation of the time variable t, and is therefore somewhat extraneous to the formulation. The definition of t is arbitrary; $f(t)$ can be the development factor from t to $t + 1$, or alternatively $f(t)$ can be the development factor from $t - 1$ to t. Similarly, the beginning of the accident year can be $t = 0$ or $t = 1$ (or even $t = -1$ or $t = 1.7275$).

The above comment is not intended to suggest that the selection of the time scale embodied in the variable t is trivial; a different result will be obtained for each scale chosen. However, to simplify discussion, we can express Mr. Sherman's equation as

$$f(t) = 1 + \frac{a}{t^b} \tag{2}$$

where we are searching for the best a, b, <u>and</u> scale t that fits the data.

Like the author, these reviewers have found it useful in many circumstances to fit parametric equations to incomplete, erratic or irregular loss development data. This review will expand slightly on Mr. Sherman's paper by offering some alternative equations, and discussing some desirable characteristics for loss development models of this kind. In addition, we will offer some specific comments and point out some pitfalls associated with Mr. Sherman's approach.

ALTERNATIVE MODELS

The parametric equation in (2) above is referred to by the author as the inverse power curve. We refer to this equation as the polynomial decay model. As the author points out in Section II, this equation has the property that the initial development, a, decays at a rate of

$$1 - \left(1 - \frac{1}{t}\right)^{b} \tag{3}$$

over the interval from $(t - 1)$ to t. For example, if $b = 1$, then the following decay rates would apply.

t	Development from $t - 1$ to t	Rate of Decay
1	$1 + a$	
2	$1 + a \times \frac{1}{2}$	50%
3	$1 + a \times \frac{1}{2} \times \frac{2}{3}$	33%
4	$1 + a \times \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4}$	25%
5	$1 + a \times \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} \times \frac{4}{5}$	20%

An alternative model to the polynomial decay is one involving exponential decay:

$$f(t) = 1 + \frac{a}{b^t} \tag{4}$$

In this model the initial development, a, decays at a constant rate, $1 - b$, over each interval.

Viewing loss development as a decay process is intuitively appealing. It is certainly reasonable that, as an ever increasing proportion of losses are paid, their propensity to develop must decline.

Both the polynomial and the exponential decay models can be expanded by the addition of a third parameter involving a squared term.

$$f(t) = 1 + \frac{a}{t^b} + \frac{c}{t^{b^2}} \tag{5}$$

$$f(t) = 1 + \frac{a}{b^t} + \frac{c}{b^{t^2}} \tag{6}$$

There are also a variety of mixed models that might prove useful.

$$f(t) = 1 + \frac{at^c}{b^t} \tag{7}$$

$$f(t) = 1 + \frac{a + ct}{t^b} \tag{8}$$

All of these six models have been used by the reviewers to fit emergence data of one form or another.

Equations (5) and (6) are interesting as they can be conceptualized as modelling two different kinds of development taking place simultaneously, but decaying at different rates. For example, if the data were accident year reported losses, the a term might represent development caused by newly reported claims, while the c term might represent development on existing claims.

A specific instance where this approach is very useful is in the case of subrogation and salvage. The following table compares actual loss development factors for Auto Physical Damage to those obtained using the three parameter polynomial decay model, presented in equation (5).

Year of Development	Actual Development	Model Development
2:1	1.240	1.240
3:2	.993	.992
4:3	.996	.997
5:4	.998	.999
6:5	.999	.999
7:6	1.000	1.000

In this instance the parameters of the model are $a = -.07$, $b = 3$ and $c = .31$. In this instance the model has a nice intuitive appeal. The positive development of losses embodied by the c term decays very quickly, leaving the slower negative development of subrogation and salvage embodied by the a term.

CHOOSING A MODEL

As we have noted, all of the models described previously have proven useful in fitting various kinds of emergence data. We suspect that the reader could easily conjure up other models that would also prove useful.

Each of the models that we have described is "well-behaved", but only over a limited range of parameter values. It is worthwhile to consider what kinds of constraints on the parameters are necessary for a model to be reasonable.

In traditional applications, we want the development factors to be positive, decreasing, and approaching one. These can be expressed mathematically as

1. $f(t) \geq 1 \qquad \lim_{t \to \infty} f(t) = 1$

2. $f'(t) < 0 \qquad \lim_{t \to \infty} f'(t) = 0$

3. $f''(t) > 0 \qquad \lim_{t \to \infty} f''(t) = 0$

While the constraints on the limits are probably necessary in all situations, special circumstances may require the relaxing of one or more of the constraints on values of $f(t)$, $f'(t)$ or $f''(t)$. For example, to produce the auto physical damage factors cited earlier, it was necessary to violate the first constraint. Similarly, the third constraint restricts us to curves that are concave upward over the entire domain of t. In some instances a curve that starts out concave downward may be desired.

For Sherman's two parameter polynomial decay model

$$f'(t) = \frac{-ba}{t^{b+1}}$$

$$f''(t) = \frac{b(b+1)a}{t^{b+2}}$$

We see that all conditions are satisfied when $a > 0$ and $b > 0$ (and $t > 0$).

For the two parameter exponential decay model

$$f'(t) = \frac{a}{b^t}\left(\ln \frac{1}{b}\right)$$

$$f''(t) = \frac{a}{b^t}\left(\ln\frac{1}{b}\right)^2$$

Here all conditions are satisfied when $a > 0$ and $b > 1$.

Similar calculations to these should be performed on any proposed model before its use, so that a clear understanding of the properties and limitations of the model is obtained.

A much more critical property of any model used to estimate report-to-report development factors is whether the product of the infinite series converges. While arbitrary truncation of the series at some point (such as 80 years) may be acceptable from a practical standpoint, it would be more desirable to restrict the model by requiring that it produces a less-than-infinite development factor to ultimate.

Unfortunately, testing for convergence of the product of an infinite series is often difficult, as it usually involves intractable series of logarithms.

Such is the case with Mr. Sherman's equation. Several quick attempts failed to produce an algebraic solution to the question of whether the product series converges for all values of a and b, or some limited set. The reviewers are, however, convinced that with further effort (perhaps by someone more adept at real variable analysis) a solution to this question is obtainable.

Our investigations did lead us to the following conclusion, however. Consider the following hypothetical loss development data.

Maturity(t)	Reported Losses	Report-to-Report Development Factor
1	\$ 100	2.000
2	200	1.500
3	300	1.333
4	400	1.250
5	300	1.200
6	600	1.167
7	700	1.143
8	800	1.125
9	900	1.111
10	1,000	.
.	.	.
.	.	.

The reader should readily recognize that if the loss development continues at its present rate of \$100 per interval, losses will be infinite. It follows that the loss development factor product series must not converge.

However, it is also true that the development factors above can be produced identically using Mr. Sherman's equation by setting both a and b equal to one. This strongly suggests that the parameter b should be restricted to values greater than one in order to guarantee convergence.

We were led to raise the question of convergence by the discussion in Section II of Mr. Sherman's paper. In that section he derives the rate of decay for his model and points out that the rate of decay (as we have defined it) declines towards zero as t increases. (His "decay ratio" approaches unity.) This is in strong contrast to the exponential decay model, under which the rate of decay is constant for all values of t.

Upon initial reading of this section of the paper we were concerned that a declining rate of decay implied non-convergence of the ultimate development factor. However, upon reflection this does not appear to be the case.

It seems reasonable that there should be a relationship between the rate of decay of the development and the convergence or non-convergence of the development factor to ultimate. Clearly this question should be resolved before any model gains widespread use.

FITTING THE FUNCTION TO ACTUAL DATA

In Section I of his paper, Mr. Sherman suggests a simple procedure for fitting his equation to loss development factor data. The technique uses only natural logarithms, exponentials and linear regression, and therefore has the distinct advantage of requiring only a (reasonably sophisticated) pocket calculator to perform the calculations.

While the technique is handy, any prospective user should be aware that it does suffer from several problems. First, under the proposed transformation, an actual loss development factor of 1.000 is inadmissable because the natural logarithm of zero is undefined. What does one do under such circumstances? One possibility is to substitute a factor "sufficiently" close to 1.000.

A similar problem exists with observed development factors less than 1.000. These must be ignored or somehow smoothed out of the data.

Another problem is that the fitting technique minimizes the errors of $\ln(f(t) - 1)$ and not the errors of $f(t)$. The result is that, in the fitting process, differences between actual and fitted values are more significant when the development factors are close to 1.000 than when the development factors are significantly greater than 1.000. This bias in the errors is not necessarily bad; it simply needs to be understood as a part of the fitting process.

A related problem is that, since the measured errors are of the logs of the factors rather than the factors themselves, the coefficient of determination that results directly from the computation is inaccurate, and usually overstates the goodness of the fit.

For example, the coefficient of determination of the fit presented in Exhibit I is described as .99887. This is the coefficient of determination of a straight line through column (4) and not the coefficient of determination of the inverse power curve through column (2). This latter coefficient of determination is .971, which is still good, but less favorable than the author suggests (especially considering that there are only three data points being fit).

Obviously, the proper measurement of errors, and the decision as to what errors to minimize is key to any curve fitting procedure.

A particular problem with fitting Mr. Sherman's inverse power curve (or any of the other alternative curves that we have proposed) to the report-to-report development factors is that the resulting fitted factors will be multiplied together, compounding the errors. This can be a particular problem when the errors are not random. In such cases a significant error in the development factors to ultimate can accumulate.

For example, in Section II of his paper, Mr. Sherman uses his model to extrapolate general liability report-to-report development factors, using only the first few development factors to obtain the equation's parameters. While expressing some caution about the reliability of the resulting factors, the author does suggest that the extrapolated report-to-report development factors compare relatively favorably when compared to the actual factors over each interval.

The comparison is considerably less favorable if one compares the compounded, rather than the report-to-report, development factors. The errors in the IBNR reserve that would result from using the extrapolated factors range from 16% (1.667 vs. 1.575), to 112% (1.495 versus 1.234).

Year of Development	Extrapolated Development Factors Based On			Actual Factors
	First 2 Factors	First 3 Factors	First 4 Factors	
3 to 15	1.667			1.575
4 to 15	1.455	1.670		1.329
5 to 15	1.331	1.495	1.322	1.234

An alternative fitting approach that avoids the compounding of errors would be to fit the curve that results from compounding the factors to the actual loss emergence data, measuring the errors between actual and fitted losses reported at each valuation point. In essence, this alternative approach "dollar weights" the fitted factors.

An outline of this approach can be stated as follows. Minimize

$$\sum_{\forall(p,t)} (\hat{L}_{(p,t)} - L_{(p,t)})^2$$

Where $L_{(p,t)}$ is a valid point in the loss triangle, with p representing the exposure period of the losses (accident year, for example) and t representing the valuation point; and

$$\hat{L}_{(p,t)} = L^*_{(p,t^*)} \frac{\prod_{v=1}^{t} f_v}{\prod_{v=1}^{t^*} f_v}$$

where $L^*_{(p,t^*)}$ is some base value for the accident year in question at some time t^* (e.g., the latest valuation point), and f is the chosen decay model.

The problem so stated can be solved using partial derivatives and non-linear programming techniques.

CONCLUSION

Mr. Sherman's paper provides an excellent introduction to a timely topic. The paper presents practical ideas and approaches for the solution of problems encountered with increasing regularity in reserve analysis: incomplete, immature or fluctuating loss development data. We wholeheartedly agree with the author that the fitting of loss development data to curves such as the inverse power function often provides a practical solution to these problems.