A SIMULATION TEST OF PREDICTION ERRORS OF LOSS RESERVE ESTIMATION TECHNIQUES

JAMES N. STANARD

Abstract

This paper uses a computer simulation model to measure the expected value and variance of prediction errors of four simple methods of estimating loss reserves. Two of these methods are new to the *Proceedings*. The simulated data triangles that are tested are meant to represent sample sizes typically found in individual risk rating situations.

The results indicate that the commonly used age-to-age factor approach gives biased estimates and is inferior to the three other methods tested. Theoretical arguments for the source of this bias and a comparison of two of the methods are presented in the Appendices.

I. INTRODUCTION

The purpose of this paper is to measure the expected value and variance of prediction errors of four simple methods of estimating loss reserves. This is done by using a computer simulation model to generate several thousand different sets of known loss data, applying each estimation method to predict ultimate losses, and then calculating the difference between the predicted and the actual (simulated) ultimate values.[1,2]

Various reserve estimation techniques based on accident year data triangles are described in [2], [5], [6], [7], [20], and [21]. [21] contains a very extensive bibliography. However, the only paper to test the efficiency of the technique it proposes is [6] (and a sample size of only 50 iterations was used).

---

[1] The expected value of the prediction error is referred to as the "bias"; the bias and variance of the prediction error are together referred to as the "efficiency" of the estimation technique.

[2] Results from a previous version of this simulation model were described in [19]. The new computer model is written in Forth and assembly language on an IBM-PC, and is over twenty times faster than the old version written in APL on an IBM 5110 (each iteration now takes 11 to 15 seconds). This allows many more iterations and, therefore, much higher precision in measurements of bias.

The simulated data triangles that are tested here are meant to represent an amount of data that is typically found in individual risk rating, either self insurance programs, or working excess reinsurance treaties (expected values of 40 claims per year and \$10,400 per claim). For projecting loss reserves on much larger amounts of data, the statistical variations that are measured with this model will obviously be much less important.

II. AN OVERVIEW OF THE MODEL

View the loss process as follows: a given insured's losses during an accident year, a, are random variables drawn from some probability distribution determined by a vector of parameters, $\underline{\theta}_a$. Let $\underline{\theta}$ represent a vector (of vectors) containing all the parameters from the first accident year of the experience period through the latest year under consideration (denoted y). So

$$\underline{\theta} = (\underline{\theta}_1,\ldots,\underline{\theta}_y).$$

Let $\underline{K}$ be a vector representing the insured's known loss experience during the experience period.[3] $\underline{K}$ is a random sample drawn from the distributions determined by $\underline{\theta}$.

Let the ultimate losses that a particular insured will have for accident year a be a random variable L_a. The loss reserving and rate making processes both seek to find the "best" estimate of $E(L_a)$.[4,5] $E(L_a)$ is some function of the $\underline{\theta}_a$, whereas the experience $\underline{K}$ was drawn from distributions determined by $\underline{\theta}_1,\ldots,\underline{\theta}_y$. In order for $\underline{K}$ to be useful in estimating $E(L_a)$, there must be some relationship between the $\underline{\theta}$'s for different accident years.

The simplest assumption would be that $\underline{\theta}_1 = \ldots = \underline{\theta}_y$, that is that an insured's loss potential is constant over the experience period. A more refined model would be that the severity and frequency components of the $\underline{\theta}$'s would be

---

[3] Later in the paper K will be used to denote the familiar loss development triangle matrix, which is a particular way of summarizing the information in $\underline{K}$. K_{aj} denotes the a,j element of K, where a is the accident year.

[4] This paper will only consider estimates of $E(L)$. One might also want to estimate other attributes of the distribution of L, such as Var(L) or 95% percentile of L.

[5] Actually the loss reserving process seeks to find $E(L_a - B_{a\cdot}|K_{a\cdot},B_{a\cdot})$ where $K_{a\cdot}$ is the total known dollars of loss for accident year a (* denoting the latest known column) and $B_{a\cdot}$ is total paid dollars of loss. Footnote 7 shows that this distinction does not affect the methodology of this paper.

influenced by inflationary trends and by changes in a measurable exposure base, and that, after proper adjustments for these, the parameters would be stable over time. Examples of these type of adjustments are given in [2].

Any experience rating or reserving procedure is an estimator[6] of $E(\underline{L})$; it is some function R of the insured's past known loss and exposure information $\underline{K}$. A perfect reserve estimation procedure for accident year a would be a function R_a such that $R_a(\underline{K}) = E(L_a)$. However, $\underline{K}$ is also a random variable, so fulfilling this condition is not possible, except by chance. We can, however, hope that $R_a(\underline{K})$ is an unbiased estimator of $E(L_a)$, that is, that $E(R_a(\underline{K})) = E(L_a)$.

We would also like $R(\underline{K})$ to be close to $E(L)$, on the average. One common way of expressing this is to minimize $E((R(\underline{K}) - E(L))^2)$, the mean square error, which for an unbiased estimator is equivalent to minimizing $\text{Var}(R(\underline{K}))$. For many simple statistical models, the form of estimator R that satisfies these criteria can be explicitly calculated. This is referred to as a Uniform Minimum Variance Unbiased (UMVU) estimator.[7]

For large samples, the Maximum Likelihood Estimator (MLE) usually satisfies these properties (asymptotically). However, there are reasons why we cannot always use the MLE, the main one being that in order to calculate it we must explicitly know the forms of the probability distributions that generate L. Of course, we can specify a model of the process that we believe is "reasonable" (as is done later in this paper), but there still are several problems. First, the

---

[6] An estimator is a function of a random sample and is therefore a random variable; an estimate is the result of the estimator function applied to a particular realization of the random variable, and is therefore itself a particular number. This paper will use the term prediction as a synonym for estimate. Also, note that $\underline{L}$ denotes the vector $(L_1, \ldots, L_n)$; $\underline{R}$ is defined similarly.

[7] In the computer model that follows, the quantities actually being measured are the expected value and variance of the prediction error $(R(\underline{K}) - L)$. Note that:

1. The error of any prediction $R_a(\underline{K})$ of ultimate losses L_a is identical to the error of using $R_a(\underline{K}) - B_a$ to predict necessary loss reserves $L_a - B_a$ so the expected values and variances measured in this paper apply equally well to loss reserves.
2. $E(R(\underline{K}) - L) = E(R(\underline{K})) - E(L) = \text{Bias of } R(\underline{K})$
3. $\text{Var}(R(\underline{K}) - L) = \text{Var}(R(\underline{K})) + \text{Var}(L) - 2\,\text{Cov}(R(\underline{K}),L)$
   If L pertains to an accident year for which there is no known experience, then $\text{Cov}(R(\underline{K}),L) = 0$ and we are measuring $\text{Var}(R(\underline{K}))$ plus a constant that does not depend on R. If there is some known experience for the accident year—as is typical for loss reserving—then we are not actually measuring $\text{Var}(R(\underline{K}))$; however the variance of the prediction error is actually what we are interested in.

Note that we have dropped the subscript a when not referring to a specific accident year.

MLE can be very difficult to calculate; second, although it is known to have good properties for large samples, it may be a bad estimator for smaller samples (it is usually biased); third, while it may be a good estimator if the model we assume is in fact the true one, it may be a bad estimator for a different model—that is, it may not be robust.

III. COMPUTER MODEL

The computer generates six accident years of known loss experience ($\underline{K}^{(i)}$ for the i^{th} iteration) from distributions with fixed parameters. It then applies four estimation techniques to this set of known losses, arriving at four different predictions of $\underline{L}^{(i)}$. The differences between each of the predictions and the actual ultimate losses are stored. This whole process (generating experience, then calculating predictions) is repeated several thousand times—using the same underlying distributions and parameters. It can then be determined how well the estimates $\underline{R}(\underline{K}^{(i)})$ fared as "guesses" of $\underline{L}^{(i)}$ and which estimator function R does the best.

Each iteration produced a set of loss experience for six accident years—($a = 0,\ldots,5$) where five years of development are known for accident year 0, four years of development for accident year 1, etc. Not only was the ultimate experience generated for each of these years, but also the portion of it that would be known at any point in time.

For a single accident year a, a single iteration was generated as follows:[8]

A random number of losses, N, was drawn from a normal[9] distribution with mean = 40, variance = 60.
For each of the N claims, the following random variables were drawn ($i=1,\ldots,N$):
M_i = Month of loss within accident year (uniform with minimum = 0, maximum = 11)

[8] The forms of the distributions chosen are somewhat arbitrary, but are consistent with actuarial literature. For negative binomial frequency see [1], [8] and [17]; for lognormal severity see [4], [10], [13], [14], [16] and [18]; for exponential report lags see [15] and [22]. However, it is important to note that, as demonstrated later in the paper, the conclusions are not particularly sensitive to the choice of the underlying loss generation model.

[9] The normal distribution was chosen as a good approximation for the negative binomial, which is more difficult to simulate. Also, N was restricted to be greater than zero.

Q_i = Report lag in months (waiting time between accident date and report date) (exponential with mean = 18 months)

All experience was viewed as being analyzed as of year-end, so a claim would first become known in $\lceil \frac{(M_i + Q_i - 1)}{12}$ years after the accident year.[10]

P_i = Payment lag in months (waiting time between report date and payment date) (exponential with mean = 12 months)[11]

Then the following dates are calculated:

m_i = accident month = $12a + M_i$

r_i = report month = $12a + M_i + Q_i$

p_i = payment month = $12a + M_i + Q_i + P_i$

Note that m_i, r_i, and p_i are fixed dates (where the first month of the first accident year is taken to be 0). M_i, Q_i, and P_i are lags relative to the accident year (a = 0,...,5) in which the simulated claim occurs, and relative to each other.

The random untrended payment amount, C_i, was drawn from a lognormal distribution with mean = \$10,400 and variance = (\$34,800)^2.

The final settlement value of the claim is calculated as $C_i T(m_i, p_i)$, where $T(m,p)$ is an inflation factor equal to $\left(\frac{I_{m+p}}{I_o}\right)^{\alpha}\left(\frac{I_m}{I_o}\right)^{1-\alpha}$ and I_k is an inflation index at month k. This inflation model was suggested by Robert Butsic in [9].

So far, the number of claims, and (for each of these claims) the report date, the payment date, and the final payment amount have been determined. The last thing to do is set the reserve on each open claim. Each reserve was set as an unbiased guess of what the claim would settle for, if it closed in the month for which the reserve was being set.

[10] The APL symbol $\lceil$, referred to as "ceiling," means "the smallest integer greater than or equal to." Note that if $M_i + Q_i < 12$ the claim is reported during the accident year, "zero" years after the accident year.

[11] These parameters for M, P and Q result in the following average age-to-ultimate factors:

	12-ult	24-ult	36-ult	48-ult	60-ult
Incurred	3.72	1.60	1.24	1.11	1.05
Paid	14.29	2.94	1.69	1.30	1.15

For each claim a random Reserve Error, V_i, was drawn from a lognormal distribution with mean = 1, and variance = 2. To calculate the reserve amount, this was multiplied by $C_iT(m_i,r_i)$ where r_i is the month that the claim was first reported (and therefore reserved). Two things should be noted about this model of case reserving: (1) the reserve error is only chosen once for each claim, regardless of how many years it remains open; and, (2) this system, on the average, leads to under-reserving—by (I_p/I_r), the amount of inflation between the report month and the payment month.[12]

The known loss amount at the end of year t on the i^{th} loss from accident year a is

$$k_i(a,t) = \begin{cases} 0 & \text{if } r_i > 12t + 11 \\ C_iV_iT(m_i,r_i) & \text{if } r_i \leq 12t + 11 < p_i \\ C_iT(m_i,p_i) & \text{if } p_i \leq 12t + 11 \end{cases}$$

So the actual ultimate losses are

$$L = \sum_{i=1}^{N} C_iT(m_i,p_i)$$

The full experience matrix known at the end of year four for an insured would be

$$\begin{pmatrix} \sum_{i=1}^{N_0} k_i(0,0) & \cdots & \sum_{i=1}^{N_0} k_i(0,4) \\ \vdots & \ddots & 0 \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{N_4} k_i(4,4) & 0 \cdots & 0 \end{pmatrix}$$

This represents the familiar "loss development triangle." We will denote such an experience matrix by K (for known data).

---

[12] The author admits that this is a crude model of the case reserving process; however, it is unlikely that a more sophisticated model would significantly affect the results—unless it was one that allowed for changes in relative reserve adequacy along the diagonal. A method of setting reserves at V times the ultimate payment, which does not lead to under-reserving, was tested in [19], and it did not make a significant difference in the results. Also, see Section VI on sensitivity tests.

The matrix K is the statistic that we will use to estimate the vector of expected final loss amounts $E(\underline{L})$. Note that there are many other possible statistics we could have chosen (such as a triangle of claim counts, or a triangle of losses truncated at some "basic limits" point). Other such statistics would probably allow us to construct more efficient estimators—in fact, they definitely would unless K happened to be a "sufficient statistic" for $E(\underline{L})$, and there is no reason to believe that it is sufficient.

IV. RATING METHODS

Once the experience matrix K is calculated for one iteration, it is used as input for four different rating techniques (estimators of $E(L)$).

Let K_{aj} = Losses for accident year a known through period j (in other words, the aj element of matrix K)

K_{a*} = Latest known losses for accident year a

f_a = The age-to-ultimate factor for accident year a[13]

R_a = The estimate of expected ultimate losses, $E(L_a)$

1: Age-to-Age Factors

This is the very common procedure of projecting each accident year to its ultimate value by age-to-age factors (also known as the "chain ladder" method). So

$$R_a = K_{a*}f_a \qquad a = 0,\ldots,4$$

R_5 undefined (because $K_{5*} = 0$)

2: Modified Bornhuetter-Ferguson

This is a modified version of a commonly used method first presented in [5].[14]

$$R_a = K_{a*} + R_5\left(1 - \frac{1}{f_a}\right) \qquad a = 0,\ldots,4$$

---

[13] Age-to-age factors throughout this paper are calculated by summing corresponding elements in two adjacent columns of the triangle, then dividing these two sums. This is usually superior, as shown in [12], to taking a straight average of the individual age-to-age factors, which is likely to produce substantial additional bias.

[14] In [5] R_5 was obtained from external sources, rather than as shown here.

$$R_5 = (1/5) \sum_{h=0}^{4} K_{h*} f_h$$

3: Adjustment to Total Known Losses

This method (also referred to as the "Cape Cod method")is described in [7] and [19]. Appendix B presents a theoretical comparison of R_5 under this method with R_5 under method 2. It consists of averaging the known losses first, then applying an adjustment factor to the sum.

$$R_a = K_{a*} + R_5\left(1 - \frac{1}{f_a}\right) \qquad a = 0,\ldots,4$$

$$R_5 = \left(\sum_{a=0}^{4} K_{a*}\right) \div \left(\sum_{a=0}^{4} (1/f_a)\right)$$

4: Additive Model

Let K' denote the matrix of known loss experience where each cell is the losses incurred *during* a particular period (rather than cumulative losses *through* the period, as the matrix K denotes). The elements of K' are the differences of adjacent columns of K.

Project the unreported losses for an accident year as the sum of the expected unreported losses during each future period. Estimate the expected unreported losses by period as the average of the known losses by row. Specifically,

$$R_a = K_{a*} + \sum_{h=5-a}^{4} \frac{1}{a} \sum_{g=0}^{a-1} K'_{gh} \qquad a = 1,\ldots,5$$

$$R_0 = K_{0*}$$

This additive method is suggested by Hans Bühlmann [7]; he refers to it as the complementary loss ratio method.

V. RESULTS

Each of the four rating methods was tested under each of the following progressively more complex loss generation models. Exhibits I through V display the results for each model. These exhibits show the mean and standard deviation of the prediction error for each rating method for each accident year. The prediction error is $R_a - L_a$ (the estimated ultimate result minus the actual ultimate result). The "% of actual" is the prediction error divided by the true expected losses.

We would expect any rating technique based on known data to (on the average) under-predict by the expected amount of development between the most mature known data amount and ultimate $E(K_{a4} - L_a)$. Therefore, each of the expected prediction errors has been adjusted by this amount, so the exhibits actually show $E(R_a - L_a) - E(K_{a4} - L_a) = E(R_a - K_{a4})$. That is, we do not expect the estimation techniques to be able to predict beyond the triangle.[15]

EXHIBIT I—Claim Counts Only, No Inflation

In this version of the model, C_i was not randomly chosen, but was set at \$1. The inflation index I_m was also held constant. The results show that simple age-to-age factors produced biased results and higher standard deviations. Methods 2 and 3 have very slight biases while method 4 is unbiased. Methods 3 and 4 have slightly smaller standard deviations than method 2.

What is interesting here is not the amount of the bias (which for practical purposes is negligible), but the fact that there *is* a bias. This fact was greeted with surprise and skepticism by many actuaries when it was first presented in [19]. Appendix A gives a technical argument to support this result.

EXHIBIT II—Random Claim Size, No Inflation

In this version, C_i is randomly chosen from a lognormal distribution with mean = \$10,400 and variance = $(\$34,800)^2$. The inflation index I_m was held constant. Here we see that method 1 is clearly inferior—it is significantly biased upward and has very high standard deviations in years 3 and 4. An interesting result from the older version of the model is that the median prediction error for method 1 was usually negative—that means that in over half of the cases method 1 under-predicted the actual (simulated) results, but a few cases of large over-predictions made the mean prediction error (the bias) positive. This is because the distribution of prediction errors for method 1 was very positively skewed. Method 3 has the lowest standard deviation. Methods 3 and 4 do not appear to have significant biases.

---

[15] A technique of estimating the parameters of the distribution of Q_i directly, such as described in [22], would allow prediction beyond the triangle.

EXHIBIT III—Constant 8% Inflation, $\alpha = 0.5$

In this version, an 8% per year inflation was assumed, with 50% applying to date of accident and 50% applying to date of settlement. Here we expect that methods 2, 3 and 4 will under-predict, because they all implicitly assume that expected losses by accident year are the same, which, with inflation, is not true. Method 1 does not rely on such an assumption.

The addition of inflation accentuates the bias in method 1, making its predictions 35% above the actual values (after the "tail adjustment"). Method 2 does very well on this example because the upward bias inherent in each age-to-ultimate prediction is balanced by the fact that the method assumes no inflation. Once again methods 3 and 4 do the best in terms of standard deviation, but, as expected, they are somewhat biased downward.

EXHIBIT IV—Constant 8% Inflation, $\alpha = 0.5$,
Adjust Rating Methods for 8% Inflation

This version was run with the same loss parameters and inflation assumptions as model III. However, each of the rating methods was modified as follows:

> Each element of each row, where an arbitrary row is row a, was divided by an assumed inflation index I'_a. The rating method was applied to the resulting triangle, then each projected ultimate result was multiplied by its respective I'_a. In this case I'_a was set as 1.08^a, $a = 0,\ldots,5$. This obviously represents perfect clairvoyance about the underlying past and future inflation rate.[16]

This slightly improves the standard deviation of method 1, but does not improve the bias, which is still quite high. However, this adjustment completely removes the bias on method 4, and leaves only a slight upward bias in method 3.

EXHIBIT V—10% Inflation Dropping to 6%, $\alpha = 0.5$,
Adjust Rating Methods for 10% Inflation

In this version, the actual inflation rate was 10% for 60 months (which covers the entire known claim period), then it drops to 6%. The index assumed by the rating methods is $(1.10)^a$.

---

[16] Note that a similar adjustment can be made when dealing with a triangle where the exposure varies by accident year, i.e., (1) divide each row by the corresponding exposure, (2) apply the rating method, then (3) multiply each estimate by its exposure. This could be further improved by using credibility weighted averages in the rating method, where a row's credibility was a function of its exposure; however, developing such a system is beyond the scope of this paper.

This results in only a slight bias in method 4, and a fairly small one in method 3.

VI. SENSITIVITY TESTING

As a test of the sensitivity of the results to the specific distributions used to generate loss experience, the following additional three scenarios were run. Note that these were all run with an assumption of no inflation, so they are meant to be compared with the results on Exhibit II (which will be referred to as the "standard model").

EXHIBIT VI—No Reserve Development

The standard model was used except that the reserve error, V, was always set equal to one.

EXHIBIT VII—Uniform Frequency and Severity

The standard model was used except that the frequency, N, was distributed discrete uniform [1,79] and severity, C, was distributed continuous uniform [0,20800]. This results in an ultimate aggregate loss distribution with about the same mean and variance, but much less skewness, than the standard model.

EXHIBIT VIII—Uniform Report and Payment Lags

The standard model was used except that the report lag, Q, was distributed discrete uniform [0,36] and payment lag, P, was distributed discrete uniform [0,24]. This results in the same average lags, but with a higher percentage of claims being reported and paid within the five columns of the experience triangle than the standard model.

Although the magnitudes of the biases and standard deviations differ in Exhibits VI through VIII from Exhibit II, conclusions about the existence of bias and about the relative efficiency of the four rating methods remain substantially unchanged.

VII. CONCLUSIONS

These results indicate that for data triangles of the size tested:

1. The common age-to-age factor approach (method 1) is clearly inferior to the other three methods.

2. The additive method 4 and the average-then-adjust method 3 have significantly lower variances than methods 1 and 2, and small biases (if adjusted for inflation). In fact, method 4 may be completely unbiased.

It is important to emphasize that the bias of the various methods is heavily influenced by a few large prediction errors. This means that in practical ratemaking situations it would usually be wrong to use method 1 and then do a judgment "bias adjustment"—doing so in most cases would result in under predicting. Instead, the practitioner simply should not put much credibility in predictions based on highly leveraged age-to-ultimate factors.

One may object that allowing accurate knowledge of the underlying inflation rate gives an unfair advantage to methods 2 through 4, because it allows all of the rows of the triangle to be used in estimating any particular row's ultimate value. However, one will normally have exogenous knowledge of past inflation rates and forecasts of future rates, and using this information should improve one's ability to predict. Also, in [19] it was shown that attempts to estimate the trend rate solely from data samples of this size by fitting lines to projected ultimate values produced terrible results—extreme bias, variance, and skewness.

The above major conclusions concern the relative ranking of techniques and the existence in some cases of bias. These conclusions were found to be robust to an extreme change in the form of the underlying distributions; this robustness was also found in [19]. Of course, the specific numerical results on Exhibits I through VIII should not be considered to be any more than examples—changing the parameters or the form of the loss generating model will change these in unpredictable ways.

One way in which numerical results from a model such as this would be of interest is if the parameters of the loss generation model were estimated from an actual data set which had been projected to ultimate by a specific loss reserving technique. Simulating the distribution of prediction error would give an estimate of the potential variability of the reserve estimate—which could be used to calculate confidence intervals (containing both "parameter" and "process" risk) for the loss reserve.

REFERENCES

[1] Robert A. Bailey and LeRoy J. Simon, "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," *PCAS* XLVI, 1959, p. 159.

[2] James R. Berquist and Richard E. Sherman, "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," *PCAS* LXIV, 1977, p. 123.

[3] Peter Bickel and Kjell Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics,* Holden-Day, 1977.

[4] David R. Bickerstaff, "Automobile Collision Deductibles and Repair Cost Groups: The Lognormal Model," *PCAS* LIX, 1972, p. 68.

[5] Ronald L. Bornhuetter and Ronald E. Ferguson, "The Actuary and IBNR," *PCAS* LIX, 1972, p. 181.

[6] Hans Bühlmann, Rene Schnieper, and Erwin Straub, "Claims Reserves in Casualty Insurance Based on a Probabilistic Model," *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker 80,* 1980.

[7] Hans Bühlmann, "Estimation of IBNR Reserves by the Methods Chain Ladder, Cape Cod and Complementary Loss Ratio," unpublished, 1983.

[8] Hans Bühlmann, *Mathematical Methods in Risk Theory,* Springer-Verlag, 1970.

[9] Robert Butsic, "The Effect of Inflation on Losses and Premiums for Property-Liability Insurers," *Inflation Implications for Property-Casualty Insurance,* 1981 Casualty Actuarial Society Discussion Paper Program, p. 58.

[10] Robert J. Finger, "Estimating Pure Premium by Layer—An Approach," *PCAS* LXIII, 1976, p. 34.

[11] George Fishman, *Concepts and Methods in Discrete Event Digital Simulation,* John Wiley & Sons, 1973.

[12] Charles A. Hachemeister and James N. Stanard, "IBNR Claims Count Estimation with Static Lag Functions," unpublished, 1975.

[13] Charles C. Hewitt, Jr. and Benjamin Lefkowitz, "Methods for Fitting Distributions to Insurance Loss Data," *PCAS* LXVI, 1979, p. 139.

[14] Charles C. Hewitt, Jr., "Credibility for Severity," *PCAS* LVII, 1970, p. 148.

[15] Charles L. McClenahan, "A Mathematical Model for Loss Reserve Analysis," *PCAS* LXII, 1975, p. 134.

[16] Robert S. Miccolis, "On the Theory of Increased Limits and Excess of Loss Pricing," *PCAS* LXIV, 1977, p. 27.

[17] Gary Patrik and Russell John, "Pricing Excess-of-Loss Casualty Working Cover Reinsurance Treaties," *Pricing Property and Casualty Insurance Products*, 1980 Casualty Actuarial Society Discussion Paper Program, p. 399.

[18] Sheldon Rosenberg and Aaron Halpert, "Adjusting Size of Loss Distributions for Trend," *Inflation Implications for Property-Casualty Insurance*, 1981 Casualty Actuarial Society Discussion Paper Program, p. 458.

[19] James N. Stanard, "Experience Rates as Estimators: A Simulation of their Bias and Variance," *Pricing Property and Casualty Insurance Products*, 1980 Casualty Actuarial Society Discussion Paper Program, and Review by John Robertson, p. 485.

[20] A. Tenenbein and D. Covall, "Determination of Development Factors for Automobile Insurance Rates," New York University Graduate Business School Working Paper #79-129, 1979.

[21] J. Van Eeghan and G. de Wit, *Loss Reserving Methods*, Surveys of Actuarial Studies No. 1, Nationale-Nederlanden N.V., 1981.

[22] Edward W. Weissner, "Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood," *PCAS* LXV, 1978, p. 1.

EXHIBIT I

MODEL I—CLAIM COUNTS ONLY, NO INFLATION
5000 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
		Counts	% of Actual	Counts	% of Actual
1	0	0	0%	1.5	4%
	1	0.1	0	2.6	6
	2	0.2	1	3.7	9
	3	0.3	1	5.8	14
	4	0.9	2	11.6	29
	5	—	—	—	—
2	0	0	0%	1.5	4%
	1	0.1	0	2.5	6
	2	0.2	1	3.6	9
	3	0.2	1	5.1	13
	4	0.3	1	7.2	18
	5	0.4	1	8.8	22
3	0	0	0%	1.5	4%
	1	0.1	0	2.5	6
	2	0.2	0	3.5	9
	3	0.1	0	5.0	13
	4	0.1	0	7.1	18
	5	0.2	0	8.6	22
4	0	0	0%	1.5	4%
	1	0	0	2.5	6
	2	0.1	0	3.5	9
	3	0	0	5.0	13
	4	0	0	7.2	18
	5	0.1	0	8.6	22

EXHIBIT II

MODEL II—RANDOM CLAIM SIZE, NO INFLATION 5000 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$ — Mean: Dollars	Mean: % of Actual	Standard Deviation: Dollars	Standard Deviation: % of Actual
1	0	$ 0	0%	$ 88,600	21%
	1	9,892	2	182,206	44
	2	24,680	6	252,951	61
	3	49,766	12	392,435	94
	4	113,397	27	823,429	198
	5	—	—	—	
2	0	$ 0	0%	$ 88,600	21%
	1	9,354	2	177,605	43
	2	16,234	4	412,028	99
	3	29,183	7	303,322	73
	4	32,183	8	377,037	90
	5	36,314	9	372,499	89
3	0	$ 0	0%	$ 88,600	21%
	1	5,712	1	163,078	39
	2	13,138	3	212,171	51
	3	14,501	3	263,962	63
	4	4,662	1	320,142	77
	5	4,370	1	322,794	77
4	0	$ 0	0%	$ 88,600	21%
	1	−894	0	170,483	41
	2	−4,787	−1	438,705	105
	3	−3,986	−1	290,293	70
	4	−11,622	−3	338,545	81
	5	−7,490	−2	341,970	82

EXHIBIT III

MODEL III—8% INFLATION, $\alpha = 0.5$
15,000 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
		Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$132,643	28%
	1	13,325	3	233,786	45
	2	40,012	7	528,989	95
	3	75,972	13	674,655	113
	4	225,406	35	1,636,846	254
	5	—	—	—	—
2	0	$ 0	0%	$132,643	28%
	1	18,162	4	281,171	54
	2	35,581	6	376,524	68
	3	37,095	6	498,673	83
	4	15,500	2	639,790	99
	5	−62,654	−9	609,556	87
3	0	$ 0	0%	$132,643	28%
	1	9,766	2	194,158	38
	2	15,783	3	280,995	51
	3	−607	0	385,999	65
	4	−49,904	−8	451,253	70
	5	−138,589	−20	450,961	64
4	0	$ 0	0%	$132,643	28%
	1	−2,462	−1	185,358	36
	2	−8,613	−2	273,372	49
	3	−32,982	−6	363,169	61
	4	−80,318	−13	423,457	66
	5	−158,472	−23	441,974	63

EXHIBIT IV

MODEL IV—8% INFLATION, $\alpha = 0.5$, 8% INDEX USED IN RATING
12,750 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
		Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$108,748	23%
	1	17,024	3	226,684	44
	2	41,313	7	368,332	66
	3	81,257	14	627,023	104
	4	214,678	33	1,545,088	240
	5	—	—	—	—
2	0	$ 0	0%	$108,748	23%
	1	17,021	3	242,467	47
	2	34,663	6	328,942	59
	3	56,512	9	486,315	81
	4	75,782	12	597,070	93
	5	83,162	12	640,675	92
3	0	$ 0	0%	$108,748	23%
	1	12,228	2	209,716	41
	2	22,240	4	284,919	51
	3	30,927	5	398,680	66
	4	27,951	4	451,586	70
	5	24,978	4	496,771	71
4	0	$ 0	0%	$108,748	23%
	1	1,546	0	228,070	44
	2	3,571	0	289,117	52
	3	6,014	1	405,582	68
	4	4,862	1	433,385	67
	5	6,569	1	492,804	71

EXHIBIT V

MODEL V—10% INFLATION DROPPING TO 6%, α = 0.5, 10% INDEX USED IN RATING
8000 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
		Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$120,911	25%
	1	13,748	3	260,039	48
	2	34,538	6	397,028	69
	3	79,547	13	569,751	90
	4	227,292	33	1,331,666	193
	5	—	—	—	—
2	0	$ 0	0%	$120,911	25%
	1	12,627	2	243,327	45
	2	27,992	5	344,815	60
	3	54,273	9	446,620	70
	4	89,787	13	577,771	84
	5	108,456	15	617,252	83
3	0	$ 0	0%	$120,911	25%
	1	8,522	2	225,986	42
	2	17,345	3	310,728	54
	3	30,093	5	393,830	62
	4	42,842	6	481,436	70
	5	49,802	7	519,365	70
4	0	$ 0	0%	$120,911	25%
	1	−1,386	0	230,404	43
	2	−672	0	320,637	56
	3	6,152	1	393,860	62
	4	19,540	3	469,071	68
	5	31,185	4	516,946	69

EXHIBIT VI

SENSITIVITY TEST—PERFECT CASE RESERVING
6522 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$ Mean Dollars	Mean % of Actual	Standard Deviation Dollars	Standard Deviation % of Actual
1	0	$ 0	0%	$ 50,557	12%
	1	5,092	1	117,095	28
	2	12,428	3	147,644	35
	3	19,602	5	231,091	56
	4	67,650	16	504,934	122
	5	—	—	—	—
2	0	$ 0	0%	$ 50,557	12%
	1	4,726	1	109,725	26
	2	10,880	3	130,699	31
	3	14,206	3	186,497	45
	4	22,904	6	251,845	61
	5	17,081	4	305,529	73
3	0	$ 0	0%	$ 50,557	12%
	1	2,898	1	104,576	25
	2	6,574	2	120,845	29
	3	6,057	1	172,281	41
	4	8,546	2	229,368	56
	5	942	0	281,352	67
4	0	$ 0	0%	$ 50,557	12%
	1	−542	0	101,240	24
	2	899	0	116,160	28
	3	−658	0	168,233	41
	4	2,361	1	225,834	55
	5	−3,461	−1	278,649	65

EXHIBIT VII

SENSITIVITY TEST—UNIFORM FREQUENCY AND SEVERITY
3049 ITERATIONS

Rating Method	Accident Year	Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
		Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$ 34,704	8%
	1	9,824	2	175,539	43
	2	8,593	2	155,670	37
	3	14,083	3	195,329	46
	4	44,232	11	331,371	79
	5	—	—	—	—
2	0	$ 0	0%	$ 34,704	8%
	1	7,410	2	128,228	31
	2	7,579	2	143,738	35
	3	8,791	2	168,244	40
	4	15,059	4	222,905	53
	5	20,915	5	287,287	70
3	0	$ 0	0%	$ 34,704	8%
	1	3,690	1	74,962	18
	2	2,208	1	97,500	23
	3	1,123	0	130,512	31
	4	3,353	1	204,091	49
	5	7,643	2	263,368	64
4	0	$ 0	0%	$ 34,704	8%
	1	342	0	65,336	16
	2	−2,937	−1	94,379	23
	3	−4,874	−1	133,405	32
	4	−2,227	−1	209,730	50
	5	3,629	1	264,368	65

EXHIBIT VIII

SENSITIVITY TEST—UNIFORM REPORT AND PAYMENT LAGS
6000 ITERATIONS

		Prediction Error: $(R_a - L_a)$ minus $E(K_{a4} - L_a)$			
		Mean		Standard Deviation	
Rating Method	Accident Year	Dollars	% of Actual	Dollars	% of Actual
1	0	$ 0	0%	$ 33,898	0%
	1	8,958	2	154,927	38
	2	40,911	10	371,052	89
	3	94,371	23	667,695	162
	4	261,076	63	1,627,441	393
	5	—	—	—	—
2	0	$ 0	0%	$ 33,898	8%
	1	7,452	2	183,330	45
	2	39,655	10	348,887	84
	3	66,107	16	447,484	109
	4	79,065	19	499,807	121
	5	81,733	20	481,770	117
3	0	$ 0	0%	$ 33,898	8%
	1	5,097	1	153,015	37
	2	21,913	5	279,507	67
	3	23,616	6	321,873	78
	4	14,848	4	301,852	73
	5	13,765	3	297,400	72
4	0	$ 0	0%	$ 33,898	8%
	1	−3,670	−1	195,110	47
	2	4,794	1	276,092	66
	3	3,478	1	306,744	75
	4	−1,347	0	292,257	71
	5	1,321	0	293,734	71

APPENDIX A

AN ANALYTICAL ARGUMENT FOR BIAS OF AGE-TO-AGE FACTORS

Consider Model 1, (i.e., claim counts only and no inflation). Each row (accident year) of the data triangle K is independently and identically distributed with each other row.

This implies that $E[g(X_{ij+1},X_{ij})] = E[g(X_{kj+1},X_{kj})]\ \forall i,k$ for any function g. However $E[g(X_{ij+1},X_{ij})] \neq g(E[X_{ij+1}],E[X_{ij}])$ unless g is linear.

Let $g(X_{ij+1},X_{ij}) = \dfrac{X_{ij+1}}{X_{ij}}$

Let f_{ij} be an age-to-age factor estimated from row i. Age to age factors attempt to estimate $E[X_{kj+1}|X_{kj}]$ with $X_{kj}f_{ij}$. If this estimate were unbiased it would mean that

$$E[E[X_{kj+1}|X_{kj}]] = E[X_{kj}f_{ij}]$$

But this becomes

$$E[X_{kj+1}] = E[X_{kj}]E[f_{ij}]$$

$$E[X_{kj+1}] = E[X_{kj}]E\left[\frac{X_{ij+1}}{X_{ij}}\right]$$

$$E[X_{kj+1}] = E[X_{kj}]E\left[\frac{X_{kj+1}}{X_{kj}}\right]$$

or

$$\frac{E[X_{kj+1}]}{E[X_{kj}]} = E\left[\frac{X_{kj+1}}{X_{kj}}\right]$$

which is not true in general.[17]

[17] A similar derivation was arrived at independently by John Robertson.

APPENDIX B

COMPARISON OF "ADJUSTING, THEN AVERAGING" VERSUS "AVERAGING, THEN ADJUSTING"

Let X_i be a random variable representing observed losses for accident year i.

Assume that these losses arise from distributions with expected values that are constant over time, except for an adjustment factor. This adjustment factor can represent either a loss development factor or a trend factor or both.

So $X_i = \dfrac{\mu}{a_i} + e_i \qquad i = 1,\dots,n$

where μ = underlying expected losses
a_i = non-random adjustment factor
e_i = random error $E(e_i) = 0$, $\mathrm{Var}(e_i) = \sigma_i^2$

We wish to estimate μ.

Let $\hat{\mu}_1 = \dfrac{1}{n} \sum_{i=1}^{n} X_i a_i$

This represents trending (and/or developing) known losses for each year and averaging the results.

Let $\hat{\mu}_2 = \left(\sum_{i=1}^{n} X_i\right) \div \left(\sum_{i=1}^{n} \dfrac{1}{a_i}\right)$.

This represents the "adjustment to total known losses method."

It is easy to see that both $\hat{\mu}_1$ and $\hat{\mu}_2$ are unbiased, i.e. $E(\hat{\mu}_1) = E(\hat{\mu}_2) = \mu$. (It is important to note that this only holds if a_i is non-random, which is not the case in real estimation problems.)

Calculate the Best Linear Unbiased Estimate (B.L.U.E.)[18] of μ. That is, find weights c_i, such that $\hat{\mu}\left(= \sum_{i=1}^{n} c_i X_i\right)$ is unbiased and has minimum variance. So, minimize $\mathrm{Var}\left(\sum_{i=1}^{n} c_i X_i\right)$ subject to $E\left[\sum_{i=1}^{n} c_i X_i\right] = \mu$

$$\mathrm{Var}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i^2 \sigma_i^2$$

[18] The approach of calculating the B.L.U.E. was suggested by Aaron Tenenbein.

$$E\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} \frac{c_i}{a_i}\mu = \mu \Rightarrow \sum_{i=1}^{n} \frac{c_i}{a_i} = 1$$

Let

$$L = \sum_{i=1}^{n} c_i^2\sigma_i^2 + \lambda\left(1 - \sum_{i=1}^{n} \frac{c_i}{a_i}\right)$$

$$\frac{\partial L}{\partial c_i} = 2c_i\sigma_i^2 - \frac{\lambda}{a_i} = 0 \quad i = 1,\ldots,n$$

So $c_i = \dfrac{\lambda}{2a_i\sigma_i^2}$, $\quad \displaystyle\sum_{i=1}^{n} \frac{c_i}{a_i} = \sum_{i=1}^{n} \frac{\lambda}{2a_i^2\sigma_i^2}$

So $\lambda = \dfrac{2}{\displaystyle\sum_{i=1}^{n} \frac{1}{a_i^2\sigma_i^2}}$

So $c_i = \dfrac{1}{a_i\sigma_i^2}\,\dfrac{1}{\displaystyle\sum_{j=1}^{n} \frac{1}{a_j^2\sigma_j^2}}$

Now consider various possibilities for σ_i^2:

1. Let $X_i a_i = \mu + \epsilon_i$ where $\text{Var}(\epsilon_i) = \sigma^2\ \forall_i$

   This means that $e_i = \dfrac{\epsilon_i}{a_i}$, so $\sigma_i^2 = \dfrac{1}{a_i^2}\sigma^2$, so $c_i = a_i/n$

   Therefore $\hat{\mu}_1$ is the BLUE.

2. Let $\dfrac{\text{Var}(X_i)}{E[X_i]} = k \quad \forall_i$

   So $\dfrac{\sigma_i^2}{\mu/a_i} = \dfrac{\sigma_i^2 a_i}{\mu} \Rightarrow \sigma_i^2 a_i = k\mu$

   This means that $c_i = 1 \Big/ \displaystyle\sum_{i=1}^{n} \frac{1}{a_i}$

   Therefore $\hat{\mu}_2$ is the BLUE.

As was discussed in the results section, $\hat{\mu}_2$ performed better than $\hat{\mu}_1$ in the simulation.