Some Extensions of J. N. Stanard's Simulation Model for Loss Reserving

Richard L. Vaughan, FCAS, FSA

ABSTRACT

The loss process model and simulation procedures proposed by James M. Stanard in 1985 are extended in numerous ways, including provision for serial autocorrelation of parameters, mixtures of claim types, conditional selection of sample points, and a much greater variety of reserving methods. The extended model is used to explore many questions arising in practical loss reserving and to assist the loss reserver in choosing the best estimator for particular data conditions.

I. INTRODUCTION

A recurrent theme in the literature of loss reserving is a debate over the relative merits of traditional versus statistical estimators for aggregate reserves and ultimate losses. Both approaches start with grouped data consisting of "triangles" of reported and closed claim counts and of incurred and paid losses. Both approaches seek to complete the paid loss triangle to a parallelogram, from which reserves, ultimate losses, and future cash flows may be derived. The two approaches differ, broadly speaking, in that the traditional techniques emphasize certain intermediate quantities as aids to the loss reserver's judgement, and produce point estimates, while the statistical techniques emphasize formal models, incorporate judgement through the process of successive model building, fitting, and validation, and produce both point estimates and measures of variability.

Proponents of both approaches recognize that a model with too many parameters may fit historical data well yet produce unstable forecasts. However, they address this problem in different ways. Users of traditional techniques often start by estimating a large number of parameters, such as development factors and accident-year claim levels, and then apply judgement constraints and adjustments to achieve forecast stability. By contrast, users of statistical techniques generally start with few parameters and add new ones only as necessary to achieve a satisfactory fit.

Recently a fusion of sorts has occurred as studies have shown that the most popular of traditional loss reserve techniques, the chain-ladder, in fact has desirable properties as an estimator for certain statistical models [14][16], that each of several methods of averaging development factors may be appropriate under the right conditions [14], and that measures of variability may be calculated for the chain-ladder [13]. However, the question remains open as to which estimators are best suited to any given case arising in practice, when the occurrence, reporting, investigation, settlement, and payment of claims may be complex and may not produce loss triangles fitting simple models.

In 1985, James N. Stanard [20] attacked this problem by (a) modelling the loss process rather than the loss triangles, (b) generating synthetic data using particular parameter values and a particular number of accident and development years, (c) applying several estimators to the synthetic data and iterating the process thousands of times, (d) tabulating the differences between forecast and “actual” (i.e. simulated) ultimate losses, and (e) calculating the means and standard deviations of these prediction errors. Although Stanard’s use of simulation to compare reserving methods is the core of his paper, his examples are also interesting. They prove just how large the variance of prediction errors can be under reasonable assumptions, they strongly suggest that the conventional chain-ladder technique is both biased and unstable, they demonstrate that reserving techniques that do not attempt to estimate accident year effects may outperform those that do, and they show the desirability of not giving undue weight to the experience of immature accident years.

Stanard's lead has been followed by Daniel M. Murphy [14], who used Stanard's simulation model to compare different averages of development factors with development by linear regression; by Edward F. Peck [16], who, in a 1995 discussion of Stanard's paper, studied various weighted averages of development factors; by John W. Rollins [18], who compared paid loss development with a "structural" reserving method involving closed claim counts and average severities; and by Prakash Narayan and Thomas V. Warthen III [15], who used several different simulation models to compare traditional reserving methods with a regression technique.

In this paper, we extend Stanard's model in several directions and use it to test the efficiency of numerous reserving techniques and variations. The results reported here may help the loss reserver select those estimators likely to produce the best estimates of reserves, or of individual accident-year ultimates, for actual cases. More generally, the procedures described here may be used to obtain best estimates, by first setting up an appropriate model of the loss process, then running simulations using a set of candidate estimators, and finally applying the best-performing estimators to the actual data. The author will provide software, on request, enabling the interested reader to reproduce all of the associated exhibits and to conduct further investigations on his or her own.

II. MODEL EXTENSIONS

We extend Stanard's methodology by using a more complex simulated loss process, by refining the simulation procedure itself so as to allow approximate conditional distributions and more comprehensive reports, and by using a much larger number of reserving techniques and adjustments.

Extensions to the Simulated Loss Process

Stanard [20] postulated a five-stage model of the loss process: claim frequency, report lag, payment lag, claim severity, and case-basis reserve error. For his main exhibits, he assumed a normal distribution of claim frequency by year with a uniform distribution to months within the year; exponential distributions of report and payment lags; lognormal distribution of severities; and lognormal distribution of reserve errors, applied multiplicatively to the severities. For some of his exhibits, he also postulated both a severity trend and a fraction of inflation persisting through the interval between occurrence and payment, following Butsic [4]. To test the sensitivity of his results, Stanard also ran tests with some of the distributions rendered constant or otherwise altered.

While this is a simple model of the loss *process*, it is richer than the models of loss *triangles* that underlie many statistical loss reserving techniques. Furthermore, it is more than adequate for a preliminary investigation of the properties of reserve estimators, which, as pointed out by Robertson [17], "should work in artificially simplified situations" if they work in real life. However, we believe that certain extensions to Stanard's model can make it more realistic in ways that are essential to a proper comparison of reserving techniques.

The most important of these extensions is provision for random shifts, over time, in the means (or other location parameters) of the various distributions. This is accomplished by drawing the means from a second-level distribution having the form of a first-order autoregressive process, or, in the limiting case, a random walk process. We believe that such serial autocorrelation of parameters achieves similar results to modeling the insurance process in greater detail: for example, by assuming that some of the risks exposed each month allow their coverage to lapse and are replaced with new risks drawn from a wider population, while the others continue with parameters unchanged. We also believe that this accords with the observation that a given volume of recent experience, in insurance or other time series, has more predictive utility than the same volume of more remote experience.

A second important extension is provision for a mixture of claim types, each with its own distribution of frequency, severity, lags, and errors. Such a mixture comes closer to the demands placed on real-world loss reserving methods than does a model with a single claim type.

Other extensions to Stanard’s model include provision for seasonality factors for claim frequency; additional choices regarding the form of each distribution; provision for maximum report and payment lags; provision for a deductible and/or limit; provision for a non-zero probability that a claim (whose size as drawn from the severity distribution is used as the starting point for the case reserve) will actually close without payment; corresponding provision for an assumed probability of closure without payment to be used to discount the case reserve itself; provision for an initial reserve applicable to all claims upon reporting; and provision for specifying the number of months after reporting at which the first regular valuation takes place.

One final difference from Stanard’s model is that we generate claim frequencies monthly, rather than generating them annually and then distributing them to months. We believe that the latter approach introduces an artificial negative correlation between pairs of months within the same year.

Extensions to the Simulation Procedure

Stanard reports the mean of the prediction errors across the entire sample of simulated claim histories, thereby approximating the mean of the unconditional distribution of prediction errors. What is usually of greater interest, however, is the conditional distribution of prediction errors given a particular known claim history.

In most of what follows, we also use the entire sample and report the unconditional mean and standard deviation of the prediction errors. However, when we wish to investigate conditional distributions, we have available three alternative approaches. First, we may define a measure of distance on the space of all known claim histories, and then include in our sample only those simulated claim histories whose known portion falls within a given "acceptance radius" of some particular claim history. Second, we may define a Boolean condition on the known claim history, and include in our sample only those simulated claim histories satisfying the condition. Third, we may define a numeric function of the known claim history, and cross-tabulate one of the results of the simulation process (such as the prediction error for total reserves) against the value of this function.

Stanard's emphasis is on the prediction of ultimate losses by accident year; ours is about equally divided between ultimate losses by accident year and total reserves over all accident years. Also, Stanard expresses prediction errors as percentages of mean actual ultimate losses for each year, including the losses which are known at the evaluation date. We follow Stanard's practice when dealing with individual accident years, but we relate the prediction error of total reserves to mean runoffs rather than mean total claims.

A word is in order regarding Stanard's tail adjustment, which corrects for the fact that none of the reserving methods he studied can be expected to project incurred loss development beyond the lag represented by the latest known development period in the first accident year. The adjustment consists of adding to all results the *sample mean* of the tail beyond that lag. This completely eliminates the measured bias attributable to the failure to project the tail, but leaves the variance and standard deviation unchanged.

In effect, the tail contributes to the process variance, which is already a large part of the total variance of the prediction errors, and makes it even larger; Stanard's adjustment lets this larger variance stand. An alternative adjustment would be to compare predicted ultimate losses with actual (simulated) losses excluding development beyond the last known lag. We have retained Stanard's adjustment for the sake of comparability, but we prefer to reduce the importance of the tail by using more development years than did Stanard. Note that no adjustment is needed in situations where the loss reserving techniques themselves include a tail projection.

Finally, we compare the performance of all the reserving methods included in each simulation run, as measured by their absolute prediction errors on a particular item (such as total reserves) for each trial, giving both absolute rankings and pairwise comparisons.

Extensions to the Set of Reserving Techniques

Stanard considers only estimators that are functions of the matrix of known case-basis incurred losses, while acknowledging that other statistics, such as matrices of claim counts or limited losses, might lead to more efficient estimators. Since Stanard's simulation model automatically generates triangles of reported and closed claim counts and paid losses, in addition to case-basis incurred losses, we drop his restriction and consider estimators that are functions of these four basic triangles.

We require our estimators to predict not only ultimate losses by accident year, but also the complete matrix of estimated future payments by accident year versus development year. Therefore the triangle of known paid losses is always necessary, to calculate the reserves and to distribute them to future payment periods, even when ultimate losses are projected from other data.

Some of our estimators also make use of exposures, though these are not treated as random variables.

By thus broadening the range of estimators considered, we are able to compare the relative merits of paid versus incurred loss development and of reserving techniques involving separate consideration of frequency and severity. We do restrict attention to estimators using data grouped into discrete cells, but we have considerable flexibility in the choice of the incurral and development intervals defining these cells. In what follows, the term “accident year” refers to the incurral interval and “development year” to the development interval, whether or not these intervals are actually twelve months in length.

III. STANARD’S SIMULATIONS REVISITED

As a preliminary application of our system, we replicate Stanard’s main results and then consider the effect of changing some of his loss process assumptions.

Replication of Stanard’s Results

Exhibit 1 repeats the calculations summarized in Stanard’s Exhibits I through V, using the same four estimators (chain-ladder, modified Bornhuetter-Ferguson [2], Cape Cod [3], and additive [3]) and the same structure and process parameters, with the exception, mentioned above, that our claim frequencies are generated monthly rather than annually. Each set of calculations was iterated 10,000 times, using the same initial random seed and therefore drawing the same sets of sample data.

In each section of the exhibit, the “reserves” row displays the prediction error of ultimate claims less claims paid as of the end of year 4, summed over years 0 through 4, i.e. not including year 5 which is entirely in the future as of the evaluation date.

The purpose of this exercise is to confirm that our model and loss reserving algorithms are substantially the same as Stanard’s, giving results differing from his within a range which might reasonably be attributed to chance. This appears to be the case for the prediction errors in sections I through IV of Exhibit 1, with about the expected number of differences between them and Stanard's corresponding prediction errors being significant at the 5% level, using a normal approximation.

However, ten of the 19 available differences for section V are significant, and all 19 of these differences are greater than zero. It appears that here the two models and/or sets of reserving techniques do differ systematically. Provisionally, we ascribe this to differences between the methods of adjusting incurred losses for inflation. Stanard deflates the incurred losses themselves to a constant-dollar level, performs the reserve calculations, and restores inflation to the results. We follow the same procedure for paid losses, but we separate incurred losses into paid losses and reserves and deflate these two components separately, paid losses from dates of payment and reserves from date of valuation.

As a general rule, the mean prediction error for year 0 is zero after Stanard's tail adjustment, since the estimation methods being considered do not attempt to project the tail beyond the last known lag for the first year. In sections IV and V of Exhibit 1, however, positive means are shown for these prediction errors. This is also a result of the difference in methods of inflation adjustment. We deflate the triangle of known incurred losses to a constant-dollar basis, apply the estimation method to generate a matrix of known and projected *paid* losses, and then reflate the paid loss triangle; Stanard deflates and reflates only the incurred losses. The last step in our procedure brings in the expected inflation between the valuation date and the dates the remaining losses for the first accident year are paid, which is not accounted for by Stanard's procedure. As the tail adjustment itself is unaffected, the net result is a difference in mean prediction errors after tail adjustment.

In summary, there are minor non-random differences between our model and Stanard's. But the conclusions which may be drawn from an examination of the means and standard deviations of the prediction errors are essentially the same.

Effect of Modifying the Loss Process Assumptions

Stanard ran his simulations with three different trend assumptions, and performed sensitivity tests in which uniform distributions were substituted for the original normal, exponential, etc., but he did not report the effect of letting some of the other key assumptions vary. With an eye toward using a more complex model later in this paper, we here investigate the effects of (a) increasing the number of accident and development years, (b) increasing the expected claim frequency, (c) using a mixture of claim types with different distributions, (d) applying a fixed initial "fast track" reserve value to all claims at the date of reporting and for a certain number of months thereafter, (e) modeling some of the parameters using a second-level distribution in the form of an autoregressive process, and (f) introducing an abrupt change in frequency between two of the accident years.

Increasing the number of accident and development years. If all other parameters are unchanged, then increasing the number of development years reduces the magnitude of Stanard's tail adjustment or eliminates it entirely. It also reduces or eliminates the differences between estimates based on paid losses and estimates based on incurred losses, to the extent that these differences simply reflect the relative immaturity of the paid loss data. Increasing the number of accident years reduces sampling error, especially when the loss process is stationary. These are all advantages both for simulation studies and for practical loss reserving.

Exhibit 2 shows results like those of section II of Exhibit 1 (incurred losses, no trend) but this time with ten accident and ten development years, rather than the six and five years, respectively, of Stanard. Note that the latest year (year 9) is 12 months emerged as of the valuation date; no year is entirely in the future. As expected, the standard deviations of the prediction errors are reduced.

The difference between the modified Bornhuetter-Ferguson standard deviations and the Cape Cod standard deviations, which was quite pronounced when using just 5 years' data, is now largely eliminated. This is because the modified Bornhuetter-Ferguson estimates of expected losses give equal weight to each years' chain-ladder developed losses, which are increasingly variable the more recent and immature the year. When only five years are used, the averages are dominated by immature years, whereas when the number of years is doubled to ten, all of the newly added years are mature and the highly variable immature years become a smaller fraction of the whole.

There is also a reduction in the bias of the chain-ladder and Bornhuetter-Ferguson estimates; possible reasons for this are discussed below.

Increasing the expected claim frequency. The model used by Stanard to illustrate his simulation procedures has expected frequency of 40 claims per accident year, with variance 60. This practically assures that every year will have at least twenty or so incurred losses. But Stanard's 18-month mean reporting lag and 12-month mean payment lag lead to a good chance that one or more cells in the first column of the paid or even the incurred loss triangle will be zero, making it impossible to calculate unweighted arithmetic averages of development factors. In fact, when drawing a sample of size 10,000 or so there is some chance that the paid loss triangle in some trial will be zero in all relevant cells in the first column, making it impossible to calculate weighted or unweighted paid loss development factors at that lag.

Having zero cells in the first column is commonplace and is not a significant drawback to practical loss reserving. It simply calls for the exercise of judgement in selecting either the first development factor or the ultimate losses for the latest year, which is usually the only one affected. Indeed, as pointed out by Gogol [11], the non-zero probability of a zero denominator renders the expected value of development factors infinite, and even when the development factors are capped there may remain an upward bias, so the loss reserver must be prepared in any case to adjust or replace outlying factors.

In a simulation study, however, the failure of a reserving method to generate results in one iteration invalidates the means and standard deviations for the entire run. Thus using a small model may limit the scope of a simulation study to those reserving methods that are unlikely to be affected by any zero cells, such as incurred loss development methods using dollar-weighted average development factors. Increasing the expected claim frequency makes it possible to use simulations to compare a wider range of reserving methods, with the understanding that any special advantages or disadvantages of reserving methods in dealing with thin data will have to be considered separately.

Exhibit 3 shows the effect on section II of Exhibit 1 of multiplying the mean and variance of the expected claim frequency by a factor of 10. The increased volume of experience at all lags reduces the standard deviation of prediction errors by removing much of the sampling error; the remaining prediction errors are largely caused by unavoidable process variance.

Mixture of claim types. It would be unusual in practice for a collection of losses to be homogeneous in the sense that its frequencies, severities, lags, and case reserve errors could all be described by simple distributions. But it may well be possible to identify two or more types of claims, each describable by simple distributions, that together account for most of the losses.

Exhibit 4 shows the effect on section II of Exhibit 1 of introducing a second claim type, with the mean and variance of the frequency doubled, the mean and standard deviation of the severity multiplied by .2, and the mean report lag reduced from 18 to 6 months. The payment lag and reserve error distributions are left unchanged. The revised model is therefore a mixture of low frequency, slowly reported, major claims and higher frequency, quickly reported, minor claims.

As in Exhibit 3, but to a lesser extent, the mean prediction errors and standard deviations have generally been reduced, as has the difference between the Modified Bornhuetter-Ferguson and the Cape Cod estimators.

Fast-track reserves. It is customary to assign an average reserve to each claim on first notice and to postpone detailed evaluation for a few weeks to let information accumulate. If the claim closes sooner, no detailed evaluation will be necessary.

To illustrate the effect of this practice, Exhibit 5 shows the results in section II of Exhibit 1 modified by letting each claim have reserve $10,000 for the first three months following reporting. Thereafter, it is reserved in the same manner as in Stanard's paper, including the random reserve error. Note that $10,000 is close to the actual average severity of $10,400 assumed by Stanard.

The result is slightly lower bias and smaller standard deviations than were observed with the original model, especially for years 3 through 5 and for the chain-ladder and modified Bornhuetter-Ferguson estimators.

Second-level distribution of parameters. As set out in the first part of his paper, Stanard's model is quite general, allowing each accident year to be governed by a separate vector of parameters, of unspecified length and values. But, when implemented in his examples, the model is necessarily reduced to specifics, which may favor one estimator over another.

In particular, Stanard's illustrations treat all parameters as fixed or trended; they do not consider the possibility of random shifts in the parameters over time. As a consequence, detrended runoff patterns are stationary. This makes it advantageous to include all accident years in averages. It also gives an edge to those reserve methods, such as Bornhuetter-Ferguson, Cape Cod, and additive, that project the same runoff at each lag independent of accident year.

Exhibit 6 illustrates the effect of allowing the mean frequency, severity, and reserve error factors to vary, following a first-order autoregressive process. This is the same structure that we shall use in Section IV below. Specifically,

$y(1) = m(1)+e(1)$

$y(t+1) = m(t+1)+\phi(y(t)-m(t))+e(t+1)$

$e(t) \sim N(0,s)$, identically and independently of each other and of the y's

Here t is a time index measured in months, y is the vector of means for the parameter in question, ϕ is the lag-1 autocorrelation coefficient of the y's (our program allows ϕ to be between 0 and 1, inclusive; when ϕ equals 1, the model becomes a random walk), m is an assumed vector of means for the second-level distribution, and s is the standard deviation of the second-level distribution. The standard deviations of the parameter itself are a separate assumption, and, if these vary from month to month, the term involving ϕ in the equation for $y(t+1)$ is multiplied by the ratio of the successive standard deviations.

In Exhibit 6, the values of s for frequency, severity, and reserve errors are .1, 400, and .02, respectively, while the value of ϕ is .9816 for all three distributions. This value of ϕ gives an annual, or lag 12, autocorrelation coefficient of about 0.8.

As expected, the chain-ladder method, with its responsiveness to change from accident year to accident year, improves considerably, while the Cape Cod and additive methods perform slightly less well than before. The modified Bornhuetter-Ferguson method improves somewhat, apparently because of the reduced bias and error variance of the chain-ladder at the latest years. In short, using this particular model with shifting parameters brings the methods closer to each other but does not change their ordering.

Abrupt change in conditions. While Exhibit 6 shows the robustness of the Bornhuetter-Ferguson, Cape Cod, and additive methods in the face of a moderately shifting loss process, it is possible that an abrupt change in the accident-year frequencies or severities might so invalidate these methods as to make the chain-ladder perform relatively better. Exhibit 7 shows the results of a simulation like that in Exhibit 6, but with the mean frequency reduced 40% between years 2 and 3. This introduces a large positive bias in the projected ultimate claims for years 3, 4, and 5 and in the reserves, for the estimators that work "across" accident years, while leaving the bias of the chain-ladder estimator essentially unchanged. Even with a shift of this magnitude, the chain-ladder method only outperforms the others in some of the cells.

Many of the above adjustments to Stanard's model reduce the bias as well as the standard errors of prediction for the chain-ladder estimator. Adjustments with this effect include increasing the number of accident years, increasing the claim frequency, using a mixture of claim types, and using fast-track reserves. Gogol [11] attributes the *existence* of the chain-ladder bias to the positive probability of very small denominators in the link ratios. The *magnitude* of the bias appears to be related to the sampling distributions of the denominators of these ratios. When this distribution is tightly clustered relative to a non-zero mean, the bias is small; when the distribution is more diffuse, so that the sample includes more values close to zero, the bias is greater.

In all of the foregoing Exhibits, the link ratios have been estimated from dollar-weighted averages of the observed development factors. Increasing the number of accident years tightens the sampling distribution of each denominator (relative to its mean), as does increasing the claim frequency or introducing a second claim type, simply by increasing the expected number of reported claims. Using fast-track reserves tightens the sampling distribution at the early lags by reducing the variance of the case reserves.

IV. DETAILED MODEL FOR SUBSEQUENT SIMULATIONS

Exhibit 8 displays the model of the loss process used in the comparisons of reserving techniques to follow. It is more complex in several respects than the model employed by Stanard for his illustrations.

First, the new model involves a greater number of average annual claims, a shorter mean report lag, and a longer mean payment lag, so as to resemble primary coverage on a midsized policyholder rather than a working excess reinsurance treaty. Second, the new model involves a mixture of two claim types, with different frequency, severity, and lag distributions. Third, the means of the frequency, report lag, and reserve error distributions for both claim types are themselves generated by autoregressive processes as described above. Fourth, the number of accident years through the evaluation date has been increased from five to eight, making the effect of Stanard's tail adjustment negligible. Following Stanard, we number the years 0 through 7.

Finally, the year beginning beyond the evaluation date has been eliminated. While several reserving techniques can indeed project future incurral periods, such periods do not contribute to the loss reserves, nor is it customary to rely entirely on the reserve estimator to project losses for a future rating period.

Occasionally, we may introduce trend assumptions or other variations of the above model to illustrate the sensitivity of the techniques being discussed to particular conditions in the data.

V. LOSS RESERVING QUESTIONS ADDRESSED VIA SIMULATION

Now we use our extended simulation technique to investigate several questions that arise in practical loss reserving. While simulation can only answer such questions approximately, and only for a particular loss model, the results below, and the underlying procedure, may help guide the loss reserver in choosing the most promising estimators and judgement adjustments for actual cases.

Paid versus Incurred Loss Development

When data permits, the first question is often whether to estimate ultimate losses from the paid or from the incurred loss triangle. Exhibit 9 addresses this question using the same four reserving methods tested by Stanard, applied to the new model of the loss process.

In Exhibit 9 and those that follow, there are two new lines, labeled “# best” and “rank sum”.

The “# best” row reports the number of trials in which each reserving method produced the least absolute prediction error, for total reserves, among the methods considered. While interesting, this is only a rough guide to the best-performing methods. When two or more methods are similar, they may split the majority of the “best” rankings and allow another, less successful, method to appear better than it would if only one of the similar methods were included. Since the runoff process itself has a large variance, a biased or erratic reserving method that tends to produce the highest or lowest estimates in the set may still rank “best” in a substantial fraction, albeit less than half, of the trials. Also because of the large process variance, if one reserving method tends to be bracketed by two other methods, it will seldom produce the best estimate for a single trial but may nevertheless have the smallest mean squared error.

The “rank sum” row reports the sum, over all ranks *r*, of *r* times the number of times each method had rank *r* in absolute prediction error for reserves. This is a more useful measure of performance than the foregoing.. Lower scores are better, with a minimum possible score of 1 times the number of trials, here 10,000. Ordering the methods on this measure is equivalent to ordering them on the number of “wins” in 10,000 sets of all possible pairwise comparisons.

With the higher claim frequency, smaller frequency standard deviation, and larger number of accident years, the biases in Exhibit 9 are smaller than those observed in Stanard’s model, as are most of the standard deviations of the prediction errors. For the early accident years, the standard deviations are nearly equal for all the incurred-loss reserving methods. This does not mean that each reserving method gives identical predictions for each sample data set, but simply that the variance of the predictions is insignificant compared to the unavoidable variance of the claims process itself.

Some of the paid-loss reserving methods have non-zero bias even for year 0, because they use less mature data than the incurred methods and Stanard’s tail adjustment, based on incurred losses, does not remove all of the paid tail. The first three of the paid-loss methods are adjusted for development in the tail beyond the last known lag, using the ratio of incurred to paid losses at that lag in year 0, which for this simulation averages about 1.01; the fourth paid-loss method (additive) does not contain this adjustment. Its projections are therefore slightly understated relative to the other paid-loss methods. If the paid-loss additive results were adjusted by this factor, they would exhibit biases only slightly worse than those of the best-performing method, the incurred-loss additive.

For this particular model, the estimates based on incurred losses generally outperform those based on paid losses. But if case reserve adequacy were to change abruptly, then reserves based on incurred losses might be less accurate than reserves based on paid losses. Even if case reserving were simply biased and/or highly variable, much of the advantage of incurred loss estimators might be overcome by noise. To illustrate the latter possibility, Exhibit 10 repeats the calculations of Exhibit 9 but with the mean and standard deviation of severities increased by 25%, a probability of 20% of closure without payment introduced, the mean reserve error factor changed from 1 to 1.25, the standard deviation of the reserve errors changed from about 1.414 to 5, and the reserve on first notice changed from $4,000 to $5,000.

The probability of closure without payment together with the increased severity leave the mean severity over all claims unchanged, but the absence of a corresponding assumed probability of closure without payment in the case reserving formula means that the case reserves are conservative by 25%. Additional conservatism is provided by the mean reserve error factor's being greater than 1 and by the increase in the reserve on first notice, and the whole case reserving process is made more erratic by the increase in its standard deviation.

Comparing Exhibit 10 with Exhibit 9, we see a general worsening of the bias and standard deviation for the incurred loss estimators, reversing, in many cells, the advantage of the incurred over the paid loss estimates in Exhibit 9. The reordering of the "rank sum" row is especially striking.

Methods of Averaging Development Factors

The question of how best to average paid or incurred development factors has received considerable attention in the literature [14][16].

Exhibit 11 compares four different methods of averaging: arithmetic, dollar-weighted, least-squares, and geometric. The first two of these averages have been in common use for many years; the last two were suggested by Murphy [14]. The arithmetic average produces greater bias than the others and somewhat greater standard deviation. There is little difference among the other three, except that least-squares averaging appears to reverse the usual positive biases for recent accident years.

Depth of Averaging

Another common question facing the loss reserve practitioner is how many years' development factors to include in each average. It is commonly believed that development factors change over time with changes in claims handling practices and case reserving guidelines and personnel. Therefore the use of recent years only in the averages may produce factors closer to those presently being experienced than the use of all available years, albeit at the expense of greater sampling error.

Our present model is not ideally suited to illustrate the possible superiority of smaller but more recent averages, as it treats the reporting lag pattern as changing by accident year, and the payment lag pattern by report year, rather than allowing both to change by calendar year. Exhibit 12 illustrates four different depths of averaging, with the rest of the assumptions the same as in the incurred chain-ladder column of Exhibit 9. In this case, despite the shifting report-lag parameter, the all-year average remains best, the biases, standard errors of prediction, and rank sums all increasing as the number of years in the averages decreases.

Exclusion of Extreme Values from Averages

Excluding the greatest and least values from each column of development factors is commonly thought to make the resulting averages more robust estimators. This is borne out by comparing

Exhibit 13, which excludes extreme values, with Exhibits 11 and 12. Note that Exhibit 13 includes the dollar-weighted average, with which the exclusion of extreme values is not often combined. The standard deviations of the reserves and most of the accident year ultimates are reduced modestly, and the biases are reduced substantially. This technique should be used with caution, however, for reasons given by Wu [24].

Judgement Maximum Development Factors

In his discussion of Murphy [14], Gogol [11] assumes that the loss reserver would replace infinite or very large average development factors with judgement maximum values, to avoid the problem of infinite expectations. He goes on to show that, under mild conditions and reasonable assumptions regarding the loss process, there is still an upward bias in single or linked development factors.

Exhibit 14 shows the results of the incurred chain-ladder, modified Bornhuetter-Ferguson, and Cape Cod estimators when the development factors are capped at 5.00 for 12 to 24 months and at 2.00 for other lags. The effect on bias and standard deviation is slight in this case, because the size of the case makes these maxima difficult to pierce, but the principle of stabilizing the estimates by capping the development factors appears sound.

Weighted Average of Calculated and Reference Development Factors

Actuaries often append companywide or industrywide development factors to loss development printouts as a guide in choosing “selected” factors. This judgement procedure may be automated in a natural way by taking, at each lag, a weighted average of the calculated and reference development factors, where the weights z for the calculated factors are related to the volume v of the data by a credibility formula $z = v/(v+k)$. Exhibit 15 illustrates this procedure applied to the same three estimators as in Exhibit 14. Here, the measure of volume v is the total losses available, to the selected depth of averaging, in the denominators of the observed development factors; *k* is $600,000, chosen purely for illustrative purposes; and the reference factors are .42, .83, .95, .98, and 1.00. These reference factors were determined by sampling with a larger claim frequency and with fixed lag patterns to simulate "industry" data.

Note that, when dollar-weighted averaging of development factors is used, this procedure automatically handles the infinite-expectation problem, as lags with zero denominators have credibility zero, causing the reference factor to be substituted.

The stabilizing effect of the reference development factors is quite evident in Exhibit 15, where the bias of all three tested estimators has nearly been eliminated and the standard deviations reduced.

Graduation of Development Factors

When triangles of paid or incurred losses are completed using the chain-ladder technique, the results are sometimes faulted for perpetuating chance fluctuations in the observed experience in later lags where the data is thin. A possible solution to this problem is to graduate the later development factors, constraining them to bear a reasonable relationship to each other. Some authors [19][23] have proposed fitting curves to do this and/or to extrapolate into the tail beyond the known data.

Exhibit 16 illustrates the fitting of Richard Sherman's inverse power curve [19], to development factors using both paid and incurred chain-ladder estimators. Each graduation is applied first to all development factors, and second to all development factors after the first. The results support the latter procedure, at least for paid losses; fitting to all points tends to depress ultimate claims and reserves except for the latest year, apparently because the inverse power curve decays too rapidly to fit the data well at the early lags.

Choice of Incurral and Development Intervals

Although loss triangles for P/C lines are commonly presented accident year versus development year (with the latter possibly offset from the former by some months), data may actually be available in finer intervals, such as accident year versus development quarter or accident quarter versus development quarter. Finer development intervals are obviously useful for making off-anniversary loss projections, and finer accident intervals for accommodating rapidly changing exposures. More interesting is the question whether smaller intervals on one or both axes improve the quality of predictions when exposures are not changing.

Exhibit 17 addresses this question by applying the paid and incurred chain-ladder estimators to the same portfolio of claims as in Exhibit 9, but recompiled into cells, first by accident year versus development quarter, and second by accident quarter versus development quarter. The results are mixed. Shortening the development interval to three months while leaving the accident interval at one year has practically no effect on the prediction errors. However, shortening both the accident and development intervals to three months makes the biases and standard deviations *worse* with the paid chain-ladder estimator, but slightly *better* with the incurred chain-ladder estimator.

Pure Bornhuetter-Ferguson Estimates

In their 1972 paper, Ronald Bornhuetter and Ronald Ferguson assumed that the loss reserver would use *a priori* values for expected losses or expected loss ratios, rather than deriving these from the loss data itself. Exhibit 18 compares this "pure" Bornhuetter-Ferguson method with the same Cape Cod calculations shown in Exhibit 9, which estimate expected losses from the data and have only a small bias. The pure Bornhuetter-Ferguson estimates are better than the Cape Cod estimates both for paid and for incurred loss projections. Of course the accuracy of the *a* *priori* value of expected losses is crucial. Here we used $480 per unit of exposure, a slight understatement.

Gluck's Decay Factor

In 1997, Spencer Gluck [10] proposed using an exponential decay factor, between 0 and 1, to give less weight the more remote the accident year when estimating expected losses for a given year in the Cape Cod method. Thus estimated expected losses are not constant across accident years but may drift over time, making them responsive to trends or to other variations in the underlying data. Gluck points out that a decay factor of 1 is equivalent to the straight Cape Cod method while a decay factor of 0 is equivalent to the chain-ladder. He gives a technique for estimating an optimal decay factor from the data but generally uses judgement values from 50% to 100% with a default value of 75%.

Exhibit 19 shows the results of setting Gluck's decay factor to .9, .8, .7, and .6 in paid and incurred Cape Cod calculations; these may be compared with the pure Cape Cod and chain-ladder results shown in Exhibit 9. In this case, there is apparently not enough drift to the mean frequency to cause any of the estimators using Gluck's weights to outperform the pure Cape Cod. The biases and standard deviations generally rise, though slowly, as D decreases from 1 (Cape Cod) to 0.6.

Mixtures of Bornhuetter-Ferguson and Chain-ladder Estimates

Exhibit 20 investigates the proposition that the chain-ladder estimates might be better for mature accident years and Bornhuetter-Ferguson (actually Cape Cod) estimates better for immature years. It uses a weighted average of the two estimates, with the weight given to the Cape Cod estimate equalling the expected fraction of each accident year's ultimate claims that is unemerged as of the latest known lag. For this data, the performance of the mixed method is much better than the straight chain-ladder but it shows no advantage over the straight Cape Cod estimator.

Linear Combination of Estimates Based on Paid and on Incurred Losses

Exhibit 21 illustrates the results of using weighted averages of paid and incurred chain-ladder and of paid and incurred Cape Cod reserve estimates. As seems to be common in practice, weights of .5 and .5 have been used in both cases. The average of the paid and incurred Cape Cod estimators generally outperforms either separately, in terms of standard deviations and rankings, but the average of the paid and incurred chain-ladder estimators performs only about as well as the incurred estimator alone.

Performance in the Presence of Undetected Trend

Exhibit 22 illustrates the performance of all eight reserving methods shown in Exhibit 9, but this time with the data incorporating a severity trend of 8%, which is assumed to be undetected (or ignored) and is not adjusted for in any of the estimators.

Severity trend affects the paid or incurred chain-ladder estimators in two ways. First, *accident-year* inflation tends to give greater weight to the more recent development factors in dollar-weighted averages. In principle, this makes the averages more responsive to shifts in lag patterns, but, with the modest trend assumed here, the effect is minor. Second, *post-incurral* inflation increases all development factors. For paid-loss development, with a constant trend, the factors increase in parallel with inflating losses; there is little impact on percentage prediction errors. For incurred-loss development, however, the factors increase only to the extent that claims are closed at each lag, since our model (in this respect unchanged from Stanard) postulates no change in the case reserve until settlement. As some claims are not settled by the last lag in our data, the calculated ultimate development factors do not increase in parallel with the inflating losses, and percentage prediction errors are forced somewhat downward, as can be seen in Exhibit 22.

Our severity trend assumption has a greater impact on those estimators that average losses across accident years. The modified Bornhuetter-Ferguson and Cape Cod expected losses are averages over all years, and are therefore too low for the later ones, which account for most of the reserves. The additive projected losses at each lag are averages over the earlier years only, and are therefore too low for all years. Both patterns are evident in Exhibit 22.

Methods of Accommodating Trend

For the above reasons, it may be felt desirable to adjust for trend when using those reserving methods that take averages across accident years.

It is possible to adjust any reserve estimator for inflation by deflating the losses entering the calculation to a common index level and then reflating the projected losses. But simpler means are sometimes available. For methods that relate losses to exposure, it may be sufficient to use an inflating exposure measure, such as premium, or a synthetic measure incorporating trend, such as units of exposure multiplied by a trend factor. For the Bornhuetter-Ferguson method and its variations, it is also possible to trend the underlying chain-ladder losses, loss ratios, or loss rates internally, before averaging to obtain expected losses, and then reverse the trending at the end of the calculations.

Exhibit 23 compares the paid and incurred Cape Cod variation of the Bornhuetter-Ferguson method with no inflation adjustment, with internal adjustment, and with deflation and reflation of losses. The simulated trend is 8% per annum with 50% of the trend persisting between occurrence and settlement; the adjustments within the reserving methods assume 8% trend. The unadjusted columns repeat the biases shown in Exhibit 22. Both of the adjustment methods greatly reduce these biases. Surprisingly, the use of trending internal to the Cape Cod calculation slightly outperforms the more complex deflating-reflating procedure.

Frequency x Severity Methods

When claim counts are available, one may estimate ultimate frequency and ultimate severity separately, using a chain-ladder estimator for each, and then combine them to obtain an estimate of ultimate losses. Exhibit 24 shows the results of using this "frequency x severity" estimator. When applied to paid losses, there is some improvement over the chain-ladder estimator, but when applied to incurred losses, there is practically no change. The improvement in the paid loss case may be due to the fact that our implementation of the paid frequency x severity estimator uses reported rather than closed claim counts, and thus is, in a sense, a hybrid of paid and incurred estimators.

Average Payment Development and Case Reserve Development

Robert J. Finger's average payment development method [7] projects ultimate claims from ultimate average payments, estimated by developing average payments per closed claim (assuming constant ratios of ultimate average payments to average payments at each lag), *or* from trended average "implied reserves" per open claim (assuming that average reserves change from accident year to accident year at a known inflationary rate), *or* as a weighted average of these two estimates. Calculations proceed stepwise from the earliest accident year, making use at each step of the ratios and reserves implied by the ultimate average payments for earlier accident years.

Finger's somewhat similar case reserve development method [8] assumes that the redundancy (or inadequacy) in the case reserves for a given accident year is a function of the lag, and may be estimated from the average redundancy of the reserves for earlier accident years at the same lag.

Redundancy includes both development of known claims and emergence and development of IBNR claims. Calculations proceed stepwise from the earliest accident year, making use at each step of the reserves for all earlier accident years implied by those years' paid losses and calculated ultimate losses.

Exhibit 25 illustrates these two estimators. Average payment development exhibits a modest positive bias but is better than the paid chain-ladder estimator, especially for recent accident years and especially in terms of standard deviations. Case reserve development, on the other hand, does not perform as well as the incurred chain-ladder estimator, nor does it show any particular advantage over the paid chain-ladder estimator.

Clarke's London Market Method

Harold E. Clarke [5] fits three-parameter curves, of the form A(1-exp(-t/B)^C), to partial loss ratios at successive stages of development of each incurral period and obtains preliminary estimates of ultimate loss ratios from the asymptotes of these curves. He then obtains final estimates of ultimate loss ratios by regression of the preliminary estimates against partial loss ratios at each lag. Partial loss ratios are emerged claims divided by emerged premium, appropriate in the London Market reinsurance context; our implementation uses ultimate premium or any other measure of exposure, or an assumption of constant exposure, so that "loss ratio" may also mean pure premium or total losses.

Exhibit 26 shows that Clarke's method has a pronounced negative bias when applied to paid losses and a pronounced positive bias when applied to incurred losses, at least for this sample of data sets; the standard deviations are moderate.

Clarke emphasizes that his approach is interactive and he uses curve plots to aid the loss reserver in the exercise of judgement. This interactivity is lost in a simulation, where explicit formulas must substitute for judgement; we expect the results of applying Clarke's method interactively would be better than those shown here.

Loglinear Models

R.J. Verrall [22] (see also [12]) and Benjamin Zehnwirth [6][25][21] have created statistical estimators which apply linear models to the logarithms of the incremental paid losses or loss ratios. As a consequence, the fitted mean for a cell depends on the variance as well as the mean of the transformed value for that cell, and the predicted mean may be much greater than the mean simulated by a process which does not distribute the cell values lognormally. These loglinear models are also somewhat awkward in dealing with zero or negative incremental cells.

Verrall fits a linear model with row and column (lag) effects to the transformed data, using maximum likelihood, Bayesian, and empirical Bayesian estimation. Zehnwirth fits three-parameter Hoerl curves to the lag patterns, using adaptive filtering to let the parameters vary over time. His procedure is intended to be interactive, with successive adjustments of the second-level variances of the parameters to achieve an appropriate balance between fit and parsimony. As with Clarke's method, this interactivity is lost in a simulation.

Exhibit 27 shows Verrall's maximum-likelihood, Bayesian, and empirical Bayesian chain-ladder models, and two runs of Zehnwirth: one using fixed parameters and the other allowing the parameters to vary modestly for the first five years. For this particular simulation, Verrall's empirical Bayesian model and Zehnwirth's fixed-parameter model do well (among paid loss estimators) in terms of both bias and standard deviations.

Berquist-Sherman Adjustments

Berquist and Sherman [1] introduced several means of adjusting loss development calculations for changes in conditions, two of which, the disposal rate adjustment and the case reserve adequacy adjustment, may be applied to the data entering any reserve calculation. The disposal rate adjustment uses interpolation [9] to place paid claims or other quantities on a level consistent with the latest ratio at each lag of closed claim counts to projected ultimate claim counts. The case reserve adequacy adjustment adjusts average case reserves per open claim to the latest level at each lag, de-trended at an assumed annual rate, derives the total case reserves, and adds these to paid claims to obtain adjusted case-basis incurred claims.

Exhibit 28 illustrates the disposal rate adjustment applied to the paid chain-ladder estimator and both adjustments applied singly and together to the incurred chain-ladder estimator. While the disposal rate adjustment does not improve the results, the case reserve adequacy adjustment, used by itself, does reduce the bias in the incurred chain-ladder estimator, at the expense of somewhat higher standard deviations for most years.

Combination of Adjustments

In the foregoing paragraphs we have seen that the use of incurred loss development, dollar-weighted or least-squares averages, exclusion of extreme values, a weighted average of calculated and reference development factors, or graduation of development factors each improve the efficiency of the basic chain-ladder estimator, at least for our particular sample of simulated data sets. In Exhibit 29, we examine the effect of various *combinations* of these selections and adjustments on the incurred chain-ladder estimator. Of those shown, the most successful combines dollar-weighted averages and reference factors. Exhibit 29 also shows the result of this same combination on the Cape Cod estimator: exceptionally small biases at all but the latest accident years and exceptionally low standard deviations both for the latest accident years and for the reserves.

Means and Standard Deviations of Prediction Errors Conditional on the Known Data

To use Monte-Carlo simulation to investigate the distribution of reserve errors given a particular set of *known* data, it would be most useful to be able to draw a sample all of whose points contained exactly the same known data and differed only in their runoffs. Unfortunately, this is not practical; the same loss process governs both the known and the runoff data, and the probability of obtaining precisely the same known data on two runs is exceedingly small. One way around this problem is to draw a sample of points close, if not identical, to the given set of known data.

Exhibit 30 returns to the estimators of Exhibit 9 and illustrates the effect of screening each sample point for proximity to a particular set of “case data”, here the first sample data set with dollar values rounded to the nearest thousand. Proximity is defined by a Euclidean metric, the square root of the sum of squares of the differences between the two data sets in the latest cumulative paid losses for each accident year. This is actually only a partial metric as it does not distinguish between sample data sets that differ in other ways than the latest cumulative paid losses.

Even with such a weak definition of closeness, we must still allow considerable latitude if we wish to include a reasonable proportion of the trial sample points in the final sample. Defining an acceptance radius as a multiple of the norm of the “case data” using the metric described above, we find that a radius of .4 results in the acceptance of some one-seventh of the trial sample points. For this particular sample, “case data”, and acceptance radius, the most of the positive reserve biases are reduced or negative biases increased, and the standard deviations of the reserve errors are generally reduced.

The above procedure is a special case of sampling the distribution conditional on satisfaction of an arbitrary Boolean relation on the known data. Exhibit 31 illustrates the incurred-loss estimators from Exhibit 9, applied to those sample data sets where the latest accident year incurred losses at the first evaluation are within 25% of the average incurred losses at first evaluation of the earlier accident years. This excludes those situations where an unusually high or low latest accident year would attract the attention of the loss reserver and lead to a judgement adjustment. As might be expected, the main effect of this particular Boolean screen is to eliminate completely, indeed reverse, the positive bias in the chain-ladder ultimate claims for the latest accident year. It also reduces the positive bias, or increases the negative bias, in the methods which involve averages of losses or loss ratios across accident years.

Detailed Distributions of Prediction Errors

Stanard [20] remarks that, with the chain-ladder estimator, a few sample points create very large positive prediction errors and contribute to the observed positive bias, whereas most of the sample points actually produce negative prediction errors. This is borne out by our tests.

When generating Section II of Exhibit 1, we obtained details on the distribution of prediction errors for year 4. While the mean error for the chain-ladder estimator is about $142,000 (before Stanard's tail adjustment), the median is about -$64,000, and 5,935 of the 10,000 sample points have prediction errors less than zero. The three largest positive errors are about $81,300,000, $56,800,000, and $26,500,000, while the three largest negative errors are about -$5,300,000, -$3,000,000, and -$2,900,000. The mode of the distribution, when grouped into bands of width

$100,000, is the band from -$200,000 to -$100,000. The other three estimators in Exhibit 1 produce more symmetrical error distributions.

When generating Exhibit 9, we obtained details on the distribution of prediction errors for total reserves. While the mean error is about $75,000 for the incurred chain-ladder estimator, the median is about -$8,300 and 5,072 of the sample points have prediction errors less than zero. The three largest positive errors are about $9,000,000, $7,400,000, and $5,900,000, while the three largest negative errors are about -$3,400,000, -$2,500,000, and -$2,400,000. The mode of the distribution, when grouped into bands of width $200,000, is the band from -$200,000 to $0. The other three incurred-loss estimators produce more symmetrical and concentrated distributions than this, while the four paid-loss estimators follow a pattern similar to the incurred-loss estimators, but considerably less concentrated.

For these two tests, we also defined an independent variable against which to tabulate the prediction errors; this variable was the incurred losses at first evaluation for the latest accident year (year 4 in Exhibit 1 and year 7 in Exhibit 9).

Since the chain-ladder estimate of the latest accident year is a multiple of the claims at the first evaluation, a linear regression of the prediction errors of ultimate claims against the claims at first evaluation should certainly be significant. For the simulation in Exhibit 1, the coefficient of variation R^2 is 39.8%, the t statistic for the slope is 81.35 (df=9,998) and the slope itself is 4.868; in effect this is an average ultimate development factor for lag 12 months. This relationship is obviously significant despite the fact that the prediction errors do not display equal variances when grouped by ranges of values of the independent variable, but instead increase almost monotonically. The other reserving methods also produce extremely significant regressions, though with smaller slopes, as would be expected given their indirect dependence on the emerged claims for the latest accident year.

For the simulation in Exhibit 9, where the dependent variable is the reserve rather than ultimate claims for the latest accident year, we expect the strength of the relationship to be lower though it should still be significant and the slope positive. Indeed, R^2 is now just 19.5%, t is 49.3 and the slope is 1.983. The assumption of equal variances is again violated, with variance increasing at the upper end of the range of the independent variable. The regressions for the other estimators considered, including all the paid methods, are significant but weaker than for the incurred chain-ladder. Interestingly, the regressions for the paid Cape Cod and additive estimators have a small negative slope.

VI. CONCLUSIONS

Several judgement adjustments to the chain-ladder and related reserving methods have been shown to improve its efficiency when tested using simulated loss histories involving mixtures of claim types and serial autocorrelation of parameters. In most scenarios, the chain-ladder method with these adjustments does not show any clear superiority to Buehlmann’s additive method, the Cape Cod variation of the Bornhuetter-Ferguson method or the Cape Cod variation with the same adjustments. Scenarios have been suggested, however, in which the relative performance of the methods might be reversed. Other reserving methods, including loglinear models, have been tested and found promising. In some cases, their promise may only be fully realized when they are used interactively rather than reduced to rules suitable for a simulation. Finally, our procedure allows the loss reserver to choose the best reserve estimator for a given case, from a set of candidates, by modeling the loss process for the case in some detail and then using simulation to compare the performance of the different estimators.

References

[1] James R. Berquist and Richard E. Sherman, “Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach”, *PCAS* LXIV, 1977, pp. 123-184. In addition to their disposal rate and case reserve adequacy adjustments, it is from Berquist and Sherman’s paper that we adopted the term “reserve test” for each separate estimate of reserves and incurred claims in a set of such estimates.

[2] Ronald L. Bornhuetter and Ronald E. Ferguson, “The Actuary and IBNR”, *PCAS* LIX, 1972, pp.181-195

[3] Hans Buehlmann, “Estimation of IBNR Reserves by the Methods Chain Ladder, Cape Cod and Complementary Loss Ratio”, unpublished, 1983

[4] 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

[5] H. E. Clarke, "Recent Developments in Reserving for Losses in the London Reinsurance Market", *PCAS* LXXV, 1988, pp. 1-48

[6] Piet de Jong and B. Zehnwirth, "Claims Reserving, State Space Models and the Kalman Filter" *JIA* 110, no. 1, 1983, pp. 157-181

[7] Robert J. Finger, “Average Payment Development”, private communication, 1987

[8] Robert J. Finger, “Case Reserve Development”, private communication, 1987

[9] Edward W. Ford and Mark Doepke, "5B Basic Case Study", *1986 Casualty Loss Reserve Seminar Transcript*, pp. 716-760. In their Case Study No. 1, Ford and Doepke apply a Berquist-Sherman disposal rate adjustment using linear interpolation, rather than the more involved selection of exponential curves used by Berquist and Sherman themselves. Our disposal rate adjustments follow Ford and Doepke’s approach.

[10] Spencer M. Gluck, “Balancing Development and Trend in Loss Reserve Analysis”, presented at the 1997 Annual Meeting of the Casualty Actuarial Society, 1997

[11] Daniel F. Gogol, Discussion of Murphy [14], *PCAS* LXXXII, 1995, pp. 72-77

[12] D.V. Lindley and A.F.M. Smith, “Bayes Estimates for the Linear Model”, *Journal of the Royal Statistical Society*, Series B, Volume 34, 1972, pp. 1-41

[13] Thomas Mack, “Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates”, *ASTIN Bulletin* 23, number 2, 1993, pp.213-226

[14] Daniel M. Murphy, “Unbiased Loss Development Factors”, *PCAS* LXXXI, 1994, pp. 154-222

[15] Prakesh Narayan and Thomas V. Warthen, III, “A Comparative Study of the Performance of Loss Reserving Methods through Simulation”, *Casualty Actuarial Society Forum*, Summer 1997, Volume 1, pp. 175-195

[16] Edward F. Peck, Discussion of Stanard [20], *PCAS* LXXXII, 1995, pp. 104-120

[17] John P. Robertson, Discussion of Stanard [20], *PCAS* LXXII, 1985, pp. 149-153.

[18] John W. Rollins, “Performance Testing Aggregate and Structural Reserving Methods: A Simulation Approach”, *Casualty Actuarial Society Forum*, Summer 1997, Volume 1, pp. 137-173

[19] Richard E. Sherman, "Extrapolating, Smoothing, and Interpolating Development Factors", *PCAS* LXXI, 1984, pp. 122-155. Sherman introduces the inverse power curve y=1+a(t+c)^(-b) used by our system as one means of graduating and extrapolating development factors.

[20] James N. Stanard, “A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques”, *PCAS* LXXII, 1985, pp. 124-148.

[21] G.C. Taylor, *Claim Reserving in Non-Life Insurance*, 1986, pp. 188-195

[22] R.J. Verrall, “Bayes and Empirical Bayes Estimation for the Chain Ladder Model”, *ASTIN Bulletin* 20, number 2, 1990, pp. 217-243

[23] Alfred O. Weller, “Generalized Bondy Development”, *ARCH* 1989.2, pp. 1-25

[24] Cheng-Sheng Peter Wu, “Downward Bias of Using High-Low Averages for Loss Development Factors”, *Casualty Actuarial Society Forum*, Summer 1997, Volume 1, pp. 197-240

[25] B. Zehnwirth, *ICRFS Manual*, Version 3.0, Volume II, 1986

Exhibit 1. Replication of Stanard's Results Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size =10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
I. Claim Counts Only								
0	0	3.6	0	3.6	0	3.6	0	3.6
1	.2	6.3	.2	6.3	.2	6.2	0	6.1
2	.3	9.2	.2	8.8	.1	8.7	-.1	8.6
3	.7	14.1	.4	12.3	.2	12.0	0	11.9
4	1.4	29.4	.6	17.8	.3	17.3	.1	17.3
5	n.a.	n.a.	.3	21.4	0	20.8	-.2	20.8
Reserve	1.1	16.7	.6	12.9	.3	12.3	0	12.1
II. Incurred Losses								
0	0	23.4	0	23.4	0	23.4	0	23.4
1	3.0	54.7	3.4	67.7	2.2	49.3	.2	51.9
2	7.3	126.3	6.4	76.5	3.5	53.8	.3	53.6
3	12.9	106.3	9.8	95.0	4.5	65.2	.5	63.6
4	38.7	334.7	12.1	116.2	3.6	77.9	-.1	74.6
5	n.a.	n.a.	12.1	112.1	2.5	76.1	-.1	75.3
Reserve	26.4	185.7	13.4	126.8	5.8	76.1	.3	75.1
III. Severity Trend 8%; a=.5								
0	0	24.5	0	24.5	0	24.5	0	24.5
1	3.2	56.9	4.7	78.3	2.6	50.9	-0.2	51.9
2	7.6	137.5	7.5	86.6	2.8	53.8	-1.7	51.9
3	13.3	111.1	8.1	102.6	.3	64.4	-5.2	61.6
4	39.6	360.5	2.5	116.4	-8.9	73.3	-13.8	69.5
5	n.a.	n.a.	-7.7	107.3	-19.7	70.6	-22.7	69.4
Reserve	27.2	201.0	9.0	133.0	-1.9	69.4	-9.1	65.5
IV. Severity Trend 8%; a =.5; Inflation Adjustment in Estimates 8%								
0	.3	25.1	.3	25.1	.3	25.1	0.3	25.1
1	3.6	58.1	3.9	72.2	2.6	52.4	.5	55.0
2	8.0	134.4	7.0	81.3	3.9	57.2	.5	56.8
3	13.7	111.8	10.6	100.7	5.1	68.7	1.0	68.0
4	39.4	340.5	12.9	121.8	4.1	80.2	.2	76.4
5	n.a.	n.a.	12.8	115.7	3.0	76.7	.3	75.9
Reserve	27.7	192.6	14.3	128.9	6.4	76.1	1.0	74.7
V. Severity Trend 10%, then 6%; a =.5; Inflation Adjustment 10%								
0	.4	25.4	.4	25.4	.4	25.4	.4	25.4
1	3.8	59.2	4.2	73.7	2.9	53.3	.7	56.0
2	8.9	137.9	7.9	83.2	4.7	57.8	1.1	57.5
3	15.8	115.1	12.4	103.3	6.7	69.5	2.2	74.2
4	44.1	355.2	16.4	126.7	7.3	82.2	3.1	78.6
5	n.a.	n.a.	19.4	122.4	8.9	78.8	5.9	78.2
Reserve	31.2	201.3	17.1	133.7	9.0	78.4	3.1	78.1

Exhibit 2. Stanard's Model Extended to 10 Years by 10 Years Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size =10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	0	.6	0	.6	0	.6	0	.6
1	0	2.2	0	2.2	0	2.1	0	3.3
2	0	4.9	0	4.9	0	4.9	-.1	5.8
3	.2	12.2	.2	11.9	.2	11.8	0	12.3
4	.4	21.4	.3	21.5	.3	21.4	.1	21.7
5	1.0	26.2	.5	26.0	.3	25.0	-.1	26.1
6	.8	58.5	.4	57.2	.1	56.7	-.5	57.1
7	2.1	52.1	.8	44.4	.2	43.1	-.6	43.0
8	4.1	95.6	1.3	62.5	.2	59.7	-.7	58.3
9	18.8	186.8	3.9	68.5	1.8	62.4	1.0	61.1
Reserve	11.1	100.8	3.0	63.9	1.2	58.0	-.3	57.8

Exhibit 3. Stanard's Model with Frequencies Increased by a Factor of 10 Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size =10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	0	7.6	0	7.6	0	7.6	0	7.6
1	0.3	14.0	.2	13.8	.1	13.7	-.2	13.7
2	1.1	19.8	.8	18.4	.6	18.1	0	17.9
3	2.2	34.5	1.5	27.3	.9	26.4	.2	25.0
4	5.1	56.7	1.7	29.4	.7	26.1	0	24.8
5	n.a.	n.a.	2.0	27.4	.7	24.8	.2	24.5
Reserve	3.7	35.4	1.8	27.8	1.0	25.8	0	24.8

Exhibit 4. Stanard's Model with Mixture of Two Claim Types Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size =10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	0	18.1	0	18.1	0	18.1	0	18.1
1	1.4	28.2	1.1	29.7	.9	27.3	0	30.4
2	2.6	43.1	2.0	39.4	1.5	37.7	-.2	38.1
3	4.2	57.7	2.1	50.5	1.1	47.5	-1.1	47.8
4	12.5	109.8	4.4	61.9	2.1	55.9	0	54.6
5	n.a.	n.a.	3.5	59.1	.5	54.0	-.9	53.8
Reserve	9.8	77.7	4.5	63.8	2.7	57.5	-.6	58.5

Exhibit 5. Stanard's Model with Fast Track Case Reserve of $10,000 for First Three Months Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size =10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	0	23.4	0	23.4	0	23.4	0	23.4
1	2.8	53.8	3.0	66.0	2.0	49.3	.3	48.7
2	6.5	117.9	4.7	73.0	2.8	56.0	0	55.4
3	10.4	97.6	6.6	82.7	3.3	63.5	0	62.1
4	19.9	206.2	7.5	93.6	2.5	72.2	-.3	70.5
5	n.a.	n.a.	7.7	95.8	1.9	75.4	-.2	74.6
Reserve	16.8	141.6	9.3	111.8	4.5	74.8	.0	72.9

Exhibit 6. Stanard's Model with Mean Frequency, Severity, and Reserve Errors Drawn from Second-level Distributions; ϕ=.9816 Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size =10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	19.7	.0	19.7	.0	19.7	.0	19.7
1	2.6	50.9	2.4	46.2	1.6	41.2	-.6	47.6
2	6.1	68.4	5.3	65.5	3.2	57.9	-.1	58.3
3	13.3	135.8	9.1	96.3	4.9	85.6	.7	82.6
4	35.8	251.7	12.8	106.1	5.4	83.3	1.5	79.4
5	n.a.	n.a.	10.4	106.5	2.0	90.2	-.6	89.7
Reserve	24.5	146.0	12.6	103.8	6.4	82.9	.6	82.4

Exhibit 7. Stanard's Model with Mean Frequency, Severity, and Reserve Errors Drawn from Second-level Distributions; ϕ=.9816 Mean Frequency Reduced 40% Starting with Year 3 Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size =10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	24.4	.0	24.4	.0	24.4	.0	24.4
1	1.9	39.5	.5	39.6	.6	37.4	-.7	43.6
2	6.6	72.6	2.1	60.6	2.0	57.3	.2	61.5
3	12.4	220.4	19.5	158.3	19.2	150.5	17.6	143.7
4	33.4	292.7	32.2	126.4	30.8	104.0	29.8	103.8
5	n.a.	n.a.	39.5	124.5	38.6	105.7	37.1	107.1
Reserve	21.2	155.4	20.1	115.0	19.5	102.4	16.8	108.4

Exhibit 8. Loss Process for Subsequent Simulations		
	Claim Type I	Claim Type II
Frequency distribution (monthly)	Normal	Normal
Second-level distribution of means	AR(1)	AR(1)
Mean	4	8
Standard deviation	.1	.2
Autocorrelation coefficient	.9816	.9816
Standard deviation	2	4
Report lag distribution	Exponential	Exponential
Second-level distribution of means	AR(1)	AR(1)
Mean	12	6
Standard deviation	.2	.1
Autocorrelation coefficient	.9816	.9816
Maximum lag	60	60
Payment lag distribution	Exponential	Exponential
Mean	24	12
Maximum lag	60	60
Severity distribution	Lognormal	Lognormal
Mean	10,000	1,000
Standard deviation	40,000	2,000
Butsic's trend parameter a	.8	.8
Probability that claim is closed without payment	0	0
Deductible	0	0
Limit	1,000,000	1,000,000
Case reserve error distribution	Lognormal	Lognormal
Second-level distribution of means	AR(1)	AR(1)
Mean	1	1
Standard deviation	.02	.02
Autocorrelation coefficient	.9816	.9816
Standard deviation	1.4142136	1.4142136
Assumed probability that claim is closed without payment	0	0
Reserve at first notice	4000	4000
Number of months until first evaluation	2	2

Exhibit 9. Comparison of Paid and Incurred Loss Development
Prediction Errors as Percent of Mean Actual Ultimate Losses
Sample Size = 10,000

	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Estimators Applied to Paid Losses								
0	.0	5.0	.1	6.5	.1	6.3	-1.0	4.1
1	.5	14.1	.5	12.8	.4	12.4	-.9	9.5
2	1.1	20.2	1.0	19.0	.7	18.4	-.9	9.5
3	1.8	27.0	1.7	25.0	1.0	23.9	-.8	22.1
4	1.9	34.7	1.9	31.3	.9	29.7	-1.1	28.2
5	3.2	46.8	2.8	37.8	1.2	35.2	-.8	33.9
6	5.1	69.8	3.0	45.5	.7	41.3	-1.2	40.0
7	16.5	170.2	4.3	52.1	1.4	46.4	-.4	45.5
Reserve	11.6	87.2	5.9	58.1	2.4	49.9	-2.7	42.7
# best	1431		767		1300		1913	
rank sum	55934		47135		43830		43040	
Estimators Applied to Incurred Losses								
0	.0	5.0	.0	5.0	.0	5.0	.0	5.0
1	.2	8.3	.2	8.4	.2	8.3	.0	9.1
2	.5	14.2	.4	14.1	.4	13.9	-.1	14.7
3	1.0	20.7	.7	20.2	.7	20.1	.0	20.7
4	1.1	26.6	.6	25.6	.6	25.4	-.2	25.8
5	1.7	34.3	1.1	31.7	.9	31.4	.0	31.6
6	2.8	47.6	1.2	39.6	.9	39.1	-.1	38.6
7	6.1	73.2	2.5	46.7	1.8	45.3	.9	44.7
Reserve	5.1	47.3	2.6	40.6	2.1	39.5	.2	40.5
# best	1839		388		673		1689	
rank sum	46382		42139		40715		40825	

Exhibit 10. Comparison of Paid and Incurred Loss Development Case Basis Reserve Errors Increased from Basic Model Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Estimators Applied to Paid Losses								
0	.0	7.2	.2	9.6	.2	8.9	-1.3	4.5
1	.7	17.4	.7	16.4	.6	15.6	-1.2	11.2
2	1.0	23.7	1.1	22.2	.7	21.1	-1.4	17.6
3	1.4	31.0	1.6	29.2	.7	27.6	-1.5	24.8
4	2.4	39.3	2.3	34.7	1.0	32.5	-1.4	30.1
5	3.3	50.5	2.6	41.5	.6	38.4	-1.8	36.3
6	7.3	80.1	4.5	49.3	1.6	44.7	-.7	42.8
7	19.2	182.0	4.5	56.9	1.0	51.0	-1.2	49.7
Reserve	13.6	96.7	6.7	66.9	2.5	57.6	-4.0	46.1
# best	1783		951		1492		2422	
rank sum	50307		41303		38397		37642	
Estimators Applied to Incurred Losses								
0	.0	7.2	.0	7.2	.0	7.2	.0	7.2
1	.6	15.6	.6	15.7	.6	15.6	.1	16.9
2	1.6	26.4	1.2	26.6	1.2	26.3	.2	27.9
3	2.5	37.7	1.7	37.4	1.8	37.0	.1	38.5
4	3.0	44.8	1.7	44.0	1.7	43.5	-.5	45.0
5	3.9	55.1	1.8	53.4	1.8	52.8	-.6	53.9
6	6.9	69.6	3.5	61.7	3.3	60.9	.7	61.4
7	9.6	92.8	4.0	62.1	2.9	60.1	.6	59.6
Reserve	10.8	72.0	5.5	65.7	5.1	63.9	.3	68.3
# best	1344		303		492		1213	
rank sum	49694		48206		46175		48276	

Exhibit 11. Comparison of Methods of Averaging Development Factors Incurred Chain-ladder Method Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	Arithmetic		Dollar-weighted		Least Squares		Geometric	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0	.0	5.0	.0	5.0
1	.2	8.3	.2	8.3	.2	8.3	.2	8.3
2	.7	14.0	.5	14.2	.4	14.4	.5	13.9
3	1.7	20.5	1.0	20.7	.4	21.2	1.2	20.2
4	2.6	26.4	1.1	26.6	-.1	27.1	1.5	25.9
5	5.1	34.9	1.7	34.3	-1.0	34.6	2.7	33.5
6	10.1	50.3	2.8	47.6	-3.0	46.7	4.6	47.1
7	24.1	83.5	6.1	73.2	-7.9	67.7	8.3	72.2
Reserve	17.1	50.6	5.1	47.3	-4.3	46.6	7.3	45.4
# best	4140		720		3705		1435	
rank sum	25772		24638		26659		22931	

Exhibit 12. Comparison of Depths of Averaging Development Factors Incurred Chain-ladder Method with Dollar-weighted Averages Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	All Years		5 Years		4 Years		3 Years	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0	.0	5.0	.0	5.0
1	.2	8.3	.2	8.3	.2	8.3	.2	8.3
2	.5	14.2	.5	14.2	.5	14.2	.5	14.2
3	1.0	20.7	1.0	20.7	1.0	20.7	1.0	20.7
4	1.1	26.6	1.1	26.6	1.1	26.6	1.2	26.9
5	1.7	34.3	1.7	34.3	1.8	34.6	2.1	35.5
6	2.8	47.6	3.1	48.3	3.5	49.2	4.1	50.6
7	6.1	73.2	7.5	76.4	8.7	79.5	10.7	85.0
Reserve	5.1	47.3	5.8	48.9	6.4	50.5	7.6	53.6
# best	3047		1938		2050		2965	
rank sum	22835		24498		25687		26980	

Exhibit 13. Results of Excluding Extreme Values in Averaging Development Factors Incurred Chain-ladder Method Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000										
	Dollar-weighted All Years		Dollar-weighted 5 Years		Arithmetic All Years		Least Squares All Years		Geometric All Years	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0	.0	5.0	.0	5.0	.0	5.0
1	.2	8.3	.2	8.3	.2	8.3	.2	8.3	.2	8.3
2	.5	14.2	.5	14.2	.7	14.0	.4	14.4	.5	13.9
3	1.0	19.8	1.0	19.8	1.2	19.6	.8	20.0	1.0	19.6
4	1.0	25.0	1.0	25.0	1.2	24.9	.8	25.1	1.1	24.8
5	1.0	31.6	1.0	31.6	1.7	31.6	.6	31.7	1.3	31.5
6	.8	44.7	1.0	45.3	2.5	45.2	.0	45.2	1.6	45.1
7	1.6	67.9	2.4	69.6	6.6	70.3	-.3	68.6	2.9	69.3
Reserve	2.4	41.7	2.7	42.7	5.4	42.3	.9	42.7	3.3	42.1
# best	1902		485		3048		2662		1903	
rank sum	28660		30803		28329		32050		30158	

Exhibit 14. Use of Judgement Maximum Development Factors Estimators Applied to Incurred Losses Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000						
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0	.0	5.0
1	.2	8.3	.1	8.3	.1	8.2
2	.5	14.1	.4	14.1	.4	13.9
3	1.0	20.7	.7	20.2	.7	20.1
4	1.1	26.6	.6	25.6	.6	25.4
5	1.7	34.2	1.1	31.7	.9	31.4
6	2.8	47.5	1.2	39.5	.9	39.1
7	6.0	73.2	2.5	46.7	1.8	45.3
Reserve	5.1	47.3	2.6	40.6	2.0	39.4
# best	3970		659		5371	
rank sum	22010		19623		18367	

Exhibit 15. Weighted Average of Calculated and Reference Factors Estimators Applied to Incurred Losses Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000						
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0	.0	5.0
1	.0	7.4	.0	7.4	.0	7.4
2	.0	12.7	.0	12.6	.0	12.6
3	.0	18.9	-.2	18.6	-.2	18.6
4	-.1	24.5	-.3	23.9	-.3	23.8
5	-.1	31.9	-.4	30.0	-.5	29.9
6	.3	44.4	-.4	37.9	-.5	37.7
7	3.2	68.6	1.3	45.1	.9	44.1
Reserve	1.3	41.1	.0	35.4	-.2	34.8
# best	3989		635		5376	
rank sum	21954		19500		18546	

Exhibit 16. Results of Graduating Development Factors Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	Inverse Power All Lags Paid Chain-ladder		Inverse Power Lags > 24 Months Paid Chain-ladder		Inverse Power All Lags Inc'd Chain-ladder		Inverse Power Lags > 24 Months Inc'd Chain-ladder	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0	.0	5.0	.0	5.0
1	.4	11.5	.2	11.6	.4	7.0	.4	7.0
2	-2.1	16.1	-2.4	16.3	.6	11.8	.7	11.8
3	-6.9	21.4	-7.2	21.6	.4	17.5	.4	17.5
4	-11.2	28.4	-11.1	28.5	-.3	22.7	-.2	22.8
5	-15.9	37.7	-12.5	38.2	-2.1	29.3	-1.4	29.7
6	-16.6	55.7	8.9	72.8	-5.9	41.0	.0	45.4
7	26.6	182.8	20.7	174.7	5.1	78.8	3.4	70.6
Reserve	-10.0	82.0	-1.4	82.5	-.7	40.4	1.3	39.5
# best	1318		2170		2946		3566	
rank sum	30397		26648		22125		20830	

Exhibit 17. Use of Incurral and/or Development Intervals Smaller than Annual Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	Accident Years Devel. Quarters Paid Chain-ladder		Accident Quarters Devel. Quarters Paid Chain-ladder		Accident Years Devel. Quarters Inc'd Chain-ladder		Accident Quarters Devel. Quarters Inc'd Chain-ladder	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.2	9.1	.0	5.0	.1	5.5
1	.5	14.1	.6	15.7	.2	8.3	.2	8.5
2	1.1	20.2	1.1	22.1	.5	14.2	.5	14.1
3	1.8	27.0	1.9	27.7	.9	20.7	.9	20.6
4	1.9	34.7	1.9	35.2	1.1	26.6	1.1	26.4
5	3.2	46.8	3.2	46.8	1.7	34.3	1.8	34.4
6	5.0	69.8	5.9	74.8	2.8	47.6	2.9	47.5
7	16.4	170.2	35.7	501.8	6.0	73.2	5.5	67.5
Reserve	11.5	87.2	19.4	201.8	5.1	47.3	5.0	46.4
# best	2217		1989		2708		3086	
rank sum	27014		28046		22860		22080	

Exhibit 18. Cape Cod versus Pure Bornhuetter-Ferguson Methods Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	Paid Loss Cape Cod		Paid Loss Pure B-F		Incurred Loss Cape Cod		Incurred Loss Pure B-F	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.1	6.3	-.1	3.2	.0	5.0	.0	5.0
1	.4	12.4	-.1	9.5	.2	8.3	.0	8.5
2	.7	18.4	-.1	15.4	.4	13.9	-.1	14.2
3	1.0	23.9	.0	20.7	.7	20.1	-.1	20.2
4	.9	29.7	-.3	26.3	.6	25.4	-.4	25.2
5	1.2	35.2	.1	31.4	.9	31.4	-.3	30.8
6	.7	41.3	-.2	36.7	.9	39.1	-.4	37.4
7	1.4	46.4	.9	40.8	1.8	45.3	.7	41.6
Reserve	2.4	49.9	.1	32.8	2.1	39.5	-.2	35.2
# best	2104		3343		1931		2622	
rank sum	27500		23293		26353		22854	

Exhibit 19. Use of Gluck's Decay Factors in Cape Cod Method
Prediction Errors as Percent of Mean Actual Ultimate Losses
Sample Size = 10,000

	D=.9		D=.8		D=.7		D=.6	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Estimators Applied to Paid Losses								
0	.1	6.2	.1	6.1	.1	6.0	.1	5.8
1	.4	12.4	.4	12.4	.4	12.5	.4	12.5
2	.7	18.4	.7	18.4	.7	18.5	.8	18.6
3	1.0	23.9	1.1	24.0	1.2	24.1	1.2	24.3
4	1.0	29.8	1.1	30.0	1.2	30.3	1.3	30.6
5	1.4	35.5	1.6	35.9	1.8	36.5	2.0	37.2
6	.9	41.7	1.2	42.4	1.5	43.4	1.8	44.9
7	1.6	46.8	1.8	47.7	2.1	49.3	2.6	51.7
Reserve	2.7	50.6	3.0	51.7	3.4	53.1	3.9	55.0
# best	2372		373		357		1720	
rank sum	43464		45360		48053		51205	
Estimators Applied to Incurred Losses								
0	.0	5.0	.0	5.0	.0	5.0	.0	5.0
1	.1	8.2	.1	8.2	.1	8.2	.1	8.2
2	.4	13.9	.4	13.9	.4	13.9	.4	13.9
3	.7	20.1	.7	20.1	.7	20.1	.7	20.1
4	.6	25.4	.7	25.4	.7	25.5	.8	25.5
5	1.0	31.5	1.1	31.7	1.2	31.8	1.3	32.0
6	1.1	39.4	1.3	39.8	1.5	40.4	1.7	41.0
7	2.0	45.9	2.3	46.9	2.7	48.3	3.1	50.1
Reserve	2.3	40.0	2.5	40.6	2.8	41.4	3.1	42.2
# best	2564		257		278		2079	
rank sum	39540		41759		44148		46471	

Exhibit 20. Weighted Average of Chain-ladder and Cape Cod Estimates
Weights Equal Expected Fractions Unemerged
Prediction Errors as Percent of Mean Actual Ultimate Losses
Sample Size = 10,000

	Paid Loss Estimates		Incurred Loss Estimates	
Year	Mean	Std Dev	Mean	Std Dev
0	.0	5.8	.0	5.0
1	.5	13.1	.2	8.3
2	.9	19.3	.5	14.1
3	1.5	25.3	.9	20.6
4	1.4	31.8	1.0	26.2
5	1.8	38.5	1.6	33.3
6	1.2	45.1	2.1	44.0
7	1.4	47.8	2.7	52.7
Reserve	3.4	52.0	3.5	41.7
# best	4581		5419	
rank sum	15419		14581	

Exhibit 21. Weighted Average of Paid and Incurred Estimates Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000						
	Paid		Incurred		Average	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Chain-ladder estimator						
0	.0	5.0	.0	5.0	.0	5.0
1	.5	14.1	.2	8.3	.4	8.5
2	1.1	20.2	.5	14.2	.8	13.7
3	1.8	27.0	1.0	20.7	1.4	19.0
4	1.9	34.7	1.1	26.6	1.5	24.8
5	3.2	46.8	1.7	34.3	2.5	32.8
6	5.1	69.8	2.8	47.6	3.9	47.6
7	16.5	170.2	6.1	73.2	11.3	102.3
Reserve	11.6	87.2	5.1	47.3	8.3	55.7
# best	1341		1863		1067	
rank sum	43745		35679		34642	
Cape Cod Estimator						
0	.1	6.3	.0	5.0	.1	5.6
1	.4	12.4	.2	8.3	.3	7.9
2	.7	18.4	.4	13.9	.5	12.8
3	1.0	23.9	.7	20.1	.8	17.4
4	.9	29.7	.6	25.4	.7	22.2
5	1.2	35.2	.9	31.4	1.0	27.4
6	.7	41.3	.9	39.1	.8	34.1
7	1.4	46.4	1.8	45.3	1.6	42.7
Reserve	2.4	49.9	2.1	39.5	2.2	37.3
# best	2297		2134		1298	
rank sum	35127		32179		28628	

Exhibit 22. Comparison of Paid and Incurred Loss Development Severity Trend = 8%, a = .5; No Inflation Adjustments Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Estimators Applied to Paid Losses								
0	.0	4.8	.5	8.3	.4	7.3	-1.0	4.9
1	.6	15.3	1.5	16.2	.9	14.6	-1.1	10.6
2	1.2	21.6	2.7	22.2	1.2	20.0	-1.9	16.8
3	2.1	28.7	3.2	27.6	.5	24.8	-3.6	22.3
4	2.2	35.4	1.8	32.2	-2.1	29.0	-6.8	27.0
5	3.5	46.7	-1.1	37.2	-6.5	33.2	-11.4	31.6
6	5.2	67.0	-7.8	42.5	-15.0	37.5	-19.5	36.3
7	16.6	151.2	-16.4	46.5	-25.1	40.7	-28.5	39.9
Reserve	11.8	81.9	-7.2	53.1	-18.1	41.4	-27.3	34.5
# best	1570		834		1051		1191	
rank sum	54087		47029		46421		53321	
Estimators Applied to Incurred Losses								
0	.0	4.8	.0	4.8	.0	4.8	.0	4.8
1	.2	8.4	.3	9.1	.3	8.8	.0	8.7
2	.5	13.8	.6	14.4	.5	13.9	-.3	13.4
3	.8	19.9	.9	19.5	.6	19.1	-.7	18.4
4	.9	25.0	.3	23.6	-.2	23.0	-1.8	22.3
5	1.2	31.8	-.8	28.3	-1.6	27.7	-3.6	26.8
6	1.2	43.1	-4.8	34.1	-6.1	33.3	-8.4	32.4
7	-1.0	67.1	-15.3	40.3	-17.7	38.9	-19.8	38.2
Reserve	1.2	43.2	-7.7	34.4	-9.7	32.3	-13.4	30.3
# best	2179		577		847		1751	
rank sum	43435		38550		37762		39395	

Exhibit 23. Inflation Adjustments Severity Trend = 8%, a = .5; Trend for Inflation Adjustments 8% Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000												
	Paid Loss Cape Cod Estimates						Incurred Loss Cape Cod Estimates					
	No Adjustment		Loss Rates Trended Inside Cape Cod Calculations		Losses Deflated, then Reflated After Calculations		No Adjustment		Loss Rates Trended Inside Cape Cod Calculations		Losses Deflated, then Reflated After Calculations	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.4	7.3	.1	5.9	.2	6.2	.0	4.8	.0	4.8	.0	5.0
1	.9	14.6	.4	13.5	.5	13.7	.3	8.8	.2	8.3	.2	8.6
2	1.2	20.0	.8	19.6	.8	19.8	.5	13.9	.3	13.5	.3	13.9
3	.5	24.8	1.1	25.2	1.2	25.3	.6	19.1	.5	19.0	.6	19.7
4	-2.1	29.0	1.1	30.2	1.2	30.3	-.2	23.0	.4	23.5	.5	26.1
5	-6.5	33.2	1.6	35.0	1.6	35.1	-1.6	27.7	.4	28.7	.5	30.0
6	-15.0	37.5	1.2	40.0	1.3	40.1	-6.1	33.3	-.3	34.7	-.3	36.1
7	-25.1	40.7	2.2	43.8	2.3	43.9	-17.7	38.9	-.7	40.9	-1.2	41.6
Reserve	-18.1	41.4	3.0	46.1	3.1	46.4	-9.7	32.3	.1	34.8	.1	36.0
# best	1755		1440		1617		2198		1321		1669	
rank sum	40989		35224		35265		34339		30774		33409	

Exhibit 24. Frequency x Severity Methods Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000				
	Paid Loss Frequency x Severity		Incurred Loss Frequency x Severity	
Year	Mean	Std Dev	Mean	Std Dev
0	-1.0	4.1	.0	5.0
1	-.6	11.2	.2	8.3
2	-.1	18.4	.5	14.1
3	.7	25.7	1.0	20.8
4	.8	33.7	1.1	26.6
5	2.1	45.3	1.8	34.3
6	4.1	69.2	2.9	47.7
7	15.5	168.2	6.2	73.3
Reserve	8.3	83.8	5.2	47.4
# best	3778		6222	
rank sum	16222		13778	

Exhibit 25. R. J. Fingers' Methods Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000				
	Average Payment Development		Case Reserve Development	
Year	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0
1	.5	13.6	1.9	19.0
2	1.1	19.6	4.0	30.0
3	1.8	25.9	5.5	35.7
4	1.9	32.7	6.0	38.1
5	2.7	41.0	6.9	46.1
6	3.0	49.0	8.1	60.7
7	4.4	52.8	11.2	84.5
Reserve	5.9	54.7	16.8	66.1
# best	5192		4808	
rank sum	14808		15192	

Exhibit 26. H.E. Clarke's London Market Method Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	Estimator Applied to Paid Losses				Estimator Applied to Incurred Losses			
	Curves Fit to All Lags		Curves Fit to All Lags Except First		Curves Fit to All Lags		Curves Fit to All Lags Except First	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.7	13.4	2.7	14.3	4.3	13.4	6.2	14.0
1	-.8	16.1	1.8	17.1	4.6	15.5	6.9	16.1
2	-2.7	18.9	-.1	19.2	4.7	16.8	7.4	18.0
3	-6.2	23.9	-3.0	24.3	5.3	22.2	8.9	24.9
4	-8.8	30.6	-5.2	30.9	5.0	25.8	10.3	32.2
5	-10.6	37.1	-2.9	50.7	5.2	32.1	9.3	39.7
6	-9.8	57.7	-4.1	59.4	5.3	46.2	9.9	50.5
7	-8.6	80.7	-2.4	90.5	7.4	65.1	11.3	82.0
Reserve	-18.0	51.6	-5.1	57.9	16.1	43.8	27.1	53.5
# best		2133		2368		3103		2396
rank sum		27017		25128		22512		25343

Exhibit 27. Loglinear Models Estimators Applied to Paid Losses Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000										
	R.J. Verrall's Bayesian Model						B. Zehnwirth's Adaptive Model			
	Maximum Likelihood		Pure Bayesian		Empirical Bayesian		Fixed Parameters		Varying Parameters	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	-1.0	4.1	-1.0	4.1	-1.0	4.1	-1.0	4.1	-1.0	4.1
1	2.2	18.3	.7	7.7	.4	12.7	-.4	8.1	.0	8.5
2	4.1	25.6	.6	13.5	.0	17.9	-2.1	14.1	-.1	16.7
3	6.8	30.7	-.1	19.3	-.9	22.8	-4.6	19.8	-.1	26.7
4	10.2	38.4	1.2	26.1	-2.2	28.3	-4.1	26.1	6.4	44.4
5	20.8	54.3	3.2	33.2	-2.0	34.0	-.6	32.4	10.4	51.8
6	48.5	103.4	8.1	43.0	-.7	40.9	4.3	40.9	14.2	59.9
7	165.3	413.1	14.9	53.0	2.7	46.8	13.7	49.2	21.6	68.2
Reserve	98.6	188.2	10.6	43.7	-1.4	47.0	1.9	40.0	19.7	82.5
# best	1226		2956		2198		1924		1696	
rank sum	39575		25081		27175		25726		32443	

Exhibit 28. Estimators with Berquist-Sherman Adjustments Prediction Errors as Percent of Mean Actual Ultimate Losses Sample Size = 10,000								
	Paid Chain-ladder, Disposal Rate Adjustment		Incurred Chain-ladder, Disposal Rate Adjustment		Incurred Chain-ladder, Case Reserve Adequacy Adjustment		Incurred Chain-ladder, Disposal Rate and Case Reserve Adequacy Adjustments	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0	.0	5.0	.0	5.0
1	.7	12.9	.2	8.0	.4	13.5	.1	12.9
2	1.0	18.7	.6	13.7	.8	19.2	.3	17.9
3	1.7	25.6	1.1	20.4	1.1	24.7	.1	24.1
4	1.9	33.2	1.3	26.5	1.0	30.7	.0	29.8
5	3.7	45.8	2.4	34.3	1.2	37.1	.5	36.5
6	10.1	71.0	5.8	48.9	.9	44.0	2.9	44.1
7	23.9	178.0	18.5	84.3	1.9	52.2	13.3	57.6
Reserve	16.5	88.9	11.5	50.3	2.8	51.2	6.6	51.2
# best	1855		3724		2216		2205	
rank sum	28666		24106		23594		23634	

Exhibit 29. Combinations of Options and Adjustments
Estimators Applied to Incurred Losses; Extreme Values Excluded from Averages
Prediction Errors as Percent of Mean Actual Ultimate Losses
Sample Size = 10,000

	Chain-ladder, Dollar-weighted, Ref. Factors		Chain-ladder, Dollar-weighted, Graduation		Chain-ladder, Least Squares, Ref. Factors		Chain-ladder, Least Squares, Graduation		Cape Cod, Dollar-weighted, Ref. Factors	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	5.0	.0	5.0	.0	5.0	.0	5.0	.0	5.0
1	.0	7.4	.4	7.0	.0	7.4	.4	7.0	.0	7.3
2	.0	12.7	.6	11.8	-.1	12.8	.6	11.8	.0	12.6
3	.0	18.4	.3	17.4	-.1	18.5	.3	17.4	-.1	18.1
4	-.1	23.6	-.4	22.6	-.2	23.7	-.5	22.6	-.2	22.9
5	-.6	30.2	-1.8	29.0	-.9	30.2	-2.1	29.0	-.7	28.5
6	-1.1	42.5	-1.7	42.7	-2.1	42.3	-2.9	42.3	-1.4	36.5
7	.1	65.1	-.7	66.7	-2.9	64.1	-4.6	65.2	-.3	43.2
Reserve	-.6	37.3	-1.2	36.3	-2.4	37.2	-3.4	35.7	-1.1	31.9
# best	1144		1577		1032		1769		4478	
rank sum	29801		30435		31551		31978		26235	

Exhibit 30. Comparison of Paid and Incurred Loss Development
Sample Points Screened for Proximity to Case Data
Prediction Errors as Percent of Mean Actual Ultimate Losses
Sample Size = 10,000 (after 72,530 trials)

	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Estimators Applied to Paid Losses								
0	.0	5.7	.3	7.0	.3	7.0	-1.2	5.0
1	.3	13.3	.8	14.3	.7	14.0	-1.3	9.9
2	.0	18.9	1.0	19.3	.7	18.8	-2.2	15.5
3	2.5	23.3	1.4	21.6	.9	21.0	-2.7	18.5
4	12.6	27.0	.9	20.0	.4	19.3	-2.9	17.6
5	-5.5	40.9	2.3	38.8	.9	37.0	-3.2	35.1
6	.8	57.3	3.6	43.2	1.9	40.8	-1.5	39.2
7	4.2	117.2	2.1	47.7	-1.0	44.8	3.3	43.5
Reserve	9.5	67.7	5.5	50.6	3.0	45.7	-6.4	35.2
# best	1453		805		1476		1815	
rank sum	55285		46647		43634		42257	
Estimators Applied to Incurred Losses								
0	.0	5.7	.0	5.7	.0	5.7	.0	5.7
1	.2	9.2	.2	9.6	.2	9.5	.0	9.8
2	.2	15.0	.1	15.6	.1	15.4	-.4	15.7
3	.8	19.2	.4	18.7	.4	18.6	-.4	18.6
4	1.7	19.7	.1	17.4	.0	17.3	-.7	17.2
5	-.2	36.4	.0	34.8	-.1	34.6	-1.3	34.4
6	1.2	45.4	.6	38.8	.3	38.4	-.7	37.8
7	3.3	68.2	1.7	45.5	1.1	44.3	.2	43.6
Reserve	3.1	44.5	1.2	38.6	.8	37.6	-1.3	37.3
# best	1651		389		754		1657	
rank sum	48106		43083		41429		39559	

Exhibit 31. Comparison of Paid and Incurred Loss Development
Sample Points Screened for Regularity of Latest Accident Year at First Evaluation
Estimators Applied to Incurred Losses
Prediction Errors as Percent of Mean Actual Ultimate Losses
Sample Size = 10,000 (after 24,755 trials)

	Chain-ladder		Modified Bornhuetter-Ferguson		Cape Cod		Additive	
Year	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
0	.0	40.7	.0	39.1	.0	38.6	.0	40.1
1	.2	5.2	.1	5.2	.1	5.2	.0	5.2
2	.5	8.4	.4	8.2	.4	8.1	.0	9.1
3	.8	14.6	.4	14.5	.4	14.4	-.2	15.3
4	.6	21.0	.1	20.5	.1	20.5	-.7	21.1
5	1.9	26.3	1.0	25.3	1.0	25.2	.1	25.7
6	2.7	33.3	.8	30.9	.7	30.8	-.1	30.9
7	-3.0	45.7	-.8	37.8	-1.0	37.5	-1.6	37.2
Reserve	1.4	43.3	.8	42.6	.7	42.5	-1.0	42.5
# best	4052		300		1408		4240	
rank sum	26327		25030		24046		24597	