EXPERIENCE RATES AS ESTIMATORS: A SIMULATION OF THEIR BIAS AND VARIANCE
By James N. Stanard

REVIEWED BY John P. Robertson

Mr. Stanard's paper opens a new area of actuarial reasearch, namely the use of simulation to investigate the reliability of commonly used pricing (and related) models. He is not using simulation to forecast insurance results directly, but rather to determine how well a given technique for such forecasting can be expected to perform. I believe this is not a paper to read to find final answers, but rather to find groundbreaking results from a technique which should become more widely used.

In this review, I will comment on the interpretation of the results from the standpoint of bias and variance, clarify (I hope) the algebra underlying the derivation of the "Adjustment to Total Known Losses Method" and conclude with some comments on other possible areas of application of this technique.

BIAS OF RESULTS

On an initial reading of the paper, I was surprised by the biases developed by the various experience rating procedures. These range from +30% to -10% of the expected losses under the first set of parameters and are statistically significant since they are quite a bit larger than the standard error of the estimate of the expected losses. The techniques used are similar to commonly used ones which are not normally assumed to be biased. It is possible, however, to see the causes for the bias in the procedures used.

Mr. Stanard notes the underreserving of claims by the amount of future inflation. This combined with the use of loss development factors that only go to the fifth year, produce a downward bias in Methods 1 through 3. Methods 4 through 6 are given downward biases both by the use of claim development factors that only go through year 5, and by use of the actual average known claim size. Given the reserving method used, the average actual known claim size for a given accident year will tend to increase as the accident year develops since new reported claims come in at a higher average amount and outstanding reserves are increased for inflation. The latter has two effects on Method 6; it serves to reduce the average claim size (as in Methods 4 and 5) and since it reduces the latest year the most and the earliest year the least, it biases the average trend downwards. Finally, the restriction of fitted slopes to be positive in Methods 1 and 6 contributes an upward bias to these methods.

The above effects are all of the significant sources of bias in the cases where the trend factor used is equal to the true inflation rate underlying the model (8% per year). (There is another minor source which I will discuss in the "Adjustments to Total Known Losses" section.) I do not agree with Mr. Stanard's comment that in this case "the bias need not be zero because the rating technique may not take inflation into account exactly the way the loss generating model does". If the ultimate losses for each accident year were projected without bias (which, of course, they're not), then any of the Methods 2 through 5 should give unbiased results. Methods 1 and 6 would also give unbiased results if exponential fits were used (the straight line gives a very slight downward bias) and if the slope of the fitted line were allowed to become negative. Clearly, if the trend factor used is not equal to the underlying rate of inflation, then a bias will be introduced into any of the above methods.

I remain surprised by the positive bias of $50,000 reported for Method 2c' under the first set of parameters. The only major source of bias in Method 2 is the downward one due to lack of full development. The possible likely error in the bias, as noted by Mr. Stanard in his section on "Validity of the Results" is about $50,000. The direct interpretation of the simulation is, therefore, that Method 2 is very unlikely to have a negative bias. This seems to me to be in conflict with the "a priori" expectation of Method 2's bias. The fact that the second set of parameters gives a negative bias does not help explain this conflict since under the second set of parameters all of the biases (including method 2's) have moved down about $90,000 from the biases under the first set of parameters.

The fact that biases exist in the methods under consideration is interesting, but I hope my discussion has shown that their existence is not too surprising, since reasons can be found for expecting bias. Of more interest is the relative magnitudes of the biases, since these are harder to predict in advance. Also Methods 1 and 6 have both positive and negative sources of significant bias, and predicting which will win out would not be easy by a priori methods.

VARIANCE OF RESULTS

I believe that the simulated standard deviations are of far greater interest than the biases. Bias in pricing methods is something that actuaries are used to dealing with and there are obvoius techniques for eliminating the bias in Methods 2 through 5. (Under

Methods 1 and 6, it would be difficult to part with the restriction of non-negative slope; hence, eliminating bias for these methods is not as easy.)

The standard deviations are not only harder to predict than bias, but they cannot be "repaired" in the sense that a suspected non-zero bias can be repaired. I do not disagree than Method 3c' gives the best overall result of the methods tested, but I would nominate Method 5c' as having the greatest promise since it shows the least variance. In most applications of Method 5, the lack of full development of claim count would be apparent and some adjustment could be made to approximate full development. Similarly, the development to ultimate of the average claim size in each accident year could probably be addressed. Hence, eleminating the bias in Method 5 could likely be achieved. As the underlying parameters are changed, I think Method 5's advantages become clear.

Moving from the first set of parameters to the second (that is, missing a change in the trend) influences the bias of all the methods similarly. If all were unbiased under the first set of parameters, all would be biased by about $-90,000 under the second set. While Method 5 does not do any better than the other methods here, it doesn't do any worse either. The various methods do not react as uniformly to the introduction of an unsuspected claim count trend. The bias of Method 5 changes less from the first to the third set of parameters than any of the other methods. No matter what set of parameters is used, Method 5 shows the least variance.

Obviously, neither Mr. Stanard's paper nor my discussion will prove that some one experience rating technique is the ultimate such technique. I do hope that I have shown that Mr. Stanard's results already contain much information of value in choosing a technique and that the variance information he gives is likely to be as useful or more useful than the bias information.

ADJUSTMENT TO TOTAL KNOWN LOSSES METHOD (ATTKLM)

The ATTKLM is an interesting method for applying development factors to reported results. My purpose here is to show the algebra underlying the general application of this method, and to tie together all of the places it is actually used in Stanard's paper.

The Appendix to this discussion shows that if $A_i B_i = C$ for all i in some set, then $(\Sigma A_i) \div (\Sigma 1/B_i) = C$ and that is essentially the ATTKLM. The application shown in Standard's Appendix A (used in Methods 3, 5, and 6) is the first application in my appendix. Additionally, Methods 2c' through 5c' use the same theory but with the B_i as trend factors. Stanard justifies this latter use in his Appendix C, wherein he notes that he is also justifying the use as loss or claim development factors. While his Appendix A and my appendix show that the ATTKLM will not introduce any bias into an experience rating method, his Appendix C makes a strong argument for expecting less variance in results when this method is used.

The paper effectively compares the use of the ATTKLM to the more normal "adjust, then average" since this is the essential difference between Methods 2 and 3 and also between Methods 4 and 5. In either case, the ATTKLM shows less variance than the "adjust, then average" method. It would be interesting to compare Methods 2c through 5c to Methods 2c' through 5c' to further test this comparison. Stanard's footnote 22 notes "the simulation did not provide conclusive results either way." Since the average loss development factors are likely to be larger than the average trend factors, it is possible that the test using loss development factors is more likely to show a difference.

As a final comment on the use of the ATTKLM, I must point out that as used in methods 3c' and 5c', it introduces a slight bias. To illustrate with Method 3c', Stanard has taken

$$\left(\frac{1}{5}\Sigma k_j\right) \div \left[\left(5 / \Sigma (1/f_j)\right) \times \left(5 / \Sigma \left(1 / (1.08)^{6-j}\right)\right)\right]$$

where it is more correct to take

$$\left(\frac{1}{5}\Sigma k_j\right) \div \left[5 / \Sigma (1/f_j)\left(1/(1.08)^{6-j}\right)\right]$$

One cannot keep spinning off factors B_i since after the first time the relationship $A_iB_i=C$ fails to hold. The effect of this difference is to introduce a very slight negative bias in Methods 3c' and 5c', which coincidentally is approximately offset by the use of $(1.08)^3$ in place of $5 \div \Sigma 1/(1.08)^{6-j}$.

CONCLUSION

In most applications of Stanard's technique, one is not going to be able to specify the distributions underlying the experience (if one could, then one could estimate mean losses far more accurately than any normal experience rating method allows). Thus, the most significant conclusions to be drawn from the simulations are not in the area of what model best fits some given data, but rather are in the area of choosing a technique in the absence of any information other than the reported results.

Areas worthy of further investigation include refinements to the underlying assumptions to make the model more realistic (including application of credibility weighting techniques), use for larger models (what variation should be anticipated in Homeowners or Auto Liability indications when the standard ratemaking techniques are used?), and the use for testing possible variance in rating or ratemaking methods due to particular components of the methods. Obviously the larger the model, either in terms of number of assumptions needed or the number of claims and other items which may need to be simulated, the greater the possibility of the cost of running the simulations of becoming prohibitive. Even though the simulations presented are based on relatively simple experience rating techniques, it is clear that a great deal of work was required to achieve the results.

In summary, Mr. Stanard has provided a very interesting and useful paper, both from the standpoint of the results of the simulations given and also because of the introduction of the "average, then adjust" method of applying trend and development factors.

APPENDIX

The following lemma generalizes the result of Stanard's Appendix A. Applications include the two used in his paper and two new ones.

Lemma: If $A_i B_i = C$ for $i=1$ to n

$$\text{Then} \quad (\Sigma_i A_i) \div \Sigma_i (1/B_i) = C \tag{1}$$

$$\text{Also} \quad (\Sigma_i W_i A_i) \div \Sigma_i (W_i/B_i) = C \text{ for any } W_i \tag{2}$$

Proof: $A_i = C/B_i$

$$\Sigma A_i = C \times \Sigma (1/B_i)$$

$$(\Sigma A_i) \div \Sigma (1/B_i) = C$$

This establishes (1). (2) follows by substituting $W_i A_i$ and B_i/W_i for A_i and B_i in (1).

Applications: For these applications, think of k_j, u_j, and u as expected values rather than as actual reported values.

1) $A_j = k_j (1+i)^{6-j}$, $B_j = f_j$, $C = u = A_j \cdot B_j = k_j \cdot f_j (1+i)^{6-j}$

$$\text{Then } u = (\Sigma k_j (1+i)^{6-j}) \div (\Sigma 1/f_j)$$

This is the application given in Stanard's Appendix A. The fact that $u = k_j f_j (1+i)^{6-j}$ follows from his Assumption 3) as follows:

$IBNR_j = u_j - k_j = u_j - u_j/f_j = (1 - 1/f_j)(u_j)$ (Definitions)

$= (1 - 1/f_j) \dfrac{u}{(1+i)^{6-j}}$ (By his assumption x(3))

Thus, $u_j = \dfrac{u}{(1+i)^{6-j}} \therefore u = u_j (1+1)^{6-j} = k_j f_j (1+i)^{6-j}$

2) $A_j = u_j$, $B_j = (1+i)^{6-j}$, $C = u = A_j B_j$

$$\text{Then} \quad u = (\Sigma u_j) \div \Sigma 1/(1+i)^{6-j}$$

$$= (\Sigma u_j) \times (5 \div \Sigma 1/(1+i)^{6-j})$$

This is the method used to adjust untrended results per footnote 23 of Stanard's paper.

3) $A_j=k_j$, $B_j=f_j(1+i)^{6-j}$, $C=u=A_jB_j$

Then $u=(\Sigma k_j) \div \Sigma(1/f_j(1+i)^{6-j})$

These first three examples show that f_j and $(1+i)^{6-j}$ play symmetric roles in the projection of results.

4) "Ratemaking"

Let A_i be the reported loss ratio (developed to ultimate or not) and let B_i be the ratio of loss trend to premium trend (if any) from the i^{th} year to the period the rates apply to (times an ultimate development factor if not included in A_i). Then $(\Sigma W_i A_i) \div (\Sigma W_i/B_i)$, where the W_i are weightings of the various years such as 10-15-20-25-30, gives an estimate of the ultimate loss ratio analogous to the commonly used $\Sigma W_i A_i B_i$. Mr. Stanard's results indicate that the former may show less variation about the true mean loss ratio than the latter.