

AN ESSAY AT MEASURING THE VARIANCE OF ESTIMATES OF OUTSTANDING CLAIM PAYMENTS

BY FRANK ASHE

E. S. Knight & Co. Research Centre, Sydney

ABSTRACT

The variance of statistical estimates of outstanding claim payments for long-tailed general insurance portfolios is examined. The variance's three components are discussed. As there is no accepted technique for measuring this variance three methods are investigated empirically for its measurement—a parametric method, the jackknife method, and the bootstrap method. No method stands out as superior to the others and it is recommended that all three be evaluated and used to gauge the possible errors in the estimation of outstanding claims.

1. INTRODUCTION

1.1. *The Outstanding Claims Problem*

The estimation of outstanding claims is a major component of the determination of the profit of a general insurance company. There is no one accepted technique for this estimation and, indeed, there cannot be. For each company (and possibly for each portfolio within the company) the nature of the claims and their administration follows a sufficiently unique path as to make the application of any one method an unwise practice. Concomitant with this lack of a standard technique is the knowledge that the estimate from any one method is subject to a certain degree of error. This error arises partly because we have an inadequate model and partly because of the inherent stochastic nature of the claims process.

1.2. *Earlier Work on Second Moments*

As the computing power and statistical theory available to practitioners has increased over the years there has been an increase in the amount of work being done on the statistical nature of the errors involved in the estimation. The natural step beyond the first moment of the distribution of outstanding claims (that is the amount of claims) is the second moment and this has received the most attention. Investigation of the asymmetry and other esoteric features of the distribution have not been feasible in general.

Two approaches to determining the second moment which are not demanding in their data requirements and are quite flexible are given by TAYLOR and ASHE (1983) and DE JONG and ZEHNWIRTH (1980, 1982). TAYLOR and ASHE use least-squares linear regression for their model estimation and a straightforward parametric method for the determination of the second moment. DE JONG and ZEHNWIRTH use a more flexible statistical structure, the Kalman Filter, for their

estimation. Again a parametric method is used to calculate the second moment of their estimate.

This paper will follow the approach of TAYLOR and ASHE in that the models considered will be multiple linear regressions fitted using the linear regression package GLIM (BAKER and NELDER, 1978). However there are significant differences in the choice of error distribution. This is discussed in Section 3.1.1. Although the Kalman filter method includes the linear regression as a subset of its possible models the lack of widely available, easy to use packages inhibits its use.

1.3. Sources of Error in Prediction

The errors involved in a statistical prediction can be categorised in a three-fold way. Using the terminology of BARTHOLOMEW (1975) they are the statistical error, the estimation error and the specification error. These are usually measured as variances around the mean so that under the assumption of independence of the generation of those errors we may decompose the total error into exclusive and exhaustive parts.

The specification error arises when we specify the model of reality to be used for our prediction. Almost surely this will be a simplification of reality as we try to concentrate on the major features of the claims process. Further our model may be wrong because of our ignorance of the important factors of the process being modelled, and because of the limitations of the available data. The very nature of the specification error implies a great difficulty in estimating its size. Little more will be said of this error.

The estimation error arises because even with an accurately specified model of the process under study we must estimate the parameters governing the model. For instance in a dice tossing experiment we may have specified the multinomial model correctly but must still estimate the probabilities of obtaining each outcome of a toss by experiment. Although the standard error of parameter values is a widely used part of statistical theory, their effect on prediction error has received little acknowledgement.

Statistical error is the inherent randomness associated with our model. Even when we have correctly specified our model and accurately estimated its parameters the stochastic nature of our process prevents accurate prediction. This source of error is by far the most tractable and thus is the most commonly calculated error when such things are calculated.

Within the regression framework there is a great deal of interplay between the three errors and the model fitting process. Naively, in order to reduce the mis-specification of the model it is desirable to have a complex model in the sense that we have many regression terms. As well, the more terms in the regression model the smaller will be the estimate of the residual variance, all other things being equal. As the residual variance is used to calculate the variance of the underlying process this will lead to a reduction in the statistical error. To counter this complexity there is the tendency for an increase in the number of terms in

the regression equation to be associated with an increase in the uncertainty of the estimates of the coefficient of the regression equation.

MILLER (1984) discusses the difficulties of selecting the regressor variables and of obtaining accurate estimates of the standard errors of the coefficients of the variables.

The notion of keeping the model simple in order that the estimation error does not swamp the statistical error dovetails neatly with the further notion that it is possible for the data to be "more precise" than our explanations. By this is meant that any reasonable model proposed, either *a priori* or after some model fitting, will be shown to be incorrect if the data set is large enough. Given that there will almost always be a bad fit to the data when judged on the grounds of a purist statistician we need not try to fit the data to a high degree of accuracy. Here we are on the road leading to the redundancy concept of MARTIN-LOF (1974).

2. DESCRIPTION OF METHODS

2.1. *Jackknife*

The Jackknife method was first proposed as a technique for the removal of bias from statistical estimation by QUENOUILLE (1956). It can be summarised briefly as follows (a good starting paper for the practitioner is R. G. MILLER (1974)). A statistic calculated from a sample of size n has value S , say. If we remove the i th element from the sample and recalculate the statistic we obtain a value $S(i)$, say. From the n statistics $S(\cdot)$ found by removing each of the original n values we compute pseudo-values $P(\cdot)$ of the statistic by forming

$$P(i) = n \cdot S - (n-1) \cdot S(i)$$

If the observations contribute equally to the determination of the value of the statistic and the bias term of the statistic is of order $1/n$ then the average of the $P(\cdot)$ will be unbiased to order $1/n$.

TUKEY (1958) proposed that the standard error of the average of the P is the standard error of the original statistic. This suggestion has been taken up and expanded so that distribution of the $P(\cdot)$ has been used to calculate many statistics of interest regarding the second and higher order moments of the original statistic.

2.2. *Bootstrap*

The Bootstrap is a recently developed tool (for example, see, FREEDMAN and PETERS (1984)) for the non-parametric evaluation of second moments of various statistics. In brief, after some model has been fitted the error distribution of the data is estimated by the deviations of the fitted data values from the observed values. Random drawings from this observed discrete distribution are applied to the fitted data values to give pseudo-data. The fitting process is re-applied to this pseudo-data to give a new value for the statistic of interest. The random drawing

and refitting is iterated until a sufficiently good estimate of the distribution of the statistic of interest can be gained.

2.3. *Parametric Method*

TAYLOR and ASHE (1983) describe a method of estimating the second moment of the distribution of outstanding claims. Their algorithm makes use of the parametric assumption of the distribution of the residuals of their fitted model (in their case, Normal). The coefficients of this error distribution are estimated by regression techniques and these are applied to the future estimates of payments to give their estimated variance.

Two ways of estimating the amount of outstanding claims are given in Section 3. The first is very similar to that detailed in TAYLOR and ASHE but has a different assumption for the distribution of residual. Their parametric method has been adapted to this form of error distribution. The second estimation method is quite different and the parametric method has not been applied to it.

2.4. *Quantum of Outstanding Claims as a Statistic*

Once a model has been fitted to the data it is a straightforward arithmetical exercise to calculate the future expected claims payments. This amount is a function of the estimated parameters of the regression model fitted and therefore a statistic of the data. There is no reason why it cannot be treated in a purely statistical manner. The novelty of the approach to many people arises from the *ad hoc* and heuristic nature of the claims estimation procedure until recent time.

It has been stated by some that the statistical approach to the errors of prediction in the claims reserving problem is too theoretical for an area where there is still a lot of judgment needed to come to a final decision on the quantum. Further it is said that the probabilistic statements (such as quoting a 95% confidence limit on the quantum) do not make sense in this problem where we must also make judgments on future inflation rates and rates of investment return, a field where the assumption of distributions is very subjective. Some doubt has also been cast on the large standard deviations calculated by such methods as in TAYLOR and ASHE.

The author is in partial agreement with this line of reasoning, and indeed, has argued at seminars that other mathematical and operational research tools have a place in the actuarial armoury. These tools include fuzzy set theory, fuzzy logic, and scenario analysis. However the various algorithms currently used in estimation (such as the naive PPCI method described in TAYLOR and ASHE, 1983) rely on

- (i) averages of past data collapsed in different ways, and
- (ii) projections of obvious trends.

These analyses may not have an explicit statistical formulation but nevertheless they rely a great deal on simple statistical techniques.

The analyses included in this paper do have an explicit statistical formulation so it is natural to follow the statistical theory and produce the likely error in these predictions that flows from our explicit assumptions. Even for the forms of analysis which do not make explicit assumptions we may often be able to set up a statistical procedure which models the analysis and ask what is the prediction error of this model. If this error is thought to be too high by practitioners then the onus is on the practitioner to show where their more *ad hoc* procedures improve significantly on the statistical model of their procedures.

3. IMPLEMENTATION

In order to gauge the effect of the different methods of estimating the second moments the three techniques have been applied to two forms of estimating the claim payments with five long term liability insurance portfolios:

- Portfolio A compulsory insurance for injury to people involved in motor vehicle accidents who are not at fault, the portfolio is large and can be considered as having a monopoly in its area of insurance. The data are available for 10 payment years and 24 development years.
- Portfolio B a relatively small portfolio covering the same class as A and again can be considered as a monopoly. Data for the last 13 accident years are available.
- Portfolio C small portfolio in the same class as A and B, not a monopoly. Data available for the last 13 accident years.
- Portfolio D workers' compensation insurance for a large group of employees of a Government. Data for 10 payment years and 20 development years.
- Portfolio E workers' compensation claims that give rise to some excess-of-loss reinsurance payments for a given reinsurer. Data for the last 12 accident years are available.

These are all real portfolios with which the author has been involved in consulting on the amount of outstanding claims. The techniques of this paper have been evolved to give the author an understanding of the statistical uncertainties associated with the analysis of the portfolios. The techniques have thus been tried on a wide variety of real data with all its associated inadequacies, outliers and other pitfalls. The five portfolios have a variety of payment patterns and differing patterns of claims inflation.

3.1. *Estimation Methods*

Two methods are used for fitting the data; the invariant see-saw method of TAYLOR (1981), and a variation on the payment per claim incurred method. The

first method models the payment per claim finalised (PPCF) observed for each development year within each year of origin—the PPCF is defined simply as the total claim payments for a given period divided by the number of claims finalised within that period. The second method models the payment per claim incurred (PPCI) observed for the same periods as the invariant see-saw model. The PPCI is defined as the total claim payments for a development period divided by the total number of claims incurred for the relevant year of origin.

There can be large differences in the denominator of the fraction used to evaluate PPCI and PPCF as experience changes. This leads to different precisions being ascribed to these values in different cells. For example in the early development years the number of claims finalised in a year is measured in hundreds or thousands while in the later years we may have only one claim finalised per year. The average payment per claim finalised has a much lower standard error for thousands of claims finalised than for one. The weight in the regression for cells is initially set to the number of claims incurred for the year of origin for the PPCI analysis and the number of claims finalised for the PPCF analysis. This weighting is the first step in making the error distributions in the cells have equal variance.

3.1.1. *Invariant See-Saw*

The invariant see-saw method as described in TAYLOR (1981) and used in TAYLOR and ASHE (1983) has been slightly changed for this exercise. Where the previous authors used the least-squares technique to estimate the parameter coefficients the present paper makes more use of the power of GLIM and sets the error distribution of the fitted PPCFs as Gamma. GLIM then maximises the likelihood function to obtain the parameter coefficients. The choice of the Gamma distribution is based on two pragmatic points. The first is the asymmetrical distribution of the claim payments. Payments are almost always positive with a large number of small payments to claimants and infrequent very large payments. Although the true distribution of claim payments is unknown it is felt that the Gamma distribution will be of better service than the Normal distribution implicit in the approach of previous authors. The second reason is based on a feature of GLIM, viz that when a Gamma distribution of error is specified it is further assumed that the standard deviation of the distribution is a constant multiple of mean of the distribution. Thus cells with different means have different standard deviations but there is a constant coefficient of variation across all cells (subject to the weighting mentioned in the preceding section). This is in accordance with the author's observations that the variance of payments per claim finalised is generally higher when their average is high than when it is low.

3.1.2. *Payment Per Claim Incurred*

The approach taken for this model has evolved primarily in order to estimate the inflation of claim payments over time. GLIM is used to estimate the natural logarithm of the PPCI by a linear additive model, thereby fitting a multiplicative

model to the PPCI themselves. When a variable representing time is incorporated into the model then its coefficient is the estimated force of inflation acting on average payments over the time period.

Let

$$\begin{aligned} P &= \text{year of payment} \\ x_i &= \text{other variables used in the model} \\ a, b_i, c &= \text{estimated coefficients.} \end{aligned}$$

The model fitted is

$$\log(\text{PPCI}) = a + \sum b_i x_i + cP$$

$\exp(c) - 1$ is the estimated rate of inflation of claim payments.

The error distribution of the PPCI is taken to be Normal. GLIM estimates the coefficients of the parameters of the model by maximising the likelihood function determined by this logarithmic link and the Normal error distribution.

The logarithmic relationship between the fitted value and the linear model leads to difficulties in the evaluation of the estimation error of the predicted payments using the parametric method. GLIM gives the first and second moments of the coefficients of the model in units of log (\$). This distribution must then be exponentiated to give the distribution in dollars. The Normal distribution of the statistical error can then be added to this estimation error. This algorithm is not yet implemented and so no parametric estimate of the errors of prediction are given.

3.2. *Jackknife Details*

3.2.1. *Definition of an Element*

The jackknife algorithm proceeds by discarding one sample observation (datum) and re-estimating the statistic of interest. Data on individual claims is not available. The smallest piece of data is on claim payments in a given development year for a given year of origin (a cell), so in implementing the jackknife this is the datum discarded.

3.2.2. *Heteroscedasticity in the Observed Cells*

The Jackknife method is usually described for data sets where all the observations have equal weight in evaluating the final statistic. Unfortunately the real data from the above portfolios need major readjustment to their raw weights (Section 3.1) before any decent analysis can be undertaken. The most usual form this takes is to reduce heterogeneity in the residual variance between development years. This is done by defining a function of development year, say $VD(D)$, which is used to adjust the prior weight given to the cell multiplicatively. Rarely is something more complicated than a two or three valued step function used for a PPCF model. For the PPCI models VD is first taken as a function of the average payment per claim incurred for that development year. This effectively

sets the variance of the PPCI proportional to the average value. This assumption is somewhat different to that implicit in the Gamma error distribution assumed for the PPCF Model where the coefficient of variation is taken as constant for all cells.

The other form of adjustment made to the weight function is for obvious outliers. A weight of zero is typically used in practice, though non-zero weights are perhaps preferable so that the datum still has some bearing on the final statistic.

In practice the adjustment process described here is done on an *ad hoc* trial and error basis and is stopped when there is no obvious non-random pattern emerging when the standardised residuals from the latest model are fitted against explanatory variables.

The usual jackknife procedure treats all sample observations with equal weight. Our analysis has observations with markedly different weights, so it is not obvious that the usual jackknife algorithm will give an appropriate answer.

Consideration of the case where the statistic of interest is the mean of a sample and the weights represent repetitions of a sample value leads to the following algorithm.

Let

- $w(i)$ = weight given to the cell i ;
 - W = sum of all weights;
 - S = statistic using the full sample;
 - $S(i)$ = statistic re-estimated after dropping cell i from regression;
 - $P(i)$ = pseudo-value,
- (1)
$$= (W \cdot S - (W - w(i)) \cdot S(i)) / w(i),$$
- P = the jackknifed estimate,
 - $$= \sum w(i)P(i) / W.$$

However the weights are used for the fitting of the model, they do not necessarily give the importance of the datum in the determination of the projected claims.

The ambiguity as to the form that the jackknife estimate takes is a function of our ignorance of the series expansion that underlies the jackknife methodology which in its simplest form is:

$$S = \mu + a/n + \dots$$

where a is the bias term to order $1/n$.

Formula 1 is based on n being replaced by $W = \sum w(i)$. The true form is conjectural at present, indeed A. MILLER (1984) proposes that with significant competition bias the highest order term is of order $n^{-1/2}$. Both (1) and the usual jackknife algorithm are used for estimation in this paper.

3.3. Bootstrap Details

3.3.1. Evaluation of Error Distribution

The error distribution needed for the Bootstrap is formed from the scaled residuals of the fitted model i.e.

$$\text{PPCI} = (\text{observed value} - \text{fitted value}) / (\text{weight})^{1/2} / \text{scaling factor}$$

$$\text{PPCF} = (\text{observed value} - \text{fitted value}) / \text{fitted value} / (\text{weight} * \text{scaling factor})^{1/2}.$$

These residuals are all on the same standardized scale. The scaling factor is given by GLIM.

The residuals are presumed to be independent and identically distributed. The homoscedasticity is gained by the method outlined in Section 3.2.2.

No attempt is made to model slight departures from the independence assumption when refitting the residuals in the simulation. An example of such a slight departure is when we observe a set of payments in one calendar year which may all be below the fitted values. We may not wish to model this feature of the data as there may be no convincing explanation of the phenomenon and therefore no way of predicting its possible recurrence. By leaving it out of the model we increase the residual variation and thus, *ceteris paribus*, the estimated error of prediction.

3.3.2. *Heteroscedasticity in the Prediction Cells*

Once the pseudo-data have been simulated and a re-estimation of the model parameters made there is another use which can be made of the error distribution—future statistical error can be estimated by simulation techniques. The raw errors can be rescaled by the scale factor estimated by GLIM and by either the number of claims expected to be finalised (PPCF method) or the total number of claims incurred for the year of origin (PPCI method). The vector used to reduce heteroscedasticity in the fitted cells is also used to reintroduce that heteroscedasticity in the predicted payments. Thus future payments can be estimated which vary both because of realisations of the estimation error (calculated by way of fitting the pseudo data) and the statistical error (fitting the observed residuals onto predicted average payments).

The estimation error and the statistical error are assumed to be independent and so the total variance can be decomposed into these two parts and their relative importance gauged.

3.4.

From the pseudo-estimates of the jackknife and the iterations of the bootstrap we have numerous estimates of future payments. As in TAYLOR and ASHE these payments are broken down by year of occurrence of the claim and the year of payment. These estimates are produced using GLIM and are output to a temporary computer file. A post-processor has been written which takes these numerous estimates and produces a covariance matrix of the estimated future payments for each of these cells.

With the parametric method this covariance matrix is produced directly from the estimated covariance matrix of the parameter coefficients and the estimated variance of the payments.

This large covariance matrix can be collapsed in various ways to give such statistics as:

- (i) the standard deviation of the outstanding claims for each accident year and the correlation matrix of the accident year estimates, or
- (ii) the standard deviations of the total claims in each payment year.

From the covariance matrix we have directly the standard deviation of the expected payments in the next year for each accident year. This can be used to judge the suitability of the model when the claims estimation is redone in one year's time.

4. CONCLUSIONS AND DISCUSSION

4.1.

The following table shows the results of applying the second moment estimators with the two estimation methods to the different portfolios.

Port- folio	Model	Estimate			Standard deviation			
		Jackknife Weighted (a)	Equal (b)	Para- metric	Jackknife Weighted (a)	Equal (b)	Parametric	Bootstrap
		\$M	\$M	\$M	\$M	\$M	\$M	\$M
A	PPCF	1639.9	1639.9	1660.0	75.4	59.9	50.3	50.9
B	PPCF	85.1	85.0	82.3	22.6	6.6	8.8	5.6
	PPCI	60.9	61.0	65.2	20.5	6.8	—	5.2
C	PPCF	27.2	27.1	27.2	2.4	1.9	2.7	2.4
D	PPCI	140.6	140.6	144.2	52.9	14.7	—	11.1
E	PPCF	126.7	126.7	128.4	12.0	9.0	13.2	9.8

Notes:

(a) Applying formula (1).

(b) Applying the usual jackknife formula with equal weights given to all cells.

4.2.

The two methods for evaluating the jackknife outstanding claims give very different estimates for its standard deviation. The estimates deriving from the weighted jackknife are sometimes so far removed from the bootstrap and parametric estimates that it can be confidently stated the approach is wrong.

4.3.

The three methods generally give similar results which gives us greater confidence that we know the level of error in our estimate caused by the statistical procedures used. The variations in the results are not enough to discredit the high (to some) coefficients of variation calculated using the approach of TAYLOR and ASHE.

In the absence of any indication that one PPC method is superior to any other we can combine the estimates with equal weight given to each but taking care to combine like with like.

The jackknife gives an estimate of only the estimation error. We should therefore compare it to the estimation error of the parametric and bootstrap approaches. For the PPCF models we find that the estimation and statistical errors are approximately equal (it is assumed they are independent). We can then take the average of the three estimation errors and a second average of the two statistical errors which we then combine to give us a new estimate of the total error.

For example, in Portfolio C using the PPCF method we find the statistical error is 46% of the parametric methods estimate and 44% of the Bootstrap estimate.

The estimation error is

$$[(1.9)^2 + 0.54(2.7)^2 + 0.56(2.4)^2]/3 = 3.59 = (1.9)^2.$$

The statistical error is

$$[0.46(2.7)^2 + 0.44(2.4)^2]/2 = 2.94 = (1.7)^2.$$

The total error of prediction is taken as

$$3.59 + 2.94 = 6.53 = (2.6)^2.$$

4.4.

There is the special form of the specification error called competition bias by MILLER (1984). This error comes when a regression is attempted with many variables which are entered into the model one at a time with that variable giving the greatest reduction of the residual variance entering at each step. The paradigm of this technique is the stepwise regression of Efronson. If many variables of marginal significance are present then sampling variation can lead to one or more of these having the appearance of greater significance and therefore being included in the model. This tends to lead to lower residual variation and lower estimates of variance of prediction.

While the author does not use a formal stepwise regression procedure there is still the possibility that competition bias will be a significant source of error. The parametric and bootstrap methods will not measure this error. The jackknife analysis may give some indication that it may exist if there are only a few cells which cause the variable to appear significant. As these cells are deleted there may be a sufficiently large change in the coefficient of that variable so that the jackknife estimates of the values of the coefficient of the variable and the standard error of the coefficient lead to it being judged non-significant.

The statistical error is assumed to be of a certain parametric form. This assumption has little solid theoretical backing which raises doubts on the accuracy of the estimate of statistical error.

The bootstrap by using the distribution of the standardised residuals overcomes two problems of the parametric analysis. The residuals show how the data differ from the fitted values given both the parametric assumption used in the fitting algorithm and the form of the fitted function. These residuals are then added to the fitted function for future years to give realisations of future payments.

If we assume that:

- (i) the same model would be fitted when information is finally available on the future payments, and
- (ii) the observed residuals of the future payments would come from the same distribution as the past payments' residuals

then the bootstrap's realisations are non-parametric to a large extent.

This removes one possible source of specification error.

4.5.

The jackknife pseudo-estimates can give us an approximation to the influence function of Hampel (mentioned in REY, 1979). This function measures the influence on the statistic of a particular cell. By observing how the estimated outstanding claims change when a particular cell is removed we can judge the relative influence of that cell on our estimate. Very large absolute values of this function indicate that the cell is having too great an influence on our estimation process. Particular attention should be paid to such cells in order to find out anything unusual that may have occurred such as large payments to some claimants or changes in the claim handling procedures that may have led to some departure from the trend of payment around them. If too great a sensitivity is found then consideration could be given to changing the model to one that is more robust.

Plotting the residual values against various parameters of our model may not give sufficient indication of the effect of certain cells on the estimation as such plots indicate more the effect of cells on the fitting process.

4.6.

The jackknife was originally introduced as a method for removing the bias from estimators. The table indicates that the PPCF estimator does not have a large bias, being at most of the order of 2%. However the PPCI method has a larger bias, reaching 7% for the naive model fitted to Portfolio B and 2½% of Portfolio D.

4.7.

The parameter estimates and their standard deviations from the jackknife agree substantially with those from the parametric approach except for the coefficient of the inverse of the speed of finalisation (perhaps the most important variable in the model). The coefficient of this variable typically has approximately the

same estimate from the two approaches but its standard error is much higher for the jackknife approach. If the jackknife estimate is to be believed then the parametric approach will give an underestimate of the estimation error unless the changes in the covariance matrix of the parameter estimates counteracts this.

4.8.

The second moment of the claims distribution is used, *inter alia*, to judge the likelihood of sufficiency of provisions made for these claims. In the author's experience this is usually made by making some parametric assumption for the distribution of total claims (such as Normal or Lognormal) and then making a probabilistic statement on the adequacy.

The bootstrap iterations give us simulated realisations of the claim amounts and so allow us to approximate the distribution. Adequacy of provisions can then be judged from this distribution thereby bypassing the errors caused by assumption of the wrong distribution. The observed distribution is, as expected, skewed towards higher values. The degree of skewness is only slight for the portfolios considered here.

4.9.

In a recent paper, TAYLOR (1984) has advocated a formal way of combining the estimates of outstanding claims from different models. In its most pure form it requires the covariance matrix of the estimates for each year of occurrence from a given model and the covariance matrix of the years of occurrence between methods. While the first matrix could be obtained from the methods in his paper with ASHE the second covariance matrix was educated guesswork.

The jackknife offers a way to obtain this esoteric second matrix in some cases. With the PPCI and PPCF methods giving pseudo-estimates on the same data the matrix can be built up observing the correlations of the pseudo-estimates as each cell is dropped in turn. This technique can be used to give a correlation matrix for each accident year's future payment as determined by the two methods which must then be collapsed to the total accident year payment covariance matrix.

The correlation coefficient between the PPCF and PPCI estimates for all outstanding claims in Portfolio B has been calculated as 37%.

A disturbing point arises from the consideration of the two estimates for Portfolio B. The estimated standard deviation for the PPCI method is \$6.3M and for the PPCF method is \$9.0M and there is a 37% correlation between the two central estimates. Yet there is a \$24M difference between the two estimates which thus appear to be significantly different. Note that specification error has not been included in the quoted standard deviations.

For Portfolio B the PPCF model was used in practice to help estimate the outstanding claims. The PPCI model was not considered suitable and the model quoted here was fitted as an example for this paper. The fit to the past data is as adequate as the PPCF model in the statistical sense. In practice the PPCI

method has been taken as having a large specification error while the PPCF was taken as having a relatively small specification error.

This indicates the importance of analysing the data by more than one method. Although the specification error cannot be estimated formally an indication of its possible size can be gained by observing the range of estimates made by different techniques.

4.10

The PPCI bootstrap technique gives a comparatively low estimate of the statistical error viz \$2.2M for Portfolio B as compared to \$4.4M for the PPCF method, and \$1M for Portfolio D. These results are sufficiently unusual as to warrant suspicion on the technique used for their evaluation. This is currently being investigated further.

4.11

For the PPCF models the total variance of the estimated outstanding payments is approximately half estimation error and half statistical error. However for the individual accident years the statistical error rises to about 80% of the total variance or higher. This can be explained by the independence of the statistical errors between cells and therefore between accident years and the positive correlation of the estimation errors. As the accident years are combined the statistical errors add simply but the estimation errors become more important as they add together with the positive covariances.

5. CONCLUSION

No method for measuring the second moment of the estimate of outstanding claims stands out from the others as superior. Their features can be summarised thus:

Parametric	small number of calculations estimation and statistical errors available accurate if the parametric assumptions are correct
Jackknife	influence of individual data points on the estimate is available only estimation error is available estimate of outstanding claims with possible reduced bias
Bootstrap	non-parametric estimation and statistical errors available distribution of outstanding claims given.

As the estimates of the second moment are generally different for the three methods it is recommended that all three be evaluated then combined. The study of the three estimates will give greater insight to the evaluation of the quantum of outstanding claims.

REFERENCES

- BAKER, R. J. and NELDER, J. A. (1978) *The GLIM System, Release 3: Generalized Linear Interactive Modelling*. Numerical Algorithms Group: Oxford.
- BARTHOLOMEW, D. J. (1975) Errors of Prediction in Markov Chain Models. *Journal of the Royal Statistical Society B* 37, 444–456.
- DE JONG, P. and ZEHNWIRTH, B. (1980) A Random Coefficients Approach to Claims Reserving. *Research Paper no. 219: School of Economic and Financial Studies*, Macquarie University, North Ryde.
- DE JONG, P. and ZEHNWIRTH, B. (1982) Claims Reserving, State-space models and the Kalman Filter. *Journal of the Institute of Actuaries* 110, 157–182.
- FREEDMAN, D. A. and PETERS, S. C. (1984) Bootstrapping a Regression Equation: Some Empirical Results. *Journal of the American Statistical Association* 79, 97–106.
- MARTIN-LOF, P. (1974) The Notion of Redundancy and Its Use as a Quantitative Measure of the Discrepancy between a Statistical Hypothesis and a Set of Observational Data. *Scandinavian Journal of Statistics* 1, 3–18.
- MILLER, A. J. (1984) Selection of Subsets of Regression Variables. *Journal of the Royal Statistical Society A* 147, 389–425.
- MILLER, R. G. (1974) The Jackknife—A Review. *Biometrika* 61, 1–15.
- QUENOUILLE, M. H. (1974) Notes on Bias in Estimation. *Biometrika* 43, 353–360.
- REY, W. J. J. (1979) Robust Statistical Methods. *Lecture Notes in Mathematics* v690, Springer-Verlag: Berlin.
- TAYLOR, G. C. (1981) An Invariance Principle for the Analysis of Non-life Insurance Claims. *Journal of the Institute of Actuaries* 110, 205–242.
- TAYLOR, G. C. (1985) Combination of Estimates of Outstanding Claims in Non-Life Insurance. *Insurance: Mathematics & Economics* 4, 81–91.
- TAYLOR, G. C. and ASHE, F. R. (1983) Second Moments of Estimates of Outstanding Claims. *Journal of Econometrics* 23, 37–61.
- TUKEY, J. W. (1958) Bias and Confidence in not-quite large samples. *Annals of Mathematical Statistics* 29, 614.

FRANK ASHE

E. S. Knight & Co. Research Centre, 71 York Street, Sydney, N.S.W. 2000, Australia