Second Moments of Estimates of Outstanding Claims

Article *in* Journal of Econometrics · September 1983

DOI: 10.1016/0304-4076(83)90074-X

Gregory Clive Taylor
UNSW Australia

Journal of Econometrics 23 (1983) 37–61. North-Holland

SECOND MOMENTS OF ESTIMATES OF OUTSTANDING CLAIMS

G.C. TAYLOR and F.R. ASHE

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

We begin by considering the lack of methodology for obtaining second moments of outstanding claims in non-life insurance (sections 1 and 2), and give some arguments as to why this lack is to be deprecated (section 3). We suggest that part of the reason is the lack of generality of the models and methods currently in use in claims analysis, and we suggest further that some unification might be achieved through regression analysis (section 4). The claims analysis problem is formulated and solved in terms of regression methods (sections 5 to 8), estimates of both first and second moments of outstanding claims being obtained. Two specific examples of the recommended regression procedures are presented algebraically (section 9). One of these is extended to a numerical example (section 10). The paper concludes with some general observations (section 11).

1. Introduction

The estimation of outstanding claims is a significant component in the determination of operating profit of a general insurer. Commonly, the required provision for outstanding claims will be a number of times larger than annual profit. In these circumstances, errors of quite tolerable, or even statistically unavoidable, proportions in the estimation of outstanding claims can lead to such serious consequences as the reversal of sign of the emerging profit.

However efficient the method used for the estimation of outstandings, the result cannot be better than an estimator, in the statistical sense, and subject to inherent statistical error. Given this background, one finds it interesting that there has, to date, been little systematic development of techniques for examining the dispersions of the various estimates of outstandings currently in use.

There are in fact notable exceptions to this in the literature [Bühlmann, Schnieper and Straub (1980), Hachemeister (1978, 1980), Reid (1978)]. However, these exceptions have a common feature. Each requires rather more detailed claims data than is routinely available. The additional information allows the authors concerned to build up proper stochastic models of the claims process. Such stochastic models admit evaluation of the higher moments of estimated outstanding claims in a natural way.

Another exception which is not demanding in its data requirements and is highly flexible is proposed by de Jong and Zehnwirth (1980, 1982).

Sections 6.3, 6.4 and 8 are rather abbreviated versions of their originals [Taylor and Ashe (1982)]. The complete version of the paper may be obtained on application to the authors.

2. The lack of technique for obtaining second moments

There appear to be several reasons for the disinterest in second moments which has dominated outstanding claims estimation techniques so far.

Firstly, company accounts require the entry of a single figure for outstanding claims liability. In effect, they fail to come to grips with the stochastic nature of the liability. However that may be, the initial effect of this requirement is to focus the attention of management on central estimates (e.g. the mean) of outstanding claims without reference to the higher-order moments of the estimators employed.

Secondly, the rather *ad hoc* nature of the estimation methods used in many of the common techniques [chain ladder described, for example, by Taylor and Matthews (1977), separation by Taylor (1977), etc.] have a heuristic development and are not well adapted to the estimation of higher-order moments. To a large extent, the method proposed by de Jong and Zehnwirth (1980, 1982) circumvents these difficulties.

Thirdly, there is a good deal of feeling among actuaries that attempts to estimate higher-order moments are futile anyway. One does hear such statements as: 'It is pointless to concern oneself with second moments until the problem of first moments has been solved.'

We consider these three points in turn.

Firstly, although the statutory accounting requirements tend to engender a fixation on central estimates of outstanding claims, enlightened managements are coming to recognise the value of information concerning variability of these estimates. The authors' own experience is one of steadily increasing inquisitiveness in this respect on the part of management.

Requests for information on this variability arise in a range of contexts:

(i) the insurer wishes to establish a strategic margin in his provision for outstanding claims;
(ii) auditors wish to gauge the safety of the provision for outstanding claims adopted in an insurer's accounts;
(iii) a statutory body, responsible for the examination of filed requests for premium rate increases, wishes to obtain a measure of variability of the business being rated in order to assess the reasonableness of the safety margin included in the filed premium rate.

Indeed, within our own practice, the situation is now such that we have been enquiring of ourselves whether second moments, far from being a theoretical curiosity possibly to be affixed to the 'real' substance of an investigation, are dispensable at all.

As to the second of the above comments, its validity must be accepted and acted upon. There is a need for models with a more rigorously formulated structure which recognises the essentially stochastic nature of the claims process and is amenable to analysis taking this nature into account.

The third point, the one that attempts to consider second-order moments are futile when the treatment of first-order moments is so crude, is one which we thoroughly reject. The short section 3 is devoted to this issue.

3. The need for second-order moments

To us, statements such as that quoted in section 2 are as logically dubious as they are unprofitable. Because, of course, 'the problem of first moments' will never be solved except perhaps in the sense of producing estimators of outstanding claims which resemble minimum-variance estimators. In other words, even with data problems, inflation uncertainty and so on set aside, the inherently stochastic nature of the claims process will prevent the actuary from stating the amount of outstanding claims with certainty. All that can be done is to construct an estimate and indicate the margin of uncertainty. Models which provide a truer representation of the actual process will tend to reduce the margin of uncertainty associated with the estimate.

Thus one arrives at the conclusion that it is only in terms of second moments (or some similar distributional characteristics) that 'the solution' of the problem of first moments can be adjudged to be indeed the solution. On this reasoning, the statement 'It is pointless to concern oneself with second moments until the problem of first moments has been solved,' becomes a logical absurdity.

At this point we might say that we are not convinced that the crudity of technique often alleged by actuaries is quite so all-pervading as is argued. However, whether or not the allegations are correct is not material to our argument. What we have sought to argue above is this: To the extent that a method is crude, i.e., not an accurate representation of reality, this fact will manifest itself in the form of a large ratio of standard deviation to mean in the estimates of outstanding claims.

Indeed, this ratio might be adopted as the measure of what is crude and what isn't. If the ratio is small, then the model involved may be accepted as a reasonable representation of reality irrespective of how simple it might be. We do not necessarily equate crudity with simplicity.

The common practice of actuaries currently in providing estimates of outstanding claims is to quote:

(i) a single ('best' or 'central') estimate;

OR (ii) a range, determined in a subjective and undefined manner, in which the amount of outstanding claims is 'likely' (but with unspecified probability) to lie;

OR (iii) both (i) and (ii).

If several actuaries choose alternative (i), then the emergence of differences in results is inevitable. The existence of these differences is not inherently alarming, given that the process under investigation is stochastic. What is unfortunate is that the recipient of the estimates has no means of assessing the significance of the differences.

If alternative (ii) is chosen, its nebulous nature is likely to render it futile. In particular, what meaning is to be attached to the concept of 'likely'. Does it mean 'with 95% probability'? 'With 50% probability'? Or what? There may be a temptation for the actuary to adopt *tacitly* a high probability, thus expanding his likely range to the point where it is too wide to be of practical use. On the other hand, if a low probability is adopted tacitly, with the concomitant narrowing of the likely range, the term 'likely' may be misleading to the recipient of estimates.

Alternative (iii) suffers equally from the dubious value of (ii).

Now consider a situation in which several actuaries estimate outstanding claims for the same portfolio and on the basis of the same data. But suppose that, contrary to common practice, each actuary considers total outstanding claims as a random variable and provides:

(i) an estimate of the mean;
(ii) an estimate of the variance of (i).

In such circumstances the recipient of the advice can form:

(i) some idea of the significance or otherwise of differences between the several estimates of the mean;
(ii) an opinion of the quality of the estimates of mean by reference to the associated variances.

Hence the collection of estimates of mean which, in the absence of second moment estimates, might seem inconsistent, may well be seen to be consistent in the light of the second moments.

4. The use of regression methods

It is our opinion that regression methods should assume a more prominent role in the estimation of outstanding claims than has been the case hitherto.

Such an opinion does not spring merely from a desire to make actuarial work 'just like statistics' (though we do not wish to exclude this possibility). There are two more specific and, in our opinion, cogent reasons — one concerned with first moments of estimated outstanding claims, the other with second moments.

If one surveys the current literature of models and methods, one tends to find a lack of generality. The methods used in the estimation of the parameters involved are often based on heuristic reasoning, and so are rather specific to the models with which they are associated. Because of this, a simple desirable change in model structure may lead not to a correspondingly simple change in, but rather a total overhaul of, estimation procedure.

We give one example. Consider a very simple method, the payments per unit of risk method [Cumpston (1976)]. The model can be written as

$$\mathrm{E}C_{ij} = n_i q_j, \tag{1}$$

where

C_{ij} = claim payments in development period j of period of origin i expressed in constant dollar values, i.e., adjusted to remove the effect of inflation;
n_i = number of units of risk (i.e., a measure of exposure) in period of origin i;
q_j = expected claim payments in development period j of period of origin i per unit of risk in period of origin i.

The method involves estimation of q_j by

$$\hat{q}_j = \sum_i c_{ij} \Big/ \sum_i n_i, \tag{2}$$

where c_{ij} denotes a realization of the random variable C_{ij}, and summation in (2) is over all values of i for which realizations c_{ij} exist.

Suppose now that one wishes to introduce the following 'slight' modification to model (1):

$$\mathrm{E}C_{ij} = n_i \mu_i q_j, \tag{3}$$

where μ_i is a factor, dependent on period of origin, and reflecting changes in average claim size from one such period to another.

It is not immediately clear how the estimator (2) is to be changed to accommodate the new model (3). To derive a solution to this problem is not difficult. But that is not the point at issue. The point is that the change in model from (1) to (3) generates quite a new model, and a new estimation procedure has to be produced essentially *ab initio*.

Note, however, that (3) may be written as

$$\log(C_{ij}/n_i) = \log \mu_i + \log q_j + \varepsilon_{ij}, \tag{4}$$

where ε_{ij} is a stochastic error term. This provides a *linear model* of $\log(C_{ij}/n_i)$. Hence regression techniques may be used to fit the model. Similarly, regression techniques could have been applied to model (1), though such an application is trivial.

One must, of course, bear in mind that the application of regression methods needs to be accompanied by a certain degree of care. The following questions, for example arise:

(i) What is the distribution of C_{ij}?
(ii) How is this distribution changed by the transformation of C_{ij} (logarithmic in the above case) required to produce a linear model?
(iii) What is the variance of $\log(C_{ij}/n_i)$, and so what weight should be given to the (i,j)-cell in the regression fitting?

These (and perhaps others) are all questions which need proper consideration if a model is to be fitted to the data by regression methods. As will become apparent in section 10, further questions arise in connection with the modelling of parameters like $\log \mu_i$.

In general, if regression methods are to be used, each data set must be regarded as comprising, at least potentially, its own set of problems. An attempt should be made to ascertain what prominent features are contained in the data. Simple plots will usually be of assistance here. Some experimentation with different assumptions may be carried out to obtain rough answers to questions of the type raised above.

What we are *not* recommending is a sausage-machine approach consisting of some wholly automated regression procedure applied blindly to every set of data.

The second point in favour of regression methods is that they contain established techniques for estimation of second moments. Moreover, the linearity of the model will render the combination of second moments of the various parameters relatively simple computationally.

Consider, for example, model (4). Let $\hat{\mu}_i$ and $\hat{q}_j$ be estimators of $\log \hat{\mu}_i$ and $\log q_j$, respectively. Then, in the course of the regression analysis, it will be possible to obtain the covariance matrix of all the variables $\hat{\mu}_i$, $\hat{q}_j$, ε_{ij}. These covariances can then be combined quite simply to provide an estimate of the covariance matrix of the $\log(C_{ij}/n_i)$ for (i,j)-cells relating to the future. We give a general version of these statements in section 6.3.

Once again, however, care may be required. In the example of (4) dealt with above, the problem of deriving covariances of the C_{ij} from those of the $\log(C_{ij}/n_i)$ would require careful treatment. Nonlinear transformations of the basic data will introduce this type of problem. In the major example given in section 10, we use a considerably more complex (and, we think, more efficient) model than (1) or (4), but it has nevertheless the advantage of not introducing nonlinearity at any stage.

In this latter type of model, analysis of the above type will produce the covariance matrix of the future C_{ij}; and hence of any linear combination of these, e.g., total outstanding claims, outstanding claims subdivided by accident year, future cash flows generated by currently outstanding claims, etc.

5. Formulation of the regression procedure

5.1. Claims model

Conventionally, we suppose that the process generating claim payments is described by an abstract model. We suppose this model to be stochastic, and to deal with claim payments C_{ij} indexed by period of origin and development period (there are other possibilities, e.g., indexed by notification period instead of period of origin).

Thus the model takes the form

$$C_{ij}=g(\omega,i,j)+\varepsilon_{ij}, \tag{5}$$

where g is some prescribed function, ω is a vector of parameters, and ε_{ij} is a stochastic variate. The function $g(\cdot)$ is adjusted so that

$$\mathrm{E}\varepsilon_{ij}=0. \tag{6}$$

5.2. Fitting the model

It is supposed that observations c_{ij} on the random variables C_{ij} are available for all (i,j) in some set $\mathscr{I}$. It is then necessary to obtain an estimate $\hat{\omega}$ of the parameter vector ω by reference to this data.

Generally, there might be some advantage in fitting a transformed model to transformed data, thus,

$$S_{ij}=h(C_{ij})=h(g(\omega,i,j))+e_{ij}, \tag{7}$$

where h is some function and e_{ij} is again a stochastic variate with

$$\mathrm{E}e_{ij}=0. \tag{8}$$

Still working generally, one is now able to fit the model, obtaining the required estimates $\hat{\omega}$ by weighted least squares. That is, the objective function

$$Q = \sum_{(i,j)\in\mathscr{I}} w_{ij}[s_{ij} - h(g(\omega, i,j))]^2 \tag{9}$$

is set up, and the estimate $\hat{\omega}$ taken as that value of ω which minimizes Q.

In (9), the s_{ij} are realizations of the S_{ij}, and w_{ij} is the *weight* associated with the (i,j)-cell of data. It is assumed that w_{ij}^{-1} is proportional to the variance of S_{ij},

$$V[S_{ij}] = w_{ij}^{-1}\sigma^2,$$

for some constant σ^2 independent of (i,j).

It is fitting to interpolate here a remark on the use of weighted least squares. This will be especially suitable if the e_{ij} are normally distributed. In other circumstances, its use may be open to question. In the algebra which follows we use weighted least squares throughout. Parallel arguments could be developed from the other distributional assumptions. Some computer packages will perform regressions on the basis of these other distributional assumptions. The package which we use (see sections 10 and 11) certainly does. In the data sets with which we have dealt, the use of non-normal distributions does not seem to have led to any improvement in fitting a model.

Now the quadratic form Q can be written in terms of matrices and vectors. We choose an ordering of the $(i,j)\in\mathscr{I}$, and write out the s_{ij} in a column vector in this order. Call this vector $\boldsymbol{s}$. Similarly, the $h(g(\omega, i,j))$ are arranged in a column vector $\boldsymbol{h}$, and the w_{ij} in a diagonal matrix W. Then (9) becomes

$$Q = (\boldsymbol{s} - \boldsymbol{h})^T W(\boldsymbol{s} - \boldsymbol{h}), \tag{10}$$

where the upper T designates matrix transposition.

The present paper is concerned with a special case of the structure of $\boldsymbol{h}$. We consider the case in which the function $h(g(\omega, i,j))$ can be written in the form.

$$\begin{aligned} h(g(\omega, i,j)) &= \text{linear combination of the components of } \omega \\ &= \boldsymbol{x}^T(i,j)\,\omega, \end{aligned} \tag{11}$$

where $\boldsymbol{x}(i,j)$ is the column vector of coefficients in the linear combination, and may be dependent on (i,j).

If the $\boldsymbol{x}^T(i,j)$ are written out in a column (using the same ordering as established above), they produce a matrix X. Then (11) gives

$$\boldsymbol{h} = X\boldsymbol{\omega}, \tag{12}$$

and (10) becomes

$$Q = (\boldsymbol{s} - X\boldsymbol{\omega})^T W (\boldsymbol{s} - X\boldsymbol{\omega}). \tag{13}$$

The problem now is to determine the value $\tilde{\boldsymbol{\omega}}$ of $\boldsymbol{\omega}$ which minimizes Q.

This is the standard set-up of a weighted regression problem, and $\hat{\boldsymbol{\omega}}$ may be found by standard regression techniques (see section 6.2).

5.3. *Projection of outstanding claims*

Let $\mathscr{I}_i^0$ denote the subset of $\mathscr{I}$ consisting of all (i,j) (i given) such that j is a *future* development year. Then outstanding claims in respect of year of origin i are given by

$$\hat{P}_i = \sum_{j:(i,j)\in\mathscr{I}_i^0} \hat{C}_{ij}, \tag{14}$$

where $\hat{C}_{ij}$ is the *regression estimate* of C_{ij}.

Thus, outstanding claims are easily projected provided that regression estimates $\hat{C}_{ij}$ are easily found.

One case in which the $\hat{C}_{ij}$ will be easily obtainable is that in which the function $h(\cdot)$ is linear, i.e., S_{ij} is a linear transformation of C_{ij}. Then the $\hat{C}_{ij}$ are found as linear transformations of the $\hat{S}_{ij}$. The major example, dealt with in section 10, is of this type.

Brief treatment of the case of nonlinear $h(\cdot)$ is given in section 8.

5.4. *Second moments of future claims*

There are various subsets of future claims in which we might be interested, e.g., outstanding claims according to accident year, future cash flows according to payment year, etc.

For the purpose of this section we suppose merely that our interest lies in linear combinations of some subsets of the future $\hat{C}_{ij}$, i.e., in

$$\hat{\boldsymbol{P}} = U\hat{\boldsymbol{C}}, \tag{15}$$

where $\hat{\boldsymbol{C}}$ is the vector of future $\hat{C}_{ij}$ written out in a particular order, and U is a matrix each of whose rows provides a vector of the coefficients involved in one of the linear combinations of interest. These coefficients can differ from unity if, for example, their function is to include inflation in future claim payments or to allow for discounting for investment income.

Now the covariance matrix of $\hat{\boldsymbol{P}}$ is

$$V[\hat{\boldsymbol{P}}]=UV[\hat{\boldsymbol{C}}]U^T, \tag{16}$$

and so the problem is reduced to the problem of estimation of the covariance matrix of estimated future payments $\hat{\boldsymbol{C}}$.

As in section 5.3, the problem is easily solved when the function $h(\cdot)$ is linear.

6. Details of the regression procedure

6.1. Further specification of the problem

Generally we use the notation established in section 5. However, as noted in sections 5.3 and 5.4, considerable simplification of the problem is achieved if the function $h(\cdot)$ is linear. This will be assumed to be the case in the present section.

This requirement, together with (12), implies that the function $g(\cdot)$ is also a linear transformation of the parameter vector.

Under these conditions, (13) becomes

$$Q=(s-X\omega)^T W(s-X\omega), \tag{17}$$

with

$$s=Hc, \quad H \text{ diagonal}; \tag{18}$$

$$X=HG, \quad G \text{ the matrix representing the linear transformation } g(\cdot). \tag{19}$$

The parameter vector $\hat{\omega}$ is to be estimated by $\hat{\omega}$, obtained by minimization of Q.

Then estimates of some linear combinations of future claim payments are to be obtained [see (15)]. Combining this equation with (18), (12) and (19) in succession gives

$$\hat{\boldsymbol{P}}=U\hat{\boldsymbol{C}}=UH^{-1}\hat{\boldsymbol{S}}=UH^{-1}X\hat{\omega}=UG\hat{\omega}. \tag{20}$$

Finally, the covariance matrix of $\boldsymbol{P}$ is to be found. Corresponding to (16), we have

$$V[\hat{\boldsymbol{P}}]=UGV[\hat{\omega}]G^TU^T. \tag{21}$$

6.2. Estimation of first moments

This section deals with the calculation of $\hat{\boldsymbol{P}}$, given in (20). The estimator $\hat{\omega}$ is well-known [see e.g. Bibby and Toutenburg (1977, pp. 56–57)]. It is

$$\hat{\omega} = Z^{-1} X^T W s, \tag{22}$$

where

$$Z = X^T W X. \tag{23}$$

Then, by (20) and (22),

$$\hat{\boldsymbol{P}} = U G Z^{-1} X^T W s. \tag{24}$$

In practice, if some form of regression package is being used, the estimates $\hat{\omega}$ would normally be provided directly by the package, and then incorporated directly in (20) to give $\hat{\boldsymbol{P}}$.

6.3. Estimation of second moments

This section deals with the estimation of $V[P]$, given by (21). This last expression can be expanded to

$$V[\hat{\boldsymbol{P}}] = U G Z^{-1} X^T W V[S] W X Z^{-1} G^T U^T. \tag{25}$$

by means of (24).

Now, using the fact (section 5.2) that

$$V[S_{ij}] = w_{ij}^{-1} \sigma^2, \tag{26}$$

in (25) yields the following unbiased estimator of $V[\hat{P}]$:

$$\hat{V} = \hat{\sigma}^2 U G Z^{-1} G^T U^T, \tag{27}$$

where $\hat{\sigma}^2$ is an unbiased estimate of σ^2.

The standard regression estimate of σ^2 is

$$\hat{\sigma}^2 = r^{-1} s^T [I - W X Z^{-1} X^T W] s, \tag{28}$$

where

$$r = \dim s - \operatorname{rank} X. \tag{29}$$

Apart from the factor r^{-1}, (28) is of course the weighted sum of squared residuals of the regression.

Comparison of (20) with (27) indicates that

$$V[\hat{\omega}]=\sigma^2 Z^{-1}.$$

In practice, a regression package will usually provide the estimate $\hat{\sigma}^2 Z^{-1}$ of this matrix, and the user will need only to form $\hat{V}$ by multiplication by U, G in accordance with (27).

7. Estimation error and statistical error

There remains one slightly subtle point in the estimation of second moments of $\hat{\boldsymbol{P}}$.

Note the nature of $\hat{\boldsymbol{P}}$ in (20). It is an unbiased estimator of the aggregation $U\hat{\boldsymbol{C}}$ of future claim payments. Therefore, its variance, as given by (21), is the variance of this estimator about the parameter which it estimates.

It is necessary to distinguish $\hat{\boldsymbol{P}}$ from $\boldsymbol{P}=U\boldsymbol{C}$, the vector of random variables of which $\hat{\boldsymbol{P}}$ provides a forecast. Note that here $\boldsymbol{C}$ is being used to denote *future* values of C_{ij} written out as a vector, whereas in section 6.1 $\boldsymbol{C}$ denoted past values.

Since $\hat{\omega}$ is unbiased for ω, $E\boldsymbol{P}$ is linear in ω, and $\hat{\boldsymbol{P}}$ is the same linear function of $\hat{\omega}$,

$$\mathrm{E}\hat{\boldsymbol{P}}=\mathrm{E}\boldsymbol{P} \quad (\hat{\boldsymbol{P}} \text{ unbiased for } \mathrm{E}\boldsymbol{P}). \tag{30}$$

Note that in normal practical applications, e.g., when $\boldsymbol{P}$ denotes the vector of outstanding claims at a given date according to year of origin, the variance in which one will be interested is not simply $V[\hat{\boldsymbol{P}}]=E[\hat{\boldsymbol{P}}-E\boldsymbol{P}]^2$, but $E[\hat{\boldsymbol{P}}-\boldsymbol{P}]^2$. Let this last expression be denoted by $V^*[\hat{\boldsymbol{P}}]$.

Note that $V[\hat{\boldsymbol{P}}]$ measures only the variation of the *forecast* $\hat{\boldsymbol{P}}$ about its mean, this variation arising from uncertainty in the parameter estimates $\hat{\omega}$. On the other hand, $V^*[\hat{\boldsymbol{P}}]$ is concerned with the variation between the forecast $\hat{\boldsymbol{P}}$ and the corresponding *random outcome* $\boldsymbol{P}$. It therefore includes allowance for the fact that, quite apart from uncertainties arising from parameter estimation, $\boldsymbol{P}$ is stochastic. Formally,

$$V^*[\hat{\boldsymbol{P}}]=V[\hat{\boldsymbol{P}}]+V[\boldsymbol{P}]-2\operatorname{cov}[\hat{\boldsymbol{P}},\boldsymbol{P}], \tag{31}$$

the last term here representing the covariance of $\hat{\boldsymbol{P}}$, $\boldsymbol{P}$. Note that $\hat{\boldsymbol{P}}$ is a function of past observations, whereas $\boldsymbol{P}$ is a vector of future observations.

There will usually be no reason to expect correlation of such disjoint sets of observations, in which case (31) becomes

$$V^*[\hat{\boldsymbol{P}}] = \underset{\substack{\text{estimation}\\ \text{error}}}{V[\hat{\boldsymbol{P}}]} + \underset{\substack{\text{statistical}\\ \text{error}}}{V[\boldsymbol{P}]}. \tag{32}$$

In the terminology introduced here, the *estimation error* is a measure of the uncertainty in the forecast $\hat{\boldsymbol{P}}$ due to estimation error in the parameters $\boldsymbol{\omega}$ underlying $E\boldsymbol{P}$; the *statistical error* is the additional uncertainty arising from the fact that, even if $\hat{\boldsymbol{P}}$ were known to estimate $E\boldsymbol{P}$ with complete accuracy, the stochastic nature of $\boldsymbol{P}$ would still generate uncertainty [Bartholomew (1975)].

An estimator of $V[\hat{\boldsymbol{P}}]$ was obtained in section 6.3. An estimator of $V[\boldsymbol{P}]$ is now required.

This can be obtained using (20) with the hats removed, (26) and an argument like that of section 6.2, giving the unbiased estimator of $V[\boldsymbol{P}]$,

$$\hat{\sigma}^2 U H^{-1} W^{-1} H^{-1} U^T, \tag{33}$$

where $\hat{\sigma}^2$ is the unbiased estimator of σ^2 given in (28).

Combining the unbiased estimates (27) (estimation error) and (33) (statistical error), one finally arrives at the following unbiased estimate of $V^*[\hat{\boldsymbol{P}}]$:

$$\hat{V}^* = \hat{\sigma}^2 U[GZ^{-1}G^T + H^{-1}W^{-1}H^{-1}]U^T. \tag{34}$$

8. The use of nonlinear transformations

The analysis of sections 6 and 7 has proceeded on the assumption that the transformation $h(\cdot)$ is a linear one, although when introduced in (7) $h(\cdot)$ was not necessarily linear. We now turn our attention very briefly to the case of nonlinear $h(\cdot)$.

Clearly, this case will be much more difficult to treat than the one dealt with in sections 6 and 7. It will be possible to obtain only approximate results. Such results as are obtained may nevertheless prove useful.

Standard approximations based on a Taylor series expansion of a vector function $\boldsymbol{\mu}(\cdot)$ operating on a random vector $\boldsymbol{Y}$ lead to

$$\mathrm{E}u_i(\boldsymbol{Y}) = u_i(\boldsymbol{\mu}) + \tfrac{1}{2}\mathrm{tr}\,\{[\mathrm{D}^2 u_i(\boldsymbol{\mu})]V[\boldsymbol{Y}]\}, \tag{35}$$

where $u_i(\cdot)$ is the ith component of the vector function $u(\cdot)$,

$$\mathrm{D}u_i(\boldsymbol{\mu})=\left(\frac{\partial u_i}{\partial y_1}(\boldsymbol{\mu}),\frac{\partial u_i}{\partial y_2}(\boldsymbol{\mu}),\ldots\right),$$

and $\mathrm{D}^2u_i(\boldsymbol{\mu})$ is the matrix whose (j,k)-element is $(\partial^2u_i/\partial y_j\partial y_k)(\boldsymbol{\mu})$.

Similarly,

$$\begin{aligned}V[u_i(\boldsymbol{Y})]&=\mathrm{E}\,\mathrm{D}u_i(\boldsymbol{\mu})(\boldsymbol{Y}-\boldsymbol{\mu})(\boldsymbol{Y}-\boldsymbol{\mu})^T[\mathrm{D}u_i(\boldsymbol{\mu})]^T\\&=[\mathrm{D}u_i(\boldsymbol{\mu})]\,V[\boldsymbol{Y}]\,[\mathrm{D}u_i(\boldsymbol{\mu})]^T,\end{aligned}\tag{36}$$

to second order.

Eq. (35) can be used as follows. Suppose that it has been found convenient to carry out a regression after a nonlinear transformation of the claims data. Accordingly, the C_{ij} have been transformed to S_{ij} by (7) in which the function $h(\cdot)$ is nonlinear.

We suppose that our interest still lies in $\hat{\boldsymbol{P}}=U\hat{\boldsymbol{C}}$ as in (15), where $\hat{\boldsymbol{C}}$ is an unbiased estimator of $E\boldsymbol{C}$.

Write $u(\cdot)=h^{-1}(\cdot)$ provided that the latter exists. Note that, by (35) $u(\hat{\boldsymbol{S}})$ will not be an unbiased estimator of $E\boldsymbol{C}$. However, it can be shown using (35) that an unbiased (to second order) estimate of $\boldsymbol{C}$ is

$$\hat{\boldsymbol{C}}=u(\hat{\boldsymbol{S}})+\tfrac{1}{2}[\Delta(\boldsymbol{S})-\Delta(\boldsymbol{S})],\tag{37}$$

where

$$\Delta_i(\boldsymbol{Y})=\tfrac{1}{2}\mathrm{tr}\,\{[\mathrm{D}^2u_i(\mathrm{E}\boldsymbol{Y})]\,V[\boldsymbol{Y}]\}.$$

The ith component of $\Delta(\boldsymbol{S})-\Delta(\hat{\boldsymbol{S}})$ is estimated by

$$\mathrm{tr}\,\{\mathrm{D}^2u_i(\hat{\boldsymbol{S}})\{V[\boldsymbol{S}]-V[\hat{\boldsymbol{S}}]\}\},$$

where estimates of the two covariance matrices will be available from the regression of $\boldsymbol{S}$.

Similarly, (36) can be used to obtain an estimate $\hat{V}[\hat{\boldsymbol{C}}]$ *of* $V[\hat{\boldsymbol{C}}]$ for $\hat{\boldsymbol{C}}$ defined as in (37),

$$\hat{V}[\hat{C}]=[\mathrm{D}u(\mathrm{E}\hat{\boldsymbol{S}})]\,V[\hat{\boldsymbol{S}}][\mathrm{D}u(\mathrm{E}\hat{\boldsymbol{S}})],\tag{38}$$

where once again an estimate of $V[\hat{\boldsymbol{S}}]$ will be available from the regression of $\boldsymbol{S}$.

Eqs. (37) and (38) provide estimators of first and second moments of $\hat{\boldsymbol{C}}$. From there it is a simple matter to obtain estimates of first and second moments of the vector of interest, $\boldsymbol{P}=U\boldsymbol{C}$.

9. Examples

9.1. Payments per claim incurred method of analysis

This model is described in section 4 [see (1)]. We take the n_i as non-stochastic even though, realistically, there will be a (usually small) stochastic component. Interpreting the model in the notation of section 5, the parameter vector $\boldsymbol{\omega}$ consists of the q_j. Then (5) is

$$C_{ij}=g(\omega,i,j)+\varepsilon_{ij}=G\omega+\varepsilon_{ij}, \tag{39}$$

where G is a diagonal matrix containing the n_i (with suitable repetitions).

The function $h(\cdot)$ [see (7)] is taken as the identity transformation. This is a special case of that dealt with in sections 6 and 7 [where it was required that $h(\cdot)$ be linear]. In this case,

$$\boldsymbol{S}=\boldsymbol{C}. \tag{40}$$

As in section 5.2, it is supposed that

$$V[C_{ij}]=w_{ij}^{-1}\sigma^2. \tag{41}$$

More specifically for this case, it is assumed that

$$V[C_{ij}]=n_i w_j^{-1}\sigma^2, \tag{42}$$

where the w_j would need to be determined by experimentation with the data.

With this translation of notation, (19) becomes

$$X=HG=G; \tag{43}$$

(24) becomes

$$\hat{\boldsymbol{P}}=UGZ^{-1}G^TWc, \tag{44}$$

where

$$Z=G^TWG; \tag{45}$$

and finally (34) becomes

$$\hat{V}^* = \hat{\sigma}^2 U[GZ^{-1}G^T + W^{-1}]U^T. \tag{46}$$

9.2. Invariant see-saw method of analysis

As a second example, a rather more complex one is chosen. The method involved is the *invariant see-saw method*. This is developed and described by Taylor (1981). A brief recapitulation of the method is given below.

Let

n_i = number of claims incurred in year of origin i;

n_{ij} = number of claims finalized in development year j of year of origin i;

f_{ij} = n_{ij}/n_i,
= speed of finalization in development year j of year of origin i;

S_{ij} = C_{ij}/n_{ij},
= payments per claim finalized (PPCF) (in constant dollar values) in development year j of year of origin i;

τ_{ij} = $\sum_{k=0}^{j} n_{ik}/n_i = \sum_{k=0}^{j} f_{ik}$,
= proportion of claims, incurred in year of origin i, finalized by end of development year j,
= 'operational time' corresponding to end of development year j of year of origin i;

$\bar{\tau}_{ij}$ = $\frac{1}{2}(\tau_{ij} + \tau_{i,j+1})$,
= average operational time during development year $j+1$ of year of origin i;

$\{u_1, \ldots, u_n\}$ = a partition of the interval $[0, 1]$.

Define *confined operational time* as

$$\bar{\tau}_{ij}^{(k)} = \bar{\tau}_{ij} \wedge t_k = \min(\bar{\tau}_{ij}, t_k),$$

where $t_1, \ldots, t_q \in (0, 1]$ are preset constants.

It is assumed that

$$\mathrm{E}[S_{ij}] = \alpha + \mu_l + \sum_{k \in \mathscr{S}} \beta_k \bar{\tau}_{ij}^{(k)} + \gamma/f_{ij}, \tag{47}$$

where $\mathscr{S}$ is a subset of $\{1, \ldots, n-1\}$, α, $\{\beta_k\}_{k \in \mathscr{S}}$, γ are constants, and μ_l is an addition to PPCF which depends on year of payment l (and thus makes allowance for variation of average claim size with year of payment).

These last constants are determined by weighted least-squares regression where the weight associated with S_{ij} is

$$w_{ij}=v_j/n_{ij}, \tag{48}$$

the v_j being a set of constants intended to remove heteroscedasticity with respect to development year.

The v_j are chosen as follows. Let $\hat{S}_{ij}$ be the regression estimate fitted to S_{ij} and m_j the number of observations S_{ij}(j fixed) entering into the regression. Define

$$s_j^2=\sum_i (n_{ij}/v_j)(S_{ij}-\hat{S}_{ij})^2/m_j. \tag{49}$$

The v_j are chosen so that the s_j^2 are approximately equal over all values of j.

It is supposed that

$$\mu_l=\alpha'+\sum_{r=1}^{p}\lambda_r(l\wedge l_r), \tag{50}$$

where $l_1,\dots,l_p$ are preset constants and $\lambda_1,\dots,\lambda_p$ are unknown parameters.

Substitution of this in (47) and absorption of α' into α gives

$$\mathrm{E}[S_{ij}]=\alpha+\sum_{r=1}^{p}\lambda_r(l\wedge l_r)+\sum_{k\in\mathscr{S}}\beta_k\bar{\tau}_{ij}^{(k)}+\gamma/f_{ij}. \tag{51}$$

The reasoning leading to a model of this form is given by Taylor (1981). Apart from the fact that in a couple of places piecewise linear approximations are adopted for functions which are not necessarily of this form, (51) gives the most general form of PPCF such that the amount of projected outstanding claims, in constant dollar values, is independent of the speed of finalization of claims projected for the future.

The above model can be expressed in terms of the notation used in sections 5 to 7.

The transformation $h(\cdot)$ is linear. It is represented by the matrix H, diagonal, and containing the elements n_{ij} written out in the same order as are the C_{ij}. Then, by (51),

$$\omega=(\alpha,\lambda_1,\dots,\lambda_p,\beta_{k_1},\beta_{k_2},\dots,\gamma)^T,$$

where the β_k included are those such that $k\in\mathscr{S}$.

Also by (51), X is a matrix containing elements equal to 0, 1, $l \wedge l_r$, $\bar{\tau}_{ij}^{(k)}$, $1/f_{ij}$ appropriately arranged.

The weight matrix W is given by (48).

10. Numerical example

The following numerical example is based on the invariant see-saw model set out in section 9.2. The data on which this example is based is given in the appendix. We now discuss the parameters of the model.

Operational time is fitted as a piecewise linear continuous function with kinks at times $t_1 = 0.55$ and $t_2 = 0.85$ ($t_3 = 1.0$). We therefore need three parameters $(\beta_1, \beta_2, \beta_3)$ to represent the slope in each interval. Confined operational time is as described in section 9, i.e., the slope between 0 and 0.55 is $\beta_1 + \beta_2 + \beta_3$, the slope between 0.55 and 0.85 is $\beta_2 + \beta_3$ and β_3 is the slope above 0.85.

The effect of the year of payment l is represented by choosing l_1 to equal 3. This might be interpreted as superimposed inflation occurring in the last two years with no superimposed inflation before that (payment year 1 is the latest payment year and l increases into the past). A parameter (λ) needs to be estimated giving the amount of inflation experienced.

We thus have a six-parameter model — a mean value, the effect of speed of finalization, three parameters for operational time, and superimposed inflation.

No strong evidence exists of heteroscedasticity and so the v_j have been taken as 1.0.

The estimation of parameters and their covariances is performed using the linear regression package GLIM [Baker and Nelder (1978)]. The calculations inherent in (34) and (16) are performed by a FORTRAN program written by one of the authors.

The parameters' estimated values and standard errors are given in table 1. The correlation matrix of these estimates is

$$\begin{pmatrix} 1.0 & & & & & \\ -0.36 & 1.0 & & & & \\ -0.87 & 0.19 & 1.0 & & & \\ -0.27 & 0.09 & 0.01 & 1.0 & & \\ -0.14 & 0.65 & 0.07 & -0.41 & 1.0 & \\ 0.17 & -0.71 & -0.07 & 0.27 & -0.99 & 1.0 \end{pmatrix}.$$

The coefficient of determination for this regression is 72%.

Table 1

Parameter	Estimated value	Standard error
α	16700	2526
γ	235.4	53.5
λ	−3607	768.4
β_1	−12710	8996
β_2	−27530	41010
β_3	35830	37980

σ is estimated at 28374, i.e., when the speed of finalization, operational time, etc. are taken into account, there is still an estimated standard deviation of this amount around the average payment per claim finalized.

The simplest form of U in (15) is that in which it is a vector consisting solely of 1's. $\hat{P}$ is then the total provision. We find that $\hat{P}$ is \$22.3M and (34) gives us a standard deviation of \$2.08M. The superimposed inflation was assumed to stop at the present. This assumption was accommodated by setting l at one for all predicted values.

An alternative form for U is one which produces estimated provisions for each accident year. The provisions and their estimated standard deviations are shown in table 2.

Table 2

Accident year	Estimated provisions (\$000)	Estimated standard deviations (\$000)
1981	5827	794
1980	4866	688
1979	4221	600
1978	2879	417
1977	1788	281
1976	1077	175
1975	745	132
1974	600	120
1973	298	79
1972	0	0

The covariance matrix is found using (16) and shows correlation coefficients of about 40% for accident year provisions.

A much simpler model which may still be considered adequate is formed by deleting the variables pertaining to operational time. There are then only three parameters: whose estimated values and standard errors are given in table 3.

Table 3

Parameter	Estimated value	Standard error
α	15080	2228
γ	300.4	34.25
λ	-3520	773.4

The coefficient of determination is 70%. The parameter correlation matrix is

$$\begin{pmatrix} 1.0 & & \\ -0.31 & 1.0 & \\ -0.97 & 0.21 & 1.0 \end{pmatrix}.$$

The provision is estimated as \$21.7M with a standard deviation of \$2.06M. Similar predictions are made for the accident years as in the six-parameter model. The work involved in moving from the six-parameter model to the three-parameter one is trivial. Numerous other sets of parameters were fitted before choosing these.

11. Concluding remarks

Sections 5 to 8 provide a framework within which first and second moments of linear functions of future claim payments may be obtained. The framework is quite general provided that the basic parameter estimation is carried out by means of regression techniques.

It is the use of such techniques which provides the generality. The generality of approach is to be distinguished from alternative methods of estimation which, in claims analysis, are usually heuristic and specific to the model under consideration.

Moreover, as noted in section 10, the use of the established body of theory surrounding regression analysis permits the elimination from the model of parameters whose estimates are statistically insignificant. Such a step adds to both the simplicity and the stability of the fitted model.

Perhaps the most obvious advantage involved in the use of regression methods is the availability of computer packages for the performance of regression analysis. This greatly expedites experimentation in the search for and selection of a model to fit the data.

The numerical example presented in section 10 gives some indication of the flexibility of structure available in the choice of model. That example was dealt with by means of the GLIM (Generalized Linear Interactive Modelling) package. This package has the additional advantage of being operable in an interactive mode. This enables new variables to be defined and inserted into the model, and others to be removed, during a single session at a computer terminal. The experimentation, selection of model, and application of model in the example of section 10 took about two hours. The cost of computer time was quite small.

The example of section 10 also indicates the large number of potentially significant variables in the model. The payments per claim finalized, S_{ij} in the notation of this paper, may well be a function (even after adjustment by some inflation index) of:

(i) operational time;
(ii) year of origin (this affecting average claim sizes);
(iii) year of claim payment (this effect including any differential between claims escalation and the inflation allowed for in the index used to adjust claim payments).

To establish which of these variables are significant when using one of the 'standard' methods of claims analysis is a virtually hopeless task. In most such methods, the estimation procedures are simply too specific to admit easily additional variables such as those listed above. Consider for example the Payments per Claim Incurred method reviewed in section 4, considering especially how one would begin to unravel the simultaneous effects of year of origin and year of payment as explanatory variables of the payments per claim incurred.

When regression methods are applied to the same type of problem, it cannot be said that the problem becomes easy. Its solution does, however, become at least feasible.

It is sometimes suggested that formal models such as the regression models which we are recommending in this paper lead to the use of too many parameters, i.e., overfitting. To any who might take such a stand, we point out that, in the example of section 10, the number of parameters involved in what appear quite acceptable models varies from three to six. We have tended to favour the six-parameter model. However, even the three-parameter model seems a reasonable representation of the data. But, whether we choose to use three or six parameters, we achieve far greater parsimony than would just about any of the 'standard' methods when applied to the data of ten years of origin. Even the payments per claim incurred method of section 4, which we regard as rather rough and insensitive when applied to the data of section 10, involves ten parameters. Most other methods would involve more.

Our experience with other sets of data is that a pruning of the number of regression parameters to about six or less is not exceptional.

Lest it appear that the authors' enthusiasm for the regression techniques discussed in this paper is overdone, we hasten to dissociate ourselves from any suggestion that this paper provides a mechanistic or fully automated algorithm for claims analysis. Our view is that claims analysis is a special case of data analysis; that therefore there are few preconceptions as to what should be done with the data; indeed, anything goes, if it leads to a model which exhibits acceptable adherence to the data and is pausible in the light of any collateral information. To us, faced with a problem of multivariate data analysis, regression analysis represents a most useful exploratory tool. We use it in this way.

Appendix

This appendix contains the data used in the example given in section 10. Some explanation of the variables is necessary:

(a) Accident year is coded with the current year being given the value 1.0 and the code increasing as the accident year is further in the past.
(b) Development year is measured from 1.0.
(c) Payment year is given the value of 1.0 for all future payment years. The numerical value has a maximum of 3.0 for the purposes of the regression to allow two years of superimposed inflation to be modelled.
(d) For cells which represent years in the future, the claims finalized number is an estimate based on the pattern of speed of finalization by development year observed in the past. The model is one which is insensitive to changes in these numbers. The payment per claim finalized is, of course, not known and is represented as 'n/a'.

The data is arranged sequentially such that the ten cells with development year measured as 1.0 are first, then those ten cells with development year 2.0, and so on. Accident year goes from one to ten repetitively.

Cell number	Claims finalized	Payment per claim finalized	Inverse speed of finalization	Operational time
1	22.000	15637.	19.091	0.02619
2	17.000	22158.	25.588	0.01954
3	43.000	8360.	11.698	0.04274
4	32.000	13776.	16.969	0.02946
5	33.000	12004.	16.727	0.02989
6	30.000	14772.	20.000	0.02500
7	41.000	7575.8	15.146	0.03301
8	35.000	8300.2	19.914	0.02511
9	37.000	9516.7	19.486	0.02566
10	40.000	8946.2	15.250	0.03279
11	101.28	n/a	4.1470	0.17295
12	92.000	10724.0	4.7283	0.14483
13	111.00	9564.4	4.5315	0.19582
14	115.00	7370.7	4.7217	0.16483
15	121.00	7744.5	4.5620	0.16938
16	187.00	3706.9	3.2086	0.20583
17	155.00	7150.0	4.0065	0.19082
18	158.00	6340.5	4.4114	0.16356
19	186.00	4752.8	3.8763	0.18031
20	124.00	6185.0	4.9194	0.16721
21	142.22	n/a	2.9532	0.46283
22	156.25	n/a	2.7840	0.43018
23	83.000	17390.	6.0602	0.38867
24	146.00	7749.3	3.7192	0.40516
25	204.00	4154.4	2.7059	0.46377
26	166.00	5975.8	3.6145	0.50000
27	218.00	3560.5	2.8486	0.49114
28	243.00	3811.6	2.8683	0.45122
29	130.00	7183.8	5.5462	0.39945
30	157.00	3888.8	3.8854	0.39754
31	80.603	n/a	5.2107	0.72809
32	88.557	n/a	4.9121	0.71156
33	138.77	n/a	3.6247	0.60912
34	103.00	10323.	5.2718	0.63444
35	87.000	9253.3	6.3448	0.72735
36	120.00	6412.4	5.0000	0.73833
37	100.00	15624.	6.2100	0.74718
38	153.00	6644.8	4.5556	0.73529
39	239.00	4951.0	3.0167	0.65534
40	93.000	5192.9	6.5591	0.60246
41	41.717	n/a	10.068	0.87371
42	45.833	n/a	9.4910	0.86604
43	71.822	n/a	7.0035	0.81845
44	82.983	n/a	6.5435	0.80569
45	37.000	19080.	14.919	0.83967
46	55.000	9179.1	10.909	0.88417
47	67.000	4066.9	9.2687	0.88164
48	48.000	15642.	14.521	0.87948
49	61.000	7307.3	11.820	0.86338
50	141.00	3739.9	4.3262	0.79426

Cell number	Claims finalized	Payment per claim finalized	Inverse speed of finalization	Operational time
51	15.656	n/a	26.826	0.94201
52	17.201	n/a	25.289	0.93849
53	26.955	n/a	18.661	0.91664
54	31.143	n/a	17.435	0.91078
55	34.054	n/a	16.210	0.90403
56	13.000	36203.	46.154	0.94083
57	17.000	20709.	36.529	0.94928
58	26.000	5650.9	26.808	0.93257
59	26.000	12346.	27.731	0.92372
60	22.000	26109.	27.727	0.92787
61	11.646	n/a	36.063	0.97452
62	12.795	n/a	33.997	0.97297
63	20.051	n/a	25.086	0.96336
64	23.167	n/a	23.439	0.96079
65	25.331	n/a	21.791	0.95782
66	20.436	n/a	29.359	0.96870
67	6.0000	34381.	103.50	0.96779
68	14.000	35428.	49.786	0.96126
69	23.000	22948.	31.348	0.95770
70	14.000	10453.	43.571	0.95738
71	3.8409	n/a	109.35	0.99295
72	4.2199	n/a	103.08	0.99252
73	6.6127	n/a	76.066	0.98987
74	7.6402	n/a	71.071	0.98916
75	8.3542	n/a	66.074	0.98834
76	6.7399	n/a	89.022	0.99134
77	13.380	n/a	46.413	0.98340
78	5.0000	56081.	139.40	0.97489
79	6.0000	44362.	120.17	0.97781
80	10.000	13995.	61.000	0.97705
81	.76417	n/a	549.62	0.99843
82	.83957	n/a	518.12	0.99834
83	1.3156	n/a	382.33	0.99775
84	1.5201	n/a	357.22	0.99759
85	1.6621	n/a	332.10	0.99741
86	1.3409	n/a	447.45	0.99808
87	2.6620	n/a	233.28	0.99631
88	11.030	n/a	63.194	0.98639
89	6.0000	70841.	120.17	0.98613
90	3.0000	75743.	203.33	0.98770
91	.27508	n/a	1526.8	0.99967
92	.30222	n/a	1439.3	0.99965
93	.47359	n/a	1062.1	0.99953
94	.54719	n/a	992.34	0.99950
95	.59833	n/a	922.58	0.99946
96	.48271	n/a	1243.0	0.99960
97	.95825	n/a	648.06	0.99923
98	3.9704	n/a	175.55	0.99715
99	7.0000	n/a	103.00	0.99514
100	2.0000	33974.	305.00	0.99180

References

Baker, R.J. and J.A. Nelder, 1978, The GLIM system, Release 3: Generalized linear interactive modelling (Numerical Algorithms Group, Oxford).

Bartholomew, D.J., 1975, Errors of prediction in Markov chains, Journal of the Royal Statistical Society B 37, 444–456.

Bibby, J. and H. Toutenburg, 1977, Prediction and improved estimation in linear models (Wiley, New York).

Bühlmann, H., R. Schnieper and E. Straub, 1980, Claims reserves in casualty insurance based on a probabilistic model, Bulletin of the Association of Swiss Actuaries, 21–45.

Cumpston, J.R., 1976, Payments per unit of risk claims models, General Insurance Bulletin 1, 8–11.

de Jong, P. and B. Zehnwirth, 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 B. Zehnwirth, 1982, Claims reserving, state-space models and the Kalman filter, Journal of the Institute of Actuaries, forthcoming.

Hachemeister, C.A., 1978, A structural model for the analysis of loss reserves, Bulletin d'Association Royale des Actuaries Belges 73, 17–27.

Hachemeister, C.A., 1980, A stochastic model for loss reserving, Transactions of the 21st International Congress of Actuaries 1, 185–192.

Reid, D.H., 1978, Claim reserves in general insurance, Journal of the Institute of Actuaries 105, 211–296.

Taylor, G.C., 1977, Separation of inflation and other effects from the distribution of non-life insurance claim delays, ASTIN Bulletin 9, 217–230.

Taylor, G.C., 1981, An invariance principle for the analysis of non-life insurance claims, Submitted for publication.

Taylor, G.C. and F.R. Ashe, 1982, Second moments of estimates of oustanding claims, Technical report (E.S. Knight and Co. Research Centre, Sydney).

Taylor, G.C. and T.G. Matthews, 1977, Experimentation with the estimation of the provision for outstanding claims in non-life insurance, Transactions of the Institute of Actuaries of Australia, 178–254.