CHAIN LADDER AND INTERACTIVE MODELLING (CLAIMS RESERVING AND GLIM)

By A. E. Renshaw, B.Sc., Ph.D.
(of the City University, London)

[Presented at the Seminar, 'Applications of Mathematics in Insurance, Finance and Actuarial Work', sponsored by the Institute of Mathematics and Its Applications, the Institute of Actuaries, and the Faculty of Actuaries, held at the Institute of Actuaries, 6–7 July 1989.]

1. Motivation

The prediction of outstanding claims amounts in non-life insurance is, by its very nature, highly speculative. Partially because of this and partially because of the variety of features suggested by various researchers for possible inclusion in the structure of the underlying prediction model, the past two decades have seen a proliferation of methodologies for making such predictions. Specific details of these developments are contained in a comprehensive and highly detailed survey conducted by Taylor (1986)[10] in which a taxonomy of methods is established. One feature common to all of these methods is the utilization of current and past records of claims amounts—invariably in the form of the familiar so-called run-off triangle or a variant thereof—to calibrate the proposed prediction model before use. Prudence dictates that diagnostic checks should then be made to establish whether or not the data are supportive of the structure imparted to the prediction model before use, a feature which apart from some notable exceptions including Zehnwirth (1985)[14] and Taylor (1983),[8] is not always emphasized in the literature.

Our purpose is not to add to the existing plethora of methodologies but rather to return to the grass roots of the subject by exploring more fully the statistical setting for the basic chain-ladder and related techniques. Essentially a deterministic technique, see for example Hossack *et al.* (1983),[5] it was left to Kremer (1982)[6] to point out that the mathematical structure underpinning the chain-ladder technique is identical to that of the linear statistical model involving a log response variable regressing on two non-interactive convariables. Yet, judging by the lack of literature, there would not appear to have been a concerted effort to develop this connection. Perhaps the answer lies partly in the realization, in some quarters, that the model is heavily parameterized, a phenomenon known to lead to predictor instability.

The aims therefore are:

(i) To develop more fully the statistical analogue of the original actuarial chain-ladder technique.

(ii) To investigate the magnitude and nature of predictor instability associated with the technique.

(iii) To suggest a method for improving predictor stability.
(iv) To make the methodology readily available to practitioners so that they may make their own judgements in these matters.

The GLIM software package, because of its user defined macro facility, is an invaluable tool in achieving these objectives. Indeed we note with interest that Taylor (1983)[(8)] and Taylor & Ashe (1983)[(9)] used the GLIM package to fit Taylor's so called 'invariant see-saw' model to run-off data.

We identify our philosophical approach to estimating claims whole-heartedly with the sentiments expressed by Taylor & Ashe (1983)[(9)] from which we quote the following passage:

> 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 plausible 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 would view this application of GLIM to run-off data as the natural extension of other applications of generalized linear models in actuarial work reported by Haberman and Renshaw (1988).[(4)]

2. CLAIMS DATA

Claims run-off data are generated when delay is incurred in settling insurance claims. Typically the format for such data is that of a triangle (Figure 1.1) in which the rows (i) denote accident years and the columns (j) delay or development years. The settlement or payment year is $k=i+j-1$. The entries in the body of the triangle are the adjusted (non-cumulative) amounts

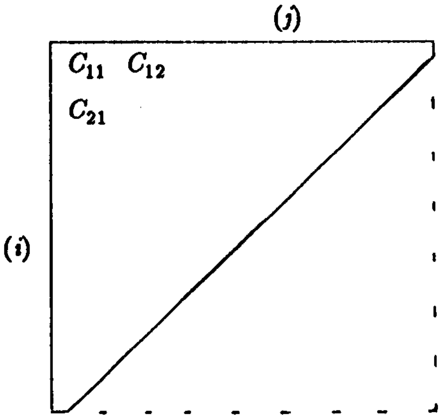

Figure 1.1.

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

$$C_{ij} = \frac{(\text{claims amount}) \times (\text{inflation factor})}{(\text{exposure})}.$$

The triangle is augmented each year with the addition of a new diagonal. Two noteworthy variations of the triangular format are induced by either truncation after a fixed period of delay or by the removal of data for the early settlement years.

Additional information in the guise of numbers of claims settled per cell is required to implement Taylor's (1983)[8] 'invariant see-saw' method.

An obvious first step in any analysis is to plot the adjusted claims against accident year, against development year and against payment year. One might even be tempted to use a three-dimensional plot. Such displays can be very informative about the type of model structure that the data might support.

The remit is essentially to predict likely claim amounts in the incomplete south-east region bounded by broken lines in Figure 1.1. A two stage modelling/predicting process is envisaged.

3. LOG-NORMAL MODELS

Let

$$Y_{ij} = \log(C_{ij})$$

and consider the class of log-normal models defined by

$$Y_{ij} = m_{ij} + \varepsilon_{ij}$$

with

$$\varepsilon_{ij} \sim \text{IN}(0, \sigma^2) \qquad \text{and} \qquad Y_{ij} \sim \text{IN}(m_{ij}, \sigma^2).$$

Here we have assumed that the normal responses Y_{ij} decompose (additively) into deterministic non-random components (means) m_{ij} and independent homoscedastic normally distributed random error components about a zero mean. It will be necessary to monitor these assumptions by displaying various residual plots on fitting specific model structures to the logarithms of the adjusted claims data.

A number of specific model structures are of interest. These include:

Case (I) $\quad$ M: $\quad m_{ij} = \mu + \alpha_i + \beta_j \qquad (3.1)$

with accident and development years treated as non-interactive covariates. This structure is identical to that used in a two-way analysis of variance (ANOVA), but based on the incomplete data sketched in Figure 3.1(a). Indeed, our brief is to estimate the incomplete south-east triangular region. The structure is identical to that associated with the traditional actuarial chain-ladder technique.

Case (II) $\quad$ M: $\quad m_{ij} = \mu + \beta_j + \gamma_k$

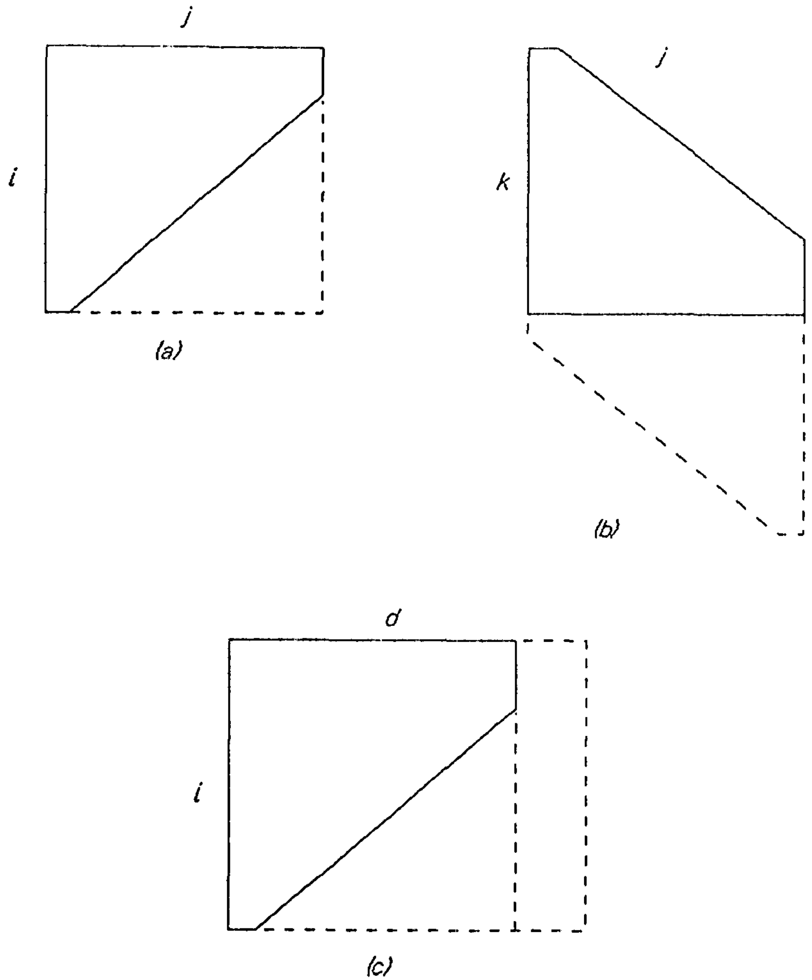

Figure 3.1. Typical run-off domains and prediction regions.

with development and settlement years treated as non-interactive covariates. The structure is motivated by the traditional actuarial so-called separation method, see, for example, Hossack *et al.* (1983)[5]; and was first treated statistically by Taylor (1979).[7] Depicting the various levels of k along the rows while still representing the levels of j as columns distorts the basic data matrix into the form

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

sketched in Figure 3.1(b). This time our brief is the seemingly difficult one of predicting values in the lower protruding triangular region.

Case (III) M: $m_i(d) = \alpha_i + \beta_i \log(1+d) + \gamma_i d$ (3.2)

with $d = j-1$ treated as a continuous regressor variable. A version of this structure is discussed by Dejong and Zehnwirth (1983)[2] in which parameters are estimated recursively using the Kalman filter. Practical implementation is possible using Zehnwirth's (1985)[14] ICRFS purpose designed software package.

The untransformed model structure is

$$\exp(m_i(d)) = K_i(1+d)^{\beta_i} \exp(\gamma_i d) \qquad (K_i = e^{\alpha_i})$$

so that $\gamma_i < 0$ ensures claims amounts ultimately decay. Referring to the data matrix sketched in Figure 3.1(c), prediction beyond the observed limit of d as well as in the south east triangular region is feasible.

Case (IV) M: $$\begin{cases} m_{ij} = \mu + \alpha_i + \beta_j & j = 1, 2, \ldots, q \\ m_i(d) = \lambda_i + v_i \log(d) + \gamma_i d, & d > q. \end{cases}$$

Here we have written d for j when j exceeds some fixed integer q. The model is clearly a mixture of Case I and Case III applied to separate parts of the data matrix.

Each of the models discussed above has obvious submodels. We concentrate on Case I.

4. MODEL FITTING

Consider the two-way ANOVA model structure

$$\text{M:} \quad m_{ij} = \mu + \alpha_i + \beta_j$$

with an incomplete experimental design dictated by the pattern of adjusted claim amounts illustrated in Figure 4.1; obviously, $g = 0$, $w = 0$ for a run-off triangle, while $j = 1, 2, \ldots, l$; $i = 1, 2, \ldots, r$ in general. It is well known that whereas this parametric representation of the model structure involves a total of $r + l + 1$ parameters, it contains only $r + l - 1$ so-called free parameters. Consequently two contraints must be imposed on the parameters before estimation can proceed. The GLIM system sets $\alpha_1 = \beta_1 = 0$ and computes maximum likelihood estimates for the parameters. As a direct consequence of the normal error structure this is equivalent to estimation by least squares.

Define indicators δ_{ij} for all cross-classified factor levels (i, j) according to

$$\delta_{ij} = 1 \text{ if } C_{ij} > 0, \qquad \delta_{ij} = 0 \text{ otherwise.}$$

Then

$$\delta_{++} = \sum_{ij} \delta_{ij}, \qquad \delta_{i+} = \sum_j \delta_{ij}, \qquad \delta_{+j} = \sum_i \delta_{ij}$$

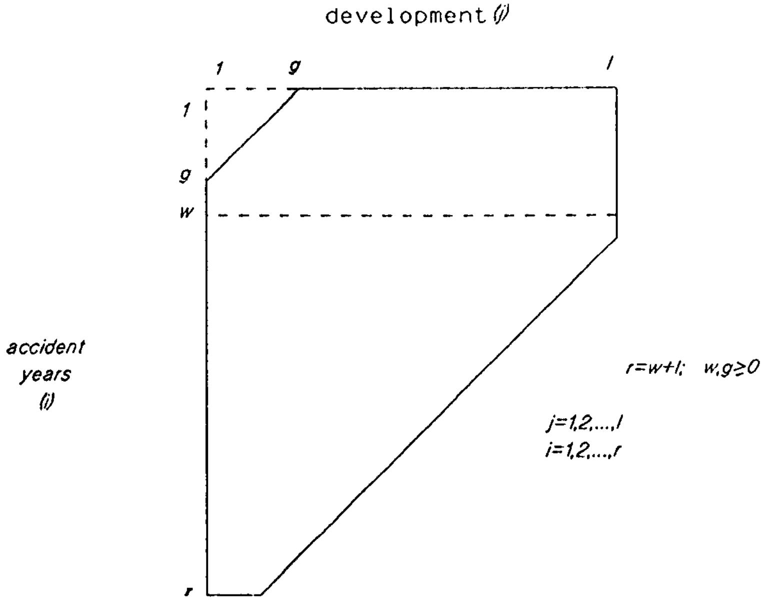

Figure 4.1. Typical claims data format.

denote the total number of observations, the number of observations in row i and the number of observations in column j respectively.

We choose $\hat{\mu}$, $\hat{\alpha}_i$, $\hat{\beta}_j$ $(i, j \neq 1)$ so as to minimize

$$\sum_{ij}(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j)^2 \qquad (\hat{\alpha}_1 = \hat{\beta}_1 = 0).$$

Partial differentiation with respect to $\hat{\mu}$, $\hat{\alpha}_i$ for each $i\,(\neq 1)$ and $\hat{\beta}_j$ for each $j\,(\neq 1)$ leads to the system of linear equations

$$y_{++} = \delta_{++}\hat{\mu} + \sum_j \delta_{+j}\hat{\beta}_j + \sum_i \delta_{i+}\hat{\alpha}_i$$

$$y_{+j} = \delta_{+j}(\hat{\mu} + \hat{\beta}_j) + \sum_i \delta_{ij}\hat{\alpha}_i, \qquad j = 2, 3, \ldots, l$$

$$y_{i+} = \delta_{i+}(\hat{\mu} + \hat{\alpha}_i) + \sum_j \delta_{ij}\hat{\beta}_j, \qquad i = 2, 3, \ldots, r$$

where $\quad y_{++} = \sum_{ij} y_{ij}, \qquad y_{i+} = \sum_j y_{ij}, \qquad y_{+j} = \sum_i y_{ij}$

denote the grand total, row totals and column total of the transformed adjusted claims. The solution of this set of non-singular linear equations yield the required estimates.

By way of illustration, the artificial data set

$j\rightarrow$	1	2	3	Totals
i 1	2	4	6	12
$\downarrow$ 2	2	3	4	9
3	3	2		5
4	2			2
Totals	9	9	10	28

$(l=3, \omega=1, r=4, g=0)$

gives rise to the system of linear equations

$$\begin{bmatrix} 28 \\ 9 \\ 10 \\ 9 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 & 3 & 2 & 3 & 2 & 1 \\ 3 & 3 & 0 & 1 & 1 & 0 \\ 2 & 0 & 2 & 1 & 0 & 0 \\ 3 & 1 & 1 & 3 & 0 & 0 \\ 2 & 1 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \\ \hat{\alpha}_4 \end{bmatrix}$$

which yield the solution

$$\hat{\mu}=2{\cdot}917,$$

$$\hat{\beta}_2=\cdot 667, \qquad \hat{\beta}_3=2{\cdot}583,$$

$$\hat{\alpha}_2=-1{\cdot}000, \qquad \hat{\alpha}_3=-\cdot 750, \qquad \hat{\alpha}_4=-\cdot 917.$$

The corresponding fitted and predicted values

$$\hat{m}_{ij}=\hat{\mu}+\hat{\alpha}_i+\hat{\beta}_j, \quad (\hat{\alpha}_1=\hat{\beta}_1=0)$$

are

2·917	3·583	5·500
1·917	2·583	4·500
2·167	2·833	4·750
2·000	2·666	4·583

Scrutiny of these fitted and predicted values reveals the true nature of the assumed non-interactive model structure which manifests itself in the constant differences between columns and between rows.

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

A noteworthy submodel is that involving development year effects only. The one-way ANOVA sub-structure is

$$\text{H:} \quad m_{ij} = \mu + \beta_j$$

where, again we define $\beta_1 = 0$ because of overparameterization. This time the incomplete nature of the data matrix (Figure 4.1) is irrelevant. The parameter estimates are determined by

$$y_{++} = \delta_{++}\hat{\mu} + \sum_j \delta_{+j}\hat{\beta}_j$$

$$y_{+j} = \delta_{+j}(\hat{\mu} + \hat{\beta}_j), \qquad j = 2,3, \ldots, l.$$

The solution is

$$\hat{\mu} = \bar{y}_{\cdot 1}, \qquad \hat{\beta}_j = \bar{y}_{\cdot j} - \bar{y}_{\cdot 1}$$

so that the fitted and predicted values are the column averages. Justification for using this simplified model is sought by examining the t-statistics associated with the parameters α_i, examination of further residual plots and through a formal ANOVA F-test based on the statistic

$$\frac{(R_H - R_M)/(l-1)}{R_M/(\delta_{++} - l - r + 1)}$$

in which R_M and R_H denote the residual sums of squares or deviance under the full model M and the submodel H respectively.

Whereas it has been established by Kremer (1982)[6] that the model structure in use here is identical to that utilized in the standard actuarial chain-ladder technique as described, for example, in Hossack *et al.* (1983)[5], the current treatment of the model differs in two important respects—namely the ways in which the model parameters are estimated and the predicted values are constructed.

5. PREDICTED VALUES

The model is fitted on the log-response scale. On this scale

$$\hat{m}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j \tag{5.1}$$

provides a point predictor for the empty (i,j)th cell in the south east triangular region. Since the $\hat{m}_{ij}$ are linear in the Y_{ij}s, they are distributed normally with

$$E(\hat{m}_{ij}) = E(\hat{\mu}) + E(\hat{\alpha}_i) + E(\hat{\beta}_j) \tag{5.2}$$

and

$$V(\hat{m}_{ij}) = V(\hat{\mu}) + V(\hat{\alpha}_i) + V(\hat{\beta}_j) + 2(\text{Cov}(\hat{\mu}\,\hat{\alpha}_i) + \text{Cov}(\hat{\mu}\,\hat{\beta}_j) + \text{Cov}(\hat{\alpha}_i\,\hat{\beta}_j)). \tag{5.3}$$

If, in keeping with common practice, the predictor is augmented by an independent additive error term, distributed as $N(0, \sigma^2)$, then σ^2 has to be added to the RHS of (5.3).

Reverting to the original (anti-log) scale, predictors $\mathring{C}_{ij}$ are needed where

$$\mathring{m}_{ij} = \log(\mathring{C}_{ij}).$$

Since the $\mathring{m}_{ij}$s are normally distributed, the $\mathring{C}_{ij}$s are log-normally distributed with

$$E(\mathring{C}_{ij}) = \exp(E(\mathring{m}_{ij}) + \tfrac{1}{2}V(\mathring{m}_{ij})) \tag{5.4}$$

and

$$\sqrt{V(\mathring{C}_{ij})} = E(\mathring{C}_{ij})\sqrt{\exp(V(\mathring{m}_{ij})) - 1} \tag{5.5}$$

One method of computing predicted values and their standard errors, apparently favoured by Zehnwirth (1985)[14], is based on (5.4) and (5.5) in which $E(\mathring{m}_{ij})$ and $V(\mathring{m}_{ij})$ are replaced by their estimated values as dictated, in this instance by (5.2) and (5.3). It should be stressed, however, that Zehnwirth is working within a Bayesian framework and would presumably seek to justify the method of prediction within this framework.

6. PREDICTED TOTALS AND THEIR STANDARD ERRORS

Practitioners have a vested interest in

(i) the predicted row totals

$$\mathring{t}_i = \sum_{j > c(i)}^{l} \mathring{C}_{ij}, \qquad i = w+2, w+3, \ldots, r$$

where l and $c(i) = l + 1 - i$ are the upper and lower limits of j;

(ii) the predicted diagonal totals

$$\mathring{t}_k = \sum_{\substack{ij \\ i+j=k+1}} \mathring{C}_{ij}, \qquad k = r+1, r+2, \ldots, r+l-1;$$

(iii) the overall predicted total

$$\mathring{t} = \sum_{i=w+2}^{r} \mathring{t}_i = \sum_{k=r+1}^{r+l-1} \mathring{t}_k$$

together with their standard errors.

Consequently, for the predicted row totals, it follows that

$$V(\mathring{t}_i) = \sum_{j > c(i)}^{l} V(\mathring{C}_{ij}) + 2\sum_{j > c(i)}^{l} \sum_{k > j} \mathrm{Cov}(\mathring{C}_{ij}\,\mathring{C}_{ik}).$$

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

Making use of the Theorem 2.4 of Aitchison and Brown (1969)[1] it can be shown that

$$\mathrm{Cov}(\mathring{C}_{ij}\mathring{C}_{ik}) = E(\mathring{C}_{ij})\, E(\mathring{C}_{ik})\, (\exp(\mathrm{Cov}(\mathring{m}_{ij}\, \mathring{m}_{ik})) - 1) \tag{6.1}$$

from which (5.5) is retrieved on setting $j = k$. Further, (5.1) implies that for $j \neq k$

$$\begin{aligned}\mathrm{Cov}(\mathring{m}_{ij}\, \mathring{m}_{ik}) = {} & V(\hat{\mu}) + 2\mathrm{Cov}(\hat{\mu}\, \hat{\alpha}_i) + V(\hat{\alpha}_i) + \mathrm{Cov}(\hat{\beta}_j \hat{\beta}_k) \\ & + \mathrm{Cov}(\hat{\mu}\, \hat{\beta}_j) + \mathrm{Cov}(\hat{\mu}\, \hat{\beta}_k) + \mathrm{Cov}(\hat{\alpha}_i\, \hat{\beta}_j) + \mathrm{Cov}(\hat{\alpha}_i\, \hat{\beta}_k).\end{aligned} \tag{6.2}$$

This time (5.3) is retrieved on setting $j = k$. Also note the useful identity

$$2\mathrm{Cov}(\mathring{m}_{ij}\mathring{m}_{ik}) = (V(\mathring{m}_{ij}) - V(\hat{\beta}_{ij})) + (V(\mathring{m}_{ik}) - V(\hat{\beta}_{ik})) + 2(\mathrm{Cov}(\hat{\beta}_j\hat{\beta}_k) - \sigma^2)$$

where we have assumed the augmented version of (5.3).

Yet more general versions of (6.1) and (6.2), namely

$$\mathrm{Cov}(\mathring{C}_{i_1j_1}\mathring{C}_{i_2j_2}) = E(\mathring{C}_{i_1j_1})E(\mathring{C}_{i_2j_2})\, (\exp(\mathrm{Cov}(\mathring{m}_{i_1j_1}\mathring{m}_{i_2j_2})) - 1)$$

and

$$\begin{aligned}\mathrm{Cov}(\mathring{m}_{i_1j_1}\mathring{m}_{i_2j_2}) = {} & V(\hat{\mu}) + \mathrm{Cov}(\hat{\mu}\, \hat{\alpha}_{i_1}) + \mathrm{Cov}(\hat{\mu}\, \hat{\alpha}_{i_2}) + \mathrm{Cov}(\hat{\alpha}_{i_1}\, \hat{\alpha}_{i_2}) + \mathrm{Cov}(\hat{\beta}_{j_1}\hat{\beta}_{j_2}) \\ & + \mathrm{Cov}(\hat{\mu}\, \hat{\beta}_{j_1}) + \mathrm{Cov}(\hat{\mu}\, \hat{\beta}_{j_2}) + \mathrm{Cov}(\hat{\alpha}_{i_1}\, \hat{\beta}_{j_2}) + \mathrm{Cov}(\hat{\alpha}_{i_2}\, \hat{\beta}_{j_1})\end{aligned}$$

catering for between row dependencies are needed to compute the variances of the predicted diagonal totals and the overall predicted total. Notice that (6.1) and (6.2) are retrieved on setting $i_1 = i_2 = i$ (together with $j_1 = j$, $j_2 = k$).

7. PREDICTOR INSTABILITY

First the comment that the adjusted claim amounts are generally characterized by significant differences between development years but only small differences across accident years.

The extent of any instability exhibited by each predicted value depends directly on the number of parameters used to make the prediction, in this case just three which is not excessive, and more importantly on the extent to which the estimates of these parameters are sensitive to fluctuations in the data. Not surprisingly in view of the nature of the model structure and data format, simulation exercises confirm that predictions are sufficiently robust to data fluctuations in the heart of and in the north-west corner of the run-off triangle; and that stability deteriorates as data points further into the other two corners of the run-off triangle are varied. However, the instability in the north-east corner is generally not a serious problem since claims amounts in this region are relatively low in comparison with the remainder of the data triangle. The position is further improved if truncation has occurred.

Consequently, it is essential to improve predictor stability for the more recent accident years. There are a number of possibilities such as the estimation of the α_is by empirical Bayes, see Verrall (1988)[12] or by Kalman filtering as proposed by Dejong and Zehnwirth (1983)[2] and applied to Case III (discussed in Section 3). *We note with particular interest in passing that were one to attempt to* generate the α_i s as a first order autoregressive process within GLIM, the facility to handle non-diagonal weight matrices recently proposed by Green (1988)[3] is needed.

Another possibility which we have been pursuing is a reduction in the total number of row parameters based on the multiple comparison t-criteria

$$\left|\frac{\hat{\alpha}_i-\hat{\alpha}_j}{\sqrt{\hat{V}(\hat{\alpha}_i-\hat{\alpha}_j)}}\right|<h \qquad \forall\, i,j\,(i\neq j).$$

The objective is to partition the set of α_is by varying the limit h. This would seem to work well, is objective, intuitively appealing, and induces the required degree of stability provided no new parameters are allocated to the more recent accident year.

8. IMPLEMENTATION

This is by user defined macros within GLIM. Essentially four primary macros are required:

(i) to create related vectors, scalars and to output data plots;
(ii) to do the model fitting and output graphical checks;
(iii) to conduct the multiple comparison t-tests;
(iv) to output further graphical checks; to compute and output the predicted claims amounts, their totals and standard errors.

It is suggested that these macros could form the basis of a more extensive suite of macros to be offered to practitioners. It is noted with interest that one such practitioner, Taylor (1988),[11] strongly recommends the use of such regression methods.

9. AN APPLICATION

Consider the non-cumulative run-off triangle with exposures (Table 9.1) computed from the data given in Taylor and Ashe (1983)[9] and used by them to illustrate their 'invariant see-saw' method. Inflation effects are not discussed so we ignore these. The plot of adjusted claims against delay (Figure 9.1) is informative, hinting that a model of the type defined by (3.2) as well as that defined by (3.1) might well be appropriate. We concentrate on the latter because of its historical interest. The remaining adjusted claims plots are relatively uninformative and are consequently not reproduced here.

Table 9.1 *Run-off claims data and exposures*

development year j	1	2	3	4	5	6	7	8	9	10
accident 1	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
year 2	352118	884021	933894	1183289	445745	320996	527804	266172	425046	
(i) 3	290507	1001799	926219	1016654	750816	146923	495992	280405		
4	310608	1108250	776189	1562400	272482	352053	206286			
5	443160	693190	991983	769488	504851	470639				
6	396132	937085	847498	805037	705960					
7	440832	847631	1131398	1063269						
8	359480	1061648	1443370							
9	376686	986608								
10	344014									

EXPOSURES

610 721 697 621 600 552 543 503 525 420

```
[o]
[o] adjusted claims vs. delay
[o]
[o]    3040. |
[o]    2880. |          Y
[o]    2720. |
[o]    2560. |                Y
[o]    2400. |
[o]    2240. |
[o]    2080. |     Y    Y
[o]    1920. |     Y          Y
[o]    1760. |     2
[o]    1600. |     Y    2     Y
[o]    1440. |     Y          2
[o]    1280. |     3    2     Y       Y
[o]    1120. |     Y                  Y
[o]     960. |          Y                     Y
[o]     800. 3                Y       2       Y     Y
[o]     640. 4                Y       Y       Y             Y
[o]     480. 3                Y       Y             Y
[o]     320. |                               Y     Y       Y
[o]     160. |                       Y       Y     Y              Y
[o]       0. |
[o] -----------:---------:---------:---------:---------:---------:---------:
[o]           0.00      2.00      4.00      6.00      8.00     10.00     12.00
[o]
[o]
```

Figure 9.1.

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

Residual plots for the two-way ANOVA Model defined by (3.1) (Figures 9.2(a)–(e)) are reasonably supportive of the model although the histogram is slightly skewed. Estimates for the model parameters and their standard errors are given in standard GLIM format (Table 9.2). Here the model parameters of (3.1) have been recoded according to 1 for μ, the general mean; DY_(j) for β_j, the development year parameters and AY_(i) for α_i, the accident year parameters. The system automatically sets $\alpha_1 = \beta_1 = 0$, a feature utilized in the development of Section 4.

```
[o]
[o]   histogram of residuals
[o]
[o]  [-1.00,-0.75)  1  X
[o]  [-0.75,-0.50)  2  XX
[o]  [-0.50,-0.25)  4  XXXX
[o]  [-0.25, 0.00) 20  XXXXXXXXXXXXXXXXXXXX
[o]  [ 0.00, 0.25) 18  XXXXXXXXXXXXXXXXXX
[o]  [ 0.25, 0.50)  8  XXXXXXXX
[o]  [ 0.50, 0.75]  2  XX
[o]
[o]
```

Figure 9.2(a).

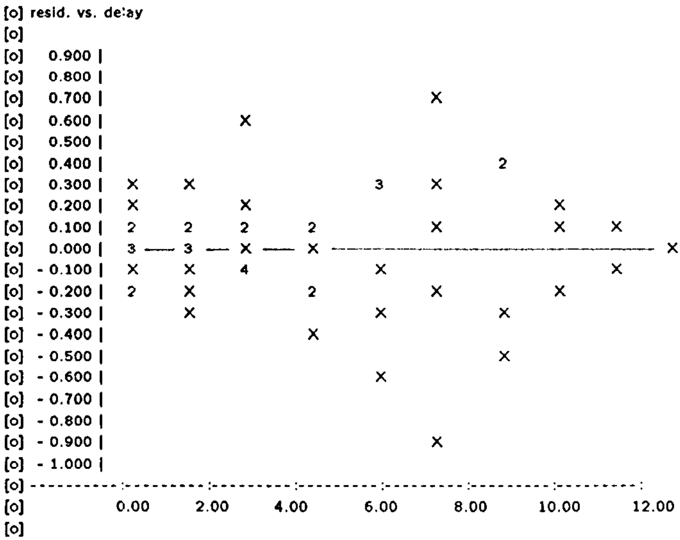

Figure 9.2(b).

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

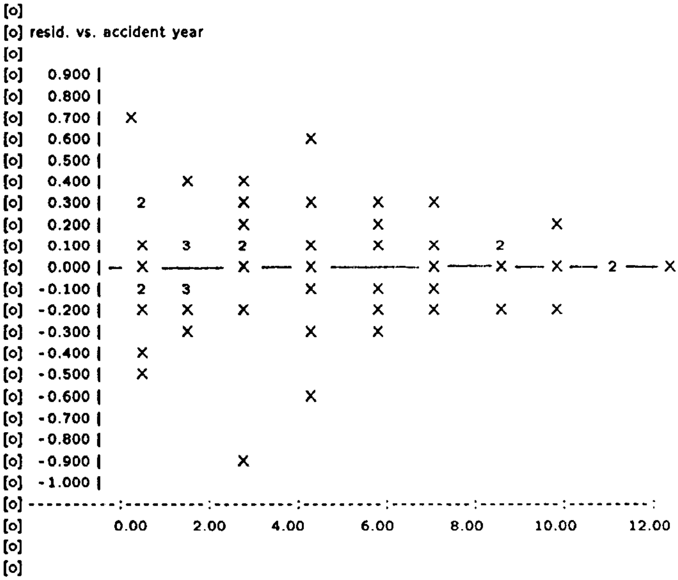

Figure 9.2(c).

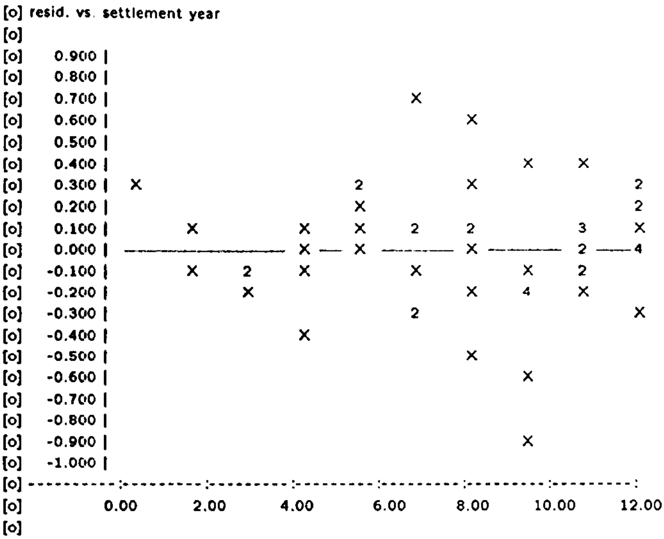

Figure 9.2(d).

```
[o]
[o] resid. vs. fitted values
[o]
[o]    0.900 |
[o]    0.800 |
[o]    0.700 |                             X
[o]    0.600 |                                              X
[o]    0.400 |                               2
[o]    0.300 |                             X     XX X X     X
[o]    0.200 |                         X         X              X
[o]    0.100 |                         X     2   X X    X 2XX X
[o]    0.000 | ----- X ---------------- X--- XX ----- X XXXX -----
[o]   -0.100 |                         X     X     X  X   3   X
[o]   -0.200 |                      X   XX         X        X 2
[o]   -0.300 |                             X     X     X
[o]   -0.400 |                                      X
[o]   -0.500 |                           X
[o]   -0.600 |                                   X
[o]   -0.700 |
[o]   -0.800 |
[o]   -0.900 |                               X
[o]   -1.000 |
[o] ---------:---------:---------:---------:---------:---------:---------:
[o]         4.000     4.800     5.600     6.400     7.200     8.000     8.800
[o]
```

Figure 9.2(e).

Table 9.2

	estimate	s.e.	parameter
1	6.106	0.1646	1
2	0.9112	0.1607	DY_(2)
3	0.9387	0.1681	DY_(3)
4	0.9650	0.1761	DY_(4)
5	0.3832	0.1857	DY_(5)
6	-0.004909	0.1978	DY_(6)
7	-0.1181	0.2142	DY_(7)
8	-0.4393	0.2387	DY_(8)
9	-0.05351	0.2806	DY_(9)
10	-1.393	0.3786	DY_(10)
11	0.1938	0.1607	AY_(2)
12	0.1489	0.1681	AY_(3)
13	0.1533	0.1761	AY_(4)
14	0.2988	0.1857	AY_(5)
15	0.4117	0.1978	AY_(6)
16	0.5084	0.2142	AY_(7)
17	0.6731	0.2387	AY_(8)
18	0.4952	0.2806	AY_(9)
19	0.6018	0.3786	AY_(10)

scale parameter taken as 0.1162

Attempted model simplification by excluding accident year effects leads to an F-statistic value of 1·481 on 9,36 degrees of freedom with an observed significance level of approximately 20%. Whereas this is supportive of the simplification, two of the residual plots (Figures 9.3(a) and (b)) under the simplified one-way development year effects model become unacceptably distorted. The explanation for this is possibly to be found in the values of the parameter estimates (Table 9.2) under the full two-way ANOVA model. The t-statistics (obtained by dividing the estimates by their standard errors) indicate that the accident year parameters from year six onwards are all in fact significant; a feature which would appear to synchronize with the residual plots (Figures 9.3a–b). Consequently, we retain the two-way ANOVA model for the time being. We also have a vested interest in investigating the extent of predictor instability for this model. The run-off claims data, their expected (fitted) values under this model, the predicted claims values and their standard errors are presented in Table 9.3 together with the predicted totals and their standard errors.

We are involved in a two stage process in which the data are first utilized to calibrate/validate the proposed model before moving to the predictive second stage. Model validation is done through scrutiny of response and residual plots coupled with attempted model simplifications where appropriate. Given a satisfactory model, both the magnitude of the standard errors of the predicted values and the degree of stability exhibited by predicted values to fluctuations in the data are important aspects of performance with which to assess the effectiveness of this process. Clearly, if relatively minor fluctuations in the data induce excessive changes in the predicted values there is cause for concern, a phenomenon which is well known in the context of predictive regression modelling.

The extent of any instability exhibited by each predicted value depends directly in the number of parameters used to make each prediction, in this case just three (and not directly on the total number of model parameters), together with the extent to which the estimates of these parameters are sensitive to fluctuations in the data. We concentrate on the latter source of possible instability since the number of parameters involved in making each prediction is low. Indeed an identical number of parameters (three) is involved in each prediction based on the model defined by (3.2) in which a much more rigid structure is imputed to development year effects.

Suppose first that $g=0$, $w=0$ so that the data are triangular in shape. Not surprisingly in view of the nature of the model structure, simulation exercise reveals that predictor stability deteriorates as data points further into the apices of the run-off triangle are varied. This is illustrated by Figure 9.4(a) in which the arrows indicate the directions of decreasing predictor stability. However, the magnitude of predictor instability induced by changes in the data would not appear to be excessive in our experience except for changes in the last few data rows and columns. This is hardly surprising as so little data are yet available to stabilize the estimates of the corresponding row and column parameters.

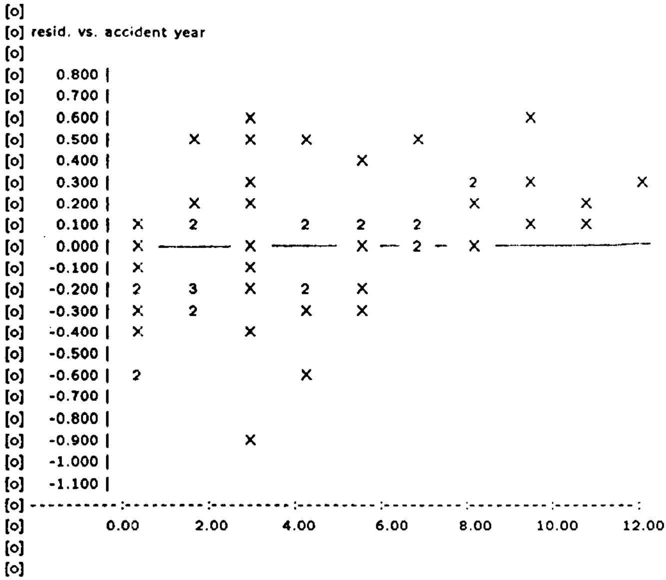

Figure 9.3(a).

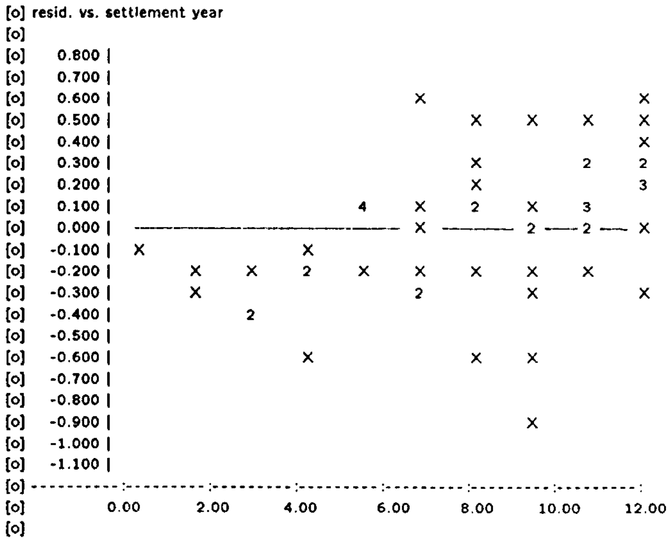

Figure 9.3(b).

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

Table 9.3

[o] The observed values and their expected values; the predicted values and their standard errors; the predicted row totals and their standard errors; the predicted diagonal totals and their standard errors; the predicted grand total and its standard error

		1	2	3	4	5	6	7	8	9	10	11
1	O_P	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948	0
	E_E	273714	680804	699807	718429	401531	272374	243232	176409	259453	67948	0
2	O_P	352118	884021	933894	1183289	445745	320996	527804	266172	425046	110927	110927
	E_E	392716	976794	1004059	1030776	576104	390793	348981	253106	372255	60216	60216
3	O_P	290507	1001799	926219	1016654	750816	146923	495992	280405	379507	102650	482157
	E_E	362971	902811	928011	952705	532470	361194	322549	233936	176652	56018	189895
4	O_P	310608	1108250	776189	1562400	272482	352053	206286	228731	340090	91988	660810
	E_E	324822	807924	830475	852574	476506	323232	288648	100679	159531	50502	210040
5	O_P	443160	693190	991983	769488	504851	470639	351067	256032	380684	102968	1090752
	E_E	362965	902795	927995	952688	532460	361188	150874	113892	180280	56952	304721
6	O_P	396132	937085	847498	805037	705960	404479	362429	264319	393004	106300	1530531
	E_E	373842	929849	955803	981237	548416	172638	158067	119225	188466	59377	401125
7	O_P	440832	847631	1131398	1063269	646825	439787	394066	287391	427310	115580	2310959
	E_E	405100	1007596	1035720	1063280	277979	191806	175485	132214	208612	65476	601535
8	O_P	359480	1061648	1443370	1268870	710402	483014	432799	315639	469311	126940	3806975
	E_E	442462	1100527	1131245	558507	316012	217870	199109	149764	235648	73541	1056661
9	O_P	376686	986608	1090099	1120656	627422	426594	382245	278770	414492	112112	4452390
	E_E	386545	961444	507415	525681	297127	204575	186611	139975	219213	67738	1375445
10	O_P	344014	973603	1001990	1030076	576709	392114	351349	256238	380989	103051	5066118
	E_E	344014	528510	546809	565520	318977	219024	199041	148448	230193	69595	2049338
11	O_P	0	5454109	4334037	3271569	2219466	1623609	1215010	797670	493102	103051	19511616
	E_E	0	1011697	900356	752136	510564	408851	358269	282618	242484	69595	3194056

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

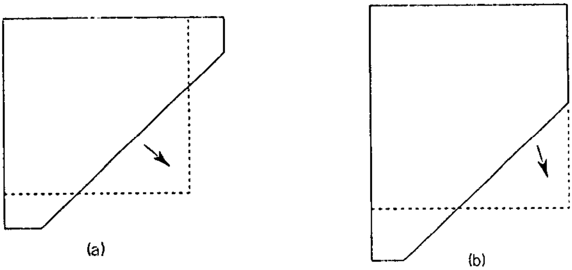

Figure 9.4. Arrows indicating direction of decreasing predictor stability.

Comparison of Tables 9.4a–b with Table 9.2 and Tables 9.5a–b with Table 9.3 give an indication of the degree of instability involved. In the construction of Tables 9.4(a) and 9.5(a) the original claims amount C'_{32} is changed approximately 10% from 1001799 to 901799 while Tables 9.4(b) and 9.5(b) are based on a substantial adjustment to the original claims amount C'_{28} from 266172 to 166172. We leave the reader to assess for his or herself the magnitude and pattern of changes induced in the predicted values by these two representative changes in the claims data by comparing Tables 9.5a–b with Table 9.3. As a further guide changes to the penultimate row or column of the run-off triangle induce some changes up to the same order of magnitude in the corresponding row or column

Table 9.4(a)

```
[o]     The parameter estimates are
[o]
[o]          estimate      s.e.     parameter
[o]   1      6.106        0.1644      1
[o]   2      0.8995       0.1604     DY_(2)
[o]   3      0.9395       0.1678     DY_(3)
[o]   4      0.9663       0.1758     DY_(4)
[o]   5      0.3852       0.1854     DY_(5)
[o]   6     -0.002226     0.1975     DY_(6)
[o]   7     -0.1145       0.2139     DY_(7)
[o]   8     -0.4345       0.2383     DY_(8)
[o]   9     -0.05308      0.2802     DY_(9)
[o]  10     -1.393        0.3780     DY_(10)
[o]  11      0.1938       0.1604     AY_(2)
[o]  12      0.1358       0.1678     AY_(3)
[o]  13      0.1539       0.1758     AY_(4)
[o]  14      0.3000       0.1854     AY_(5)
[o]  15      0.4136       0.1975     AY_(6)
[o]  16      0.5112       0.2139     AY_(7)
[o]  17      0.6772       0.2383     AY_(8)
[o]  18      0.5015       0.2802     AY_(9)
[o]  19      0.6022       0.3780     AY_(10)
[o]     scale parameter taken as  0.1158
```

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

Table 9.4(b)

```
[o] The parameter estimates are
[o]
[o]        estimate     s.e.     parameter
[o]   1    6.123       0.1663    1
[o]   2    0.9112      0.1623    DY_(2)
[o]   3    0.9387      0.1697    DY_(3)
[o]   4    0.9650      0.1779    DY_(4)
[o]   5    0.3832      0.1875    DY_(5)
[o]   6   -0.004909    0.1998    DY_(6)
[o]   7   -0.1181      0.2164    DY_(7)
[o]   8   -0.5963      0.2411    DY_(8)
[o]   9   -0.04369     0.2834    DY_(9)
[o]  10   -1.410       0.3823    DY_(10)
[o]  11    0.1415      0.1623    AY_(2)
[o]  12    0.1522      0.1697    AY_(3)
[o]  13    0.1370      0.1779    AY_(4)
[o]  14    0.2824      0.1875    AY_(5)
[o]  15    0.3953      0.1998    AY_(6)
[o]  16    0.4920      0.2164    AY_(7)
[o]  17    0.6568      0.2411    AY_(8)
[o]  18    0.4789      0.2834    AY_(9)
[o]  19    0.5854      0.3823    AY_(10)
[o]     scale parameter taken as  0.1185
```

of predicted values, with changes of a much lower order of magnitude elsewhere in the predicted values. Changes in the final row or column induce changes of a greater order of magnitude in that row or column of predicted values while leaving the remaining predicted values unchanged. We would strongly recommend that any practitioners should conduct their own simulation exercises to

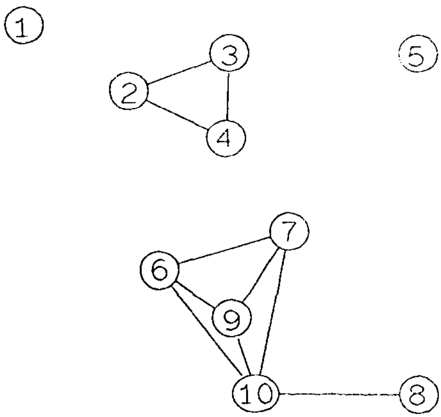

Figure 9.5. Partition of row parameters.

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

Table 9.5(a)

[o] The observed values and their expected values; the predicted values and their standard errors; the predicted row totals and their
[o] standard errors; the predicted diagonal totals and their standard errors; the predicted grand total and its standard error

		1	2	3	4	5	6	7	8	9	10	11
1	O_P	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948	0
	E_E	273597	672606	700017	719094	402155	272988	243994	177184	259453	67948	0
2	O_P	352118	884021	933894	1183289	445745	320996	527804	266172	425046	110881	110881
	E_E	392547	965032	1004361	1031732	576999	391674	350075	254218	372255	60081	60081
3	O_P	290507	901799	926219	1016654	750816	146923	495992	280405	374433	101267	475700
	E_E	358077	880292	916167	941135	526332	357281	319334	231895	173980	55162	187022
4	O_P	310608	1108250	776189	1562400	272482	352053	206286	229814	340195	92007	662016
	E_E	324886	798695	831245	853898	477545	324164	289734	100977	159296	50420	209993
5	O_P	443160	693190	991983	769488	504851	470639	352513	257404	381037	103053	1094007
	E_E	363264	893042	929436	954766	533955	362456	151230	114301	180126	56894	305043
6	O_P	396132	937085	847498	805037	705960	406075	364174	265919	393642	106462	1536272
	E_E	374412	920449	957960	984067	550342	173015	158550	119735	188435	59357	401889
7	O_P	440832	847631	1131398	1063269	649488	441904	396306	289382	428374	115855	2321309
	E_E	406073	998284	1038967	1067281	278634	192391	176171	132894	208757	65511	603148
8	O_P	359480	1061648	1443370	1274875	714208	485938	435797	318218	471060	127400	3827496
	E_E	444080	1091721	1136213	560163	317144	218803	200133	150719	236101	73670	1060426
9	O_P	376686	986608	1096943	1128391	632145	430103	385724	281654	416934	112761	4484655
	E_E	388810	955844	509694	528369	298831	205889	187972	141169	220103	68002	1382932
10	O_P	344014	961892	1002301	1031037	577605	392995	352444	257354	380963	103033	5059624
	E_E	344014	521198	545978	565009	318886	219113	199293	148818	229745	69448	2042927
11	O_P	0	5456915	4349844	3284389	2229581	1631772	1221014	801688	493724	103033	19571968
	E_E	0	1009820	902085	753019	511739	410073	359273	283586	242139	69448	3195949

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

Table 9.5(b)

[o] The observed values and their expected values; the predicted values and their standard errors; the predicted row totals and their standard errors; the predicted diagonal totals and their standard errors; the predicted grand total and its standard error

		1	2	3	4	5	6	7	8	9	10	11
1	O_P	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948	0
	E_E	278229	692032	711348	730277	408154	276866	247244	153259	266334	67948	0
2	O_P	352118	884021	933894	1183289	445745	320996	527804	166172	425046	105542	105542
	E_E	378834	942266	968567	994340	555740	376979	336645	208676	362638	57940	57940
3	O_P	290507	1001799	926219	1016654	750816	146923	495992	280405	391615	103255	494869
	E_E	370167	920709	946407	971591	543025	368354	328943	203902	184286	56986	197551
4	O_P	310608	1108250	776189	1562400	272482	352053	206286	195836	344128	90734	630698
	E_E	324822	807924	830475	852574	476506	323232	288648	87135	163197	50378	205517
5	O_P	443160	693190	991983	769488	504851	470639	351662	219218	385216	101568	1057664
	E_E	362965	902795	927995	952688	532460	361188	152765	98577	184432	56815	301042
6	O_P	396132	937085	847498	805037	705960	405156	363060	226324	397702	104860	1497101
	E_E	373842	929849	955803	981237	548416	174796	160060	103200	192821	59238	398510
7	O_P	440832	847631	1131398	1063269	647921	440552	394779	246096	432447	114021	2275816
	E_E	405100	1007596	1035720	1063280	281462	194224	177715	114455	213455	65331	600742
8	O_P	359480	1061648	1443370	1271116	711685	438908	433630	270315	475006	125242	3770902
	E_E	442462	1100527	1131245	565571	320026	220654	201674	129670	241159	73390	1060538
9	O_P	376686	986608	1092238	1122884	628691	427477	383062	238792	419613	110637	4423394
	E_E	386545	961444	513985	532508	301002	207257	189078	121234	224414	67621	1383026
10	O_P	344014	976117	1004603	1032790	578249	393179	352327	219633	385945	101760	5044603
	E_E	344014	535863	554437	573434	323459	222117	201873	128701	235890	69545	2064060
11	O_P	0	5437202	4309385	3242442	2184722	1583863	1180146	764488	496582	101760	19300592
	E_E	0	1024316	910462	759709	512367	406603	355971	276289	247916	69545	3207237

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

familiarize themselves with the nature and magnitude of such instability that exists.

Predictor instability with increasing development year ceases to be an issue (see Figure 9.4(b)) either if $w > 0$ or when using a model of the type defined by (3.2). In addition, as already stated in Section 7, instability in the north-east corner of the run-off triangle is generally not a serious problem since claims amounts in this region are relatively low in comparison with the remainder of the data matrix. One further noteworthy feature of the two-way ANOVA model when $w > 0$ is the invariance of predicted values to row permutations between the early accident years i, for which $i < w+1$ in the data matrix.

One potent way of diminishing the degree of instability to satisfactory levels in the important south-west corner of the run-off triangle is by allocating the same

Table 9.6

```
[o]
[o]        estimate     s.e.     parameter
[o]   1    6.119        0.1520   1
[o]   2    0.9024       0.1476   DY_(2)
[o]   3    0.9324       0.1528   DY_(3)
[o]   4    0.9363       0.1598   DY_(4)
[o]   5    0.3522       0.1696   DY_(5)
[o]   6   -0.01988      0.1838   DY_(6)
[o]   7   -0.1330       0.1995   DY_(7)
[o]   8   -0.4500       0.2202   DY_(8)
[o]   9   -0.05353      0.2580   DY_(9)
[o]  10   -1.406        0.3551   DY_(10)
[o]  11    0.1682       0.1267   MAY_(2)
[o]  12    0.3009       0.1746   MAY_(3)
[o]  13    0.5102       0.1467   MAY_(4)
[o]   scale parameter taken as  0.1030
[o]
[o]
```

```
[o]
[o]
[o]        histogram of residuals
[o]
[o] [-1.00,-0.75)      1   X
[o] [-0.75,-0.50)      2   XX
[o] [-0.50,-0.25)      5   XXXXX
[o] [-0.25, 0.00)     19   XXXXXXXXXXXXXXXXXXX
[o] [ 0.00, 0.25)     18   XXXXXXXXXXXXXXXXXX
[o] [ 0.25, 0.50)      8   XXXXXXXX
[o] [ 0.50, 0.75]      2   XX
[o]
[o]
```

Figure 9.6(a).

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

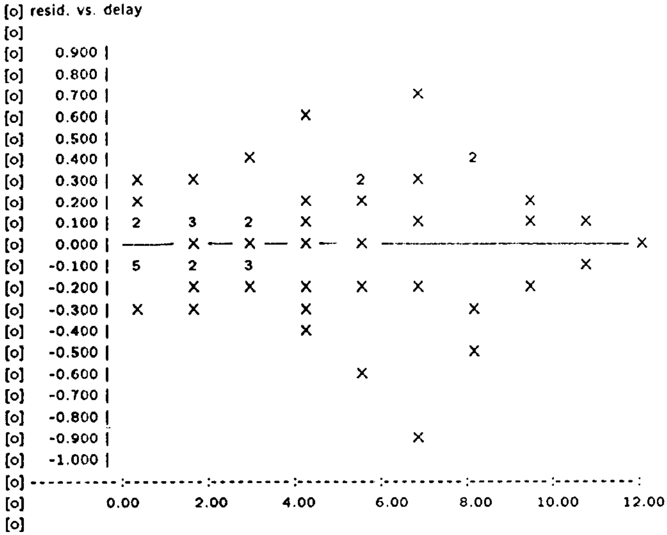

Figure 9.6(b).

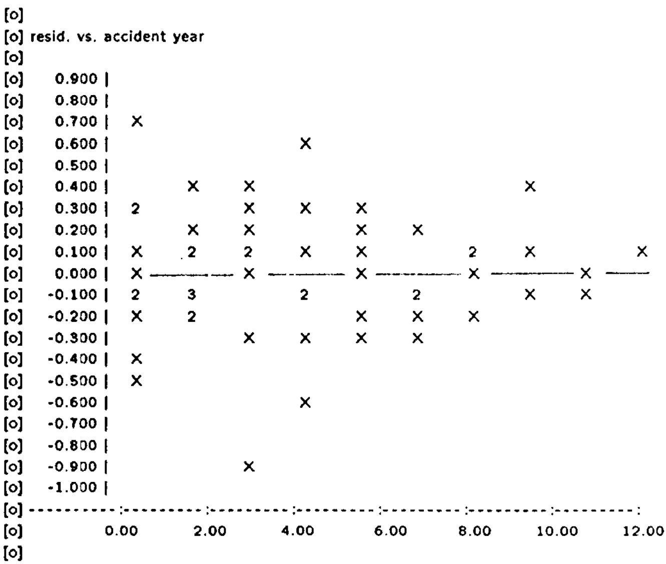

Figure 9.6(c).

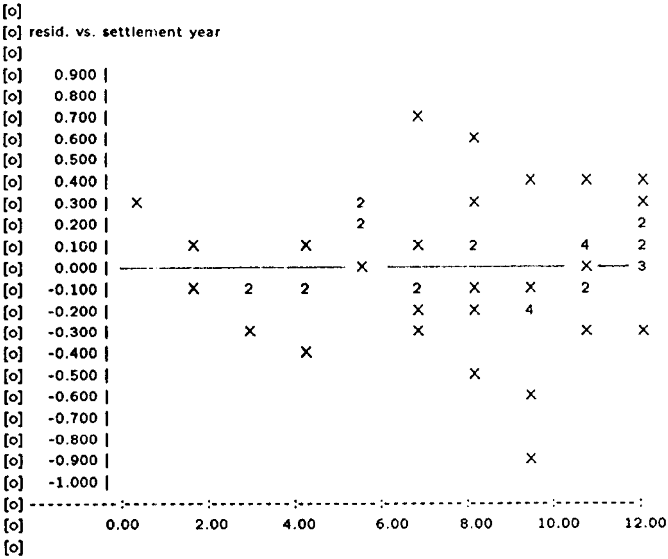

Figure 9.6(d).

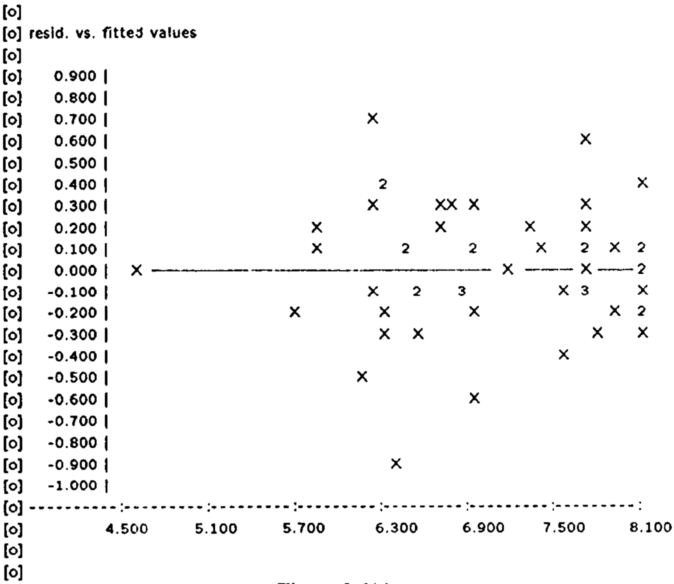

Figure 9.6(e).

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

Table 9.7

The observed values and their expected values; the predicted values and their standard errors; the predicted row totals and their standard errors; the predicted diagonal totals and their standard errors; the predicted grand total and its standard error

		1	2	3	4	5	6	7	8	9	10	11
1	O_P	357848	766940	610542	482940	527326	574398	146342	139950	227229	67948	0
	E_E	277251	683596	704356	787118	394287	271794	242714	176783	262799	67948	0
2	O_P	352118	884021	933894	1183289	445745	320996	527804	266172	425046	106178	106178
	E_E	387726	955984	985017	988881	551397	380094	339427	247225	367516	52938	52938
3	O_P	290507	1001799	926219	1016654	750816	146923	495992	280405	384580	102644	487224
	E_E	374820	924162	952229	955964	533043	367442	328128	238995	159369	51176	169274
4	O_P	310608	1108250	776189	1562400	272482	352053	206286	228721	342646	91452	662368
	E_E	333950	823393	848399	851727	474920	327376	292350	88180	141992	45595	176534
5	O_P	443160	693190	991983	769488	504851	470639	347695	254350	381537	101628	1085208
	E_E	368440	908432	936021	939693	523970	361188	139961	105542	167850	52486	283104
6	O_P	396132	937085	847498	805037	705960	438326	392606	287204	430819	114755	1663710
	E_E	417901	1030385	1061677	1065842	594310	166773	152894	115495	184237	57995	358105
7	O_P	440832	847631	1131398	1063269	623326	431180	386205	282521	423795	112884	2259910
	E_E	411088	1013585	1044367	1048464	230541	164054	150401	113612	181233	57050	443959
8	O_P	359480	1061648	1443370	1033155	577409	399417	357755	261709	392596	104568	3126589
	E_E	380805	938919	967434	374774	213559	151969	139322	105243	167882	52847	586503
9	O_P	376686	986608	1072397	1078343	602663	416886	373402	273156	409747	109142	4335736
	E_E	397461	979985	383577	391165	222899	158615	145416	109846	175225	55158	752657
10	O_P	344014	831639	857917	862674	482131	333509	298722	218525	327797	87313	4300228
	E_E	317968	294325	306862	312932	178319	126892	116333	87877	140180	44127	692499
11	O_P	0	5065567	4037095	3011151	2071740	1507170	1077338	732839	436939	87313	18027152
	E_E	0	765453	664249	532789	401484	326170	260516	213277	154305	44127	2145715

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

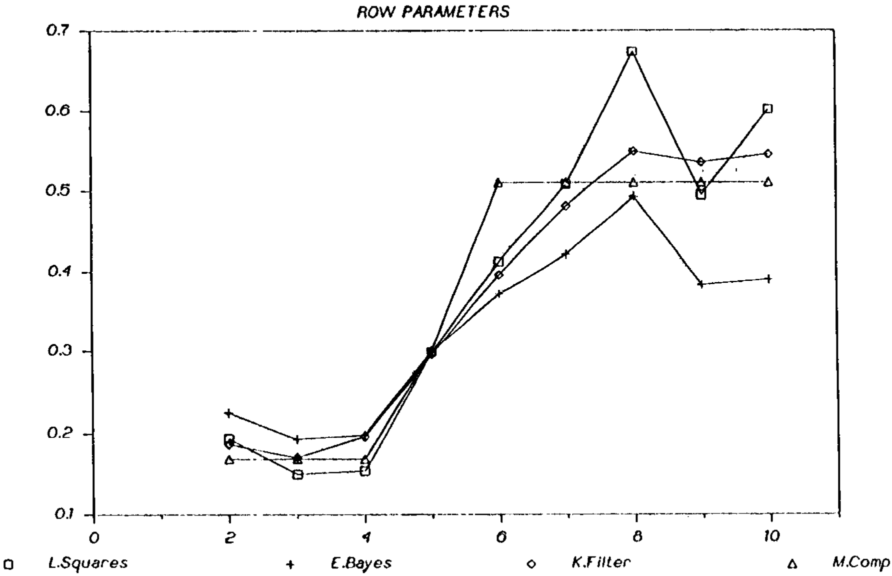

Figure 9.7.

α_i parameters to more than one accident year where appropriate. Indeed, this is vital if acceptable levels of stability are to be induced for the most recent accident years for which relatively little data are, as yet, available. We stress that this defect is also present in the traditional actuarial deterministic chain-ladder technique, giving rise to much concern about the apparant continuing esteem afforded to the technique.

A way forward is to examine all contrasts

$$\alpha_{i_1} - \alpha_{i_2}, \qquad i_1 \neq i_2$$

between row parameters. Such contrasts are invariant of the somewhat arbitrary choice of the two parameter constraints ($\alpha_1 = \beta_1 = 0$) needed to estimate the α_is.

Application of the multicomparison t-criterion

$$\left| \frac{\hat{\alpha}_{i_1} - \hat{\alpha}_{i_2}}{\sqrt{\hat{V}(\hat{\alpha}_{i_1} - \hat{\alpha}_{i_2})}} \right| < h \qquad \forall\, i_1, i_2\ (i_1 \neq i_2)$$

for $h = \cdot 5$, induces the partition in row parameters displayed in Figure 9.5 in which accident years are represented by numbered nodes; two nodes being linked if and only if the inequality is satisfied.

This allocates separate row parameters to years 1 and 5 while linking years, 2, 3

https://doi.org/10.1017/S0020268100036702 Published online by Cambridge University Press

and 4 together as well as linking years 6 to 10 inclusive; making a total of just four row parameters. For sufficiently large h, all nodes are interlinked, while linkages are shed as h is reduced.

The residual plots (Figures 9.6(a)–(e)), the parameter estimates (Table 9.6) and predicted values (Table 9.7) are presented for scrutiny.

Verrall (1989)[13] has conducted a comparative study of estimates for the α_is based on a variety of estimation methods for these data. A graphical comparison of least squares, empirical Bayes, Kalman filter and multi comparison estimators is presented in Figure 9.7.

10. POSTSCRIPT

Possible future developments for incorporating within GLIM include:

(i) alternative methods of mapping back from the logarithmic modelling space;
(ii) use of the other model structures discussed in Section 3 (partially developed);
(iii) use of methods other than the multicomparison tests to induce predictor stability.

We would like to acknowledge the financial support received from the Commercial Union Insurance Company together with the encouraging ongoing discussions held with Stavros Christofides and Peter Crane from that company. We are most interested to hear from any further practitioners interested in these highly practical developments.

REFERENCES

(1) AITCHISON, J. & BROWN, J. A. C. (1969). *The Log-Normal Distribution*. Cambridge University Press.
(2) DEJONG, P. & ZEHNWIRTH, B. (1983). Claims Reserving, State-Space Models and the Kalman Filter. *J.I.A.* **110,** 157.
(3) GREEN, P. J. (1988). Non-Diagonal Weight Matrices. *GLIM Newsletter No. 16.*
(4) HABERMAN, S. & RENSHAW, A. E. (1988). Generalised Linear Models in Actuarial Work. Presented to a joint meeting of the Staple Inn Actuarial Society and the Royal Statistical Society, General Applications Section. February 1988.
(5) HOSSACK, I. B., POLLARD, J. H. & ZEHNWIRTH, B. (1983). *Introductory Statistics with Applications in General Insurance.* Cambridge University Press.
(6) KREMER, E. (1982). IBNR Claims and the Two-Way Model of ANOVA. *Scandinavian Actuarial Journal.*
(7) TAYLOR, G. C. (1979). Statistical Testing of a Non-Life Insurance Model. *Proceedings Actuarial Sciences Institute, Act. Wetemschappen, Katholieke Univ.* Leuven, Belgium.
(8) TAYLOR, G. C. (1983). An Invariance Principle for the Analysis of Non-Life Insurance Claims. *J.I.A.* **110,** 205–242.
(9) TAYLOR, G. C. & ASHE, F. R. (1983). Second Moments of Estimates of Outstanding Claims. *Journal of Econometrics* **23,** 37–61.
(10) TAYLOR, G. C. (1986). *Claims Reserving in Non-Life Insurance.* North-Holland.

(11) TAYLOR, G. C. (1988). *Regression Models in Claims Analysis (II)*, William M. Mercer, Campbell, Cook, Knight, Sydney, Australia.
(12) VERRALL, R. (1988). Bayes Linear Models and the Claims Run-Off Triangle. *Actuarial Research Report No. 7*. The City University, London.
(13) VERRALL, R. (1989). Private communication.
(14) ZEHNWIRTH, B. (1985). *Interactive Claims Reserving Forecasting System*. Benhar Nominees Pty Ltd, Turramurra, NSW, Australia.