[D5]
REGRESSION MODELS BASED ON LOG-INCREMENTAL PAYMENTS
Contributed by S Christofides

The first article in Volume 2 of this Manual by B Zehnwirth has shown the close connection between the intuitive Chain Ladder technique and the more formal two way analysis of variance model based on the log-incremental payments.

Models initiated by this more formal definition of the basic chain ladder have recently started to gain acceptance in loss reserving work and a number of papers on the subject have now been published. These models differ from the traditional techniques by a more formal definition of both the model assumptions and the parameter estimation and testing. With the formal models statistical estimates of reserves, that is both mean estimates and the associated model standard errors, can be calculated. The basic chain ladder is deterministic and produces point estimates of reserves.

The purpose of this paper is to serve as a basic introduction to these methods for the practitioner. To facilitate this a PC spreadsheet package is used to show how run-off models of the log-incremental payments can be identified and fitted in practice using multiple regression.

The approach adopted considers the basic chain ladder technique first and shows how the intuitive chain ladder model can be made more formal. The parameters of this model are then estimated and the implied underlying payment pattern compared with the chain ladder derived pattern. Both models are used to "fill in the square" and the results compared. In the case of the formal model it is also shown how the regression results are used to derive estimates of the individual future payments and their standard errors and how accident year and overall standard errors can be calculated.

The simple example makes it easier to follow the calculations and is intended to allow the reader to focus on the more interesting modelling aspects of the later sections.

A more realistic example is then analysed. The data is first viewed graphically to identify an appropriate run-off model to fit. The identified model is fitted and tested. The model is then redefined with fewer parameters and refitted. The results, both future payments and their standard errors, from these models are calculated and compared.

The data is then adjusted for inflation and for claim volume and a series of models are identified and tested. Three of these are used to obtain estimates which are then compared.

A degree of theory is assumed. The model parameters are estimated using multiple regression and matrix operations are used to calculate the variance-covariance matrices. All the computations and graphs are done in a PC spreadsheet package, Supercalc 5 in this case although Lotus 123 could have been used equally effectively. The wide availability, ease of use and power of these packages makes these methods accessible to all. Alternatively any programming language with matrix manipulation capabilities, such as APL or SAS, could be used for this work. Programs have also been written in GLIM (see A Renshaw, 2).

A. Introduction

Almost all actuarial methods for estimating claims reserves have an underlying statistical model. Obtaining estimates of the parameters is not always carried out in a formal statistical framework and this can lead to estimates which are not statistically optimal. These traditional methods generally produce only point estimates.

The models, such as the basic chain ladder, are often overparameterised and adhere too closely to the actual observed data. This process can lead to a high degree of instability in values predicted from the model as the close adherence to the observed values results in parameter estimates which are very sensitive to small changes in the observed values. A small change in an observed value, particularly in the south-west or north-east regions of the data triangle, can result in a large change in the predicted values. In practice attempts may be made to achieve some stability in the results by using benchmark patterns, by selection of development factors and a number of other such techniques.

Formal statistical models are used extensively in data analysis elsewhere to obtain a better understanding of the data, for smoothing and for prediction. Explicit assumptions are made and the parameters estimated via rigorous mathematics. Various tests can then be applied to test the goodness of fit of the model and, once a satisfactory fit has been obtained, the results can be used for prediction purposes.

This process allows us to focus on the model being fitted and should also highlight any inadequacies in the model. The estimates of the parameters, on the basis of the model, can be made statistically optimal. Peculiarities in the data may be identified and often investigation of these can yield useful additional information to the modeller.

All modelling, whether based on the traditional actuarial techniques such as the chain ladder or on more formal statistical models, requires a fair amount of skill and experience on the part of the modeller. All these models are attempting to describe the very complex claims process in relatively simple terms and often with very little data. The advantage of the more formal approach is that the appropriateness of the model can be tested and its shortcomings, if any, identified before any results are obtained.

B. The basic chain ladder technique and the underlying stochastic model

The following simple example considers a 4 by 4 triangle of cumulative payments:

CUMULATIVE PAID CLAIMS

DEVELOPMENT YEAR

ACC YR	0	1	2	3
0	11073	17500	19339	20105
1	14799	24156	26500	
2	15636	26159		
3	16913			

The usual (weighted) basic chain ladder development factors are (see Vol 1 Section E8):

0 to 1	1 to 2	2 to 3
1.633781	1.100418	1.039609

where 1.633781 = (17500 + 24156 + 26159)/(11073 + 14799 + 15636) etc.

Using these factors the square can be completed in the usual way:

CUMULATIVE PAID CLAIMS

DEVELOPMENT YEAR

ACC YR	0	1	2	3
0	11073	17500	19339	20105
1	14799	24156	26500	27550
2	15636	26159	28786	29926
3	16913	27632	30407	31611

The actual and fitted portions of the square have been separated for illustration. It is assumed in this example that there are no payments beyond the 3rd development period so that the first (zero'th) accident year is complete.

The chain ladder produces successive cumulative losses from which the future incremental payments can be derived by subtraction. It is therefore possible to split the overall chain ladder derived reserve estimate for an accident year into its incremental or payment year values.

The underlying model is better illustrated by these incremental payments which are shown in the table below.

INCREMENTAL PAID CLAIMS

DEVELOPMENT YEAR

ACC YR	0	1	2	3	0/S
0	11073	6427	1839	766	—
1	14799	9357	2344	1050	1050
2	15636	10523	2627	1140	3767
3	16913	10719	2775	1204	14698
				Total	19515

The accident year projected future payments and the overall estimate are shown in the last column. The chain ladder estimate of future payments to development period 3 for all accident years is 19515.

Dividing each of these incremental amounts by the final, or ultimate, accident year value gives the following:

PERCENTAGE PAID CLAIMS

DEVELOPMENT YEAR

ACC YR	Ultimate	0	1	2	3
0	20105	55.08	31.97	9.15	3.81
1	27550	53.72	33.96	8.51	3.81
2	29926	52.25	35.16	8.78	3.81
3	31611	53.50	33.91	8.78	3.81

The basic chain ladder has produced the following underlying incremental payment pattern:

Development year	0	1	2	3
Incremental paid %	53.50	33.91	8.78	3.81

Note that this underlying pattern can be calculated directly from the development factors.

The basic chain ladder assumptions can be restated as follows:

a: Each accident year has its own unique level.

b: Development factors for each period are independent of accident year or, equivalently, there is a constant payment pattern.

These assumptions can now be used to define the model more formally.

Let:

A_i be the ultimate (cumulative) payments for the i-th accident year.

B_j be the percentage of ultimate claims paid during the j-th development period.

P_{ij} be the incremental paid claims for accident year i paid during development period j

The chain ladder model can thus be described by the following equations

$$P_{ij} = A_i \times B_j \quad \text{for i,j from 0 to 3}$$

and the condition

$$\sum B_j = 1 \qquad \text{where j is summed from 0 to 3}$$

The next section considers how these equations may be solved and estimates of the parameters obtained.

C. Estimating the parameters of the formal chain ladder model

As the main set of relations involves products the usual approach is first to make these linear by taking logarithms and then use multiple regression to obtain estimates of the parameters in log-space. It will eventually be necessary to reverse this transformation to get back to the original data space.

Dealing with the main set of equations is relatively easy. Taking logarithms (natural logarithms will be assumed throughout and denoted by ln) gives

$$\ln (P_{ij}) = \ln A_i + \ln B_j$$

Unfortunately taking logarithms of the second condition does not produce a linear equation as

$$\ln(\sum B_j) \neq \sum (\ln B_j)$$

It is possible to obtain estimates of these parameters using iterative procedures but this is not pursued here. It is more convenient to drop the condition and concentrate initially on obtaining the parameter estimates from the remaining, now linear, set of equations.

Dropping the condition gives rise to a singularity and so it is necessary to introduce a new condition in order to obtain the parameter estimates. This does not affect the eventual results but it does change the interpretation of the parameters.

For ease of reference the parameters are now redefined ($\ln A_i = a_i$ etc) and an error term introduced.

$$\ln(P_{ij}) = Y_{ij} = a_i + b_j + e_{ij}$$

where e_{ij} is some error term.

As indicated above without some restriction these equations are singular. Note for example that a_3 appears only in one equation which involes b_0 and an error term. An infinite number of combinations of a_3 and b_0 are possible as long as they sum to the same view.

For convenience in this example b_0 is set to zero. Another approach is to set both a_0 and b_0 equal to zero and introduce a constant, k, into the model. The chain ladder assumes each accident year has a unique level so the model to be fitted below will follow the former description. The alternative definition is considered later in Section H and the advantages of this choice outlined.

The predictions obtained by either approach will be the same so the restriction can be chosen at the convenience of the modeller.

The model to be fitted is described by:

$$\ln(P_{ij}) = Y_{ij} = a_i + b_j + e_{ij}$$

where i and j go from 0 to 3 and $b_0 = 0$

The model has seven parameters to be estimated, the same number as the basic chain ladder model.

The following table is in the form most convenient for the regression facility of any of the popular spreadsheet packages.

		Y-variate		←		Design Matrix X				→
i	j	P_{ij}	Y_{ij}	a_0	a_1	a_2	a_3	b_1	b_2	b_3
0	0	11073	9.31226	1	0	0	0	0	0	0
0	1	6427	8.76826	1	0	0	0	1	0	0
0	2	1839	7.51698	1	0	0	0	0	1	0
0	3	766	6.64118	1	0	0	0	0	0	1
1	0	14799	9.60231	0	1	0	0	0	0	0
1	1	9357	9.14388	0	1	0	0	1	0	0
1	2	2344	7.75961	0	1	0	0	0	1	0
2	0	15636	9.65733	0	0	1	0	0	0	0
2	1	10523	9.26132	0	0	1	0	1	0	0
3	0	16913	9.73584	0	0	0	1	0	0	0

Each row corresponds to a data value and its representation by the model parameters. The last but one row, for example, describes the accident year 2, development year 1, value in log-space as the sum of the a_2 and b_1 parameters. The coefficients of the other parameters are zero for this data value.

The resulting matrix of parameter coefficients, made up of ones and zeros in this case, will be referred to as the model design matrix **X**. It is governed by the model chosen.

Within the class of log-linear models changing the model just involves changing the design matrix.

The regression takes the $\ln(P_{ij})$ or Y_{ij} values as the dependent variable and each of the columns of the matrix **X** as the independent variables.

The spreadsheet regression command, which requires a colummn for the dependent values and a range for the independent values (i.e. the design matrix) is then used to carry out the regression and output the result. It is necessary to specify that the fit is without a constant and to define a results or output range. This is quite straightforward in practice and the results are produced almost instantly.

The spreadsheet output in this case will be:

```
Regression Output:
     Constant                         0
     Std Err of Y Est                 .0524
     R Squared(Adj,Raw)     .9976     .9992
     No. of Observations              10
     Degrees of Freedom               3
```

Coefficient(s)	9.288	9.591	9.692	9.736	-.4661	-1.801	-2.647
Std Err of Coef.	.0400	.0400	.0428	.0524	.04277	.05015	.06591

A brief description of this fairly standard spreadsheet regression output will be found in Appendix 2.

The results can also be obtained by matrix manipulation. An indication of how this can be done is given in Section D.

The coefficients are the parameter estimates and are in the same order as the columns of the design matrix.

So the model estimate for a_0 is 9.288, for a_1 it is 9.591 and so on until b_3 which is estimated as -2.647.

The payment pattern can be derived from this output. This is done by exponentiating the development year parameters b_j's, remembering to bring in the b_0 which was set to 0, and scaling so that the exponentiated values add up to the required 100%.

A formal proof of this is beyond the scope of this paper and the interested reader is referred to Verrall's paper (5) "Chain Ladder and Maximum Likelihood". The table below, and the comparison with the basic chain ladder result, may be sufficient to satisfy the majority of practitioners.

The following table shows these basic calculations

Parameter	b_0	b_1	b_2	b_3	
Coefficient	0	-.4662	-1.8015	-2.6472	sum
exp (b_j)	1	0.6274	0.1651	0.0709	1.8634
Payment %	53.67	33.67	8.86	3.80	100

This is very close to the basic chain ladder derived pattern.

BCL Payment %	53.50	33.91	8.78	3.81

The slight differences arise from the way the parameter estimates are derived. The same underlying model is assumed in both cases. Unfortunately however a fair amount of further manipulation is necessary to obtain estimates of ultimate values for each accident year. These cannot be derived simply from the accident year regression coefficients.

In order to progress further it is now necessary to go back and consider what assumptions were made by the spreadsheet in deriving the parameter estimates. This requires a more detailed consideration of the formal model and in particular the structure of the assumed error term.

These aspects are considered in the following section.

D. Fitting assumptions and error terms

The spreadsheet regression is fitted by least squares. That is by minimizing the sum of the squares of the error terms e_{ij}.

It is usual and convenient to assume that the error values e_{ij} are identically and independently distributed with a normal distribution whose mean is zero and variance some fixed σ^2.

i.e. $e_{ij} = \text{IID } N(0, \sigma^2)$

In matrix form it can be shown that, under these assumptions, the parameter estimates are given by

$$(\mathbf{X}^T\mathbf{X})^{-1}\,\mathbf{X}^T\mathbf{Y}$$

where **X** is the design matrix and $\mathbf{X}^T$ its transpose and **Y** is the data vector. The standard errors can also be calculated in matrix form.

These assumptions can be tested by analysis of the residual (error) terms, by plots and other diagnostic tests. Residual plots are shown and discussed later.

Recalling that the original payments were transformed by taking logarithms the error normality assumption in log-space implies that the data in the original space are log-normally distributed.

The IID assumption estimates are also the maximum likelihood estimates in this case and it can be shown that the parameter estimates so obtained are unbiased. Since maximum likelihood estimates are invariant under transformation Verrall (5) shows in "Chain Ladder and Maximum Likelihood" how maximum likelihood estimates of development factors can be obtained by direct substitution.

As the log-normal distribution is skewed with a tail to the right some extreme high values are to be expected. This is sometimes a feature of incremental claims payment triangles. The cause is usually a large claim payment in later development periods, the settlement perhaps of a particularly large claim, when the overall level of payments is low.

These assumptions are not claimed to be theoretically justified for log-incremental claims payments. They have an intuitive appeal and are chosen primarily for convenience. Alternative assumptions, which may well be more generally applicable to claim payments, can be made and results obtained. These tend to require more complex computations or iterative procedures which generally necessitate the use of specially written software.

Further comments on the error terms are to be found in the final section of this paper which also includes some suggestions for dealing with negative incremental payments.

E. Predicting future payments and their standard errors

In order to derive estimates of the model parameters it was convenient to take logarithms and work in log-space. To obtain results in the original space it is necessary to reverse this transformation.

Obtaining the parameter estimates in log-space is relatively straightforward. To revert back to the original space is not so simple and it is necessary to use the relationships between the parameters of the log-normal distribution and the underlying normal distribution.

Again for simplicity the easiest approach is adopted here. This approach is also used by Zehnwirth and by Renshaw and again the justification can be found in their papers. These estimates, in the original space, are not necessarily unbiased especially where a small number of data points are being fitted. Verrall (6) shows how it is possible to obtain unbiased estimates but the calculations are more complicated.

The estimates to be used here are given by the following

The future values $\hat{P}_{ij}$'s are calculated from the estimates obtained in the log-space $\hat{Y}_{ij}$ as follows

a) $\hat{P}_{ij} = \exp(\hat{Y}_{ij} + 0.5 \operatorname{var}(\hat{Y}_{ij}))$

Their standard errors are given by

b) $\text{s.e.}(\hat{P}_{ij}) = \hat{P}_{ij} \operatorname{sqrt}(\exp(\operatorname{var}(\hat{Y}_{ij}))^{-1})$

So the first step is to derive the predicted values and their standard errors in log-space.

The predicted values in log-space are obtained from the estimates of the parameters produced by the regression.

For example the first future value to be predicted is for accident year 1 development year 3 and this is given by

$$\begin{aligned}\hat{Y}_{13} &= a_1 + b_3\\ &= 9.591 - 2.647\\ &= 6.944\end{aligned}$$

To obtain the variance of this, and the other estimates, it is necessary to calculate the variance-covariance matrix.

This matrix is given by

$$\sigma^2\, \mathbf{X}_f\, (\mathbf{X}^T\mathbf{X})^{-1}\, \mathbf{X}_f^T$$

where σ^2 is the model variance (scalar) and depends on the data

$\mathbf{X}_f$ is the design matrix of the future values and
$\mathbf{X}_f^T$ is its transpose and

$(\mathbf{X}^T\mathbf{X})^{-1}$ is the model information matrix

with $\mathbf{X}$ the design matrix and $\mathbf{X}^T$ its transpose.

In a spreadsheet a small macro can be written to carry out this calculation. The results of each stage of this calculation for the simple example above are to be found in Appendix 1.

Note that changing data values in the original triangle only affects the scalar factor σ^2 and so the lengthy matrix calculation only need be done once for a given size model.

The usual practice therefore is to calculate the matrix product

$$\mathbf{X}_f\, (\mathbf{X}^T\mathbf{X})^{-1}\, \mathbf{X}_f^T$$

and multiply by the specific data σ^2 as necessary.

A library of these matrices could be built up for the models to be used, to cater for different sizes of triangles for instance, and stored for future use.

The design matrix of future values $\mathbf{X}_f$, following the same format as the original design matrix, is as follows:

← Future Design Matrix **X** →

i	j	a_0	a_1	a_2	a_3	b_1	b_2	b_3
1	3	0	1	0	0	0	0	1
2	2	0	0	1	0	0	1	0
2	3	0	0	1	0	0	0	1
3	1	0	0	0	1	1	0	0
3	2	0	0	0	1	0	1	0
3	3	0	0	0	1	0	0	1

The matrix

$$\mathbf{X_f}\,(\mathbf{X^TX})^{-1}\,\mathbf{X_f^T}$$

in this case is (see Appendix 1)

$$\begin{matrix}
1.66667 & .00000 & 1.33333 & .00000 & .00000 & 1.33333 \\
.00000 & 1.25000 & .75000 & .00000 & .75000 & .25000 \\
1.33333 & .75000 & 1.91667 & .00000 & .25000 & 1.41667 \\
.00000 & .00000 & .00000 & 1.66667 & 1.33333 & 1.33333 \\
.00000 & .75000 & .25000 & 1.33333 & 1.91667 & 1.41667 \\
1.33333 & .25000 & 1.41667 & 1.33333 & 1.41667 & 2.58333
\end{matrix}$$

The variance-covariance matrix of future values is calculated from the above by just multiplying through by the model σ^2 which in this case is

$$.0524^2 = .002744$$

The variance-covariance matrix is then

$$\begin{matrix}
.00457 & .00000 & .00366 & .00000 & .00000 & .00366 \\
.00000 & .00343 & .00206 & .00000 & .00206 & .00069 \\
.00366 & .00206 & .00526 & .00000 & .00069 & .00389 \\
.00000 & .00000 & .00000 & .00457 & .00366 & .00366 \\
.00000 & .00206 & .00069 & .00366 & .00526 & .00389 \\
.00366 & .00069 & .00389 & .00366 & .00389 & .00709
\end{matrix}$$

Note that these matrices are square and symmetric with each side equal to the number of future values to be projected. The diagonal elements contain the variances of each of these values and are in the same order as the future design matrix elements.

To obtain the variances to be used for projecting future values we will follow common practice and add the model variance (σ^2) to the variances calculated above. These two sources of error are the estimation and statistical errors. These variances recognise that the parameter coefficients are estimates (and subject to error) as well as the inherent noise in the process or data. We do not attempt to correct or estimate any specification or selection errors which may well be equally significant contributors to a total overall error term. Our final example gives some indications of how projected values can be affected by the choice of model parameters. For a more detailed explanation of these types of error the reader is referred to the paper by Taylor (3).

The variances for the future values in log-space are the sum of the variance-covariance matrix values obtained above and the model variance σ^2.

So the variance for the first projected value which was estimated above, Y_{13}, is

$$1.66667 \times 0.05238^2 + 0.05238^2 = .007317$$

The following table shows the various values and their variances and standard errors

i	j	$\hat{Y}_{ij}$	$Var(\hat{Y}_{ij})$	$\hat{P}_{ij}$	$var(\hat{P}_{ij})$	$se(\hat{P}_{ij})$
1	3	6.94395	.007317	1041	7953	89
2	2	7.89094	.006174	2681	44520	211
2	3	7.04521	.008003	1152	10662	103
3	1	9.26969	.007317	10650	833010	913
3	2	7.93438	.008003	2803	63122	251
3	3	7.08865	.009832	1204	14328	120

We note here that the sum of the variances is 973595 which is a value that will be used later.

F. Accident year and overall standard errors

Calculating the variances or standard errors across accident years and in total requires one further step involving the covariances. The information needed is in the last matrix above together with the values calculated for $\hat{P}_{ij}$'s and their variances.

The variance of the sum of two values A and B is given by

$$Var(A+B) = Var(A) + Var(B) + 2Cov(A,B)$$

and this extends to sums of more than two values by including all pairs of covariances. Note that Cov(A,B) = Cov(B,A).

A justification is given in Renshaw's paper that in the case of log-linear models the covariances can be calculated in the original space by the following convenient formula

$$Cov(\hat{P}_{ij}\ , \hat{P}_{kl}) = E(\hat{Y}_{ij})\, E(\hat{Y}_{kl})\, (\exp(Cov(\hat{Y}_{ij}\ , \hat{Y}_{kl}\ ) -1)$$

In practice this can be set up fairly easily in the spreadsheet once the individual values have been estimated and the variance-covariance matrix computed. It does nevertheless involve a fair amount of computation. To illustrate the calculation consider the standard error for the second accident year.

Two values are involved $\hat{P}_{22}$ and $\hat{P}_{23}$, which were estimated as 2681 and 1152.

Their standard errors obtained above were 211 and 103 respectively. The covariance, in log-space, for these estimates can be found in the variance-covariance matrix and is 0.00206. So the covariance in the original space is

$$\begin{aligned}\mathrm{Cov}(\hat{P}_{22}, \hat{P}_{23}) &= 2681 \times 1152\,(\exp(.00206) - 1)\\ &= 6363\end{aligned}$$

The required variance of the sum is then given by

$$\mathrm{Var}(\hat{P}_{22} + \hat{P}_{23}) = 211^2 + 103^2 + 2 \times 6363 = 67868$$

So the estimated standard error of the total assumed outstanding claims for this year is 261 or just under 7% of the estimated value of 3833 (2681 + 1152).

This process can be applied to obtain the standard errors for any combination of values, for instance for each accident year or each payment year and more interestingly for the overall total reserve estimate.

The total reserve estimate is the sum of all the projected values and so its variance calculation will include all possible combinations of covariances (of pairs) of values involved in the calculation. This, surprisingly, makes the spreadsheet calculation easier as there is no need to exclude or select any values. One simply sums a range.

The calculations are as in the previous example and can be tabulated fairly easily to produce the following matrix of covariances.

(i,j)	(1,3)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
(1,3)	—	0	4394	0	0	4593
(2,2)	0	—	6363	0	15481	2216
(2,3)	4394	6363	—	0	2216	5403
(3,1)	0	0	0	—	109410	47006
(3,2)	0	15481	2216	109410	—	13145
(3,3)	4593	2216	5403	47006	13145	—
					Total =	420452

Note that the diagonal elements are left blank as the values here should be the variances which were estimated previously. The matrix is symmetric, as is to be expected, and so summing the range produces the sum of covariances of all possible pairs of values. This sum of all pairs of covariances is 420452.

The sum of the variances of the projected values obtained earlier was 973595 and so the overall variance, which is the sum of these two values, is 1394047.

The overall standard error, which is the square root of this value, is therefore estimated as 1181 or just 6% of the overall reserve estimate of 19531. The overall error is relatively small in this simple example. In practice, with real data involving more accident and development years, the percentage errors tend to be higher. The table below summarizes the results.

Project values and their standard errors:

		Development	Period			
Acc Yr		0	1	2	3	Tot Acc Yr
1	Amount				1041	1041
	s error				89	89
2	Amount			2681	1152	3833
	s error			211	103	261
3	Amount		10650	2803	1204	14657
	s error		913	251	120	1118
				Overall Total		19531
				Standard Error		1181

The chain ladder overall estimate was 19515. The individual values obtained by the two methods are also close but the chain ladder estimates are point estimates whereas the regression based estimates are statistical estimates with both a mean and a standard error estimate.

All the usual information that can be produced from the traditional chain ladder can be derived from the regression chain ladder including estimates of development factors. The stochastic approach as shown above can produce additional information, based on the model assumptions, such as standard errors of parameters and reserve estimates, that the traditional approach does not. The statistical estimates obtained by the regression approach also facilitate stability comparisons across companies and classes.

This completes our consideration of the regression chain ladder. The technique does not require that we have a complete triangle of data and can work with almost any shape data as long as there are sufficient points from which to obtain estimates of the parameters.

In the next section a log-linear regression model is fitted which is motivated by the run-off shape of the data. This model has fewer parameters as the development parameters are subject to some curve fitting. This is used to project values outside the original triangle shape, that is a tail is projected. The computation approach is identical to the above. The only differences are that there are now more data points to be fitted and the design, and future design matrices are different.

G. Identifying and fitting a regression model

1. Preliminary analysis: Identifying the model.

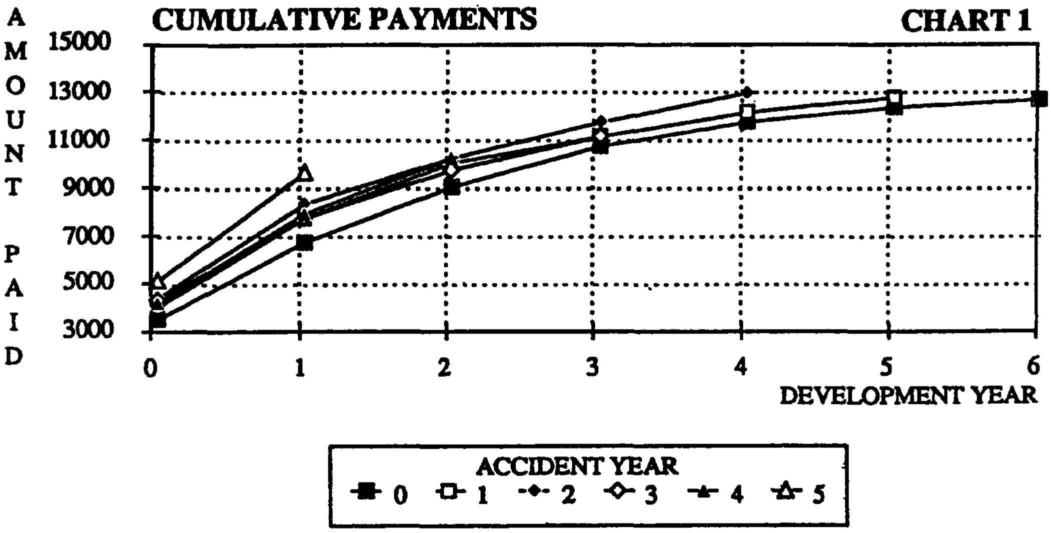
We will now consider a new data set and attempt to identify and fit an appropriate log-linear model to this data.

The first stage is a visual examination of the data. As a spreadsheet is being used it is very easy to plot the values and look at the resulting line charts rather than attempt to visualize these by looking at the data triangles.

The cumulative claims payments, which are from a UK Motor Non-Comprehensive account, are as follows:

	Development Year						
Acc Yr	0	1	2	3	4	5	6
0	3511	6726	8992	10704	11763	12350	12690
1	4001	7703	9981	11161	12117	12746	
2	4355	8287	10233	11755	12993		
3	4295	7750	9773	11093			
4	4150	7897	10217				
5	5102	9650					
6	6283						

The graph below shows these figures as line charts.



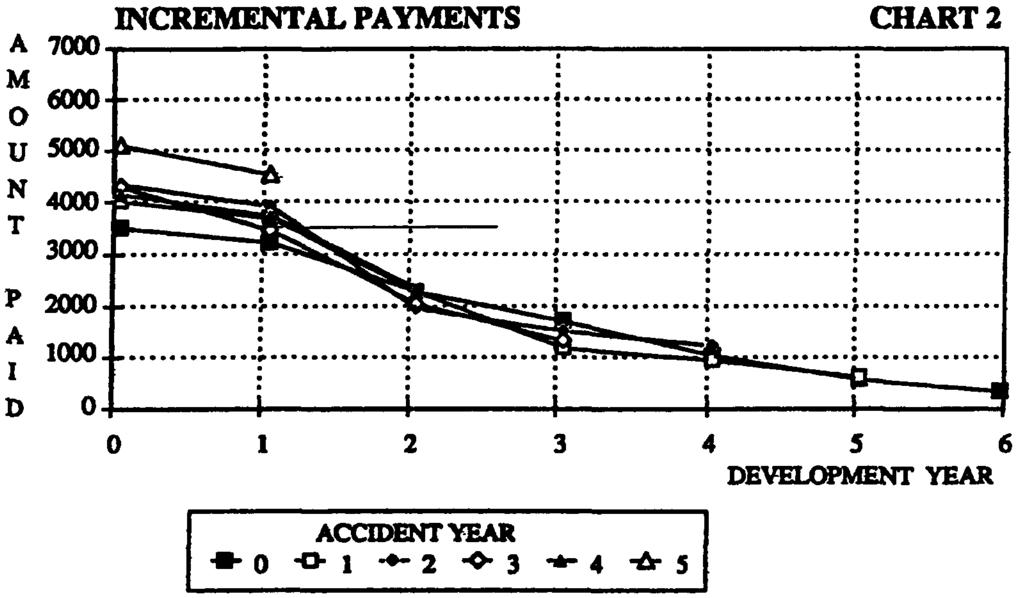
This is a useful presentation but it is hard to identify from this alone an appropriate model to use. Part of the problem arises from the fact that cumulative payments are clearly not independent. The incremental payments are expected to eventually decline but it is not easy to see any pattern or trend from this cumulative plot alone.

For these reasons the incremental data are now considered.

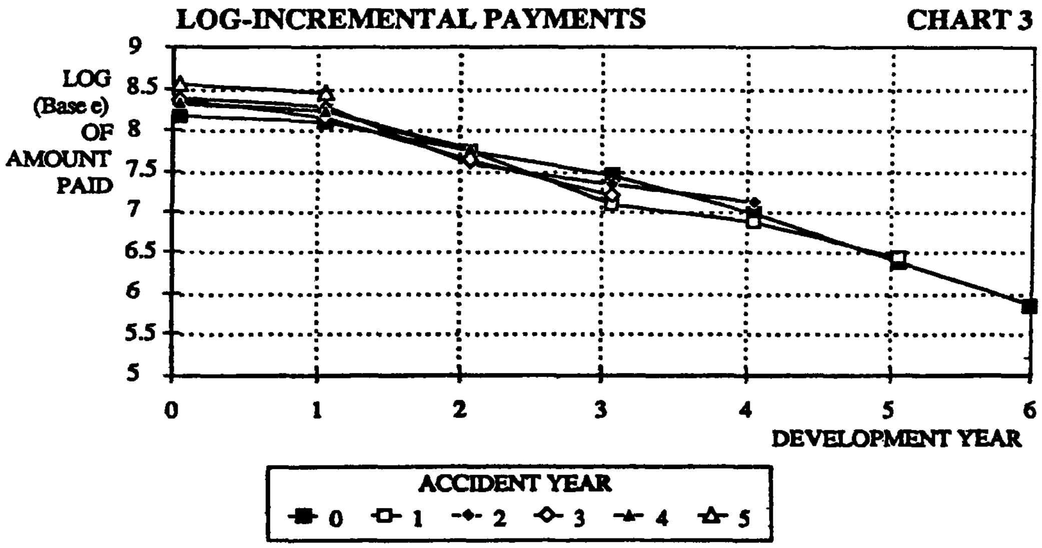
The incremental payments are

	Development Year						
Acc Yr	0	1	2	3	4	5	6
0	3511	3215	2266	1712	1059	587	340
1	4001	3702	2278	1180	956	629	
2	4355	3932	1946	1522	1238		
3	4295	3455	2023	1320			
4	4150	3747	2320				
5	5102	4548					
6	6283						

Even before these values are plotted a more promising trend can be detected across the development direction. Plotting these values we have:



Finally taking logarithms (base e) of these values and plotting as before produces the following line chart:



Looking at the accident year lines the first four or five look fairly bunched together and the last two (the last one is only a single point) appear to be at a higher level. From development year one the lines look reasonably straight and to have the same slope. These observations indicate that incremental payments from development year 1 on are decaying exponentially, as their logarithms appear to lie approximately on a straight line.

The first model to be fitted is based on these observations and will assume that each accident year has its own parameter or level. Development year zero will be assumed to have its own parameter and in line with the observation above the development parameters from d_1 on will be assumed to be linearly related or to lie along a straight line with some slope to be determined.

This is a start to the modelling process for this data set. The model is not expected to be the final or best for the data but is being used to illustrate various aspects of the modelling process. Note in particular that the plotted log-incremental data has been used to identify an appropriate model to start the process.

The techniques here can be applied in exactly the same way to more complex situations. As an example a different decay rate can be assumed for each accident year if the plot indicates that there is support for such a hypothesis. The model will then be very similar to the one described by Ajne in the second article of this volume. The only difference, apart from the decay rates, is that he fits the first two development periods before curve fitting whereas the example here curve fits from development one as this appears to be supported by the data.

The use of spreadsheets with their comprehensive graphics capabilities enables the modeller to carry out the initial stages of the data analysis phase very quickly as the above charts illustrate. Graphical presentation can also enhance reserving reports to management who may be less actuarially inclined than the writers of such reports.

H. Defining the model

The first model as identified above will now be defined more formally. There is a unique level for each accident year and a unique value for the zero development period. The parameters for development periods 1 to 6 are assumed to follow some linear relationship (straight line) with the same slope or parameter s.

Using the terminology developed earlier we have

$$Y_{ij} = a_i + d_j + e_{ij} \qquad \text{for } i, j \text{ from 0 to 6}$$

where $d_0 = d, \quad d_j = s \times j \quad \text{for } j > 0$

and e_{ij} is the error term assumed iid normal with zero mean.

Following the previous example, the spreadsheet table and design matrix are as shown below.

Table 1: Regression Table for the Full Parameter Model

				←				design matrix				→
i	j	P_{ij}	Y_{ij}	a_0	a_1	a_2	a_3	a_4	a_5	a_6	d	s
0	0	3511	8.164	1	0	0	0	0	0	0	1	0
0	1	3215	8.076	1	0	0	0	0	0	0	0	1
0	2	2266	7.726	1	0	0	0	0	0	0	0	2
0	3	1712	7.445	1	0	0	0	0	0	0	0	3
0	4	1059	6.965	1	0	0	0	0	0	0	0	4
0	5	587	6.375	1	0	0	0	0	0	0	0	5
0	6	340	5.829	1	0	0	0	0	0	0	0	6
1	0	4001	8.294	0	1	0	0	0	0	0	1	0
1	1	3702	8.217	0	1	0	0	0	0	0	0	1
1	2	2278	7.731	0	1	0	0	0	0	0	0	2
1	3	1180	7.073	0	1	0	0	0	0	0	0	3
1	4	956	6.863	0	1	0	0	0	0	0	0	4
1	5	629	6.444	0	1	0	0	0	0	0	0	5
2	0	4355	8.379	0	0	1	0	0	0	0	1	0
2	1	3932	8.227	0	0	1	0	0	0	0	0	1
2	2	1946	7.574	0	0	1	0	0	0	0	0	2
2	3	1522	7.328	0	0	1	0	0	0	0	0	3
2	4	1238	7.121	0	0	1	0	0	0	0	0	4
3	0	4295	8.365	0	0	0	1	0	0	0	1	0
3	1	3455	8.148	0	0	0	1	0	0	0	0	1
3	2	2023	7.612	0	0	0	1	0	0	0	0	2
3	3	1320	7.185	0	0	0	1	0	0	0	0	3
4	0	4150	8.331	0	0	0	0	1	0	0	1	0
4	1	3747	8.229	0	0	0	0	1	0	0	0	1
4	2	2320	7.749	0	0	0	0	1	0	0	0	2
5	0	5102	8.537	0	0	0	0	0	1	0	1	0
5	1	4548	8.422	0	0	0	0	0	1	0	0	1
6	0	6238	8.746	0	0	0	0	0	0	1	1	0

The regression output for this model is given below. For ease of reference two extra lines have been inserted in this output. Firstly the parameter labels are shown above the parameter coefficient estimates and secondly the T-Ratios are shown.

```
Regression output:
     Constant                          0
     Std Err of Y Est                  .1139
     R squared(Adj,Raw)      .9762     .9832
     No. of Observations               28
     Degrees of Freedom                19
```

	a_0	a_1	a_2	a_3	a_4	a_5	a_6	d	s
Coefficient(s)	8.573	8.574	8.665	8.554	8.637	8.846	9.042	-.296	-.435
Std Err of Coef.	.076	.072	.069	.070	.076	.091	.134	.070	.018
T-ratios	113.3	119.9	124.9	121.8	113.8	97.6	67.6	-4.2	-23.5

The development parameters, d and s are significantly different from zero as their T-Ratios (parameter estimate divided by its standard error estimate) are −4.2 and −23.5 respectively which are well outside the usual 95% confidence interval (critical) range of −2 to 2.

The accident year parameters are also all significantly different from zero, as they surely have to be with this model's assumptions (all accident year levels are significantly above zero), but they do look close to one another. In order to test whether these are distinct it is necessary to redefine the model by dropping the a_0 parameter and replacing it with a constant. The only change to the design matrix is that the first column is now made up of ones.

The regression output of the redefined model is almost identical:

Regression Output:		
Constant		0
Std Err of Y Est		.1139
R Squared (Adj,Raw)	.9762	.9832
No. Of Observations		28
Degrees of Freedom		19

	k	a_1	a_2	a_3	a_4	a_5	a_6	d	s
Coefficient(s)	8.573	.001	.092	-.019	.064	.273	.469	-.296	-.435
Std Err of Coef.	.076	.064	.069	.075	.084	.098	.132	.070	.018
T-ratios	113.3	.0	1.3	-.2	.8	2.8	3.6	-4.2	-23.5

The output clearly shows a much better definition of the same model as it identifies that the accident years 1,2,3 and 4 parameters are not significantly different from zero or, in comparison to the previous definition, significantly different from the zero'th accident year parameter which has now become the constant level value k. Based on this definition the model parameters for accident years 0,1,2,3 and 4 can be set to zero and be effectively estimated by a new common value k. This new constant of the reduced parameter model should now be an average value for the five accident years whose individual parameters have been dropped from the model.

A theoretically more appealing approach for inducing a partition in the accident year parameters, based on the multicomparison t-criterion test, can be found in Renshaw (2).

Setting a_0 to a_4 to zero reduces the model parameters to just the five parameters k, a_5, a_6, d and s which we expect to be significant.

The design matrix is now simpler as can be seen from Table 2 below.

Table 2: Regression Table for the Reduced Parameter Model

				←	design matrix			→
i	j	P_{ij}	Y_{ij}	k	a_5	a_6	d	s
0	0	3511	8.164	1	0	0	1	0
0	1	3215	8.076	1	0	0	0	1
0	2	2266	7.726	1	0	0	0	2
0	3	1712	7.445	1	0	0	0	3
0	4	1059	6.965	1	0	0	0	4
0	5	587	6.375	1	0	0	0	5
0	6	340	5.829	1	0	0	0	6
1	0	4001	8.294	1	0	0	1	0
1	1	3702	8.217	1	0	0	0	1
1	2	2278	7.731	1	0	0	0	2
1	3	1180	7.073	1	0	0	0	3
1	4	956	6.863	1	0	0	0	4
1	5	629	6.444	1	0	0	0	5
2	0	4355	8.379	1	0	0	1	0
2	1	3932	8.277	1	0	0	0	1
2	2	1946	7.574	1	0	0	0	2
2	3	1522	7.328	1	0	0	0	3
2	4	1238	7.121	1	0	0	0	4
3	0	4295	8.365	1	0	0	1	0
3	1	3455	8.148	1	0	0	0	1
3	2	2023	7.612	1	0	0	0	2
3	3	1320	7.185	1	0	0	0	3
4	0	4150	8.331	1	0	0	1	0
4	1	3747	8.229	1	0	0	0	1
4	2	2320	7.749	1	0	0	0	2
5	0	5102	8.537	1	1	0	1	0
5	1	4548	8.422	1	1	0	0	1
6	0	6283	8.746	1	0	1	1	0

The regression output for this reduced parameter model is

Regression Output:

Constant		0
Std Err of Y Est		.1119
R Squared(Ajd,Raw)	.9770	.9804
No. of Observations		28
Degrees of Freedom		23

	k	a_5	a_6	d	s
Coefficient(s)	8.608	.244	.441	-.303	-.440
Std Err of Coef.	.052	.085	.122	.068	.017
T-ratio	167.1	2.9	3.6	-4.5	-26.4

As expected the constant has now changed as it is an average value for the first five accident years. The other parameters have also changed slightly.

All the parameters are now significantly different from zero, with t-ratios exceeding absolute 2, as expected. The quality of fit is still good and the number of parameters has been reduced from nine to five. The model looks reasonable enough to warrant further investigation.

The next section considers some basic testing using residual analysis plots of the first (all parameter) model and this reduced parameter model.

Projections from both these models will be calculated and compared after this analysis.

I. Testing the models by residual analysis plots

The parameter estimates from the regressions can now be used to calculate the model estimates, in log-space, which can then be compared with the observed values in log-space. It is usual to use standardized residuals, defined as the difference between observed and fitted values divided by the model standard error, and considering these in graphical form. Under the IID assumptions used to derive the model estimates these residuals should exhibit a fair degree of randomness.

Testing now turns to the analysis of these standardized residuals. In practice these are plotted against development, accident and payment year and also against the fitted values. Working in a spreadsheet makes this process very easy as each chart can be defined as an X-Y chart with Y the standardized residuals and X the other variable in turn.

Table 3 below shows the actual values, their logarithms and the model fitted values in log-space for the full parameter model as defined in Table 1. The residuals are just the differences between the observed and fitted values in log-space and the standardized residuals are the residuals divided by the model standard error, which was .1139 for this model.

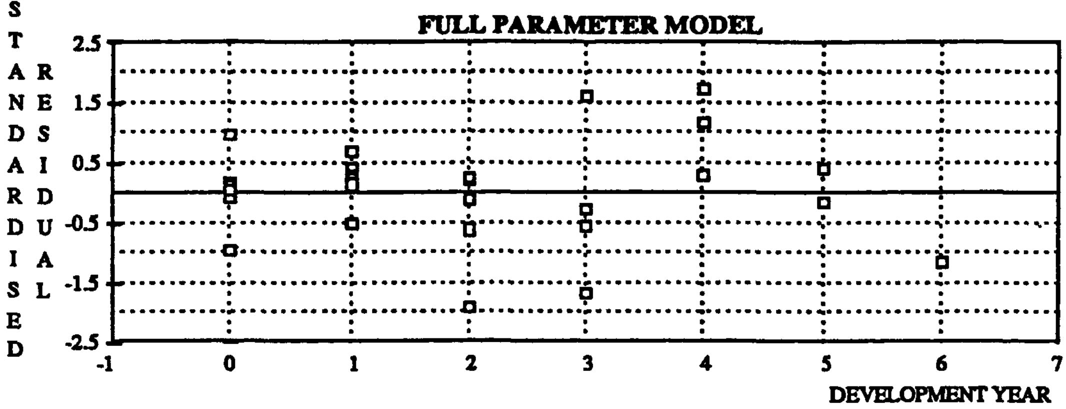
Table 3: Residuals Table for the Full Parameter Model.

Acc	Dev	Pay					Stand
i	j	i+j	P_{ij}	Y_{ij}	$\hat{Y}_{ij}$	Resid	Resid
0	0	0	3511	8.164	8.277	-.113	-.991
0	1	1	3215	8.076	8.138	-.062	-.547
0	2	2	2266	7.726	7.703	.023	.201
0	3	3	1712	7.445	7.268	.177	1.557
0	4	4	1059	6.965	6.833	.132	1.159
0	5	5	587	6.375	6.398	-.023	-.202
0	6	6	340	5.829	5.963	-.134	-1.177
1	0	1	4001	8.294	8.278	.017	.147
1	1	2	3702	8.217	8.139	.078	.683
1	2	3	2278	7.731	7.704	.027	.239
1	3	4	1180	7.073	7.269	-.196	-1.717
1	4	5	956	6.863	6.834	.029	.253
1	5	6	629	6.444	6.399	.045	.396
2	0	2	4355	8.379	8.369	.010	.091
2	1	3	3932	8.277	8.230	.047	.412
2	2	4	1946	7.574	7.795	-.221	-1.943
2	3	5	1522	7.328	7.360	-.032	-.283
2	4	6	1238	7.121	6.925	.196	1.722
3	0	3	4295	8.365	8.258	.107	.942
3	1	4	3455	8.148	8.119	.028	.249
3	2	5	2023	7.612	7.684	-.072	-.631
3	3	6	1320	7.185	7.249	-.064	-.560
4	0	4	4150	8.331	8.340	-.010	-.084
4	1	5	3747	8.229	8.202	.027	.237
4	2	6	2320	7.749	7.767	-.017	-.153
5	0	5	5102	8.537	8.549	-.012	-.104
5	1	6	4548	8.422	8.411	.012	.104
6	0	6	6283	8.746	8.746	.000	.000

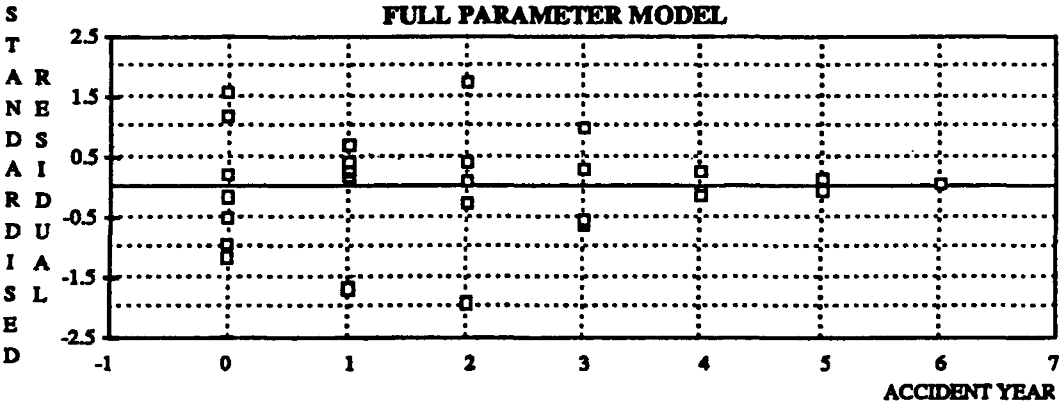
To produce the residual plots in X-Y chart form the standardized residuals column is defined as the Y-variate and the first three columns in turn as the X-variate for the accident year, development year and payment year plots. For the final plot the fitted values column is picked instead.

The various residual plots from this model are shown below in Charts 4 to 7.

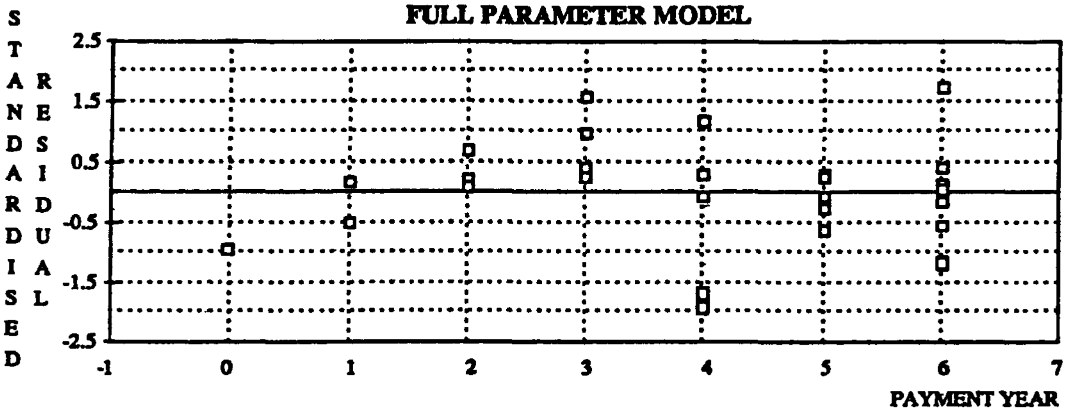
RESIDUAL ANALYSIS **CHART 4**

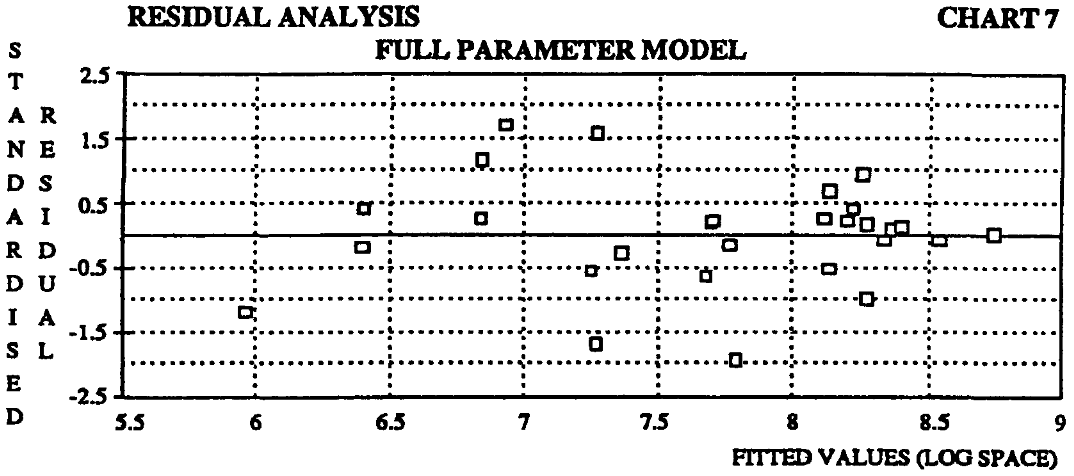


RESIDUAL ANALYSIS **CHART 5**



RESIDUAL ANALYSIS **CHART 6**



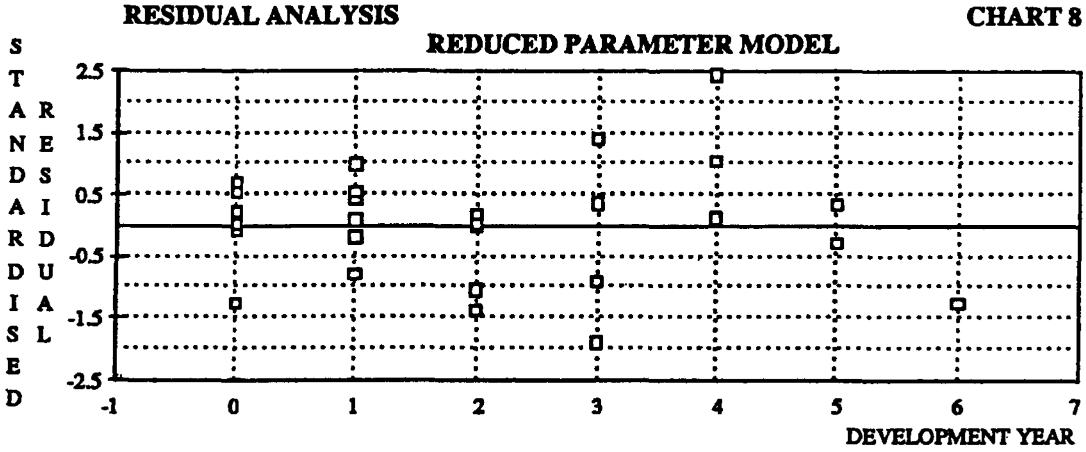


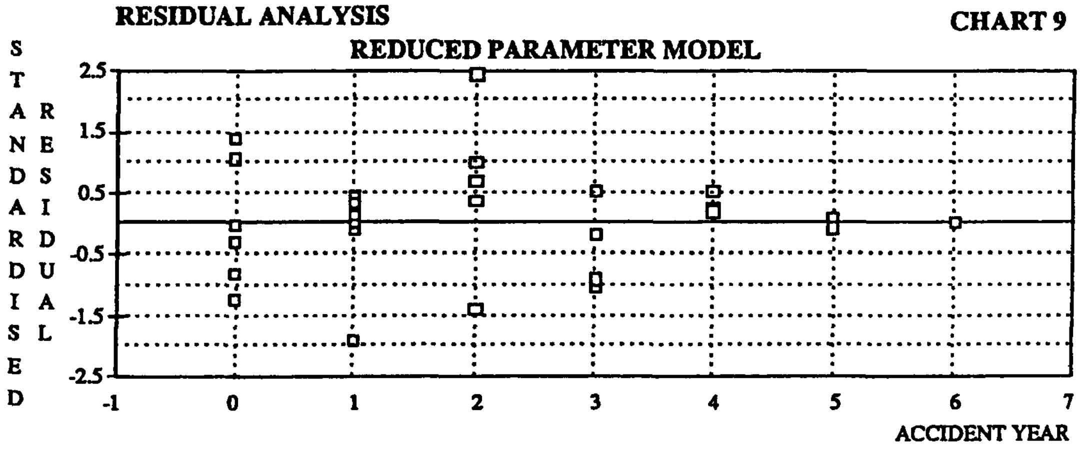
The residuals for the Reduced Parameter Model, which is defined in Table 2 (common level value for the first five accident years), are shown in Table 4 below.

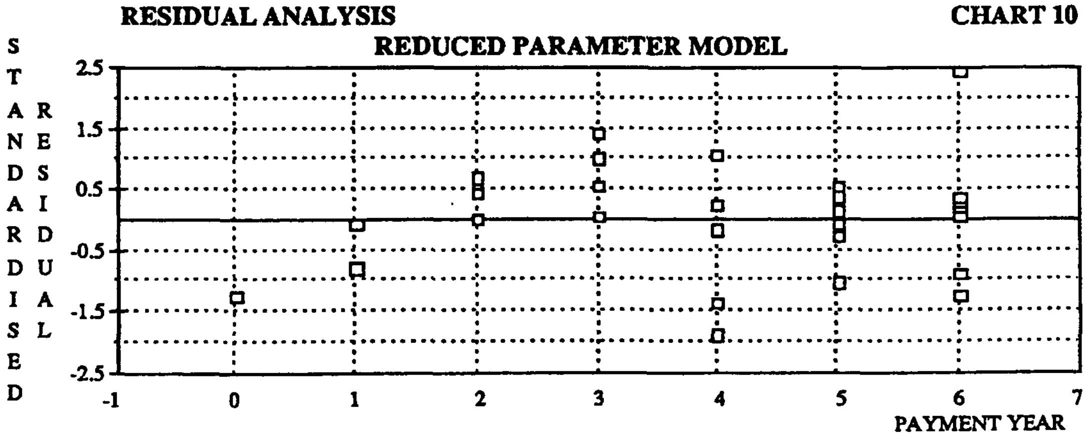
Table 4: Residuals Table for the Reduced Parameter Model.

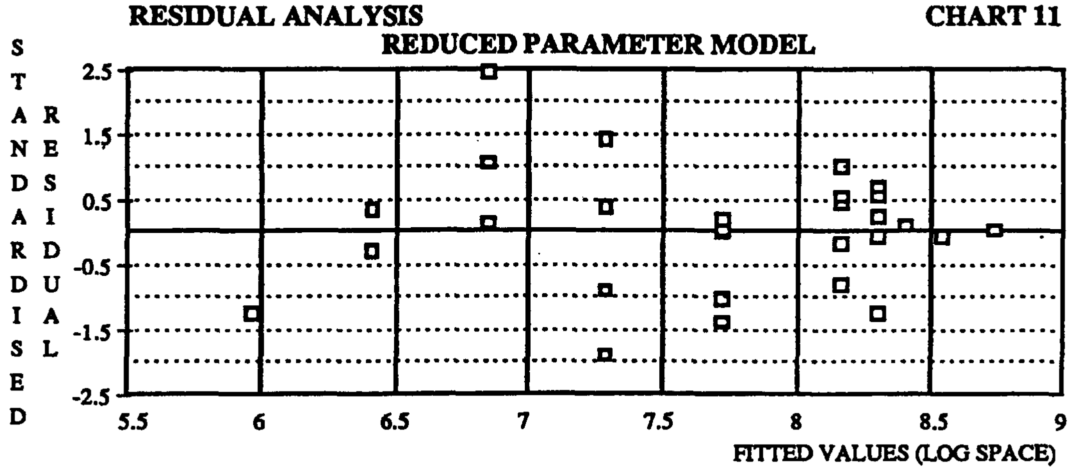
Acc i	Dev j	Pay i+j	P_{ij}	Y_{ij}	$\hat{Y}_{ij}$	Resid	Stand Resid
0	0	0	3511	8.164	8.304	-.141	-1.259
0	1	1	3215	8.076	8.168	-.093	-.828
0	2	2	2266	7.726	7.729	-.003	-.025
0	3	3	1712	7.445	7.289	.156	1.398
0	4	4	1059	6.965	6.849	.116	1.035
0	5	5	587	6.375	6.410	-.035	-.309
0	6	6	340	5.829	5.970	-.141	-1.260
1	0	1	4001	8.294	8.304	-.010	-.091
1	1	2	3702	8.217	8.168	.048	.432
1	2	3	2278	7.731	7.729	.002	.022
1	3	4	1180	7.073	7.289	-.216	-1.927
1	4	5	956	6.863	6.849	.013	.121
1	5	6	629	6.444	6.410	.035	.309
2	0	2	4355	8.379	8.304	.075	.667
2	1	3	3932	8.277	8.168	.109	.971
2	2	4	1946	7.574	7.729	-.155	-1.386
2	3	5	1522	7.328	7.289	.039	.347
2	4	6	1238	7.121	6.849	.272	2.431
3	0	3	4295	8.365	8.304	.061	.543
3	1	4	3455	8.148	8.168	-.021	-.185
3	2	5	2023	7.612	7.729	-.116	-1.039
3	3	6	1320	7.185	7.289	-.104	-.925
4	0	4	4150	8.331	8.304	.026	.236
4	1	5	3747	8.229	8.168	.060	.540
4	2	6	2320	7.749	7.729	.021	.185
5	0	5	5102	8.537	8.548	-.011	-.095
5	1	6	4548	8.422	8.412	.011	.095
6	0	6	6283	8.746	8.746	.000	.000

The various residual plots from this model are shown below in Charts 8 to 11.









This reduced parameter model has a standardized residual for accident year 2, development period 4, of 2.431 as the maximum (absolute) standardized residual value. The full parameter model had a lowest standardized residual of -1.943 (i=2, j=2). The second model has a slightly smaller standard error of .1119 compared to the .1139 of the full parameter model. There is however little difference overall between these models detectable from the above tables. Both seem to fit the data fairly well.

The next stage is to consider these residuals in graphic form to examine whether any unmodelled trends are detectable.

In all these residual plots, according to the model error assumptions, we expect a set of fairly random points bounded in about 95% of cases within the -2 to 2 range.

As Table 3 and Table 4 above indicate, all the standardized residuals for the full parameter model are just in this range (Table 3) with just one value outside the range in the case of the reduced parameter model (Table 4). Values outside this range will sometimes occur and often identify outliers that may warrant further investigation.

The development year plots (Charts 4 and 8) will generally be the most interesting and particularly where, as in these cases, it has been assumed that there is some relationship connecting the development parameters. A particular feature worth looking out for in these plots is any tendency for the residuals to spread or fan out with development. This is not too noticeable in these examples. Note however that the residuals for development periods 4 to 6 in both cases do not appear very random. There are however only a few values involved and these may well be impacted by the outlier identified earlier (i=2, j=4). We have used a very simple shape to describe the run-off from development period 1 and these residual plots are quite reasonable in the circumstances.

The accident year residual plots are shown in Chart 5 for the full parameter model and in Chart 9 for the reduced model. Considering the former first, as each accident year has its own parameter in this model, the plot should be boringly predictable with the residuals balanced about the zero horizontal. Chart 5 shows this quite clearly.

The reduced model accident year residuals, Chart 9, look very similar although here the first five accident years have effectively been fitted by a single parameter. The only visible differences are the accident 4 residuals which are all greater than zero. In a fuller analysis this parameter should be added back to the reduced model and tested for significance. It is possible that it may become more significant if measured against the average for accident years 0 to 3 although this turns out not to be the case in this instance.

In both cases the accident year residuals appear to get closer to the zero horizontal line, with increasing accident year, resulting in the left half of both charts diverging from this line. This is due, at least in part, to over-parameterisation. In the extreme right, for example, as only one point is fitted and with its own parameter a perfect fit is obtained and the residual has to be zero. For accident year 5 two points are fitted and so the accident year parameter is again effective in ensuring a close fit. The values in these late accident years are also relatively large, as they are from earlier development periods when payments tend to be higher, and they may be relatively more stable. This is considered later.

The payment year residuals (Charts 6 and 10) can be interesting but more difficult to interpret. Inflationary forces are expected to operate along this direction but as accident year levels have been assumed independent this may mask any such influences. The plots for both models look very similar, which is not very surprising, as neither model considers this direction in its definition. Both these charts appear to show a definite non-random shape for the early payment years and this would warrant further investigation. Changes in claims inflation rates during the period concerned, which are not incorporated in the model, may well be the cause. This is not pursued here. The regression analysis at least identifies areas that would warrant further investigation in practice.

It was indicated earlier that higher values, generally in earlier development periods, may be relatively more stable than later, generally lower, values. This can be tested by plotting residuals against fitted values as is shown in Charts 7 and 11. In both these charts the last few residuals on the extreme right look close to the horizontal zero line but these points are the same points identified earlier as the last two or three accident year values. The residuals show a tendency to increase (in absolute terms) as values decrease. This effect, generally known as heteroscedasticity, is also detectable from the development year plots as incremental payments eventually decrease with development. No attempt is made here to overcome any heteroscedasticity.

The error term normality assumption can also be tested more formally within the spreadsheet if required. It is possible for instance to use the Data Distribution command to calculate and tabulate a frequency distribution of the residuals and compare values in this table with preset values calculated from the standard normal distribution.

The residual analysis indicates that these models have some weakness along the payment year direction and there are sufficient reasons to doubt some of the model assumptions. A full analysis would follow these up. In particular some inflation adjustment should be made to the data and the modelling process repeated to see whether this adjustment removes the non-random look of these residuals along the payment year direction. However for the time being it will be assumed that both these models are satisfactory and the regression results will be used in the next section to project the future payments and their standard errors from these two models.

A later section will consider a model with inflation and claim volume adjustment to see if a better model can be found.

J. Using the models to project future payments and standard errors

When the basic chain ladder model with independent development parameters is fitted it is not possible to extend the projections beyond the latest development contained in the triangle without resorting to some form of external curve fitting of development factors such as the Sherman inverse power curve for example.

In these examples as a curve (straight line) has been fitted to the development parameters it is possible to extend the model projections to development periods beyond those contained in the data triangle.

The model has a natural stop as the payments are decaying exponentially and so become small relatively quickly. So we could simply sum the implied geometric series or take the values to some development period beyond which we would expect no more payments in practice.

In what follows it is assumed that there are no payments beyond development 12, as this is sufficient for purposes of illustration and cuts down the values to be projected. In practice this will need to be decided on the merits of each case and knowledge of the likely run-off period of the particular class being investigated.

The data triangle contained 28 values and our completed rectangle has a total of 91 data points (7×13). There are therefore 63 individual payments and their standard errors to calculate.

The design and future design matrices are first produced and these are used to produce the variance-covariance matrix of the future values. This is now a 63 × 63 matrix and should be within the capability of a reasonable spreadsheet. Both Lotus 123 Version 2.2 and SuperCalc5 Version 5.0 can handle square matrices of around 89 × 89.

For producing the future values and their associated (individual) standard errors only the diagonal elements of this matrix are needed. The calculations from here are fairly simple and are shown in the Tables 5 and 7 below. These tables are set in the way one would normally produce them in a spreadsheet. The values are arranged by accident year first, as this is how the future design matrix was set out. The accident year order was adopted here as this order facilitates the computation of the accident year standard errors.

The second table in each set (Tables 6 and 8) show the projected values and standard errors in a more traditional format and also include accident year and overall totals for both values and standard errors. These calculations are also set out in the spreadsheet as explained in Section F. In view of the size of the matrices involved they have not been shown here.

The various matrix products needed to calculate the variance-covariance matrix (as set out in Appendix 1 for the earlier chain ladder example) took under two minutes on a 12MHz PC fitted with a maths co-processor.

Table 5: Projection for the Full Parameter Model: Part a

i	j	$\hat{Y}_{ij}$	$\text{var}\hat{Y}_{ij}$	$\hat{P}_{ij}$	se $\hat{P}_{ij}$	% error
0	7	5.528	.0195	254	36	14.0%
0	8	5.093	.0223	165	25	15.0%
0	9	4.658	.0258	107	17	16.2%
0	10	4.223	.0301	69	12	17.5%
0	11	3.788	.0350	45	8	18.9%
0	12	3.353	.0405	29	6	20.3%
1	6	5.964	.0185	393	54	13.7%
1	7	5.529	.0210	255	37	14.6%
1	8	5.094	.0241	165	26	15.6%
1	9	4.659	.0280	107	18	16.8%
1	10	4.224	.0325	69	13	18.2%
1	11	3.789	.0377	45	9	19.6%
1	12	3.354	.0436	29	6	21.1%
2	5	6.490	.0179	664	89	13.4%
2	6	6.055	.0199	431	61	14.2%
2	7	5.620	.0227	279	42	15.2%
2	8	5.185	.0261	181	29	16.3%
2	9	4.750	.0302	117	21	17.5%
2	10	4.315	.0350	76	14	18.9%
2	11	3.880	.0405	49	10	20.3%
2	12	3.445	.0467	32	7	21.9%
3	4	6.814	.0177	919	123	13.3%
3	5	6.379	.0193	595	83	14.0%
3	6	5.944	.0216	386	57	14.8%
3	7	5.509	.0246	250	39	15.8%
3	8	5.074	.0283	162	27	17.0%
3	9	4.639	.0327	105	19	18.2%
3	10	4.205	.0378	68	13	19.6%
3	11	3.770	.0435	44	9	21.1%
3	12	3.335	.0499	29	7	22.6%
4	3	7.332	.0181	1542	209	13.5%
4	4	6.897	.0193	999	139	14.0%
4	5	6.462	.0212	647	95	14.6%
4	6	6.027	.0237	419	65	15.5%
4	7	5.592	.0269	272	45	16.5%
4	8	5.157	.0308	176	31	17.7%
4	9	4.722	.0354	114	22	19.0%
4	10	4.287	.0407	74	15	20.4%
4	11	3.852	.0466	48	11	21.8%
4	12	3.417	.0532	31	7	23.4%
5	2	7.976	.0202	2939	420	14.3%
5	3	7.541	.0208	1903	276	14.5%
5	4	7.106	.0220	1232	184	14.9%
5	5	6.671	.0240	798	124	15.6%
5	6	6.236	.0266	518	85	16.4%
5	7	5.801	.0298	336	58	17.4%
5	8	5.366	.0338	218	40	18.5%
5	9	4.931	.0385	141	28	19.8%
5	10	4.496	.0438	92	19	21.2%
5	11	4.061	.0498	59	13	22.6%
5	12	3.626	.0565	39	9	24.1%

Table 5: Projection for the Full Parameter Model: Part b

i	j	$\hat{Y}_{ij}$	var $\hat{Y}_{ij}$	$\hat{P}_{ij}$	se $\hat{P}_{ij}$	% error
6	1	8.607	.0296	5550	962	17.3%
6	2	8.172	.0290	3592	616	17.1%
6	3	7.737	.0290	2325	399	17.2%
6	4	7.302	.0298	1506	262	17.4%
6	5	6.867	.0313	975	174	17.8%
6	6	6.432	.0334	632	116	18.4%
6	7	5.997	.0362	410	79	19.2%
6	8	5.562	.0397	266	53	20.1%
6	9	5.127	.0439	172	36	21.2%
6	10	4.692	.0487	112	25	22.4%
6	11	4.257	.0543	73	17	23.6%
6	12	3.822	.0605	47	12	25.0%

TOTAL = 34377

Table 6: Projected values and Standard Errors.

Full Parameter Model.

		Development Year												
Yr		1	2	3	4	5	6	7	8	9	10	11	12	Total
0	£							254	165	107	69	45	29	669
	se							36	25	17	12	8	6	79
1	£						393	255	165	107	69	45	29	1063
	se						54	37	26	18	13	9	6	119
2	£					664	431	279	181	117	76	49	32	1830
	se					89	61	42	29	21	14	10	7	196
3	£				919	595	386	250	162	105	68	44	29	2559
	se				123	83	57	39	27	19	13	9	7	265
4	£			1542	999	647	419	272	176	114	74	48	31	4324
	se			209	139	95	65	45	31	22	15	11	7	443
5	£		2939	1903	1232	798	518	336	218	141	92	59	39	8274
	se		420	276	184	124	85	58	40	28	19	13	9	890
6	£	5550	3592	2325	1506	975	632	410	266	172	112	73	47	15659
	se	962	616	399	262	174	116	79	53	36	25	17	12	2158

Overall Total	34377
Standard Error	2742
Percent. Error	7.98

Table 7: Projection for the Reduced Parameter Model: Part a

i	j	$\hat{Y}_{ij}$	var$\hat{Y}_{ij}$	$\hat{P}_{ij}$	se $\hat{P}_{ij}$	% error
0	7	5.530	.0182	255	35	13.6%
0	8	5.091	.0209	164	24	14.5%
0	9	4.651	.0241	106	17	15.6%
0	10	4.211	.0279	68	11	16.8%
0	11	3.772	.0322	44	8	18.1%
0	12	3.332	.0371	29	6	19.4%
1	6	5.970	.0161	395	50	12.8%
1	7	5.530	.0182	255	35	13.6%
1	8	5.091	.0209	164	24	14.5%
1	9	4.651	.0241	106	17	15.6%
1	10	4.211	.0279	68	11	16.8%
1	11	3.772	.0322	44	8	18.1%
1	12	3.332	.0371	29	6	19.4%
2	5	6.410	.0146	612	74	12.1%
2	6	5.970	.0161	395	50	12.8%
2	7	5.530	.0182	255	35	13.6%
2	8	5.091	.0209	164	24	14.5%
2	9	4.651	.0241	106	17	15.6%
2	10	4.211	.0279	68	11	16.8%
2	11	3.772	.0322	44	8	18.1%
2	12	3.332	.0371	29	6	19.4%
3	4	6.849	.0136	950	111	11.7%
3	5	6.410	.0146	612	74	12.1%
3	6	5.970	.0161	395	50	12.8%
3	7	5.530	.0182	255	35	13.6%
3	8	5.091	.0209	164	24	14.5%
3	9	4.651	.0241	106	17	15.6%
3	10	4.211	.0279	68	11	16.8%
3	11	3.772	.0322	44	8	18.1%
3	12	3.332	.0371	29	6	19.4%
4	3	7.289	.0132	1474	170	11.5%
4	4	6.849	.0136	950	111	11.7%
4	5	6.410	.0146	612	74	12.1%
4	6	6.970	.0161	395	50	12.8%
4	7	5.530	.0182	255	35	13.6%
4	8	5.091	.0209	164	24	14.5%
4	9	4.651	.0241	106	17	15.6%
4	10	4.211	.0279	68	11	16.8%
4	11	3.772	.0322	44	8	18.1%
4	12	3.332	.0371	29	6	19.4%
5	2	7.972	.0195	2927	411	14.0%
5	3	7.532	.0199	1886	267	14.2%
5	4	7.093	.0208	1216	176	14.5%
5	5	6.653	.0224	784	118	15.0%
5	6	6.213	.0244	506	79	15.7%
5	7	5.774	.0270	326	54	16.6%
5	8	5.334	.0302	210	37	17.5%
5	9	4.894	.0340	136	25	18.6%
5	10	4.455	.0382	88	17	19.7%
5	11	4.015	.0431	57	12	21.0%
5	12	3.575	.0485	37	8	22.3%

Table 7: Projection for the Reduced Parameter Model: Part b

i	j	$\hat{Y}_{ij}$	var $\hat{Y}_{ij}$	$\hat{P}_{ij}$	se $\hat{P}_{ij}$	% error
6	1	8.609	.0285	5562	946	17.0%
6	2	8.170	.0279	3582	603	16.8%
6	3	7.730	.0279	2308	388	16.8%
6	4	7.290	.0284	1487	252	17.0%
6	5	6.851	.0295	959	166	17.3%
6	6	6.411	.0311	618	110	17.8%
6	7	5.971	.0333	399	73	18.4%
6	8	5.532	.0361	257	49	19.2%
6	9	5.092	.0394	166	33	20.0%
6	10	4.652	.0432	107	23	21.0%
6	11	4.213	.0476	69	15	22.1%
6	12	3.773	.0526	45	10	23.2%

TOTAL = 33847

Table 8: Projected values and Standard Errors

Reduced Parameter Model.

		Development Year												
Yr		1	2	3	4	5	6	7	8	9	10	11	12	Total
0	£							255	164	106	68	44	29	666
	se							35	24	17	11	8	6	75
1	£						395	255	164	106	68	44	29	1060
	se						50	35	24	17	11	8	6	106
2	£					612	395	255	164	106	68	44	29	1672
	se					74	50	35	24	17	11	8	6	146
3	£				950	612	395	255	164	106	68	44	29	2622
	se				111	74	50	35	24	17	11	8	6	200
4	£			1474	950	612	395	255	164	106	68	44	29	4096
	se			170	111	74	50	35	24	17	11	8	6	275
5	£		2927	1886	1216	784	506	326	210	136	88	57	37	8173
	se		411	267	176	118	79	54	37	25	17	12	8	851
6	£	5562	3582	2308	1487	959	618	399	257	166	107	69	45	15558
	se	946	603	388	252	166	110	73	49	33	23	15	10	2101

Overall Total	33847
Standard Error	2545
Percent. Error	7.52

K. Overall standard error and accident year standard errors

The calculations necessary to produce the accident year and overall standard errors shown in Tables 6 and 8 above are a repeat of those shown in Section G. The only complication is that in the above cases there are more values to project (63 rather than 6) so there is a lot more to calculate.

The results are very close. The full model produces estimated future payments of 34377 with a standard error of 2742 or 7.98%. The reduced parameter model produces estimated future payments of 33847 with a standard error of 2545 or 7.52%. The two estimated values are not significantly different but the second model has a proportionately smaller standard error. This is purely due to the smaller number of parameters used in defining this model. The second model may therefore be considered to have the slight advantage over the first.

The closeness of these results is not particularly surprising as the two models are very similar. Most of the future payments relate to the last two accident years and here both models have assumed these years to have independent levels (just like the chain ladder model) and so any smoothing from the reduced parameter model affects only the earliest accident years where the projected future payment values are not so large.

In fact assumptions about the most recent accident years are crucial to any reserve analysis. The base data used in this example is unadjusted for inflation and claim volume and the levels for the various accident years are not normally expected to be as close as those of the first five accident years above.

The next section will consider modelling the inflation and volume adjusted data.

L. Adjusting for inflation and claim volumes

It is possible to reduce the model parameters further by using an inflation index to bring all payments to current value and a claims volume adjustment or weight for each accident year so as to normalize these payments.

The claim volume values to be used in this example are based on the number of claims reported by the end of the first development period. They are scaled for convenience.

Accident Year	0	1	2	3	4	5	6
Claim Volume	1.43	1.45	1.52	1.35	1.29	1.47	1.91

An earnings index for the relavant period will be used in this case to bring payment values to payment year 6 (the latest payment year) values. In practice case is needed to ensure that the index used is the most appropriate index for the class of claims under investigation.

Payment year	0	1	2	3	4	5	6
Index	1.55	1.41	1.30	1.23	1.13	1.05	1

The inflation adjusted, volume normalized incremental payments (shown in integer format but calculated and used to many decimal places) are now as follows:

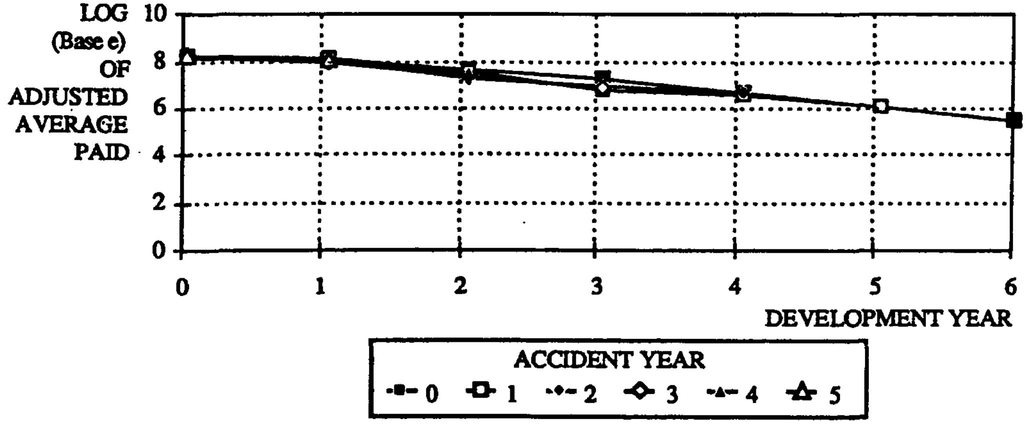
	Development Year						
Acc Yr	0	1	2	3	4	5	6
0	3806	3170	2060	1473	837	431	238
1	3891	3319	1932	920	692	434	
2	3725	3182	1447	1051	814		
3	3913	2892	1573	978			
4	3635	3050	1798				
5	3644	3094					
6	3290						

Even before any further analysis is carried out it is clear from this triangle that there is a fair amount of consistency and stability in the adjusted data.

Plotting the log-incremental adjusted data, as can be seen from Chart 12 below, appears to confirm this observation. The various lines, each representing an accident year, look closely grouped together for at least the first couple of development periods.

LOG-INCREMENTAL PAYMENTS **CHART 12**

DATA ADJUSTED FOR INFLATION AND CLAIM VOLUME



The chart indicates that accident year effects may have been reduced or eliminated and the first test will be to confirm whether this is the case. As the shape of these lines is as before the same assumptions will be made in modelling the shape.

The design matrix is initially exactly as in the previous example which assumed accident years 1 to 6 as independent variates and had an independent first development level (d) and then a linear trend with common slope s.

The regression output using the adjusted values and including the extra two lines as before is:

Regression Output: Full Parameter Model.

Constant		0
Std Err of Y Est		.1153
R Squared(Adj,Raw)	0.9788	.9851
No. of Observations		28
Degrees of Freedom		19

	k	a_1	a_2	a_3	a_4	a_5	a_6	d	s
Coefficient(s)	8.627	-.087	-.114	-.175	-.120	-.110	-.237	-.292	-.505
Std Err of Coef.	.077	.065	.070	.076	.085	.099	.133	.071	.019
T-Ratios	112.6	-1.3	-1.6	-2.3	-1.4	-1.1	-1.8	-4.1	-27.0

Accident year 3 turns out to be the only one whose parameter has a T-ratio whose absolute value exceeds 2 and may be considered significant.

So the next stage is to eliminate all the accident years with T-Ratios less than absolute 2 and refit. There are now four parameters namely

k a_3 d and s

The regression output of this model is:

Regression Output:

Constant		0
Std Err of Y Est		.1157
R Squared(Adj,Raw)	.9787	.9810
No. Of Observations	28	
Degrees of Freedom	24	

	k	a_3	d	s
Coefficient(s)	8.523	-.088	-.296	-.493
Std Err of Coef.	.054	.063	.068	.017
T-Ratios	157.2	-1.4	-4.3	-28.7

The parameters of this model can still be reduced as the accident year three parameter is now not significant. What has happened is that it is now being measured against the "average" of all the other accident year levels rather than just the first accident year level and this has been sufficient to make this last accident year parameter close enough to the average value. Care needs to be taken to ensure that none of the other parameters have become significant in the new model.

So this remaining accident year parameter will be dropped, leaving only three parameters, one for the common level k, and the two shape parameters d and s.

The regression output of this three parameter model is:

```
Regression Output:
     Constant                              0
     Std Err of Y Est                      .1179
     R Squared(Adj,Raw)      .9779         .9795
     No. Of Observations                   28
     Degrees of Freedom                    25

                          k            d            s

Coefficient(s)          8.501        -.286        -.489
Std Err of Coef.         .053         .069         .017
T-Ratios                161.3         -4.1        -28.3
```

This is an interesting stage. There are now only three parameters and all are significant. The model has a high R-squared value and appears to describe the data reasonably well. It is now tempting to use this model to project future payments.

The process is as before with the minor irritation of scaling the estimated values for claim volumes and using some future inflation index to take the projected payments to final values. The inflation rate to be used here is 7.5% p.a. which is chosen as it is close to the average annual historic rate implied by the index used to adjust the historic payments and will facilitate the comparison of the results. In practice a more appropriate prospective rate or rates will normally be utilized and a number of these used to obtain estimates.

Table 10 below shows the results derived from the full parameter model and inflation at 7.5% p.a.

Table 10: Projected values and Standard Errors.

Full Parameter Model with inflation at 7.5%

		Development Year												
Yr		1	2	3	4	5	6	7	8	9	10	11	12	Total
0	£							253	164	107	70	45	29	669
	se							36	25	18	12	9	6	80
1	£						389	253	164	107	70	45	29	1058
	se						54	37	26	18	13	9	6	120
2	£					658	427	278	181	117	76	50	32	1820
	se					89	61	43	30	21	15	10	7	198
3	£				911	592	384	250	162	106	69	45	29	2547
	se				123	84	58	40	28	19	14	10	7	267
4	£			1524	990	643	418	271	177	115	75	49	32	4292
	se			209	140	95	65	45	32	22	15	11	7	445
5	£		2910	1889	1226	797	518	336	219	142	93	60	39	8229
	se		421	277	185	126	86	59	41	29	20	14	10	896
6	£	5544	3596	2334	1515	984	639	415	270	176	114	74	48	15709
	se	972	624	406	267	177	119	81	55	38	26	18	12	2191

Overall Total	34324
Standard Error	2779
Percent. Error	8.10

The results are very close to those obtained earlier (Table 6) from the almost identical model without explicit inflation assumptions.

Increasing the inflation rate to 8.5% p.a. increases the overall estimate to 35210 with a standard error of 2858. So the one percentage change in the assumed future inflation rate impacts the estimated future payments by 2.6%.

Turning now to the reduced parameter model, that is the three parameter model with no accident year effects apart from the common level we obtain the following results assuming future inflation at 7.5% p.a.

Table 11: Projected values and Standard Errors.

Reduced Parameter Model with inflation at 7.5%.

		Development Year												
Yr		1	2	3	4	5	6	7	8	9	10	11	12	Total
0	£							249	165	109	72	47	31	673
	se							36	25	18	13	9	6	79
1	£						412	272	179	118	78	52	34	1145
	se						55	39	27	19	14	10	7	120
2	£					703	464	306	202	134	88	58	39	1994
	se					90	62	44	31	22	16	11	8	184
3	£				1018	671	443	292	193	127	84	56	37	2921
	se				1125	86	59	42	29	21	15	11	7	235
4	£			1585	1045	690	455	300	198	131	87	57	38	4586
	se			192	129	88	61	43	30	21	15	11	8	323
5	£		2945	1942	1280	845	557	368	243	160	106	70	46	8563
	se		358	235	158	108	75	52	37	26	19	13	9	541
6	£	6241	4114	2712	1788	1180	778	514	339	224	148	98	65	18201
	se	777	500	329	220	151	105	73	52	37	26	19	13	1090

Overall Total	38083
Standard Error	1725
Percent. Error	4.53

The results now look, and are, different. The overall estimate is significantly up on the previous estimates and the standard error is much reduced. The reduction in the overall standard error is due to the smaller number of parameters left in the reduced model and reflects the increased degree of smoothing that this parameter reduction has produced.

The increase in the overall projection, at just under 11%, is however too high to be explained by the derived standard errors. The main contributor can be clearly identified from the tables as the last accident year. This is not too surprising with hindsight. There is only a single data point from which to project. If it is assumed, as in the first case, that each year has an independent level then this point alone determines the level of the last accident year. The accident year residual plot for the latter model (Chart 14) shows the standardized residual for accident year 6 at around −1. Although this will not generally be considered statistically significant its impact, in a reserving context, has become significant.

Assuming a common level (in the adjusted figures) substantially reduces the influence of this last data point on its accident year estimate of future payments. As the adjusted triangle figures show, the one and only value for this last accident year is substantially below the corresponding values of the prior years. Using the same average value for all accident years gives the last accident year an average value which is now just under 16% higher than the value estimated from its own single data point.

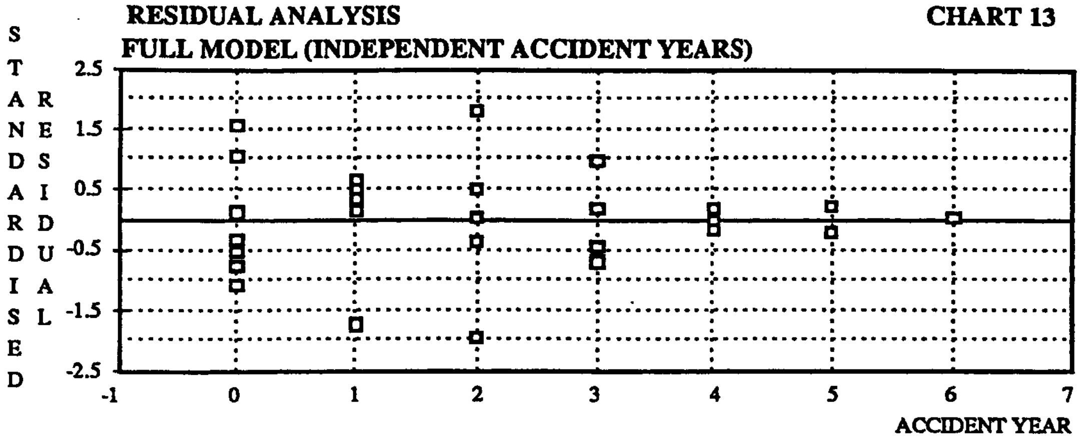
Putting the last accident year back into the model will produce results which will broadly match the full model overall estimate but with a reduced standard error. These are shown below.

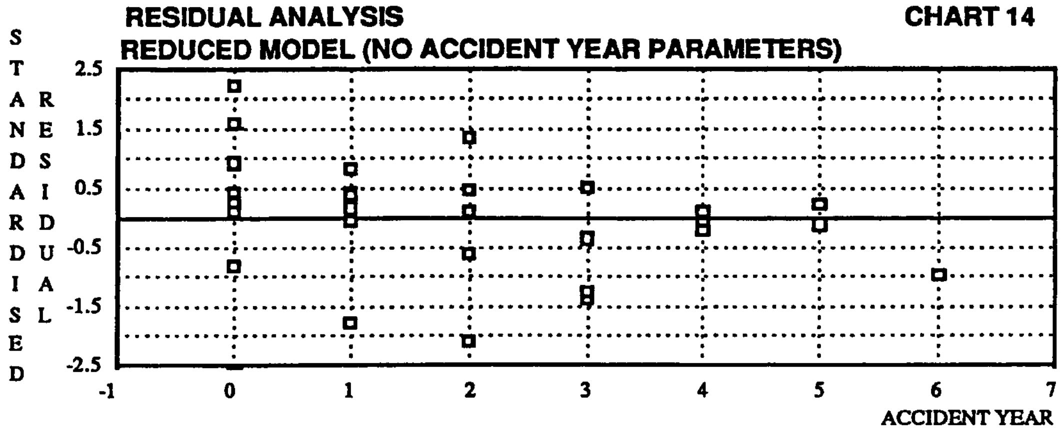
Table 12: Projected values and Standard Errors.

Reduced Parameter Model with Acc Yr 6, inflation at 7.5%.

		Development Year												
Yr		1	2	3	4	5	6	7	8	9	10	11	12	Total
0	£							249	165	109	72	47	31	673
	se							35	25	18	13	9	6	79
1	£						412	272	179	118	78	52	34	1145
	se						55	39	27	19	14	10	7	120
2	£					703	464	306	202	134	88	58	39	1994
	se					90	62	44	31	22	15	11	8	183
3	£				1017	671	443	292	193	127	84	56	37	2921
	se				125	85	59	42	29	21	15	11	7	234
4	£			1585	1045	689	455	300	198	131	87	57	38	4585
	se			192	128	88	61	43	30	21	15	11	8	322
5	£		2945	1942	1280	845	557	368	243	160	106	70	46	8562
	se		357	235	157	108	75	52	37	26	19	13	9	540
6	£	5494	3621	2387	1574	1038	685	452	299	197	130	86	57	16021
	se	981	639	421	280	188	127	87	60	41	28	20	14	2258

Overall Total	35902
Standard Error	2609
Percent. Error	7.27





(NOTE HERE THE RESIDUAL FOR ACC YR 6)
(IT RESULTS IN A 7.3% INCREASE IN THE OVERALL ESTIMATE)

Both these models are reasonable. They fit the data well and the standard errors are quite small. The results are quite different and these differences are clearly not explained by the standard errors, and are primarily due to the choice of parameters. As we know little about the underlying account it will be very difficult to choose between these models. In practice additional information, and informed views, will need to be sought to assist in this choice. This can then be used directly in deciding which parameters are to be left in the model.

A theoretically more appealing approach is to use some form of external or prior distribution and estimate in a Bayesian framework. This is explained in more detail by Verrall (6). It is possible to carry out the necessary calculations in the spreadsheet but more computation is necessary. The Bayesian approach combines formal statistical theory and informed prior estimates (knowledge and expertise!) and would appear to represent almost an ideal combination of theory and practice for reserving work. In practice more work is necessary in order to understand how sensitive the results are to these prior estimates, especially as these are made in log-space which, while convenient, are nevertheless somewhat alien from the immediate everyday experience of practitioners.

M. Final comments

This section will briefly consider some other aspects of these models which were deliberately avoided in earlier sections as the main emphasis has been a practical rather than a theoretical one.

a. Standard errors of reserve estimates

In practice, and as an approximation, as long as a sufficiently large number of future values are being projected it may be assumed that the distribution of the overall estimate obtained is normal with mean and standard error as calculated above.

That is we can use normal probability tables to establish approximate confidence intervals around the model reserve estimate. In the last example shown in Table 12 above for instance and under the conditions of the model, we have (approximately) a 95% probability that the required reserve will be less than 40194 (35902 + 1.645 × 2609). Recall however that the error estimate may be incomplete and future inflation is assumed fixed reducing the possible error further.

In practice the specific variability of a particular class reserve estimate may be less important to management than the variability of the overall company claims reserve Balance Sheet figure.

The individual class standard errors may be used to obtain estimates of this overall variability. For example if mutual independence of reserve estimates by class is assumed the overall variance may be obtained as the sum of the individual variances. Under these circumstances the percentage error in the overall reserves can drop to low figures.

Much work remains to be done in this area. At least these methods provide a start point to such considerations.

There will clearly be other factors, not incorporated into the model, that in practice will add to the error terms. There was no attempt to explicitly adjust for inflation in the first examples although the models incorporated an implicit assumption which is then implicitly projected into the future.

In the later examples values were adjusted for past inflation, using an index that may or may not have been the most appropriate, and projected values calculated using an assumed future rate of inflation, or more correctly claims cost escalation. The examples assumed a future rate which was based on the average past inflation used in adjusting the data.

Relatively small changes in these assumed future rates can lead to relatively large changes to the overall projected values. These models can be used to produce a series of results, with varying future claims escalation assumptions, from which it may be possible to derive a measure of the additional variability that may arise from this source.

These models do not attempt to allow for changes in the speed of settlement of claims. Payment developments may appear stable due to a combination of accelerating costs counteracted by a slowdown in settlements. Clearly under such circumstances estimates from a regression model on log-incremental payments, or a chain ladder projection based on cumulative payments, are likely to produce estimates which may be seriously biased.

Finally there will generally be a lot more information available to management than that used in fitting any statistical model. It is just possible that a combination of statistical derived estimates with informed estimates based on specific and detailed knowledge of the particular business, its environment and claims, may produce final estimates that have reduced variability. This will be however difficult to prove.

b. Negative values in incremental data sets

One particular problem with log-linear models is the occasional negative value in the original space.

Negative values occur in practice especially in net of reinsurance incremental payments and in classes of business subject to large subrogation or salvage recoveries. Various alternative approaches are available to the modeller to deal with negative values in practice. One approach, adopted in a commercial package (ICRFS), is to add a sufficiently large constant to all the incremental values, so that they all become positive, before the logarithmic transformation and an adjustment made in the projected values.

An alternative approach, that may be acceptable in practice, is to shift payments from one period to an adjacent one so as to eliminate a negative value. This may be justified if it is known, or suspected, that the negative value is the result of some serial correlation, for example when preceded by a relatively large value. Another possibility, which may be tried where the negative value is small is to ignore the value totally or to set it to some small positive value such as 1 ($\log_e 1=0$).

No particular approach is recommended here as ideal for dealing with negative values. In practice the reason for such negative values has to be investigated and this process often helps identify an appropriate approach to deal with the problem. Clearly one should not ignore a genuine feature of the data for the sake of convenience.

c. Parsimony

The chain ladder model is sometimes considered overparameterised as it involves a parameter for each accident year and each development period. Too many parameters can lead to model instability. Increasing the number of explanatory variables improves the quality of the fitted data but such slavish adherence to the data often results in unstable projections. At the extreme one can always obtain a perfect fit by including enough parameters in the model. Such a model fails to achieve any smoothing of the data and will be very poor for prediction purposes. Parsimonious models, that is with fewer parameters are to be preferred for this reason. This is explained in more detail in the first article in this Volume of the Manual.

d. Serial Correlation and Heteroscedasticity

The triangular shaped incremental payments data tend to decrease as the development years increase and there is usually some serial correlation present in these payments for a particular accident year. Such correlation may occur when a low payment period, due to administrative problems for example, is followed by a catching up high payment period or vice versa. On net paid claims data this may happen when a gross payment is made in one period with the incoming associated reinsurance processed during the following period.

The decline in values in the development direction tends to result in the residuals increasing with development period. This characteristic is an example of heteroscedasticity. In effect the IID assumption implies that the error terms in the original space are subject to the same percentage variation irrespective of their absolute values. Experience with payments triangles indicates that as payments diminish in the tail the percentage variation of these payments tends to be much higher than that seen in the first few development periods when a greater volume of payments is usually being made. This may be more pronounced in net rather than gross payments.

Methods to overcome this are being developed. One approach followed by Zehnwirth in the ICRFS package (Interactive Claims Reserving and Forecasting System) is to use weights. Alternative error assumptions, which may well turn out to be more appropriate, are being investigated by others. The main disadvantage of these approaches is the difficulty of obtaining the parameter estimates compared to the comparatively easy spreadsheet regression approach.

e. The Hoerl run-off curves

A particularly useful family of curves for run-off patterns is the Gamma family defined by

$$P_{ij} = K_j (1 + j)^b \exp(aj)$$

Each curve has a level parameter K_j and two shape parameters b and a the latter being an exponential. They have the immediate advantage of becoming linear in log-space and can be fitted simply by multiple regression using the techniques of this article. These curves form the start point in the ICRFS package.

As the example above illustrated these curves do not always produce good fits for all development periods. They can be particularly poor in fitting the first few development periods which clearly have a significant influence on the reserves projected for the most recent accident years where a substantial amount of the overall reserve is generally to be found.

It is possible to use the simple techniques outlined in this article to fit "mixed" models where some shape is fitted for later development periods and independent parameters fitted for the earlier periods. The example above fitted an independent first development parameter and an exponential decay curve thereafter. Any shape that can be expressed linearly (in log-space) can be tried even if in practice restrictions in "allowable" shapes will inevitably be necessary to keep any package to reasonable size.

f. Conclusion

Regression techniques are now beginning to dominate developments in claims reserving methodology. The formal approach adopted, whether utilizing maximum likelihood and IID normal errors or any other error model, at least enables the modeller to test the reasonableness of the assumptions. The model testing phase itself can often reveal interesting aspects of the data which may not be immediately obvious from looking at the cumulative payments.

These models can be very useful for inter-company comparisons and for comparing the stability of run-off triangles. Some results along these lines are to be found in Section E of the Claims Run-Off Patterns Working Party report presented to the 1989 GISG (General Insurance Study Group) Conference in Brighton.

This article is intended to give a practical introduction to these techniques and does not claim any original theoretical developments. The writer is particularly grateful to Arthur Renshaw and Richard Verrall of City University for their invaluable and patient explanations on this subject. The hope is that other practitioners can now begin to benefit by experimenting with these techniques.

References

(1) 1982: Kremer, E. "IBNR claims and the Two-Way Model of ANOVA", *Scandinavian Actuarial Journal*, 1, 47-55.

(2) 1989: Renshaw, A.E. "Chain Ladder and Interactive Modelling", *Journal of the Institute of Actuaries*, Vol. 116, 559-587

(3) 1988: Taylor, G.C. "Regression Models in Claims Analysis (II)". *William M.M. Campbell, Cook, Knight*, Sydney, Australia

(4) 1988: Verrall, R. "Bayes Linear Models and the Claims Run-Off Triangle". *Actuarial Research Paper No. 7*, The City University, London

(5) 1989: Verrall, R. "The Chain Ladder and Maximum Likelihood", *Actuarial Research Paper No. 12*, City University, London.

(6) 1989: Verrall, R. "On the Unbiased Estimation of Reserves from Loglinear Models", *Actuarial Research Paper No. 13*, City University, London.

(7) 1985: Zehnwirth, B. "Interactive Claims Reserving Forecasting System (ICRFS)" *Benhar Nominees Pty Ltd.* Tarramurra NSW Australia.

Appendix 1

Matrix calculations for the formal chain ladder example

Design matrix **X**

$$\begin{matrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0
\end{matrix}$$

Design matrix **X** transposed $\mathbf{X}^T$

$$\begin{matrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{matrix}$$

Product of $\mathbf{X}^T\mathbf{X}$

$$\begin{matrix}
4 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 3 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 2 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 3 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 2 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1
\end{matrix}$$

Inverse of $\mathbf{X}^T\mathbf{X}$ i.e. $(\mathbf{X}^T\mathbf{X})^{-1}$

.58333	.25000	.16667	.00000	-.33333	-.41667	-.58333
.25000	.58333	.16667	.00000	-.33333	-.41667	-.25000
.16667	.16667	.66667	.00000	-.33333	-.16667	-.16667
.00000	.00000	.00000	1.00000	.00000	.00000	.00000
.33333	-.33333	-.33333	.00000	.66667	.33333	.33333
.41667	-.41667	-.16667	.00000	.33333	.91667	.41667
.58333	-.25000	-.16667	.00000	.33333	.41667	1.58333

Future design $\mathbf{X}_f$

0	1	0	0	0	0	1
0	0	1	0	0	1	0
0	0	1	0	0	0	1
0	0	0	1	1	0	0
0	0	0	1	0	1	0
0	0	0	1	0	0	1

Transpose of Future Design Matrix $\mathbf{X}_f^T$

0	0	0	0	0	0
1	0	0	0	0	0
0	1	1	0	0	0
0	0	0	1	1	1
0	0	0	1	0	0
0	1	0	0	1	0
1	0	1	0	0	1

Product of Future Design $\mathbf{X}_f$ and Inverse of $\mathbf{X}^T\mathbf{X}$

i.e $\mathbf{X}_f\,(\mathbf{X}^T\mathbf{X}\,)^{-1}$

.33333	.33333	.00000	.00000	.00000	.00000	1.33333
.25000	-.25000	.50000	.00000	.00000	.75000	.25000
.41667	-.08333	.50000	.00000	.00000	.25000	1.41667
.33333	-.33333	-.33333	1.00000	.66667	.33333	.33333
.41667	-.41667	-.16667	1.00000	.33333	.91667	.41667
.58333	-.25000	-.16667	1.00000	.33333	.41667	1.58333

Final product ($\mathbf{X}_f\,(\mathbf{X}^T\mathbf{X})^{-1}$) and $\mathbf{X}_f^T$

i.e. $\mathbf{X}_f\,(\mathbf{X}^T\mathbf{X})^{-1}\,\mathbf{X}_f^T$

$$\begin{array}{rrrrrr}
1.66667 & .00000 & 1.33333 & .00000 & .00000 & 1.33333 \\
.00000 & 1.25000 & .75000 & .00000 & .75000 & .25000 \\
1.33333 & .75000 & 1.91667 & .00000 & .25000 & 1.41667 \\
.00000 & .00000 & .00000 & 1.66667 & 1.33333 & 1.33333 \\
.00000 & .75000 & .25000 & 1.33333 & 1.91667 & 1.41667 \\
1.33333 & .25000 & 1.41667 & 1.33333 & 1.41667 & 2.58333
\end{array}$$

And finally the data specific Var-Cov matrix is derived from the above values by multiplying by σ^2.

So the first entry is $1.66667 \times .0524^2 = .00457$ etc.

The Variance-Covariance matrix in this case is then

i.e. $\sigma^2\,\mathbf{X}_f\,(\mathbf{X}^T\mathbf{X})^{-1}\,\mathbf{X}_f^T$

$$\begin{array}{rrrrrr}
.00457 & .00000 & .00366 & .00000 & .00000 & .00366 \\
.00000 & .00343 & .00206 & .00000 & .00206 & .00069 \\
.00366 & .00206 & .00526 & .00000 & .00069 & .00389 \\
.00000 & .00000 & .00000 & .00457 & .00366 & .00366 \\
.00000 & .00206 & .00069 & .00366 & .00526 & .00389 \\
.00366 & .00069 & .00389 & .00366 & .00389 & .00709
\end{array}$$

Appendix 2

Spreadsheet Regression Output tables

The raw spreadsheet regression output table for the first example (4×4 chain ladder) was

```
Regression Output:
     Constant                           0
     Std Err of Y Est                   .05238
     R Squared(Ajd,Raw)       .99758    .99919
     No. of Observations                10
     Degrees of Freedom                 3

Coefficient(s)    9.2884   9.5911   9.6924   9.7358   -.4662   -1.801   -2.647

Std Err of Coef.  .0400    .0400    .0428    .0524    .0428    .0502    .0660
```

This is very typical of all the spreadsheet regression output.

A brief description of this output is given below:

a) Constant (=0)

The spreadsheet regression command usually has an option of either fitting through the origin or calculating a constant. In the case above the model was fitted through the origin so the constant calculated is zero. In the models described in the article a parameter is used in place of this constant as this makes the analysis more convenient. The calculated values will be the same but in the latter case the regression shows the standard error associated with this constant.

b) Std Err of Y Est (0.0524)

This is the estimated standard error of the residuals. It is the square root of the estimated model variance σ^2.

It is in other words the estimate of the standard deviation of the assumed underlying normal error term.

This value plays a very significant role in the estimates of future values and their standard errors.

c) R Squared (Adj, Raw) (0.9976 0.9992)

This is a statistic ranging from 0 to 1 which indicates how much variation in the data is explained by the model. The closer to 1, the more variation explained by the model. The difference in the two values is from a correction for the degrees of freedom.

In crude terms it indicates that the model explains 99.76% of the values, in the log-space.

d) No of Observations (10)

The 4 by 4 triangle contained ten values all of which were used in the fitting process.

e) Degrees of Freedom (3)

The model assumed 7 independent parameters (including the constant) and used 10 observations to estimate these. The difference, (10-7), is the number of degrees of freedom.

Note that in this case there are a lot of parameters in relation to the number of data values in the triangle. This tends to produce a high quality of fit, i.e. a high R^2 but forced adherence to the actual data by incorporating many parameters in the model can lead to a model with poor predictive qualities.

f) Coefficient(s) (9.288 9.591 etc.)

These values are the estimates of the model parameter values. They appear in the order defined by the Design Matrix one for each independent variable.

Least squares are being used to calculate these values and the solution is given by

$(\mathbf{X}^T\mathbf{X})^{-1}\,\mathbf{X}^T\,\mathbf{Y}$ where $\mathbf{Y}$ is the vector of data values.

g) Std. Err of Coef. (0.0400 0.0400 etc...)

These are the estimated standard errors of the coefficient estimates. They are the square roots of the diagonal elements of the variance-covariance matrix of the coefficients

$\sigma^2\,(\mathbf{X}^T\mathbf{X})^{-1}$

Changing values in the data triangle does not affect the design matrix $\mathbf{X}$ and only changes the scalar element or σ^2.

So different data sets result in standard errors of the model coefficients which differ only by a constant factor which is equal to the ratio of the data specific model standard errors or σ's.

◇