ISBN 0 7325 1212 3

THE UNIVERSITY OF MELBOURNE

CALCULATIONS AND DIAGNOSTICS FOR LINK RATIO TECHNIQUES

by

Ben Zehnwirth
The University of Melbourne
and
Glen Barnett
University of New South Wales

RESEARCH PAPER NUMBER 41

November 1996

Centre for Actuarial Studies
Department of Economics
The University of Melbourne
Parkville, Victoria, 3052
Australia.

Calculations and Diagnostics for Link Ratio Techniques

Glen Barnett* and Ben Zehnwirth†

November 14, 1996

Abstract

Many of the standard techniques based on link ratios (development factors), such as the Chain Ladder, can be shown to correspond to weighted linear regressions (e.g. see Murphy 1994). In fact a number of these techniques can be encompassed under a single family of models indexed by a parameter representing the amount of weighting by volume, allowing many of the results to be derived at once for models that appear to be separate. The family is discussed by Mack (1993), who gives standard errors for the total forecast in the case of the Chain Ladder model.

Many related diagnostic calculations are available, which are useful in fitting the models, checking their assumptions and choosing between competing models. For the models discussed in this note, the diagnostics generally have a simple form.

Additionally, these diagnostics are used on three case studies, to illustrate specific problems associated with the assumptions required for models like the Chain Ladder to apply. The appropriateness of various assumptions are discussed in detail.

Keywords: Link ratio, Development Factor, Chain Ladder, Regression, Standard Errors, Diagnostics

This work was partially funded by a grant from the Institute of Actuaries of Australia

* Australian Graduate School of Management, University of New South Wales, NSW 2052, Australia

† GIO Professional Fellow in Insurance, Centre for Actuarial Studies, University of Melbourne, Parkville 3052

1. Introduction

Mack (1993) derives standard errors of development factors and forecasts (including the total) for the chain ladder and discusses some diagnostics for choosing between the chain ladder and other techniques based on different relative weighting by volume. He mentions the connection to weighted least squares regression through the origin, and some of his diagnostics indicate that an intercept term may be warranted on the data he analyses.

Working more directly in a regression framework, Murphy (1994) derives results for models without an intercept, which he call multiplicative models, and those including an intercept, which he calls linear models, though his multiplicative models are actually linear models as well.

We show how the results for these models can be derived as a family, discuss calculations and diagnostics for fitting and choosing between models, and checking assumptions. Standard errors of forecasts and diagnostics on the paid loss (uncumulated) scale are also derived for a generalisation of the models discussed by Murphy and Mack.

We analyse data presented by Mack (1993), another real data array and a simulated set of data that displays many features found in actual data. Using diagnostics with these sets of data indicate problems with the models based on development factors. Possible remedies are discussed.

2. Notation and Basic Model

For simplicity, we will detail the calculations for a full loss triangle. The discussion applies to other array shapes by changing the appropriate limits on summations. We refer to accident years, development years etc. for convenience, but the discussion applies equally well for any sampling period.

Let there be n accident years, numbering the most recent accident year as 0, and the first as $n-1$, as in Murphy (1994). Let y_{ij} be the cumulative amount paid in accident year i, development year j, $i=0,...,n-1$, $j=0,...,n-1$. This simplifies many of the formulas. See Figure 1. Let $x_{ij} = y_{i,j-1}$, so that y_{ij}/x_{ij} is the observed development factor from $j-1$ to j in accident year i.

<Insert Figure 1 about here>

For the basic model, assume $y_{ij} = \beta_j x_{ij} + u_{ij}$, and that $\mathrm{Var}(u_{ij}) = x_{ij}^{\delta}\sigma_j^2$.

This corresponds to a weighted linear regression model passing through the origin. The parameter β_j represents the underlying development factor from $j-1$ to j common to all accident years.

3. Estimating the underlying development factors

From the Gauss-Markov Theorem we obtain that the best linear unbiased estimates are:

$$\hat{\beta}_j = \frac{\sum_{i=j}^{n-1} y_{ij} x_{ij} x_{ij}^{-\delta}}{\sum_{i=j}^{n-1} x_{ij}^2 x_{ij}^{-\delta}} = \frac{\sum_{i=j}^{n-1} y_{ij} x_{ij}^{1-\delta}}{\sum_{i=j}^{n-1} x_{ij} x_{ij}^{1-\delta}} = \frac{\sum_{i=j}^{n-1} (y_{ij}/x_{ij}).x_{ij}^{2-\delta}}{\sum_{i=j}^{n-1} x_{ij}^{2-\delta}}$$

$$= \frac{\sum_{i=j}^{n-1} y_{ij} x_{ij} w_{ij}^R}{\sum_{i=j}^{n-1} x_{ij}^2 w_{ij}^R} = \frac{\sum_{i=j}^{n-1} y_{ij} w_{ij}^C}{\sum_{i=j}^{n-1} x_{ij} w_{ij}^C} = \frac{\sum_{i=j}^{n-1} (y_{ij}/x_{ij}) w_{ij}^A}{\sum_{i=j}^{n-1} w_{ij}^A}$$

where the w_{ij} are the weights for the weighted regression (R), weighted chain ladder (C), and weighted average development factor (A) formulations respectively. For array shapes with the early payment or latter accident years cut off, or with missing values, the summations will be over a smaller range, but in each case, the summations are over the observations that have values for both y and x. Consequently, the limits on summations will usually be suppressed so as not to preclude the general form – the actual limits are easily found in any particular case, though difficult to write down in complete generality.

The above shows that the best linear unbiased estimates of the development factors can be equally well thought of as weighted regression though the origin, weighted chain ladder estimates and weighted averages of the individual development factors. In each case, the weights are dependent on the actual value of δ.

For some particular values of δ, each of the estimates become "unweighted". When δ is 2 the best linear unbiased estimates of the development factors (β_j, j=1,...,n-1) are given by the unweighted average development factor, when δ is 1 we get the ordinary chain ladder estimates, and when δ is 0 we get ordinary least squares through the origin.

Best Estimates: The Gauss-Markov Theorem, however, doesn't make it clear when a linear estimator of a parameter is appropriate. Specifically, if the data are sufficiently non-normal, *any* linear estimate can be a very bad one, and only choosing the best among those may be unwise. When the data are normal, however, the above estimates are the best among all estimators, not merely linear ones.

Rather than rely on the Gauss-Markov Theorem, then, it would seem more prudent to actually make the assumption of normality explicitly, and then check whether it is appropriate. If the data indicate non-normality, it would then be sensible to consider other estimators. One such widely used diagnostic is a plot of standardised residuals against normal scores (expected normal order statistics). It is not usually necessary to consider explicit tests of normality, but a test based on the squared correlation between the standardised residuals and their normal scores, as given in Shapiro and Francia (1972), is convenient if we are already doing the normal scores plot.

4. Standard Errors of the Estimated Link Ratios

$$\begin{aligned}\mathrm{Var}(\hat{\beta}_j) &= \mathrm{Var}\left(\frac{\sum_i y_{ij} x_{ij}^{1-\delta}}{\sum_i x_{ij}^{2-\delta}}\right)\\ &= \left(\frac{1}{\sum_i x_{ij}^{2-\delta}}\right)^2 \sum_i (x_{ij}^{1-\delta})^2 \sigma_j^2 x_{ij}^{\delta} \text{ a}\\ &= \left(\frac{1}{\sum_i x_{ij}^{2-\delta}}\right)^2 \sigma_j^2 \sum_i x_{ij}^{2-\delta}\\ &= \frac{\sigma_j^2}{\sum_i x_{ij}^{2-\delta}}\end{aligned}$$

Using the usual unbiased estimate of the variance parameter,

$$\hat{\sigma}_j^2 = \frac{1}{n_j - 1} \sum_i \frac{(y_{ij} - x_{ij}\hat{\beta}_j)^2}{x_{ij}^{\delta}}$$

we estimate the variance of $\hat{\beta}_j$ by

$$\hat{\text{Var}}(\hat{\beta}_j) = \frac{\hat{\sigma}_j^2}{\sum_i x_{ij}^{2-\delta}}.$$

The estimated standard error of the estimate is just the square root of this quantity,

$$\hat{\text{s.e.}}(\hat{\beta}_j) = \frac{\hat{\sigma}_j}{\sqrt{\sum_i x_{ij}^{2-\delta}}}$$

and so we can form a t-ratio for testing whether β_j is different from some specific value, β_0 by:

$$t_{(n_j-1)} = \frac{\hat{\beta}_j - \beta_0}{\hat{\text{s.e.}}(\hat{\beta}_j)}.$$

For example, it is of special interest to see if β_j is different from 1; if not, it indicates that the previous cumulative has little predictive value for the following incremental paid loss, as discussed in Venter (1996). If that is the case, analysing the incrementals will be much more informative, as well having the added benefit of enabling the introduction of parameters to capture changing payment year trends. If the analysis of the cumulatives is retained, however, some smoothing of the development factors across years is necessary, most simply by introducing curves for the changes in the development factors over time. This represents a blending of information across years that allows the extraction of what little information there is in the cumulatives about the subsequent paid losses.

5. Residuals and Standard Errors

Let $\hat{u}_{ij} = y_{ij} - \hat{\beta}_j x_{ij}$ be the raw residual for accident year *i*, development year *j*. Recall that the variance of the error term is $x_{ij}^{\delta}\sigma_j^2$. This is often used (with estimated variance parameter) as the standard error of $\hat{u}_{ij}$, though the actual variance is always smaller than this – since (if the model is correct) the fitted model is closer to the data than the true model is. The actual variance is:

$$
\begin{aligned}
\mathrm{Var}(\hat{u}_{ij}) &= \mathrm{Var}(y_{ij} - \hat{\beta}_j x_{ij}) \\
&= \mathrm{Var}(y_{ij}) + x_{ij}^2 \mathrm{Var}(\hat{\beta}_j) - 2\mathrm{Cov}(y_{ij}, \hat{\beta}_j x_{ij}) \\
&= \mathrm{Var}(y_{ij}) + x_{ij}^2 \mathrm{Var}(\hat{\beta}_j) - 2\mathrm{Cov}(\hat{\beta}_j x_{ij} + \hat{u}_{ij}, \hat{\beta}_j x_{ij}) \\
&= \mathrm{Var}(y_{ij}) + x_{ij}^2 \mathrm{Var}(\hat{\beta}_j) - 2x_{ij}^2 \mathrm{Var}(\hat{\beta}_j) \\
&= \sigma_j^2 x_{ij}^{\delta} - x_{ij}^2 \frac{\sigma_j^2}{\Sigma_i x_{ij}^{2-\delta}} \\
&= \sigma_j^2 x_{ij}^{\delta}\left(1 - \frac{x_{ij}^{2-\delta}}{\Sigma_i x_{ij}^{2-\delta}}\right)
\end{aligned}
$$

We estimate this by:

$$
\widehat{\mathrm{Var}}(\hat{u}_{ij}) = \hat{\sigma}_j^2 x_{ij}^{\delta}\left(1 - \frac{x_{ij}^{2-\delta}}{\Sigma_i x_{ij}^{2-\delta}}\right),
$$

and so the standard error is:

$$
\widehat{\mathrm{s.e.}}(\hat{u}_{ij}) = \hat{\sigma}_j x_{ij}^{\delta/2}\sqrt{1 - \frac{x_{ij}^{2-\delta}}{\Sigma_i x_{ij}^{2-\delta}}}.
$$

A standardised residual ($\hat{e}_{ij}$) can be calculated by dividing a residual by its standard error. In practice, unless δ is less than 2 and x_{ij} is relatively large, the general appearance of a pattern in a residual plot will not be appreciably affected by just using $\hat{\sigma}_j x_{ij}^{\delta/2}$ for the standard error, though the relative spread against development year will tend to be smaller in the later years that it should be.

Residuals on the Paid Loss scale: We have just found the residuals and standard errors for cumulated data, and it would be useful to know how the fitted model looks on the paid loss scale, especially since changing trends against calendar years (which development factor models don't pick up) tend to be obscured by cumulating the data. Let $p_{ij} = y_{ij} - x_{ij}$ be the paid loss in accident year *i*, development year *j* (*j* = *1*,...,*i*). Then the paid residual,

$$\begin{aligned} p_{ij} - \hat{p}_{ij} &= y_{ij} - x_{ij} - (\hat{y}_{ij} - x_{ij}) \\ &= y_{ij} - \hat{y}_{ij} \end{aligned}$$

is the same as the cumulative residual.

6. Model Selection

We will use the Akaike Information Criterion, or AIC (Akaike, 1972) to choose the "best" model. The AIC is a measure of fit that includes a penalty for the number of parameters, and a lower AIC indicates a better model. In fact, if L is the likelihood, and p is the number of parameters,

$$\text{AIC} = -2 \log L + 2p.$$

For the models under consideration, we can write the log-likelihood as a sum over the log-likelihoods for the individual development years:

$$log\mathrm{L} \;=\; \Sigma_j \; log\mathrm{L}_j,$$

and the log-likelihood for development year j can be written as:

$$\log \mathrm{L}_j = n_j(\log \hat{\sigma}_j^2 + 1 + \log 2\pi) + \delta \Sigma_i \log x_{ij}.$$

where n_j is the number of observations in development year j used in the estimation of β_j. Since likelihoods are only defined up to a multiplicative constant, the $1+log(2\pi)$ term may be dropped from the log-likelihood.

In regression models, σ^2 is not usually counted as a parameter, and for the models we are dealing with here, δ may be regarded as another variance parameter, or just as an index parameter, indicating which of the three basic models we choose. Consequently, it is only the number of development factor parameters that are counted for the AIC. This will usually be $n-1$, but may be less if consecutive β_j's are set to be equal, for example.

7. Development factor estimates and variance parameter estimates in the tail

In the tail there are generally few observations, and at the same time the tail is usually quite flat. There is often less need for different parameters for every development year, and the lack of data can make estimating different parameters for every year in the tail risky. Indeed, in the case of estimating the variance, in the last year for a full triangle it becomes impossible. Consequently it becomes important to be able to have some combined estimation in the tail. Murphy (1994) does this with the variance term in his data analysis, for example.

There is no extra effort in combining years; it is just like doing computations for a single year. We will assume that we will only wish to get a combined estimate of β when we are also getting a combined estimate of σ^2, but that we may form a combined estimate of σ^2 when we are estimating individual β_j's. It would be unusual, and generally inadvisable, to form a combined estimate of β with different σ^2's. We omit the calculations for that case, though they are not particularly complicated.

Note that with cumulated data, it will not normally be the case that we want to combine estimates of β's across years directly, though we may do so indirectly by introducing curves. However, if we analyse paid losses instead, many data sets indicate nearly constant ratios in the tail (constant percentage decrease, or "exponential" tails). A reduction in parameters like this will generally yield better forecasts and smaller standard errors.

Combined estimate of the variance parameter: Given some estimates of the β_j's, (whether these are themselves combined estimates or not) we can form an estimate of a combined σ^2 parameter by treating the residuals used in the calculation of the combined σ^2 as if they were all from a single year.

Combined estimate of the development factor: Given we are estimating a single variance parameter over the years in the combined estimate, this is simply a matter of treating the separate sets of x's and y's as if they were a single set of x's and y's, and using the usual formula.

8. Testing consecutive β_j's for equality

In this section, we will also make the assumption that the variances for the years being tested are already equal. In the case of two consecutive development years, we get a test statistic of:

$$t_{n_j+n_{j+1}-2} = \frac{\hat{\beta}_j - \hat{\beta}_{j+1}}{\hat{\sigma}^2(1/\sum_i x_{ij}^{2-\delta} + 1/\sum_i x_{i,j+1}^{2-\delta})},$$

where $\hat{\sigma}^2$ is the common estimate of σ^2, and the test statistic has a *t*-distribution with $n_j+n_{j+1}-2$ degrees of freedom. That is, the denominator is just the sum of the variances, but with the common estimate of σ^2. If β's are estimated for blocks of years and consecutive blocks are tested, the formula is almost identical, except that the *n*'s refer to the entire block, and the sums in the denominator are over all *x*'s in the block.

9. Forecasts and Standard Errors

The first part of this section gives a reasonably concise generalisation of the approach of Murphy (1994). In the interests of space, the derivation for the forecasts has been omitted, but follows directly from the arguments given by Murphy. All of these calculations are conditional on the data.

Forecasts:

$$\hat{y}_{ij} = \hat{\beta}_j \hat{y}_{i,j-1}, \qquad j = i,\ldots,n-1$$

and $\hat{y}_{ii}$ given the data is y_{ii}, a result we use throughout the following derivations.

Standard Errors: Again, conditioning on the data,

$$\begin{aligned}\mathrm{Var}(\hat{y}_{i,i+k} - y_{i,i+k}) &= \mathrm{Var}(\hat{y}_{i,i+k} - \mu_{i,i+k} + \mu_{i,i+k} - y_{i,i+k}) \\ &= \mathrm{Var}(\hat{y}_{i,i+k} - \mu_{i,i+k}) + \mathrm{Var}(y_{i,i+k} - \mu_{i,i+k}) \\ &= v^p_{i,i+k} + v^e_{i,i+k}, \text{ say}\end{aligned}$$

The first term is what Murphy calls the parameter variance, and the second the process variance.

$$\begin{aligned}\mathrm{Var}(\hat{y}_{i,i+k} - \mu_{i,i+k}) &= \mathrm{Var}(\hat{\beta}_{i+k}\hat{y}_{i,i+k-1}) \\ &= \hat{\beta}^2_{i+k}\mathrm{Var}(\hat{y}_{i,i+k-1}) + \hat{y}^2_{i,i+k-1}\mathrm{Var}(\hat{\beta}_{i+k}) + \mathrm{Var}(\hat{\beta}_{i+k})\mathrm{Var}(\hat{y}_{i,i+k-1})\end{aligned}$$

and

$$\begin{aligned}\mathrm{Var}(y_{i,i+k} - \mu_{i,i+k}) &= \mathrm{Var}(\beta_{i+k}y_{i,i+k-1} + u_{i,i+k}) \\ &= \beta^2_{i+k}\mathrm{Var}(y_{i,i+k-1}) + \mathrm{Var}(u_{i,i+k}) \\ &= \beta^2_{i+k}\mathrm{Var}(y_{i,i+k-1} - \mu_{i,i+k-1}) + \sigma^2_{i+k}x^\delta_{i,i+k} \\ &= \beta^2_{i+k}\mathrm{Var}(y_{i,i+k-1} - \mu_{i,i+k-1}) + \sigma^2_{i+k}y^\delta_{i,i+k-1}\end{aligned}$$

It is easy to show that if $X \sim N(\mu,\sigma^2)$, and $\lfloor x \rfloor$ indicates rounding down to the integer below x, then

$$E(X^k) = \sum_{r=0}^{\lfloor k/2 \rfloor} \frac{k!}{r!(k-2r)!2^r} \cdot \mu^{k-2r}\sigma^{2r}, \qquad k = 0, 1, 2, \ldots$$

Let $f_{i,j}^{\delta} = \mathrm{E}(y_{i,j}^{\delta})$. We estimate the process variance as:

$$\hat{\mathrm{Var}}(y_{i,i+k}) = \hat{\beta}^2_{i+k}\ \hat{\mathrm{Var}}(y_{i,i+k-1}) + \hat{\sigma}^2_{i+k}\hat{\mathrm{E}}(y^{\delta}_{i,i+k-1})$$

$$= \hat{\beta}^2_{i+k}\ \hat{\mathrm{Var}}(y_{i,i+k-1}) + \hat{\sigma}^2_{i+k}\hat{f}^{\delta}_{i,i+k-1}, \quad \text{say,}$$

where

$$\hat{f}_{ij}^{\delta} = \begin{cases} 1, & \delta = 0 \\ \hat{y}_{ij}, & \delta = 1 \\ \hat{y}_{ij}^{\,2} + \hat{\mathrm{Var}}(y_{ij}), & \delta = 2 \end{cases}$$

just as with Murphy (1993); simple substitution yields values for any other non-negative integer δ. Of course, to obtain estimated standard errors, we take the square roots of these estimated variances.

Additionally, approximate results for non-integer δ are possible, via a Taylor-series expansion, $E(g(T)) \approx g(\mu_T) + \frac{1}{2}\sigma_T^2 g''(\mu_T)$, see, for example, Cox and Hinkley (1974). Expansion to further terms is possible. If a distribution other than the normal is chosen, such as the gamma, computations for non-integer δ are much simplified, but then the regression results are no longer optimal. We will not pursue these considerations in this paper.

10. Forecasts and Standard Errors of Development Year Totals

Forecasts:

Let D_j be the unknown future development year total forecast, so:

$$D_j = \sum_{i=0}^{j-1} y_{ij}\text{, and}$$
$$\hat{D}_j = \sum_{i=0}^{j-1} \hat{y}_{ij}$$

Standard Errors:

$$\begin{aligned}
\text{Var}(\hat{D}_j - D_j) &= \text{Var}(\sum_{i=0}^{j-1} \hat{y}_{ij} - y_{ij}) \\
&= \text{Var}(\sum_{i=0}^{j-1} \hat{y}_{ij} - \mu_{ij} + \mu_{ij} - y_{ij}) \\
&= \text{Var}(\sum_{i=0}^{j-1} \hat{y}_{ij} - \mu_{ij}) + \text{Var}(\sum_{i=0}^{j-1} y_{ij} - \mu_{ij}) \\
&= V_j^p + V_j^e
\end{aligned}$$

again representing parameter and process variation.

$$\begin{aligned}
V_j^p &= \text{Var}(\sum_{i=0}^{j-1} \hat{y}_{ij}) \\
&= \text{Var}(\sum_{i=0}^{j-1} \hat{\beta}_j \hat{y}_{i,j-1}) \\
&= \text{Var}(\hat{\beta}_j [\hat{D}_{j-1} + y_{j-1,j-1}]) \\
&= \hat{\beta}_j^2 \text{Var}(\hat{D}_{j-1}) + [\hat{D}_{j-1} + y_{j-1,j-1}]^2 \text{Var}(\hat{\beta}_j) + \text{Var}(\hat{\beta}_j)\text{Var}(\hat{D}_{j-1}) \\
&= [\hat{\beta}_j^2 + \text{Var}(\hat{\beta}_j)] V_{j-1}^p + [\hat{D}_{j-1} + y_{j-1,j-1}]^2 \text{Var}(\hat{\beta}_j)
\end{aligned}$$

which we estimate by

$$\hat{V}_j^p = [\hat{\beta}_j^2 + \widehat{\text{Var}}(\hat{\beta}_j)] \hat{V}_{j-1}^p + [\hat{D}_{j-1} + y_{j-1,j-1}]^2 \widehat{\text{Var}}(\hat{\beta}_j).$$

Also,

$$\begin{aligned} V_j^e &= \mathrm{Var}(\sum_{i=0}^{j-1} y_{ij}) \\ &= \mathrm{Var}(\beta_j \sum_{i=0}^{j-1} y_{i,j-1} + \sum_{i=0}^{j-1} u_{i,j}) \\ &= \mathrm{Var}(\beta_j [D_{j-1} + y_{j-1,j-1}]) + \mathrm{Var}(\sum_{i=0}^{j-1} u_{i,j}) \\ &= \beta_j^2 \mathrm{Var}(D_{j-1}) + \sigma_j^2 \sum_{i=0}^{j-1} x_{ij}^\delta \\ &= \beta_j^2 V_{j-1}^e + \sigma_j^2 (y_{j-1,j-1}^\delta + \sum_{i=0}^{j-2} y_{i,j-1}^\delta) . \end{aligned}$$

Noting that, due to independence across accident years,

$$\mathrm{E}(\sum_{i=0}^{j-2} y_{i,j-1}^\delta) = \sum_{i=0}^{j-2} \mathrm{E}(y_{i,j-1}^\delta) = \sum_{i=0}^{j-2} f_{i,j-1}^\delta$$

the process variance term is estimated by

$$\hat{V}_j^e = \hat{\beta}_j^2 \hat{V}_{j-1}^e + \hat{\sigma}_j^2 (y_{j-1,j-1}^\delta + \sum_{i=0}^{j-2} \hat{f}_{i,j-1}^\delta) .$$

The estimated standard error of $\hat{D}_j$ is then $\sqrt{\hat{V}_j^p + \hat{V}_j^e}$

11. Forecasts and Standard Errors on the Paid Scale

It is often important to have forecasts and standard errors on the uncumulated paid loss scale, rather than forecasts of cumulated data. For example, if there is an anticipated change in future inflation or discount rates, or for matching of cash flows, or for certain kinds of reinsurance.

$$
\begin{aligned}
\mathrm{Var}(\hat{p}_{ij} - p_{ij}) &= \mathrm{Var}(\hat{y}_{ij} - y_{ij} - \hat{y}_{i,j-1} + y_{i,j-1}) \\
&= \mathrm{Var}(\hat{y}_{ij} - y_{ij}) + \mathrm{Var}(\hat{y}_{i,j-1} - y_{i,j-1}) - 2\mathrm{Cov}(\hat{y}_{ij} - y_{ij}, \hat{y}_{i,j-1} - y_{i,j-1}) \\
&= \mathrm{Var}(\hat{y}_{ij}) + \mathrm{Var}(\hat{y}_{i,j-1}) + \mathrm{Var}(y_{ij}) + \mathrm{Var}(y_{i,j-1}) \\
&\quad - 2\mathrm{Cov}(\hat{\beta}_j \hat{y}_{i,j-1}, \hat{y}_{i,j-1}) + 0 + 0 - 2\mathrm{Cov}(\beta_j y_{i,j-1}, y_{i,j-1}) \\
&= \mathrm{Var}(\hat{y}_{ij}) + \mathrm{Var}(\hat{y}_{i,j-1}) + \mathrm{Var}(y_{ij}) + \mathrm{Var}(y_{i,j-1}) \\
&\quad - 2\hat{\beta}_j \mathrm{Var}(\hat{y}_{i,j-1}) - 2\beta_j \mathrm{Var}(y_{i,j-1}) \\
&= \mathrm{Var}(\hat{y}_{ij}) - (2\hat{\beta}_j - 1)\mathrm{Var}(\hat{y}_{i,j-1}) + \mathrm{Var}(y_{ij}) - (2\beta_j - 1)\mathrm{Var}(y_{i,j-1}) \\
&= v^p_{ij} - (2\hat{\beta}_j - 1)v^p_{i,j-1} + v^e_{ij} - (2\beta_j - 1)v^e_{i,j-1}
\end{aligned}
$$

where to estimate these quantities we replace the *v*s by their estimates, remembering again that we must take square roots to obtain standard errors. Note that for the first forecast in a given accident year (*i=j+1*), the variance is the same as the variance for the cumulative forecast at that point, since (conditional on the data), v^e and v^p for the previous accident year are zero.

12. Models with an intercept term

Murphy (1994) gives an argument as to why having an intercept term makes sense in this context, and the data of Mack (1993) definitely indicates a need for an intercept term, as we shall see.

Assume $y_{ij} = \alpha_j + \beta_j x_{ij} + u_{ij}$, and that $\mathrm{Var}(u_{ij}) = x_{ij}^{\delta}\sigma_j^2$.

This corresponds to a weighted linear regression model. The parameter β_j no longer represents an underlying development factor in the usual sense, though it may be regarded as a (weighted) development factor calculated after the data have all been adjusted by their own weighted means, as we shall see. The results for this section may be obtained from standard weighted regression results.

13. Estimates of the parameters, their variances and covariances

Let

$$\bar{y}_j^w = \frac{\Sigma_i w_{ij}^R y_{ij}}{\Sigma_i w_{ij}^R} = \frac{\Sigma_i y_{ij} x_{ij}^{-\delta}}{\Sigma_i x_{ij}^{-\delta}} ,$$

that is, the weighted average of the y's for development year j, and define $\bar{x}_j^w$ similarly. Also, let

$$\mathrm{v}_j^w = \frac{\Sigma_i w_{ij}^R x_{ij}^2}{\Sigma_i w_{ij}^R} - (\bar{x}_j^w)^2$$

as a kind of weighted variance. Then we have

$$\hat{\beta}_j = \frac{\Sigma_i w_{ij}^R (y_{ij} - \bar{y}_j^w)(x_{ij} - \bar{x}_j^w)}{\Sigma_i w_{ij}^R (x_{ij} - \bar{x}_j^w)^2} ,$$

$$\hat{\alpha}_j = \bar{y}_j^w - \hat{\beta}_j \bar{x}_j^w ,$$

$$\hat{\mathrm{Var}}(\hat{\beta}_j) = \frac{\hat{\sigma}_j^2}{\mathrm{v}_j^w \Sigma_i w_{ij}^R}$$

$$\hat{\mathrm{Var}}(\hat{\alpha}_j) = \frac{\Sigma_i w_{ij}^R x_{ij}^2}{\Sigma_i w_{ij}^R} \frac{\hat{\sigma}_j^2}{\mathrm{v}_j^w \Sigma_i w_{ij}^R}$$

$$= \frac{\Sigma_i w_{ij}^R x_{ij}^2}{\Sigma_i w_{ij}^R} \cdot \hat{\mathrm{Var}}(\hat{\beta}_j)$$

$$\hat{\mathrm{Cov}}(\hat{\alpha}_j, \hat{\beta}_j) = \frac{-\bar{x}^w \hat{\sigma}_j^2}{\mathrm{v}_j^w \Sigma_i w_{ij}^R} .$$

$$= -\bar{x}^w \cdot \hat{\mathrm{Var}}(\hat{\beta}_j)$$

We can form t-ratios for testing specific values for α_j and β_j in the same way as before. In particular, we will often be interested in testing $\alpha_j = 0$, or $\beta_j = 1$, especially in the tail. If we find that many of the α's or $(\beta\text{-}1)$'s are not significant, we should take it as an indication that more information will be gained from a different kind of model, such as from analysing paid losses instead of cumulative values.

Estimating σ_j^2

We use an unbiased estimate for σ_j^2:

$$\hat{\sigma}_j^2 = \frac{\sum_i w_{ij}^R (y_{ij} - \hat{y}_{ij})^2}{n_j - 2}.$$

Testing consecutive parameter values for equality.

The test statistics are formed in the same way as before, as a difference of consecutive values, divided by the sum of the variances (as estimates of β or α for different years are independent).

The degrees of freedom are $n_j+n_{j+1}-4$, assuming β and α parameters are estimated for both years being tested.

14. Residuals and standard errors for the intercept model

$$
\begin{aligned}
\hat{u}_{ij} &= y_{ij} - \hat{y}_{ij} = y_{ij} - (\hat{\alpha}_j + \hat{\beta}_j x_{ij}) \\
\mathrm{Var}(\hat{u}_{ij}) &= \mathrm{Var}(y_{ij} - \hat{y}_{ij}) \\
&= \mathrm{Var}(y_{ij}) + \mathrm{Var}(\hat{y}_{ij}) - 2\mathrm{Cov}(y_{ij}, \hat{y}_{ij}) \\
&= \mathrm{Var}(y_{ij}) + \mathrm{Var}(\hat{y}_{ij}) - 2\mathrm{Cov}(\hat{y}_{ij} + \hat{u}_{ij}, \hat{y}_{ij}) \\
&= \mathrm{Var}(y_{ij}) - \mathrm{Var}(\hat{y}_{ij}) \\
&= \sigma_j^2 x_{ij}^{\delta} - \mathrm{Var}(\hat{\alpha}_j + \hat{\beta}_j x_{ij}) \\
&= \sigma_j^2 x_{ij}^{\delta} - [\mathrm{Var}(\hat{\alpha}_j) + x_{ij}^2 \mathrm{Var}(\hat{\beta}_j) + 2x_{ij}\mathrm{Cov}(\hat{\alpha}_j, \hat{\beta}_j)] \\
&= \sigma_j^2 x_{ij}^{\delta} - \frac{\sigma_j^2}{\mathrm{v}_j^w \Sigma_i w_{ij}^R}\left[\frac{\Sigma_i w_{ij}^R x_{ij}^2}{\Sigma_i w_{ij}^R} - 2x_{ij}\bar{x}_j^w + x_{ij}^2\right] \\
&= \sigma_j^2 x_{ij}^{\delta}\left[1 - \frac{w_{ij}^R}{\Sigma_i w_{ij}^R}\left(1 + \frac{(x_{ij} - \bar{x}_j^w)^2}{\mathrm{v}_j^w}\right)\right]
\end{aligned}
$$

which we estimate by replacing σ_j^2 by its estimate. We obtain standard errors by taking square roots.

15. Forecasts and Standard Errors for the intercept model

The forecast is $\hat{y}_{ij} = \hat{\alpha}_j + \hat{\beta}_j \hat{y}_{i,j-1}$, where, again, given the data, $\hat{y}_{ii} = y_{ii}$.

Similarly to the derivation above, the variance of the forecast splits into two parts,

$$\begin{aligned}\mathrm{Var}(\hat{y}_{i,i+k} - y_{i,i+k}) &= \mathrm{Var}(\hat{y}_{i,i+k} - \mu_{i,i+k}) + \mathrm{Var}(y_{i,i+k} - \mu_{i,i+k}) \\ &= v^p_{i,i+k} + v^e_{i,i+k}\end{aligned}$$

where

$$\begin{aligned}v^p_{i,i+k} &= \mathrm{Var}(\hat{y}_{i,i+k}) \\ &= \mathrm{Var}(\hat{\alpha}_{i+k} + \hat{\beta}_{i+k}\hat{y}_{i,i+k-1}) \\ &= \mathrm{Var}(\hat{\alpha}_{i+k}) + 2\hat{y}_{i,i+k-1}\mathrm{Cov}(\hat{\alpha}_{i+k}, \hat{\beta}_{i+k}) + \mathrm{Var}(\hat{\beta}_{i+k}\hat{y}_{i,i+k-1}) \\ &= \frac{\Sigma_i w^R_{ij} x^2_{ij}}{\Sigma_i w^R_{ij}} \mathrm{Var}(\hat{\beta}_{i+k}) - 2\bar{x}^w_{i+k}\hat{y}_{i,i+k-1}\mathrm{Var}(\hat{\beta}_{i+k}) \\ &\quad + \hat{y}^2_{i,i+k-1}\mathrm{Var}(\hat{\beta}_{i+k}) + \hat{\beta}^2_{i+k}\mathrm{Var}(\hat{y}_{i,i+k-1}) + \mathrm{Var}(\hat{\beta}_{i+k})\mathrm{Var}(\hat{y}_{i,i+k-1}) \\ &= \mathrm{Var}(\hat{\beta}_{i+k})\left[\frac{\Sigma_i w^R_{ij} x^2_{ij}}{\Sigma_i w^R_{ij}} - 2\bar{x}^w_{i+k}\hat{y}_{i,i+k-1} + \hat{y}^2_{i,i+k-1}\right] + [\hat{\beta}^2_{i+k} + \mathrm{Var}(\hat{\beta}_{i+k})]v^p_{i,i+k-1}\end{aligned}$$

We estimate this by substituting the estimated variance of β_{i+k} in for the variance above. Note that, as with the models without an intercept, v^p_{ii} is zero.

$$\begin{aligned}v^e_{i,i+k} &= \mathrm{Var}(y_{i,i+k} - \mu_{i,i+k}) \\ &= \mathrm{Var}(\alpha_{i+k} + \beta_{i+k}y_{i,i+k-1} + u_{i,i+k}) \\ &= \beta^2_{i+k}\mathrm{Var}(y_{i,i+k-1}) + \mathrm{Var}(u_{i,i+k}) \\ &= \beta^2_{i+k}\mathrm{Var}(y_{i,i+k-1} - \mu_{i,i+k-1}) + \sigma^2_{i+k}x^\delta_{i,i+k} \\ &= \beta^2_{i+k} v^e_{i,i+k-1} + \sigma^2_{i+k}x^\delta_{i,i+k}\end{aligned}$$

which we estimate as:

$$\begin{aligned}\hat{v}^e_{i,i+k} &= \widehat{\mathrm{Var}}(y_{i,i+k}) \\ &= \hat{\beta}^2_{i+k}\widehat{\mathrm{Var}}(y_{i,i+k-1}) + \hat{\sigma}^2_{i+k}\hat{\mathrm{E}}(y^\delta_{i,i+k-1}) \\ &= \hat{\beta}^2_{i+k}\hat{v}^e_{i,i+k} + \hat{\sigma}^2_{i+k}\hat{f}^\delta_{i,i+k-1}.\end{aligned}$$

with f defined as before.

16. Forecasts and Standard Errors of Development Year Totals for the Intercept model

Forecasts:

Let D_j be the unknown future development year total forecast, so:

$$D_j = \sum_{i=0}^{j-1} y_{ij}\text{, and}$$

$$\hat{D}_j = \sum_{i=0}^{j-1} \hat{y}_{ij}$$

Standard Errors:

$$\begin{aligned}
\mathrm{Var}(\hat{D}_j - D_j) &= \mathrm{Var}(\sum_{i=0}^{j-1} \hat{y}_{ij} - y_{ij}) \\
&= \mathrm{Var}(\sum_{i=0}^{j-1} \hat{y}_{ij} - \mu_{ij}) + \mathrm{Var}(\sum_{i=0}^{j-1} y_{ij} - \mu_{ij}) \\
&= V_j^p + V_j^e \quad \text{, as before.}
\end{aligned}$$

$$\begin{aligned}
V_j^p &= \mathrm{Var}(\sum_{i=0}^{j-1} \hat{y}_{ij}) \\
&= \mathrm{Var}(\sum_{i=0}^{j-1} \hat{\alpha}_j + \hat{\beta}_j \hat{y}_{i,j-1}) \\
&= \mathrm{Var}(n_j \hat{\alpha}_j + \hat{\beta}_j [\hat{D}_{j-1} + y_{j-1,j-1}]) \\
&= n_j^2 \mathrm{Var}(\hat{\alpha}_j) + 2n_j [\hat{D}_{j-1} + y_{j-1,j-1}] \mathrm{Cov}(\hat{\alpha}_j, \hat{\beta}_j) + \mathrm{Var}(\hat{\beta}_j [\hat{D}_{j-1} + y_{j-1,j-1}]) \\
&= n_j^2 \mathrm{Var}(\hat{\alpha}_j) + 2n_j [\hat{D}_{j-1} + y_{j-1,j-1}] \mathrm{Cov}(\hat{\alpha}_j, \hat{\beta}_j) \\
&\quad + [\hat{D}_{j-1} + y_{j-1,j-1}]^2 \mathrm{Var}(\hat{\beta}_j) + (\hat{\beta}_j^2 + \mathrm{Var}(\hat{\beta}_j)) \mathrm{Var}(\hat{D}_{j-1}) \\
&= n_j^2 \mathrm{Var}(\hat{\alpha}_j) + 2n_j [\hat{D}_{j-1} + y_{j-1,j-1}] \mathrm{Cov}(\hat{\alpha}_j, \hat{\beta}_j) \\
&\quad + [\hat{\beta}_j^2 + \mathrm{Var}(\hat{\beta}_j)] V_{j-1}^p + [\hat{D}_{j-1} + y_{j-1,j-1}]^2 \mathrm{Var}(\hat{\beta}_j)
\end{aligned}$$

which we may simplify further by writing the variance of the α estimate and the covariance of the α and β estimates in terms of the variance of the β estimate. We estimate this by

$$\hat{V}_j^p = n_j^2 \widehat{\mathrm{Var}}(\hat{\alpha}_j) + 2n_j[\hat{D}_{j-1} + y_{j-1,j-1}]\widehat{\mathrm{Cov}}(\hat{\alpha}_j, \hat{\beta}_j)$$
$$+ [\hat{\beta}_j^2 + \widehat{\mathrm{Var}}(\hat{\beta}_j)]\hat{V}_{j-1}^p + [\hat{D}_{j-1} + y_{j-1,j-1}]^2 \widehat{\mathrm{Var}}(\hat{\beta}_j)$$

Also,

$$\begin{aligned} V_j^e &= \mathrm{Var}(\sum_{i=0}^{j-1} y_{ij}) \\ &= \mathrm{Var}(n_j\alpha_j + \beta_j \sum_{i=0}^{j-1} y_{i,j-1} + \sum_{i=0}^{j-1} u_{i,j}) \\ &= \mathrm{Var}(\beta_j[D_{j-1} + y_{j-1,j-1}]) + \mathrm{Var}(\sum_{i=0}^{j-1} u_{i,j}) \\ &= \beta_j^2 \mathrm{Var}(D_{j-1}) + \sigma_j^2 \sum_{i=0}^{j-1} x_{ij}^\delta \\ &= \beta_j^2 V_{j-1}^e + \sigma_j^2 (y_{j-1,j-1}^\delta + \sum_{i=0}^{j-2} y_{i,j-1}^\delta). \end{aligned}$$

Due to independence across accident years,

$$\mathrm{E}(\sum_{i=0}^{j-2} y_{i,j-1}^\delta) = \sum_{i=0}^{j-2} \mathrm{E}(y_{i,j-1}^\delta) = \sum_{i=0}^{j-2} f_{i,j-1}^\delta$$

so the process variance term is estimated by

$$\hat{V}_j^e = \hat{\beta}_j^2 \hat{V}_{j-1}^e + \hat{\sigma}_j^2 (y_{j-1,j-1}^\delta + \sum_{i=0}^{j-2} \hat{f}_{i,j-1}^\delta).$$

The estimated standard error of $\hat{D}_j$ is then $\sqrt{\hat{V}_j^p + \hat{V}_j^e}$.

17. Forecasts and Standard Errors on the Paid Scale

$$\begin{aligned}
\text{Var}(\hat{p}_{ij} - p_{ij}) &= \text{Var}(\hat{y}_{ij} - y_{ij} - \hat{y}_{i,j-1} + y_{i,j-1}) \\
&= \text{Var}(\hat{y}_{ij} - y_{ij}) + \text{Var}(\hat{y}_{i,j-1} - y_{i,j-1}) - 2\text{Cov}(\hat{y}_{ij} - y_{ij}, \hat{y}_{i,j-1} - y_{i,j-1}) \\
&= \text{Var}(\hat{y}_{ij}) + \text{Var}(\hat{y}_{i,j-1}) + \text{Var}(y_{ij}) + \text{Var}(y_{i,j-1}) \\
&\quad - 2\text{Cov}(\hat{\alpha}_j + \hat{\beta}_j \hat{y}_{i,j-1}, \hat{y}_{i,j-1}) - 2\text{Cov}(\alpha_j + \beta_j y_{i,j-1}, y_{i,j-1}) \\
&= \text{Var}(\hat{y}_{ij}) + \text{Var}(\hat{y}_{i,j-1}) + \text{Var}(y_{ij}) + \text{Var}(y_{i,j-1}) \\
&\quad - 2\text{Cov}(\hat{\beta}_j \hat{y}_{i,j-1}, \hat{y}_{i,j-1}) + 0 + 0 - 2\text{Cov}(\beta_j y_{i,j-1}, y_{i,j-1}) \\
&= \text{Var}(\hat{y}_{ij}) + \text{Var}(\hat{y}_{i,j-1}) + \text{Var}(y_{ij}) + \text{Var}(y_{i,j-1}) \\
&\quad - 2\hat{\beta}_j \text{Var}(\hat{y}_{i,j-1}) - 2\beta_j \text{Var}(y_{i,j-1}) \\
&= \text{Var}(\hat{y}_{ij}) - (2\hat{\beta}_j - 1)\text{Var}(\hat{y}_{i,j-1}) + \text{Var}(y_{ij}) - (2\beta_j - 1)\text{Var}(y_{i,j-1}) \\
&= v_{ij}^p - (2\hat{\beta}_j - 1)v_{i,j-1}^p + v_{ij}^e - (2\beta_j - 1)v_{i,j-1}^e
\end{aligned}$$

just as before. To estimate these quantities we replace the V's by their estimates, and of course we take square roots to obtain standard errors.

18. Examples

Example 1: The data for the first example is taken from Mack (1994). The data is in Table 1.

<Insert Table 1 about here>

In any analysis, the first step should always be to plot the data The values corresponding to the 1982 accident year are marked. Note that the 1982 values sit below the other years in the plot against development year.

<Insert Figure 2 about here>

In order to decide which of the models discussed is appropriate, we calculate the AIC for the models with and without intercept, for values of δ of 0, 1 and 2. These figures are presented in Table 2.

<Insert Table 2 about here>

The models with the intercept term are all better than the corresponding models through the origin. The best value for δ is zero in each case. The fit for the first pair of development years for models with and without intercept with δ=0 are shown in Figure 3.

<Insert Figure 3 about here>

As can be seen, the line through the origin is a poor fit to the data.

If we fit the no intercept model, and examine the residuals, we obtain the plots in Figure 4.

<Insert Figure 4 about here>

Note that the residuals for 1982 are a little high on average. Also note the strong downward trend in the plot against fitted values, indicating that forecasts of smaller values are too low, and forecasts of higher values are too high. This is because of the way the model forces a low value to be followed by a low value and a high value by a high value, whereas with the actual data, values move up and down in a more random fashion. The model cannot capture this. There is also some indication of changing trends in the payment year direction.

Consequently, we fit the model with intercepts and $\delta = 0$. The results are presented in Table 3.

<Insert Table 3 about here>

Note that we don't fit an intercept for the last two regressions. The intercept for the first pair of years is highly significant, but the intercepts for the remaining years seem less important. This is typical of many data sets, and a better model might set some of the intercepts to zero. Note also that none of the slope parameters are significantly different from 1. This means that the previous cumulative is not really of much help in predicting the next incremental paid loss.

Recalling that the regressions are independent, we eliminate some parameters from the model. Note that sometimes when we eliminate either the slope or the intercept, the remaining parameter becomes significant at the 5% level. Consequently when both together are non-significant, we try removing them one at a time, retaining whichever of the slope or intercept parameters is more significant. We find that in each case the intercept is the parameter retained; it is only non-significant for two parameters: for years 5-6 and 7-8. These are retained, however, in order that the forecasts continue to track upward, but is an indication that these kinds of models, even when supplemented with different weighting schemes and intercept terms, aren't particularly suited to the data. For consistency, we will estimate an intercept for the last year rather than a slope - note that due to there only being a single data point, we can't get standard errors or p-values unless we get a combined estimate of the variance for the latter years, an issue we'll leave aside in these examples. The results are presented in Table 4.

<Insert Table 4 about here>

Note that the AIC is lower (746.35), due to a reduction in parameters without loss of fit. This model has a substantially lower AIC than one that fits only slopes (from Table 2, 776.5), indicating that plain weighted chain ladder models are inappropriate. Note that the model with all the intercepts estimated and all the slopes set to one is just taking the incremental forecast as the average paid in that development year. This model is related to the Cape Cod approach.

The residual plots for the reduced model are given in Figure 5. Note that the fit to 1982 is much better than it was, and there is much less trend in the plot of residuals against fitted values. The better fit has made the slight changing trend against payment years more clear. This is not too bad a fit to the data.

<Insert Figure 5 about here>

As discussed at the start of the paper, when δ is 0, 1 or 2, the formulas don't require the assumption of normality to be the best linear estimates of the parameters. However, if the data are sufficiently non-normal, any linear estimates can be very poor indeed. Consequently, some assessment of the normality of the residuals is prudent. To that end, we look at a plot of the residuals against the normal scores (expected normal order statistics). If the plot deviates substantially from a straight line, a non-normal distribution of errors is indicated.

<Insert Figure 6 about here>

As can be seen in Figure 6, the plot is quite straight, indicating that the use of the regression formulae will be appropriate. The squared correlation between the residuals and the corresponding normal scores is 0.9894. If we use this as a test statistic for a test of normality, we obtain a p-value larger than 0.5. This is a Shapiro-Francia test (Shapiro and Francia, 1972).

We proceed to forecast the paid losses. Table 5 shows the forecasts and standard errors of the cumulative paid losses.

<Insert Table 5 about here>

Note that the standard errors are generally decreasing as a percentage of the accident year forecast totals as we proceed down to the later years. This usually does not happen with models lacking any intercept terms, such as the chain ladder. This happens because the model has pooled the information across accident years. Consequently, the standard errors are substantially smaller than for the chain ladder model, though the mean forecast is only a little higher.

It is useful to also see the incremental paid losses, especially if we are interested in the future cash flow. Consequently we also examine the table of forecast incremental paid losses.

<Insert Table 6 about here>

Standard errors of the payment year totals would be an important quantity, and formulae for these, as well as for other calculations will be discussed in a subsequent paper by the authors.

Example 2: This is a real data set, but the values have all been multiplied by a scaling constant to help preserve confidentiality. The data array is presented in Table 7.

<Insert Table 7 about here>

Plots of the cumulated data in the three directions is presented in Figure 7. There is some indication of possible changing trends against accident years and perhaps even in the payment year direction. The AIC for δ = 2 is better

<Insert Figure 7 about here>

than 0 or 1 across various models. A model with all the intercepts is better than a model without intercepts (AIC = 1121.4 vs 1139.0), but only the first two intercepts are significantly different from zero. The parameter estimates for the model with $\delta = 2$ and intercepts for the first two years are presented in Table 8. Again, the AIC is a little higher for this model - the minimum AIC includes a few of the parameters not significant at the 5% level.

<Insert Table 8 about here>

Examination of the residuals for this model in Figure 8 now reveal quite strong trends in the payment year direction, making the model completely inappropriate. There is little point in forecasting with this model in the presence of these trends - the forecasts will be far too low, as they miss the fact that the inflation in the data more than doubles from the early years to the end.

<Insert Figure 8 about here>

Example 3: This is a simulated data set, generated from a known model. The paid losses are generated from:

$$\ln(p_{ij}) = \alpha + \gamma j + \varepsilon_{ij}$$

or equivalently

$$p_{ij} = cr^{j}.n_{ij}$$

where r represents a proportional decrease in payments over time - the paids follow an exponential decay, with some random variation.

Here, $\alpha = 10$, $\gamma = -.3$, and the ε's are normally distributed with a variance of 0.4.

Because the model generating the paid losses is known, "correct" forecasts and standard errors can be calculated and compared with the answers from the chain-ladder type models. The cumulated data are presented Table 9.

<Insert Table 9 about here>

We select a model using a similar approach as in the other examples. The regression table for this model is presented in Table 10, and the residual displays in Figure 9.

<Insert Table 10 about here>

The residual displays indicate that the model is a reasonable fit, though there is a little overforecasting at the highest predicted values.

<Insert Figure 9 about here>

A check of the normality, in Figure 10 reveals a skewed distribution of errors. The squared correlation is .9776, with a p-value of .027. The test has correctly picked up that the data don't come from a normal distribution. This accounts for there being more significant slope parameters than the single one we might expect if the assumptions were correct.

<Insert Figure 10 about here>

If we fit the chain ladder model and forecast it, we obtain a total forecast outstanding of 254130 with a standard error of 62672. Moreover there is substantial variation in accident year total forecasts, when all years are the same under the model that generated the data. If we forecast the fitted model above, we get a forecast of 294319 and a standard error of 39497, and if we forecast the model that generated the data we get a forecast of 284125 with a standard error of 30970. The chain ladder model underforecasts a little and has an inflated standard error because it is overparameterised and there is not enough pooling of information across years. The fitted model has a forecast that is quite close, though the standard error is still a little high, for the same reasons.

References

Akaike H. (1974). A New Look at Statistical Model Identification. *IEEE Transaction on Automatic Control* AC-19, 716-722.

Cox D.R. and D.V. Hinkley (1974). *Theoretical Statistics*, Chapman and Hall, London.

Mack T. (1993). Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin*, Vol 23 No. 2. 213-225.

Mack T. (1994). Which stochastic model is underlying the chain ladder method? *Insurance Mathematics and Economics*, Vol 15 No. 2/3. 133-138.

Murphy D. M. (1994). Unbiased Loss Development Factors. Proceedings of the Casualty Actuarial Society, Vol LXXXI No. 144-155, 154-222.

Shapiro S.S. and Francia R.S. (1972). Approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67, 215-216.

Venter G. (1996). Testing assumptions of age-to-age Factors. Draft Report, Instrat/Sedgwick Re, New York.

Figure 1. Triangular loss development array of size n, with accident years labelled in reverse order.

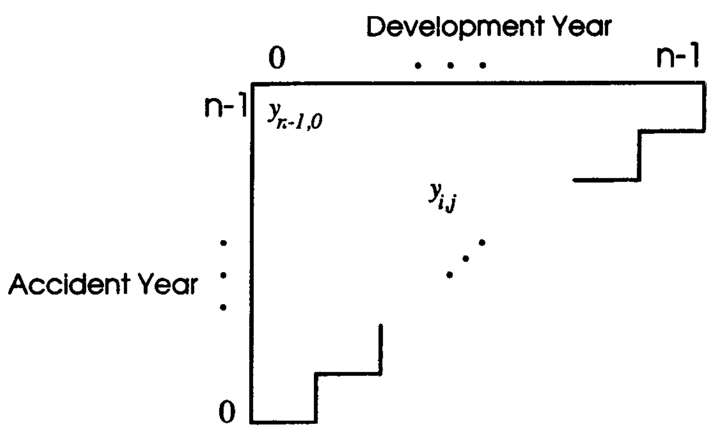


Table 1. Cumulative paid loss array for the Mack data. Rows are accident years and columns are delays.

	0	1	2	3	4	5	6	7	8	9
1981	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834
1982	106	4285	5396	10666	13782	15599	15496	16169	16704	
1983	3410	8992	13873	16141	18735	22214	22863	23466		
1984	5655	11555	15766	21266	23425	26083	27067			
1985	1092	9565	15836	22169	25955	26180				
1986	1513	6445	11702	12935	15852					
1987	557	4020	10946	12314						
1988	1351	6947	13112							
1989	3133	5395								
1990	2063									

Figure 2. Plot of cumulative paid losses against the three time directions.

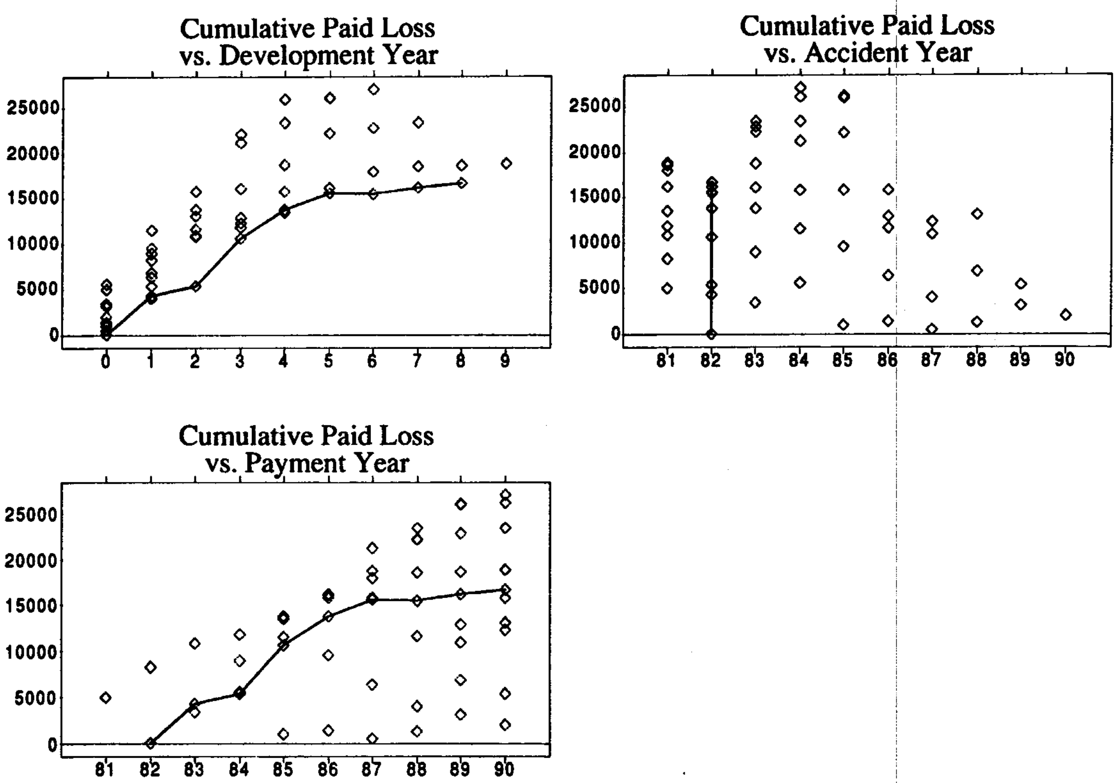


Table 2. AIC for the two main model types, at several values of δ

	Origin	Intercept
0	776.5	756.3
1	791.8	760.8
2	817.9	766.8

Figure 3. Plot showing cumulative paid losses for development year 1 against development year 0, with lines through the origin, and with an intercept term.

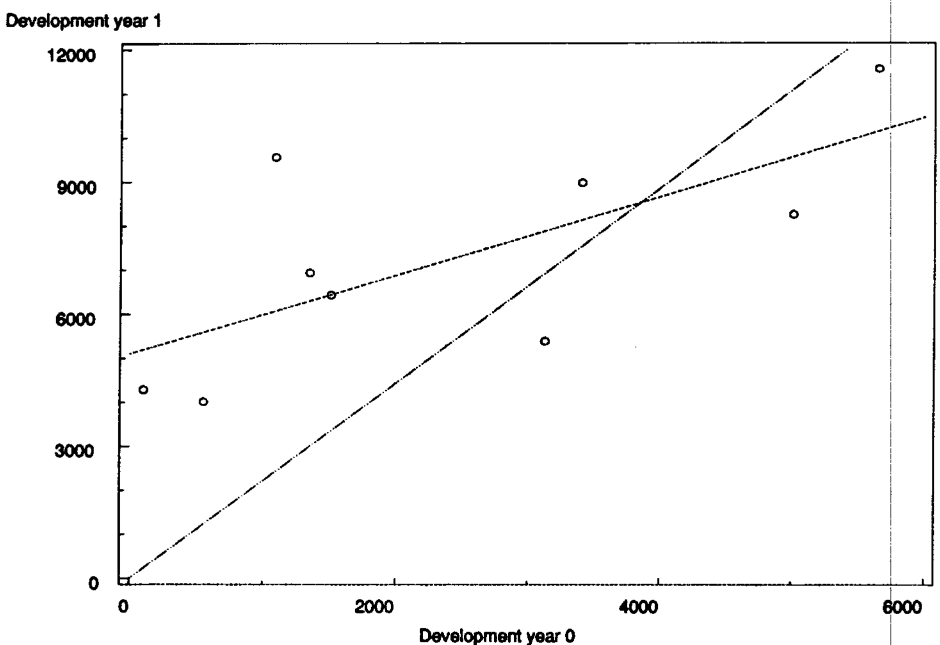


Figure 4. Residual plot for δ=0, model with no intercept. The solid line indicates values in the 1982 accident year, and the dotted line joins mean residuals

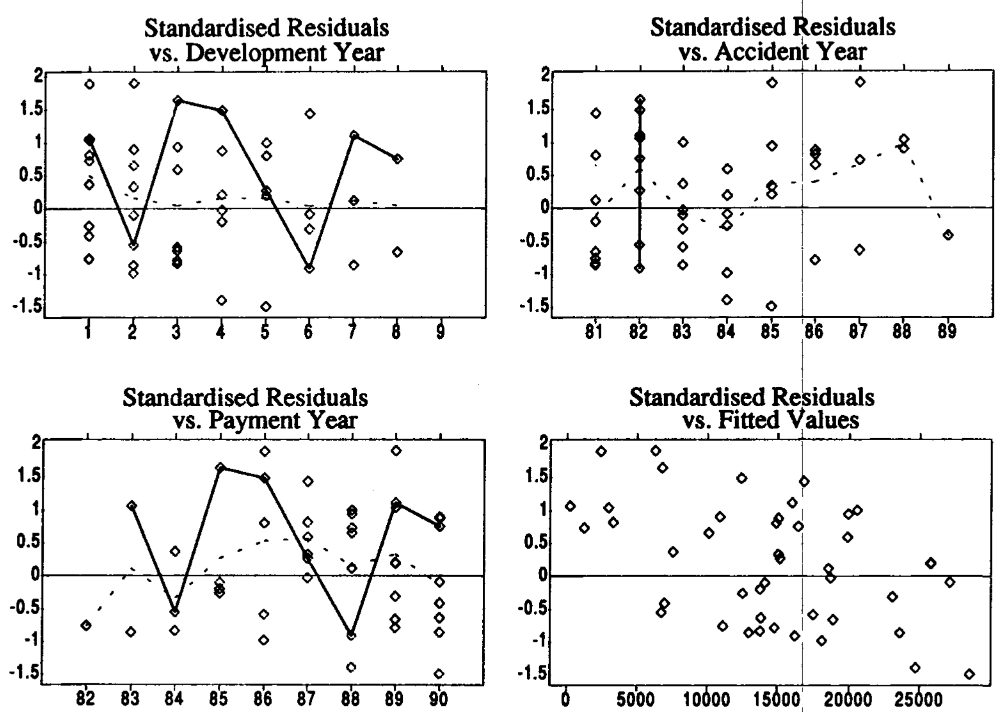


Table 3. Fit of the model with intercept and δ at 0. There is no intercept fitted for the last two years.

Link Ratio Selection - Regression Table $\delta=0$

Develop.	Intercept			Slope			
Period	Estimate	Std.Error	p value	Estimate	Slope - 1	Std. Error	p value
00-01	5113.37	1066.16	0.002	0.89114	-0.10886	0.3486	0.764
01-02	4311.47	2440.12	0.128	1.04941	0.04941	0.3091	0.878
02-03	1687.18	3543.14	0.654	1.13100	0.13100	0.2831	0.663
03-04	2061.07	1164.74	0.152	1.04148	0.04148	0.0708	0.589
04-05	4064.46	2241.92	0.167	0.90044	-0.09956	0.1136	0.445
05-06	620.43	2300.87	0.813	1.01094	0.01094	0.1123	0.931
06-07	777.33	144.68	0.117	0.99189	-0.00811	0.0076	0.479
07-08	-	-	-	1.01589	0.01589	0.0149	0.240
08-09	-	-	-	1.00922	0.00922	-	-

(AIC=756.29)

Table 4. Fit of the model with δ at 0, slopes set to 1 and all intercepts estimated.

Link Ratio Selection - Regression Table δ=0

Develop.	Intercept			Slope (Link Ratio)			
Period	Estimate	Std. Error	p value	Estimate	Estimate - 1	Std. Error	p value
00-01	4849.33	611.66	0.000	1	0	0	-
01-02	4682.50	697.98	0.000	1	0	0	-
02-03	3267.14	883.07	0.010	1	0	0	-
03-04	2717.67	296.35	0.000	1	0	0	-
04-05	2164.2	551.45	0.017	1	0	0	-
05-06	839.50	400.27	0.127	1	0	0	-
06-07	625.00	24.03	0.001	1	0	0	-
07-08	294.50	240.50	0.436	1	0	0	-
08-09	172.00	-	-	1	0	0	-

(AIC=746.35)

Figure 5. Residual plot for δ=0, model with intercepts and with slopes set to 1. The line joins mean residuals.

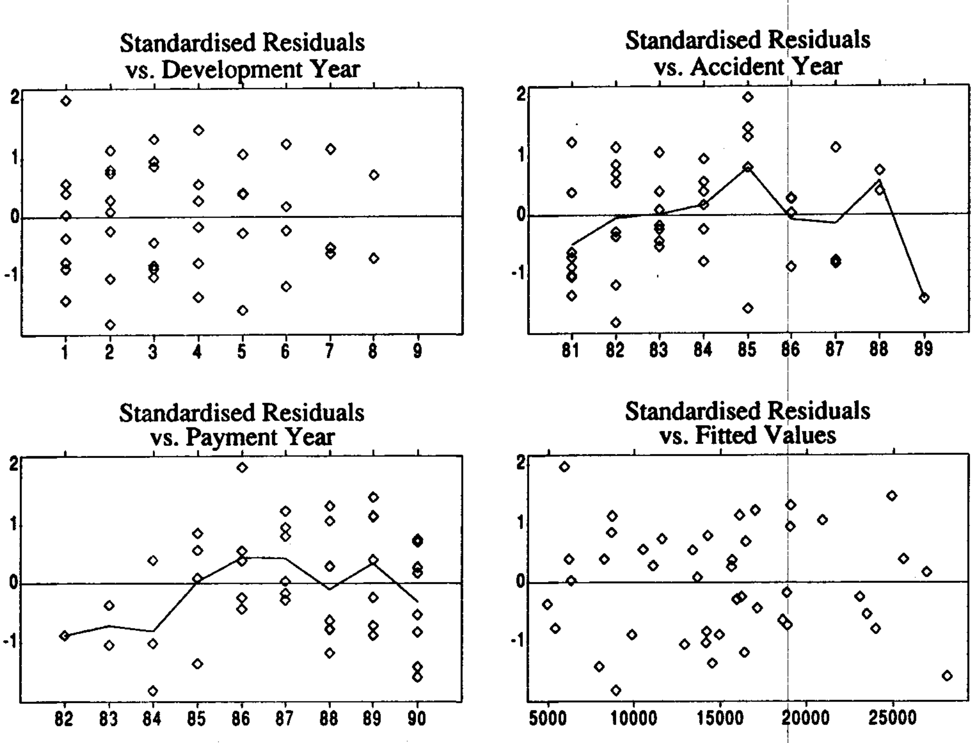


Figure 6. Normal Scores plot against standardised residuals.

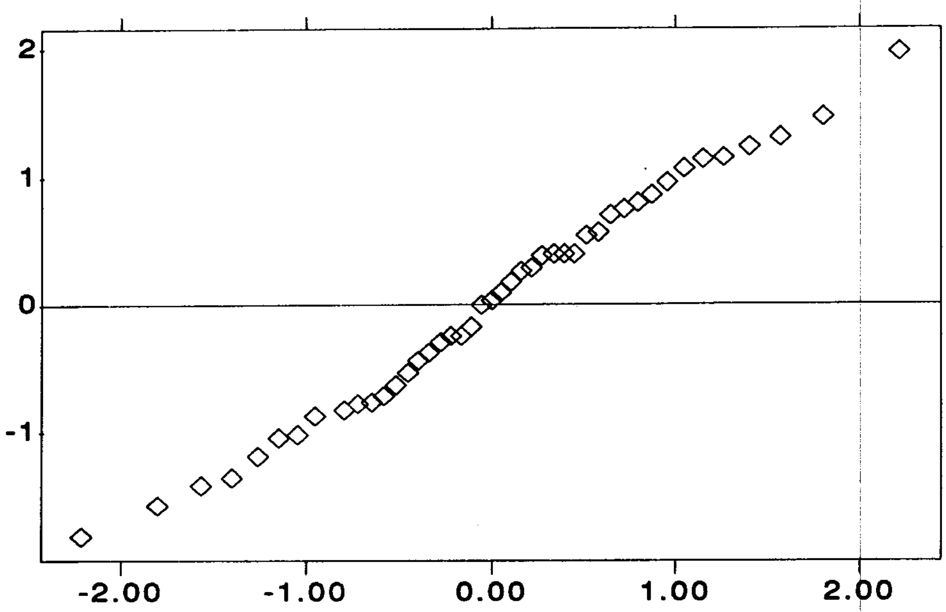


Table 5. Cumulative forecasts for the model with intercepts and slopes set to 1. Above the heavy line, the upper number is the paid amount, the lower number is the model prediction. Below the line, the upper value is the forecast, the lower value is the standard error. The row totals are accident year outstandings, the column totals are cumulative development year totals.

	0	1	2	3	4	5	6	7	8	9	Total
1981	5012	9861	12952	14174	14523	15703	17020	18634	18902	18834	0
	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834	0
1982	106	4955	8968	8663	13384	15946	16438	16121	16464	16876	172
	106	4285	5396	10666	13782	15599	15496	16169	16704	42	42
1983	3410	8259	13674	17140	18859	20899	23054	23488	23760	23932	466
	3410	8992	13873	16141	18735	22214	22863	23466	417	419	419
1984	5655	10504	16238	19033	23984	25589	26922	27692	27986	28158	1092
	5655	11555	15766	21266	23425	26083	27067	48	419	421	421
1985	1092	5941	14248	19103	24887	28119	27020	27644	27939	28111	1931
	1092	9565	15836	22169	25955	26180	895	896	988	989	989
1986	1513	6362	11128	14969	15653	18016	18856	19481	19775	19947	4095
	1513	6445	11702	12935	15852	1351	1620	1621	1674	1674	1674
1987	557	5406	8702	14213	15032	17196	18035	18660	18955	19127	6813
	557	4020	10946	12314	784	1562	1800	1801	1848	1849	1849
1988	1351	6200	11630	16379	19097	21261	22101	22726	23020	23192	10080
	1351	6947	13112	2498	2618	2946	3079	3079	3107	3107	3107
1989	3133	7982	10078	13345	16062	18227	19066	19691	19986	20158	14763
	3133	5395	2094	3259	3352	3614	3723	3724	3747	3747	3747
1990	2063	6912	11595	14862	17580	19744	20583	21208	21503	21675	19612
	2063	1934	2851	3790	3870	4099	4196	4196	4217	4217	4217
Total	-	6912	21672	44586	67770	94443	125660	157102	182924	201176	59023
St.Err	-	1934	3396	5283	5479	6134	6440	6440	6512	6513	6513

Table 6. Incremental (paid scale) forecasts for the model with intercepts fitted and slopes set to 1. Above the heavy line, the upper number is the paid amount, the lower number is the model prediction. Below the line, the upper value is the forecast, the lower value the standard error. The row totals are accident year total forecasts, the column totals are payment year totals.

	0	1	2	3	4	5	6	7	8	9	Total
1981	5012	4849	4682	3267	2718	2164	840	625	294	172	0
	5012	3257	2638	898	1734	2642	1828	599	54	172	0
1982	106	4849	4682	3267	2718	2164	840	625	294	172	172
	106	4179	1111	5270	3116	1817	-103	673	535	42	42
1983	3410	4849	4682	3267	2718	2164	840	625	294	172	466
	3410	5582	4881	2268	2594	3479	649	603	417	42	419
1984	5655	4849	4682	3267	2718	2164	840	625	294	172	1092
	5655	5900	4211	5500	2159	2658	984	48	417	42	421
1985	1092	4849	4682	3267	2718	2164	840	625	294	172	1931
	1092	8473	6271	6333	3786	225	895	48	417	42	989
1986	1513	4849	4682	3267	2718	2164	840	625	294	172	4095
	1513	4932	5257	1233	2917	1351	895	48	417	42	1674
1987	557	4849	4682	3267	2718	2164	840	625	294	172	6813
	557	3463	6926	1368	784	1351	895	48	417	42	1849
1988	1351	4849	4682	3267	2718	2164	840	625	294	172	10080
	1351	5596	6165	2498	784	1351	895	48	417	42	3107
1989	3133	4849	4682	3267	2718	2164	840	625	294	172	14763
	3133	2262	2094	2498	784	1351	895	48	417	42	3747
1990	2063	4849	4682	3267	2718	2164	840	625	294	172	19612
	2063	1934	2094	2498	784	1351	895	48	417	42	4217
Total	-	19612	14763	10080	6813	4095	1931	1092	466	172	59023
	-	-	-	-	-	-	-	-	-	-	6513

Table 7. Cumulative paid loss array for the second example

	0	1	2	3	4	5	6	7	8	9	10
1977	153638	342050	476584	564040	624388	666792	698030	719282	735904	750344	762544
1978	178536	404948	563842	668528	739976	787966	823542	848360	871022	889022	
1979	210172	469340	657728	780802	864182	920268	958764	992532	1019932		
1980	211448	464930	648300	779340	858334	918566	964134	1002134			
1981	219810	486114	680764	800862	888444	951194	1002194				
1982	205654	458400	635906	765428	862214	944614					
1983	197716	453124	647772	790100	895700						
1984	239784	569026	833828	1024228							
1985	326304	793048	1173448								
1986	420778	1011178									
1987	496200										

Figure 7. Plot of cumulative paid losses (in $000's) against the three time directions for the second example.

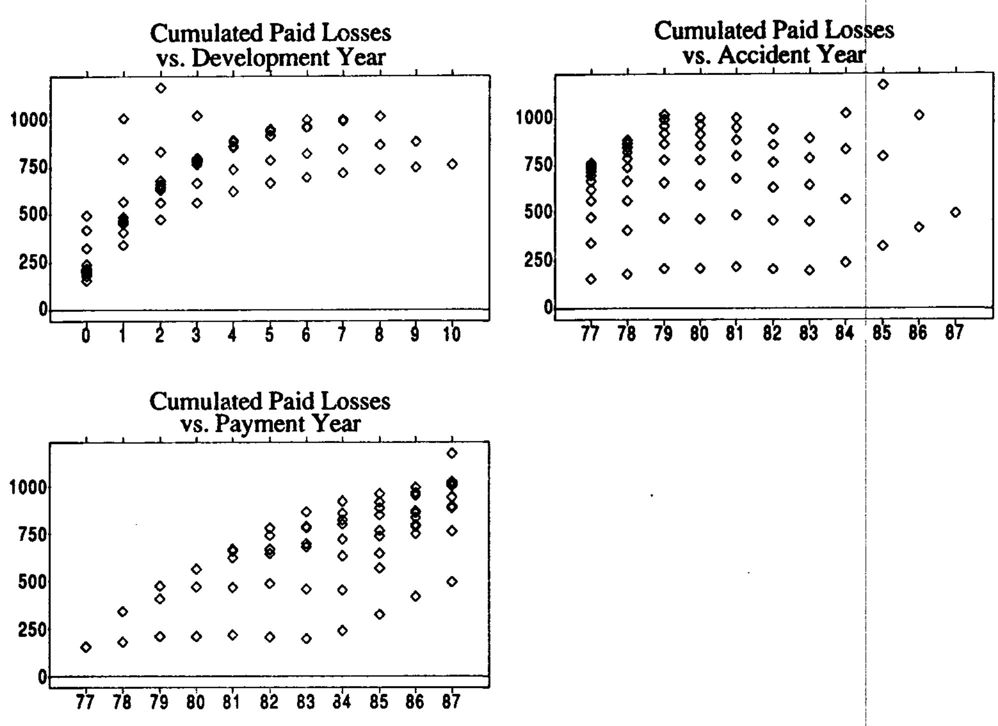


Table 8. Fit of the model with $\delta = 2$ and intercepts between the first two pairs pair of development years.

Link Ratio Selection - Regression Table δ=2

Develop.	Intercept			Slope (Link Ratio)			
Period	Estimate	Std. Error	p value	Estimate	Estimate - 1	Std. Error	p value
00-01	-56437.4	17429.24	0.012	2.54586	1.54586	0.082	0.000
01-02	-55141.5	16877.34	0.014	1.53215	0.53215	0.0366	0.000
02-03	-	-	-	1.19832	0.19832	0.0065	0.000
03-04	-	-	-	1.11307	0.11307	0.0045	0.000
04-05	-	-	-	1.07234	0.07234	0.0048	0.000
05-06	-	-	-	1.04741	0.04741	0.0020	0.000
06-07	-	-	-	1.03380	0.03380	0.0022	0.000
07-08	-	-	-	1.02581	0.02581	0.0014	0.001
08-09	-	-	-	1.02014	0.02014	0.0005	0.008
09-10	-	-	-	1.01626	0.01626	0	-

(AIC=1126.2)

Figure 8. Residual plot for δ=2, model with intercept between the first two pairs of development years. The line joins mean residuals.

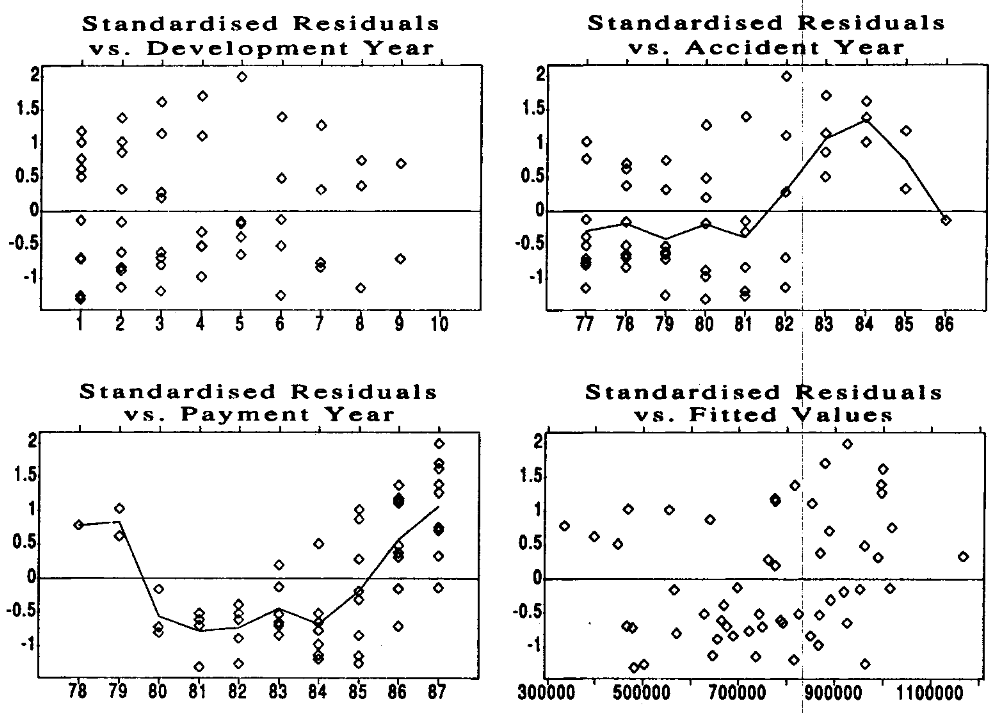


Table 9. Cumulated paid loss array for the simulated example.

	0	1	2	3	4	5	6	7	8
1978	24307	68567	72467	85860	103591	109393	111368	117662	122420
1979	19122	26125	34272	37144	42783	45750	49205	51569	55574
1980	18082	45790	73691	79390	88687	91586	99047	105932	107786
1981	80451	91862	160489	167192	173622	176315	181875	182722	187428
1982	49099	56243	68222	71703	73982	79957	84429	86495	87004
1983	33475	88192	96966	101825	107633	128383	130286	132735	133672
1984	23070	72624	78283	88192	97315	106120	113128	121762	124815
1985	14324	22676	30631	38723	44767	53309	61009	63858	64988
1986	58785	75618	81686	86913	90189	106781	109188	110096	112604
1987	9017	17016	27812	40549	44429	50965	57744	60080	
1988	12205	37185	51020	60901	65879	69259	76364		
1989	17883	23077	34506	37275	46815	51922			
1990	25584	40052	52595	58369	68783				
1991	49089	71603	95678	105269					
1992	24064	73336	76567						
1993	17858	37547							
1994	24869								

	9	10	11	12	13	14	15	16
1978	123949	126332	126493	127253	127698	128338	128839	129094
1979	57603	58973	59691	60123	61405	61861	62093	
1980	110467	111259	112597	113417	114119	114298		
1981	188673	189962	190679	191433	191869			
1982	89349	90104	91806	92402				
1983	134468	135914	137185					
1984	125308	127416						
1985	66742							

Table 10. Fit of the model with δ = 2 and non-significant parameters removed.

Link Ratio Selection - Regression Table δ=2

Develop.	Intercept			Slope (Link Ratio)			
Period	Estimate	Std. Error	p value	Estimate	Estimate - 1	Std. Error	p value
00-01	-	-	-	2.00358	1.00358	0.1874	0.000
01-02	11467.02	1930.74	0	1	0	0	-
02-03	7445.34	1076.94	0	1	0	0	-
03-04	6834.29	897.78	0	1	0	0	-
04-05	-	-	-	1.10166	0.10166	0.0182	0.000
05-06	5523.82	638.15	0	1	0	0	-
06-07	-	-	-	1.03938	0.03938	0.0078	0.000
07-08	2573.45	525.41	0.001	1	0	0	-
08-09	1837.34	208.28	0	1	0	0	-
09-10	1356.77	202.09	0.001	1	0	0	-
10-11	979.47	211.93	0.006	1	0	0	-
11-12	-	-	-	1.00619	0.00619	0.0006	0.000
12-13	1770.99	216.62	0.015	1	0	0	-
13-14	431.24	97.83	0.048	1	0	0	-
14-15	-	-	-	1.00383	0.00383	0.0001	0.007
15-16	-	-	-	1.00198	0.00198	0	-

(AIC=2516.8)

Figure 9. Residual plot for the chosen model. The line joins mean residuals.

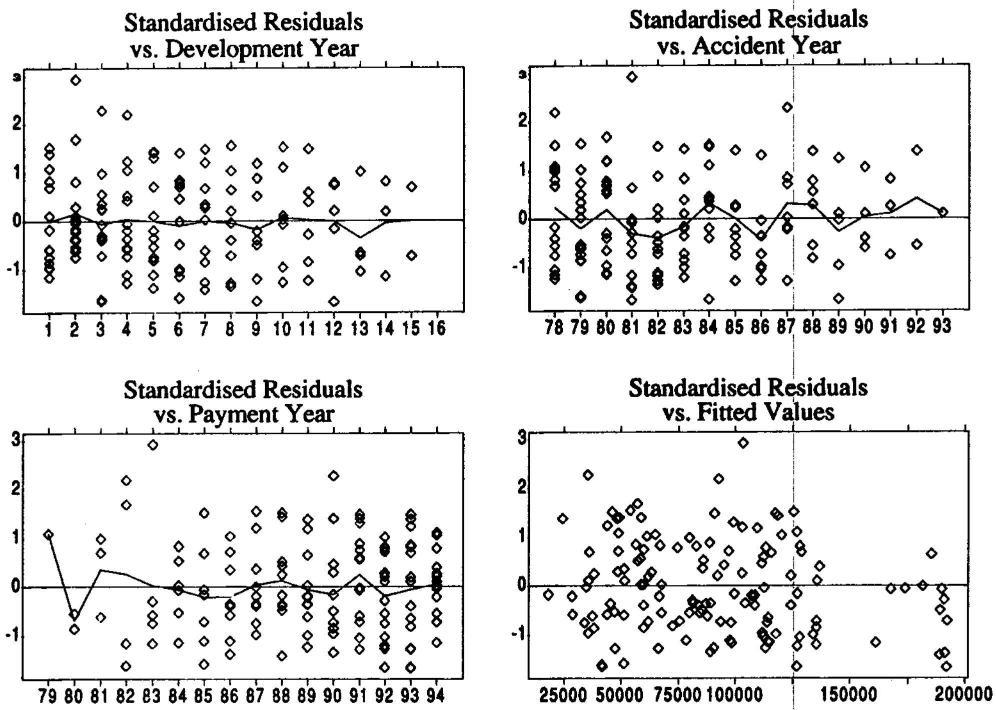
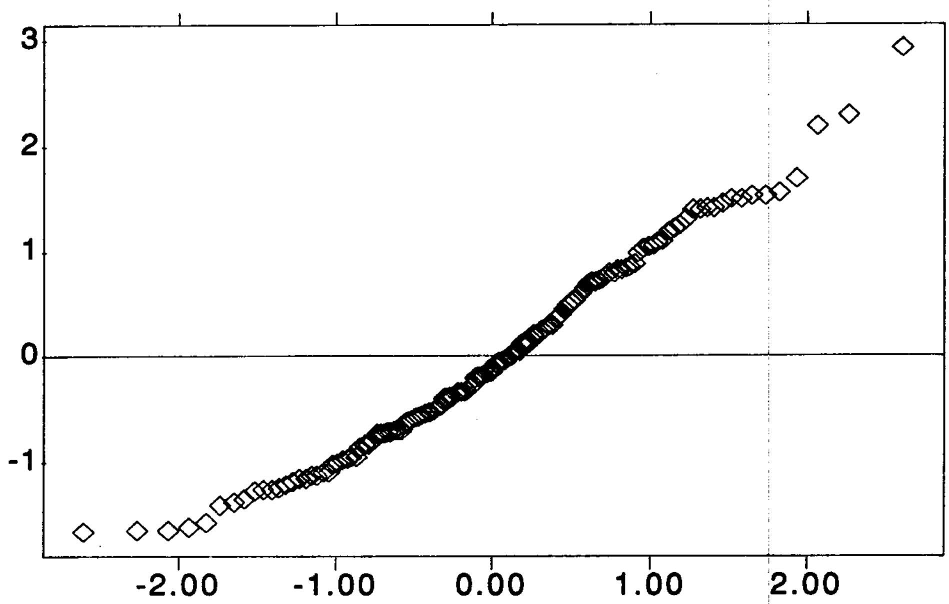


Figure 10. Normal scores plot for the chosen model.



```
# ************************************************************************************
# *                                                                                  *
# *          SPLUS CODE FOR THE CALCULATION OF DEVELOPMENT FACTORS,                  *
# *                 DIAGNOSTIC DISPLAYS AND FORECASTING TABLES                       *
# *                                                                                  *
# *     FUNCTION:     delsig(delta, model, cutoff)                                   *
# *                                                                                  *
# *     INPUT:        delta can be 0,1,2                                             *
# *                   model can be:                                                  *
# *                        0 => no intercept;                                        *
# *                        1 => intercept; or                                        *
# *                        2 => optimal model.                                       *
# *                   OPTIONAL:    cutoff is the number of development periods that  *
# *                                are cut-off at the right. By default cutoff = 0   *
# *                                                                                  *
# *     OUTPUT:       Regression table containing parameter estimates, standard errors and p-values *
# *                   Four forecast tables that contain                              *
# *                        1 observed and forecast values for cumulative data (CD)   *
# *                        2 expected values and standard errors of the forecasts for CD *
# *                        3 observed and forecast values for incremental paid losses (IPL) *
# *                        4 expected values and standard errors of the forecasts for IPL *
# *                   Four residual displays contained within one window             *
# *                        1 versus Development periods;                             *
# *                        2 versus Accident periods;                                *
# *                        3 versus Payment periods; and                             *
# *                        4 versus Fitted Values.                                   *
# *                   Normality Plot                                                 *
# *                   Box-plot                                                       *
# *                                                                                  *
# ************************************************************************************

delsig_function(delta, model, cutoff=0)
{
if ((delta < 0) || (delta > 2))
        stop(message="Delta must take the value of 0, 1 or 2")
if ((model < 0) || (model > 2))
        stop(message="Model must take the value of 0, 1 or 2")
if (cutoff < 0)
        stop(message="Cutoff must be positive")
# Read in and construct ocl Matrix
c1_scan("c:/prudmarg/abc.dat")                          # read data from file
Length_0
for(i in 0:cutoff)
        Length_Length+i
n_(sqrt(8*(length(c1)+Length)+1)-1)/2                   # n is the triangles dimensions
n_as.integer(round(n))

ObsFor_matrix(NA, n, n-cutoff, T,list(paste("A",1:n),paste("D",0:(n-1-cutoff))))
PLObsFor_matrix(NA, n, n-cutoff, T,list(paste("A",1:n),paste("D",0:(n-1-cutoff))))
forecast_matrix(NA, n, n-cutoff, T, list(paste("A",1:n),paste("D",0:(n-1-cutoff))))
PLforecast_matrix(NA, n, n-cutoff, T, list(paste("A",1:n),paste("D",0:(n-1-cutoff))))
counter_0
if (cutoff < 0)
        cutoff_0
tempCO_cutoff
for(i in 1:n)
{
        if(i<=cutoff)
```

```
	{
		for(j in 1:(n+1-i-tempCO))
		{
			counter_counter+1
			ObsFor[i,j]_c1[counter]
			if(j==1)
				forecast[i,j]_ObsFor[i,j]
		}
		tempCO_tempCO-1
	}
	else
	{
		for(j in 1:(n+1-i))
		{
			counter_counter+1
			ObsFor[i,j]_c1[counter]
			if(j==1)
				forecast[i,j]_ObsFor[i,j]
		}
	}
}
PLObsFor[,1]_ObsFor[,1]
ratio_matrix(0, n, n-1-cutoff, T,list(paste("Year",1:n), paste("Ratio",1:(n-1-cutoff))))
meanr_1:(n-1-cutoff)
for(i in 1:(n-1-cutoff))
{
	ratio[,i]_ObsFor[,i+1]/ObsFor[,i]					# individual ratios
	meanr[i]_mean(ratio[1:(n-i),i])					# mean of the ratios  ie, Chain Ladder
}
y_(ObsFor[1:n,2:(n-cutoff)])/(ObsFor[1:n,1:(n-1-cutoff)]^(delta/2))		# y
x1_1/(ObsFor[1:n,1:(n-1-cutoff)]^(delta/2))				# x1 (used for Alpha)
x2_(ObsFor[1:n,1:(n-1-cutoff)])/(ObsFor[1:n,1:(n-1-cutoff)]^(delta/2))	# x2 (used for Beta)
x_ObsFor[1:n,1:(n-1-cutoff)]						# x is used for fitted values
res_matrix(NA,n-1,n-1-cutoff, T)					# Residuals
stdres_matrix(NA,n-1,n-1-cutoff, T)					# Standised Residuals
fit_matrix(rep(0,n*(n-cutoff)), n, n-cutoff, T)			# Fitted Values
results_matrix(0,3,(n-1-cutoff),T)
stddev_matrix(0,1,(n-1-cutoff),T)
regrout_matrix(0,2,4,byrow=T)					# Regression output
CovCoeff_matrix(0,1,(n-1-cutoff),T)					# Covariances of Coefficients
VarOfCoeff_matrix(0,2,2,byrow=T)					# Variances of Coefficients
icrfsout_matrix(0,(n-1-cutoff),6,byrow=T,
	list(paste(1:(n-1-cutoff)), c("int.","st.err","p-val","slope","st.err","p-val")))
# Performing Regressions
for(i in 1:(n-3))
{
	regress.ls_lsfit(cbind(x1[1:(n-i),i],x2[1:(n-i),i]),y[1:(n-i),i], intercept=F)
	regress.print_ls.print(regress.ls,4,F)
	regress.diag_ls.diag(regress.ls)
	regrout_regress.print$coef.table					# Coefficient table
	VarOfCoeff_regress.diag$cov.unscaled*((regress.diag$std.dev)^2)
	CovCoeff[i]_VarOfCoeff[1,2]
	if (model == 0) regrout[1,4]_1
	if ((model != 1) && (regrout[1,4] > 0.05))
	{
		regress.ls_lsfit(cbind(x2[1:(n-i),i]),y[1:(n-i),i], intercept=F)
		regress.print_ls.print(regress.ls,4,F)
		regress.diag_ls.diag(regress.ls)
```

```
                regrout_regress.print$coef.table                    # Matrix of output
                icrfsout[i,4]_regrout[1,1]                          # Beta
                icrfsout[i,5]_regrout[1,2]                          # Beta Std. Error
                icrfsout[i,6]_regrout[1,4]                          # Beta p-value
                stddev[i]_regress.diag$std.dev                      # Standard deviation
        }
        else
        {
                icrfsout[i,1]_regrout[1,1]                          # Alpha
                icrfsout[i,2]_regrout[1,2]                          # Alpha Std. Error
                icrfsout[i,3]_regrout[1,4]                          # Alpha p-value
                icrfsout[i,4]_regrout[2,1]                          # Beta
                icrfsout[i,5]_regrout[2,2]                          # Beta Std. Error
                icrfsout[i,6]_regrout[2,4]                          # Beta p-value
                stddev[i]_regress.diag$std.dev                      # Standard deviation
        }
        for(j in 1:(n-i))
        {
                res[j,i]_regress.ls$res[j]                          # Residuals
                stdres[j,i]_(regress.ls$res[j])/stddev[i]           # Standised Residuals
                fit[j,i]_icrfsout[i,1]+icrfsout[i,4]*x[j,i]         # Fitted Values
                forecast[j,i+1]_fit[j,i]                            # Fit. Val. for Matrix
        }
}
if (cutoff <=1)
{
        regress.ls_lsfit(cbind(x2[1:2,n-2]),y[1:2,n-2], intercept=F)
        regress.print_ls.print(regress.ls,4,F)
        regress.diag_ls.diag(regress.ls)
        regrout_regress.print$coef.table                            # Coefficient table
        icrfsout[n-2,4]_regrout[1,1]
        icrfsout[n-2,5]_regrout[1,2]
        icrfsout[n-2,6]_regrout[1,4]
        stddev[n-2]_regress.diag$std.dev                            # Standard deviation
        for(j in 1:2)
        {
                res[j,n-2]_regress.ls$res[j]                        # Residuals
                fit[j,n-2]_icrfsout[n-2,4]*x[j,n-2]                 # Fitted Values
                forecast[j,n-1]_fit[j,n-2]                          # Fit. Val. for Matrix
        }
        if (cutoff == 0)
        {
                icrfsout[n-1,4]_ObsFor[1,n]/ObsFor[1,n-1]           # Coefficient
                forecast[1,n]_icrfsout[n-1,4]*ObsFor[1,n-1]         # Fit. Val. for Matrix
                if (stddev[n-3]^2 < stddev[n-2]^2)
                        stddev[n-1]_ stddev[n-3]^2
                else
                        stddev[n-1]_ stddev[n-2]^2
                TempStd_(stddev[n-2]^4/stddev[n-3]^2)
                if (TempStd < stddev[n-1])
                        stddev[n-1]_TempStd
                stddev[n-1]_sqrt(stddev[n-1])                       # Standard deviation
        }
}
for(i in 2:n-cutoff)
{
        PLforecast[,i]_forecast[,i] - ObsFor[,i-1]
}
```

```
PLforecast[,1]_forecast[,1]
# Forecasting
for(i in (2+cutoff):n)
{
	first_1
	PrevPara_0
	PrevProc_0
	for(j in (n-i+1):(n-1-cutoff))
	{
		ObsFor[i,j+1]_icrfsout[j,1]+icrfsout[j,4]*ObsFor[i,j]
		if (first == 1)
		{
			if ((model == 0)||(icrfsout[j,1] == 0))
			{
				parameter_(ObsFor[i,j]^2)*(icrfsout[j,5]^2)
			}
			else if (model>0)
			{
				parameter_ (icrfsout[j,2]^2 + 2*ObsFor[i,j]*CovCoeff[j]
					+ (ObsFor[i,j]^2)*(icrfsout[j,5]^2))
			}
			process_(ObsFor[i,j]^delta)*(stddev[j]^2)
			forecast[i,j+1]_(process+parameter)^(0.5)
			PLforecast[i,j+1]_(process+parameter)^(0.5)
			first_0
			PrevPara_parameter
			PrevProc_process
		}
		else
		{
			if ((model == 0)||(icrfsout[j,1] == 0))
				parameter_ ((ObsFor[i,j]^2)*(icrfsout[j,5]^2)
					+ parameter*(icrfsout[j,4]^2 + icrfsout[j,5]^2))
			else if(model>0)
				parameter_ (icrfsout[j,2]^2 + 2*ObsFor[i,j]*CovCoeff[j]
					+ (ObsFor[i,j]^2)*(icrfsout[j,5]^2)
					+ parameter*(icrfsout[j,4]^2 + icrfsout[j,5]^2))
			if(delta==0)
				fvalue_1
			else if(delta==1)
				fvalue_ObsFor[i,j]
			else if(delta==2)
				fvalue_ObsFor[i,j]^2 + process
			else
				fvalue_1			#ERROR
			process_ (icrfsout[j,4]^2)*process + (stddev[j]^2)*fvalue
			forecast[i,j+1]_ (process+parameter)^0.5
			PLforecast[i,j+1]_((process-(2*icrfsout[j,4] - 1)*PrevProc
					+parameter-(2*icrfsout[j,4] - 1)*PrevPara)^0.5)
			PrevPara_parameter
			PrevProc_process
		}
	}
}
for(i in 2:n-cutoff)
{
	PLObsFor[,i]_ObsFor[,i] - ObsFor[,i-1]
}
```

```
PLObsFor[,1]_ObsFor[,1]
cat("\n\n\t\t Regression Table \n\t\t=================\n\n")
options(digits=5)
print(icrfsout)
cat("\n\n\t\t Observed and Forecasts\n\t\t======================\n\n")
print(ObsFor)
cat("\n\n\t\t Expected and Std. Errs.\n\t\t=======================\n\n")
print(forecast)
cat("\n\n\t\t Paid Losses Obs. and Forecasts\n\t\t==============================\n\n")
print(PLObsFor)
cat("\n\n\t\t Paid Losses Exp. and Std. Errs.\n\t\t===============================\n\n")
print(PLforecast)
# Diagnostic Displays
vecfitted_fit[1:(n-1),1]
vecstdres_stdres[1:(n-1),1]
dev_rep(1,(n-1))
acc_1:(n-1)
pay_1:(n-1)
for(k in 2:(n-1-cutoff))
{
        vecfitted_c(vecfitted,fit[1:(n-k),k])            # Fitted Values
        vecstdres_c(vecstdres,stdres[1:(n-k),k])         # Std. Res.
        dev_c(dev,rep(k,n-k))                            # Development Year
        acc_c(acc,1:(n-k))                               # Accident Year
        pay_c(pay,k:(n-1))                               # Payment Year
}
# Residual Displays
win.graph()
        par(mfrow=c(2,2))
        plot( dev, vecstdres, main="Wtd. Std. Res. vs Dev. Yrs", ylab="Wtd. Std. Res.", xlab="Dev. Yr")
        plot( acc, vecstdres, main="Wtd. Std. Res. vs Acc. Yrs", ylab="Wtd. Std Res.", xlab="Acc. Yr")
        plot( pay, vecstdres, main="Wtd. Std. Res. vs Pay. Yrs", ylab="Wtd. Std. Res.", xlab="Pay. Yr")
        plot( vecfitted, vecstdres, main="Wtd. Std. Res. vs Fitted", ylab="Wtd. Std. Res.", xlab="Fitted")
# Box-plot Display
win.graph()
        boxplot(vecstdres, main="Box-plot")
# Normality Display
win.graph()
        qqnorm(vecstdres, main="Normality plot")
stop()
}
```

RESEARCH PAPER SERIES

No.	Date	Subject	Author
1	MAR 93	AUSTRALIAN SUPERANNUATION : THE FACTS, THE FICTION, THE FUTURE	David M Knox
2	APR 93	AN EXPONENTIAL BOUND FOR RUIN PROBABILITIES	David C M Dickson
3	APR 93	SOME COMMENTS ON THE COMPOUND BINOMIAL MODEL	David C M Dickson
4	AUG 93	RUIN PROBLEMS AND DUAL EVENTS	David CM Dickson Alfredo D Egidio dos Reis
5	SEP 93	CONTEMPORARY ISSUES IN AUSTRALIAN SUPERANNUATION - A CONFERENCE SUMMARY	David M Knox John Piggott
6	SEP 93	AN ANALYSIS OF THE EQUITY INVESTMENTS OF AUSTRALIAN SUPERANNUATION FUNDS	David M Knox
7	OCT 93	A CRITIQUE OF DEFINED CONTRIBUTION USING A SIMULATION APPROACH	David M Knox
8	JAN 94	REINSURANCE AND RUIN	David C M Dickson Howard R Waters
9	MAR 94	LIFETIME INCOME, TAXATION, EXPENDITURE AND SUPERANNUATION (LITES): A LIFE-CYCLE SIMULATION MODEL	Margaret E Atkinson John Creedy David M Knox
10	FEB 94	SUPERANNUATION FUNDS AND THE PROVISION OF DEVELOPMENT/VENTURE CAPITAL: THE PERFECT MATCH? YES OR NO	David M Knox
11	JUNE 94	RUIN PROBLEMS: SIMULATION OR CALCULATION?	David C M Dickson Howard R Waters
12	JUNE 94	THE RELATIONSHIP BETWEEN THE AGE PENSION AND SUPERANNUATION BENEFITS, PARTICULARLY FOR WOMEN	David M Knox
13	JUNE 94	THE COST AND EQUITY IMPLICATIONS OF THE INSTITUTE OF ACTUARIES OF AUSTRALIA PROPOSED RETIREMENT INCOMES STRATEGY	Margaret E Atkinson John Creedy David M Knox Chris Haberecht
14	SEPT 94	PROBLEMS AND PROSPECTS FOR THE LIFE INSURANCE AND PENSIONS SECTOR IN INDONESIA	Catherine Prime David M Knox

15	OCT 94	PRESENT PROBLEMS AND PROSPECTIVE PRESSURES IN AUSTRALIA'S SUPERANNUATION SYSTEM	David M Knox
16	DEC 94	PLANNING RETIREMENT INCOME IN AUSTRALIA: ROUTES THROUGH THE MAZE	Margaret E Atkinson John Creedy David M Knox
17	JAN 95	ON THE DISTRIBUTION OF THE DURATION OF NEGATIVE SURPLUS	David C M Dickson Alfredo D Egidio dos Reis
18	FEB 95	OUTSTANDING CLAIM LIABILITIES: ARE THEY PREDICTABLE?	Ben Zehnwirth
19	MAY 95	SOME STABLE ALGORITHMS IN RUIN THEORY AND THEIR APPLICATIONS	David C M Dickson Alfredo D Egidio dos Reis Howard R Waters
20	JUN 95	SOME FINANCIAL CONSEQUENCES OF THE SIZE OF AUSTRALIA'S SUPERANNUATION INDUSTRY IN THE NEXT THREE DECADES	David M Knox
21	JUN 95	MODELLING OPTIMAL RETIREMENT IN DECISIONS IN AUSTRALIA	Margaret E Atkinson John Creedy
22	JUN 95	AN EQUITY ANALYSIS OF SOME RADICAL SUGGESTIONS FOR AUSTRALIA'S RETIREMENT INCOME SYSTEM	Margaret E Atkinson John Creedy David M Knox
23	SEP 95	EARLY RETIREMENT AND THE OPTIMAL RETIREMENT AGE	Angela Ryan
24	OCT 95	APPROXIMATE CALCULATION OF MOMENTS OF RUIN RELATED DISTRIBUTIONS	David C M Dickson
25	DEC 95	CONTEMPORARY ISSUES IN THE ONGOING REFORM OF THE AUSTRALIAN RETIREMENT INCOME SYSTEM	David M Knox
26	FEB 96	THE CHOICE OF EARLY RETIREMENT AGE AND THE AUSTRALIAN SUPERANNUATION SYSTEM	Margaret E Atkinson John Creedy
27	FEB 96	PREDICTIVE AGGREGATE CLAIMS DISTRIBUTIONS	David C M Dickson Ben Zehnwirth
28	FEB 96	THE AUSTRALIAN GOVERNMENT SUPERANNUATION CO-CONTRIBUTIONS: ANALYSIS AND COMPARISON	Margaret E Atkinson
29	MAR 96	A SURVEY OF VALUATION ASSUMPTIONS AND FUNDING METHODS USED BY AUSTRALIAN ACTUARIES IN DEFINED BENEFIT SUPERANNUATION FUND VALUATIONS	Des Welch Shauna Ferris
30	MAR 96	THE EFFECT OF INTEREST ON NEGATIVE SURPLUS	David C M Dickson Alfred D Egídio dos Reis

31	MAR 96	RESERVING CONSECUTIVE LAYERS OF INWARDS EXCESS-OF-LOSS REINSURANCE	Greg Taylor
32	AUG 96	EFFECTIVE AND ETHICAL INSTITUTIONAL INVESTMENT	Anthony Asher
33	AUG 96	STOCHASTIC INVESTMENT MODELS: UNIT ROOTS, COINTEGRATION, STATE SPACE AND GARCH MODELS FOR AUSTRALIA	Michael Sherris Leanna Tedesco Ben Zehnwirth
34	AUG 96	THREE POWERFUL DIAGNOSTIC MODELS FOR LOSS RESERVING	Ben Zehnwirth
35	SEPT 96	KALMAN FILTERS WITH APPLICATIONS TO LOSS RESERVING	Ben Zehnwirth
36	OCT 96	RELATIVE REINSURANCE RETENTION LEVELS	David C M Dickson Howard R Waters
37	OCT 96	SMOOTHNESS CRITERIA FOR MULTI-DIMENSIONAL WHITTAKER GRADUATION	Greg Taylor
38	OCT 96	GEOGRAPHIC PREMIUM RATING BY WHITTAKER SPATIAL SMOOTHING	Greg Taylor
39	OCT 96	RISK, CAPITAL AND PROFIT IN INSURANCE	Greg Taylor
40	OCT 96	SETTING A BONUS-MALUS SCALE IN THE PRESENCE OF OTHER RATING FACTORS	Greg Taylor
41	NOV 96	CALCULATIONS AND DIAGNOSTICS FOR LINK RATION TECHNIQUES	Ben Zehnwirth Glen Barnett
42	DEC 96	VIDEO CONFERENCING IN ACTUARIAL STUDIES - A THREE YEAR CASE STUDY	David M Knox
43	DEC 96	ALTERNATIVE RETIREMENT INCOME ARRANGEMENTS AND LIFETIME INCOME INEQUALITY: LESSONS FROM AUSTRALIA	Margaret E Atkinson John Creedy David M Knox