general insurance seminar

Tides of Change

12-13 November 2012
Sofitel Sydney Wentworth

Actuaries Institute

Seeing the Bigger Picture in Claims Reserving

Prepared by Julie Sims

Presented to the Actuaries Institute
General Insurance Seminar
12 – 13 November 2012
Sydney

Institute of Actuaries of Australia
ABN 69 000 423 656
Level 7, 4 Martin Place, Sydney NSW Australia 2000
t +61 (0) 2 9233 3466 **f** +61 (0) 2 9233 3446
e actuaries@actuaries.asn.au **w** www.actuaries.asn.au

Abstract

It is often difficult to determine appropriate assumptions in claims reserving, particularly given the typically observed variability in claim numbers and payments. This paper describes how to produce a simple but sufficiently accurate description of the historical claims experience that:

- Smooths out variability in the claims data to provide a clearer picture of the changes in the historical experience; and
- Provides an improved foundation from which to predict the future.

An advantage of the approach is that all calculations can be done in Excel using the built-in solver.

This approach does not rely on a statistical model and so can be used on claims triangles that are difficult to model using standard distributions – for example, it can be applied to incremental incurred loss triangles with negative and zero values.

Several case studies using real data are presented to show the benefits of this approach. Its potential for use in a robotic reserving framework is also discussed.

Keywords: Reserving in Excel; outstanding claims liabilities; credibility models; varying parameter models; evolutionary reserving; adaptive reserving; robotic reserving; stochastic trend models.

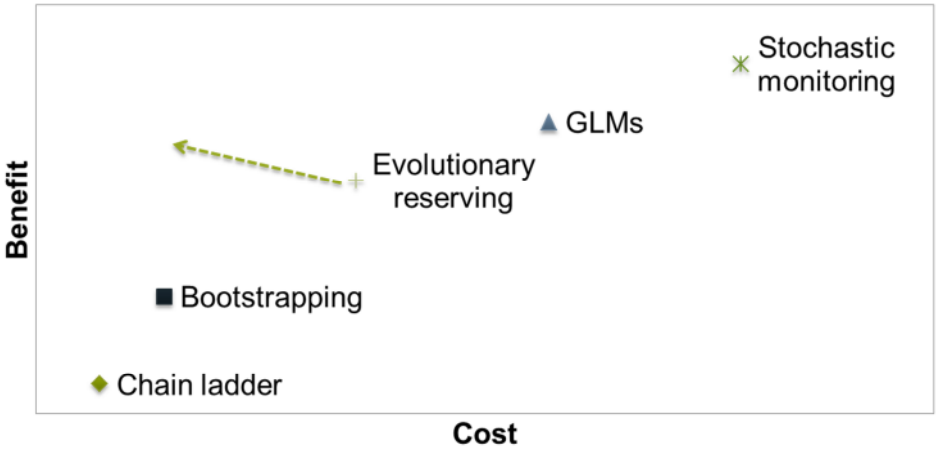
1. Introduction

The claims reserving literature contains descriptions of many models, both deterministic and stochastic. An actuary has to decide which of these models to use. The decision will be made by weighing up the costs and benefits:

- The available time, cost and other resources, such as computer software, are balanced against
- The likely benefits in better model fit and better understanding leading to more reliable predictions.

Figure 1 shows a qualitative assessment of where some of the major types of models lie in terms of cost and benefit. The motivation for this paper is to make some aspects of Evolutionary Reserving models accessible at a lower cost so that they might be more widely used and their benefits realised.

Figure 1 Qualitative cost-benefit of various types of reserving models



Deterministic methods such as Chain Ladder are widely used, but their assumptions are rarely satisfied. To compensate for this, the actuary is required to make a number of judgements, such as:

- How many accident periods should you average over?
- When should you select something different to the calculated average?
- Are there seasonality effects?
- What should you use for the multitude of ratios in the later development periods?
- Is there any superimposed inflation?

It is often difficult to make these judgements, particularly for the inexperienced actuary and especially when there is a lot of variability in claim numbers and payments. Even for an experienced actuary, it can be time consuming to make the judgements and justify them to management and

peer reviewers. Methods such as Evolutionary Reserving and Generalised Linear Models (GLMs) can reduce the reliance on judgement, but they are costly to implement and require considerable expertise. This paper describes a simplified version of these techniques that retains many of the benefits of these methods at much less cost as it can be implemented in Microsoft Excel® using the Solver. This method:

- Smooths out variability in the claims data to provide a clearer picture of the changes in the historical experience; and

- Provides an improved foundation from which to predict the future.

It does not require a statistical model be specified, although one can be. This means it can be used on claims triangles that are difficult to model using standard distributions. For example, it can be applied to incremental incurred loss triangles with negative and zero values.

Several case studies using real data are presented to show the benefits of this approach.

Evolutionary Reserving has been proposed as a foundation for a robotic reserving framework. This simplified method has a similar potential to provide low cost reviews of reserves in between full valuations.

Simple Excel macros are supplied in Appendix A to automate some of the calculations but they are not required – all models can be set up within Excel spreadsheets.

2. Background

GLMs of many types have long been applied to modelling claims reserving triangles. A review of many different models is contained in CAS Working Party on Quantifying Variability in Reserve Estimates (2005). Most require specialist software, which may be expensive and require a significant investment of time to learn.

Shapland and Leong (2010) describe a practical framework that includes much of the flexibility of GLMs and has been implemented in Excel. It allows calendar period trends, but assumes a fixed development pattern after adjustment for the calendar period trends.

Various other models allow parameters to "evolve" over time, for example:

- Zehnwirth (1994) describes log-linear models with varying accident parameters;

- England and Verrall (2001) describe a Generalised Additive Model framework that allows both accident and development parameters to vary;

- Taylor (2008) describes an approximate analytic solution to a GLM that allows development and accident parameters to vary; and

- Gluck and Venter (2009) describe a GLM with varying calendar parameters.

These models have been given many names – state space models, dynamic generalised linear models, Kalman filter, varying parameter, credibility models, adaptive reserving, stochastic trend models and, my preferred term, evolutionary reserving.

As with GLMs, all of these models require specialist software. A step towards reducing the cost of these models was taken in Sims (2011). This presentation outlined methods using Particle Filters and direct maximisation to fit the models described in Taylor (2008) that can be implemented in the free software R.

It may be feasible to combine the Excel bootstrapping of Shapland and Leong (2010) with Evolutionary Reserving models to give a variability estimate. However, this is not the intention of this paper. The focus here is what can be done to the reserving triangle to give as clear a picture as possible of the past.

This is done in a number of steps, each of which gives progressively more information:

- Look at the data triangle for evidence of changing development patterns;

- Use the Chain Ladder model to highlight deviations from a constant development pattern;

- Use a simplified development pattern to fit a model to each accident period, then see how the parameters in the development pattern are changing;

- Smooth out most of the noise between accident periods to get a clearer picture. There are four methods of doing this – grouping accident periods, a simple exponential smoother, a 2-way smoother and an evolutionary reserving model.

Finally, some implications for variability assessment are discussed.

3. What can the data tell us?

Graphs of ratios or incremental values against accident period may suggest that the development pattern is changing, whether the change is gradual or sudden and where it occurs.

Another useful indication can be given by highlighting the top few cells in each accident period. This can be done using Excel's conditional formatting, either manually, or by using the sample macro **HighlightTopTwo** in Appendix A.

Example 1 - Highlighting

This first example (and many of the subsequent examples) uses a triangle of payments per claim incurred (PPCI) of motor bodily injury data. This data was used as an example in McGuire (2007) and originally described in Taylor (2000). In this paper, it will be referred to as Dataset 1. The full triangle is shown in Appendix C. A spreadsheet using this data is available from the author illustrating the methods described in this paper.

Figure 2 shows the results of applying the macro **HighlightTopTwo** to this data. The two largest values in each row are highlighted to indicate roughly where the development pattern peaks. There is a clear shift in the location of the peak – from development years 3 to 4 to development years 5 to 6. The shift appears to be gradual rather than sudden, but it might be possible to achieve a reasonable model with a subdivision at the end of accident year 1984. The years after 1990 not shown as it is likely that the peak has not yet been reached in those years.

Figure 2 Top two values in each accident period, Dataset 1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1980	6004	12823	11228	13129	11004	7428	4949	1856	2069	803	1504	757	562	807	573	0
1981	5686	8698	13805	12720	8344	6648	4428	3243	1908	1919	2844	286	41	48	171	
1982	3271	11039	11457	15446	9826	6050	2712	1786	2403	3687	1026	331	478	401		
1983	3735	7794	8555	9675	8904	4404	4249	3336	2089	614	1233	717	381			
1984	3708	7530	10018	9912	6494	4973	12138	4265	4574	1792	961	938				
1985	2625	5366	7442	8233	7145	8352	5948	3861	1386	3271	1461					
1986	2369	6227	6452	7368	8466	6624	7291	4594	3246	944						
1987	2349	4328	6196	8053	7598	8831	6848	6726	4528							
1988	2128	6043	4177	8614	5900	6598	7851	2944								
1989	2670	4028	6082	6948	11626	10613	6027									
1990	2836	5498	5040	8578	13680	12325										

What does this achieve?

The highlighting macro takes only a few seconds to use and can often show clear evidence of a changing development pattern. It may suggest an upper limit to the number of periods that should be averaged over.

It may also indicate whether the change is gradual or sudden. A gradual change means the drift may continue in the bottom triangle. This implies greater model uncertainty should be allowed for in a variability assessment.

What next?

This test will not detect all cases where there are changes in the development pattern. In particular, for short tailed classes where most payments occur in the first development period, there is unlikely to be a shift in the peak. The next section describes a test that takes a little longer but gives more information about where changes are happening.

4. What can the Chain Ladder model tell us?

The chain ladder model applied to cumulative data, with ratios calculated using all of the data, is often a convenient way to see what a constant development pattern implies about:

- The pattern of incremental values in the development direction,

- The pattern of ultimate values in the accident direction,
- Where and how much the actual data deviates from that model.

The first item will be discussed in the next section. The third is discussed in this section.

To examine the pattern of deviations from the chain ladder model:

1. Fitted cumulative values are given by applying the calculated ratios backwards from the last diagonal of the cumulative data (see below);
2. The deviations are calculated as the difference between the actual and fitted incremental values;
3. Conditional formats are used to colour positive deviations red (dark) and negative deviations blue (light);
4. Patterns are examined for evidence of calendar trends and development pattern shifts.

The fitted values are calculated as follows. Let $C_{i,j}$ be the actual cumulative value for accident period i and development period j. The usual chain ladder ratio $f(j)$ for development period j using all accident periods is given by:

$$f(j) = \frac{\sum_{i=1}^{D-j+1} C_{i,j}}{\sum_{i=1}^{D-j+1} C_{i,j-1}} \quad \text{for } j = 2,\ldots,D \tag{1}$$

where D is the number of diagonals.

When the cumulative payment in accident period i and development period j is multiplied by $f(j+1)$, the result is the usual estimate of the cumulative payment in accident period i and development period $(j+1)$. This process is reversed to give the cumulative fitted values $D_{i,j}$:

$$\begin{aligned} D_{i,D-i+1} &= C_{i,D-i+1} \\ D_{i,j} &= \frac{D_{i,j+1}}{f(j+1)} \quad \text{for } j = 1,\ldots,D-i \end{aligned} \tag{2}$$

It is useful to set up a triangle with random values in it and apply the conditional formats described above to get a feel for what patterns can arise purely from chance. Sheet **RandomPatterns** of the example spreadsheet has such a triangle – press **F9** to see it change.

It is also possible to use the improved conditional formatting in Excel 2007 and 2010 to use colour gradients to indicate the magnitude of the deviations. This is particularly useful for seeing if the model fit is very poor anywhere.

Example 2 – Patterns in deviations

This example uses Dataset 1:

1. The cumulative payment per claim incurred is calculated from the incremental values in Appendix C.

2. The usual chain ladder ratios are calculated using all periods as in Equation (1).

3. The fitted cumulative values on the last diagonal are set equal to the actual cumulative values on the last diagonal. Fitted values on the preceding diagonals are calculated successively by dividing by the corresponding chain ladder ratio as in Equation (2).

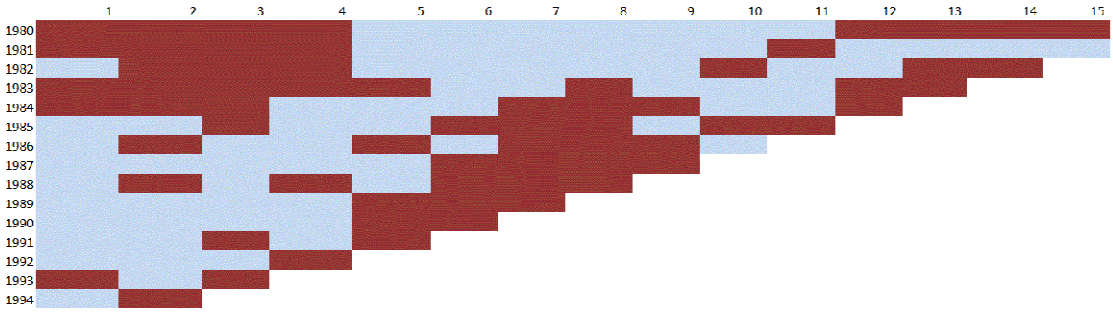
4. The fitted incremental values are calculated from the fitted cumulative values.

5. The deviations are calculated as the difference between the actual and fitted incremental values.

6. The deviations are selected and the macro **AboveBelowZero** is run.

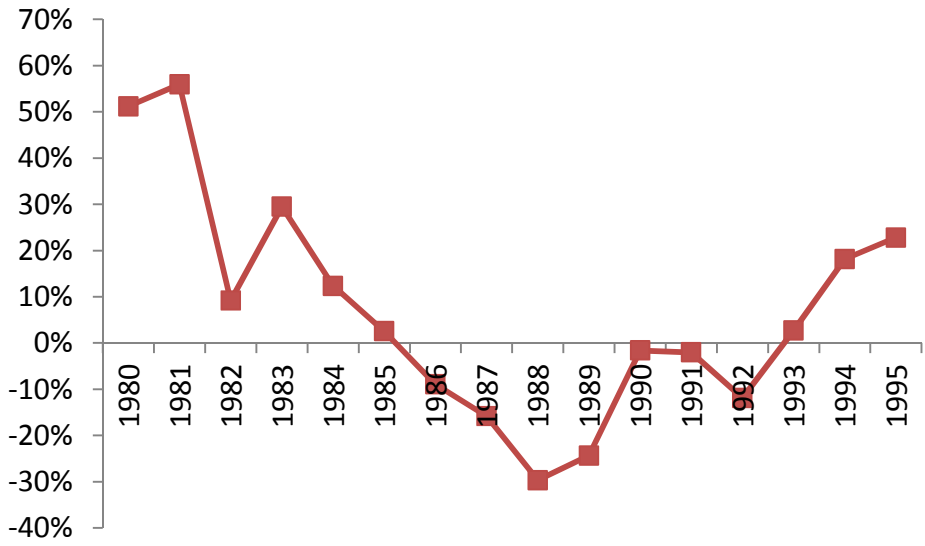
The deviations for Dataset 1 are shown in Figure 3. There are clear patterns that appear to be predominantly an accident year effect – going down each development year, the colour changes at about year 1985. However, there may also be some calendar effects – the last diagonal is predominantly red (dark).

Figure 3 Deviations from the chain ladder model, Dataset 1.



To see the size of the calendar effect, the actual and fitted values are summed for each calendar year. The difference between the actual and fitted total is divided by the fitted total to give a percentage deviation for each calendar year. These deviations are plotted in Figure 4. In the most recent calendar year, the actual values are 23% higher than the fitted values. This suggests that the forecast will be too low with this model.

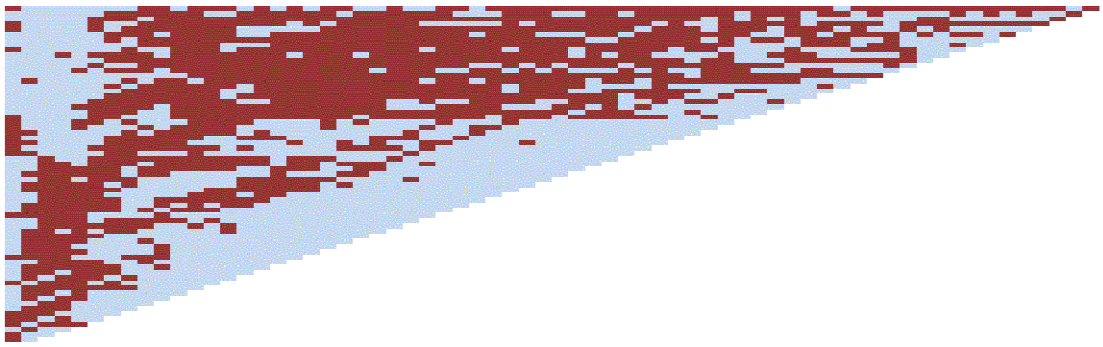
Figure 4 **Deviations from the chain ladder model by calendar year, Dataset 1.**



Example 3 – Complex patterns in deviations

A more complicated pattern is given in Figure 5. This data consists of the number of claims receiving medical payments in each quarter in a motor accident bodily injury portfolio (referred to as Dataset 2). The same process as described in Example 2 for Dataset 1 was followed for this data. The blue (light) area to the top left suggests a shift in development pattern about halfway down the triangle. However, the large blue (light) area near the diagonal indicates a calendar effect for middle development periods. This could, for example, relate to a change in rules governing eligibility for medical benefits after a certain period on benefits.

Figure 5 **Deviations from the chain ladder model, Dataset 2.**



What does this achieve?

This test takes a little longer than the previous one, but if there is a changing development pattern, it will almost always be clearly indicated in this plot. The location of the changes should suggest an upper limit to the number of periods to be averaged over.

If there are calendar effects such as changing superimposed inflation, they are very likely to be visible too. A plot like Figure 4 will show the size of the

calendar effect. If there is significant superimposed inflation, an alternative type of model to the standard techniques should be considered.

What next?

If the pattern is very complex, as in Figure 5, it may be necessary to use alternative software such as SAS to set up a GLM. However, in many cases, further exploration can provide information to guide selection for a standard model, or alternatively a satisfactory non-standard model can be developed within Excel.

The next stage in the exploration is to find a simple but sufficiently flexible development pattern that can be fitted to each accident period, or groups of accident periods, to give a clearer picture of the changes happening in the triangle.

5. Models fitted to each accident period

The following steps give more detail on how the development pattern is changing:

1. Use the fitted development pattern from the chain ladder model as a guide to choose a smaller number of parameters than the number of ratios to describe the development pattern. The pattern has to be sufficiently flexible to accommodate any movement in the peak such as seen in Figure 2. Some considerations in the choice of parameters are discussed below.

2. Fit the chosen pattern to each accident period (or groups of accident periods). See below for a discussion of fitting methods.

3. Look at plots of how the parameters change with time.

4. Look at plots of how the whole development pattern changes with time.

Parameter selection

The first step is to calculate estimates of the development pattern from the chain ladder ratios. Let $F(j_1, j_2)$ be the ratio between the estimated cumulative payment in development period j_2 and the cumulative payment in development period j_1. Then $F(j_1, j_2)$ is calculated from the chain ladder ratios $f(j)$ defined in Equation (1) by:

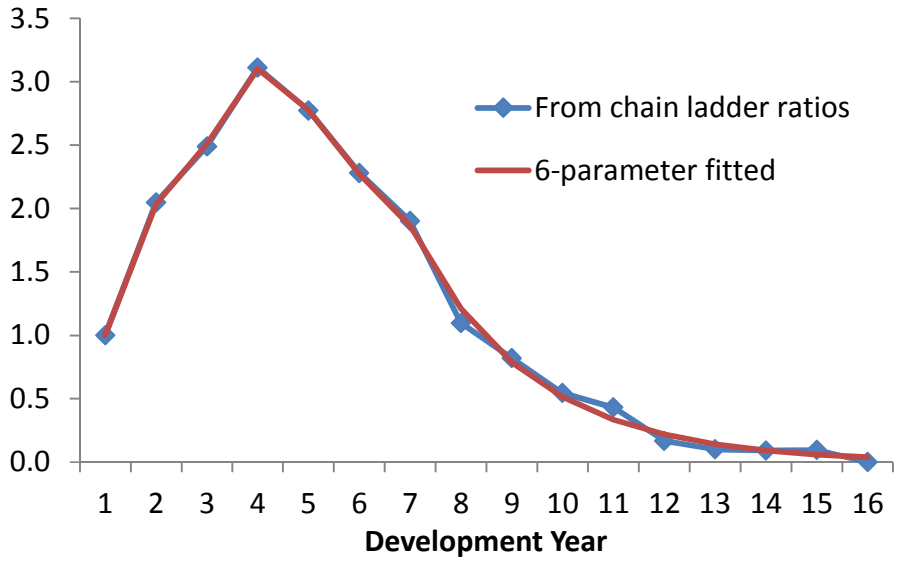
$$\begin{aligned} F(j_2, j_2) &= 1 \\ F(j_1, j_2) &= \prod_{j=j_1+1}^{j_2} f(j) \quad \text{for } j_1 < j_2 \end{aligned} \tag{3}$$

Then the levels, defined by:

$$L(1) = 1$$
$$L(j) = F(1, j) - F(1, j-1) \quad \text{for } j > 1 \tag{4}$$

are estimates of the incremental development pattern. Note that the levels are arbitrarily scaled to 1 in development period 1. See Figure 6 for an example.

Figure 6 Fitted development pattern from the chain ladder model, Dataset 1.



The next step is to decide how to describe the level curve with as small a number of parameters as will give a reasonable fit to all accident periods. My preference is to use levels and exponential trends between levels as parameters, as they have simple physical interpretations – the value of incremental payments in a cell and the percentage increase in incremental payments between cells. It is usually easy to choose the points where the exponential trends change by plotting the level estimates calculated from the chain ladder ratios then choosing a logarithmic scale for the vertical axis (see Figure 7).

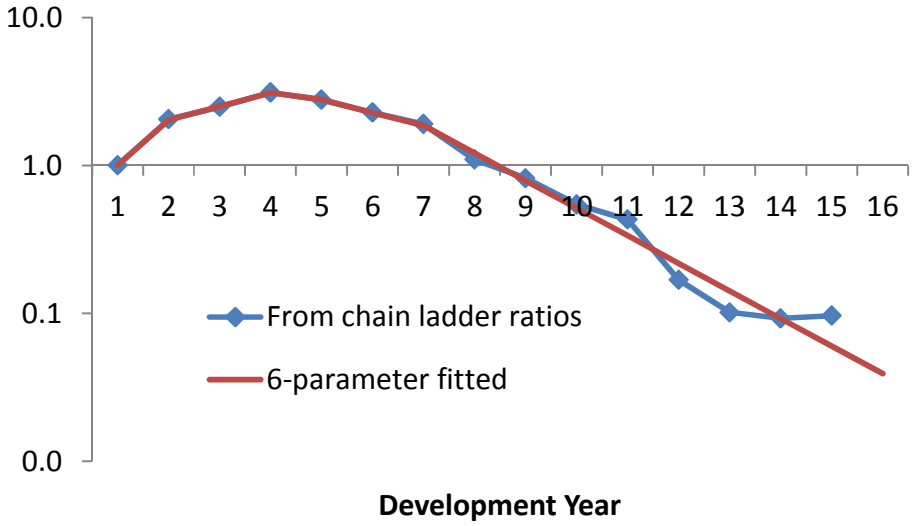
Other choices of ways to parameterise the development pattern may give fewer parameters but may be harder to interpret. The first example in McGuire (2007) (our Dataset 1) uses a modified Hoerl curve. In this case, the power term and the exponential term work in opposite directions so it is difficult to interpret them individually.

There are often unusual estimates in the last few accident periods when there are only a few data points to estimate the development pattern. If the parameters are highly correlated, as with the Hoerl curve, the estimates can be particularly extreme in the last few accident periods.

How you choose the parameters can make a difference to what you see as their values change between accident periods. It is usually better to pick features that are more stable as parameters. For example, if the amount in the first development period is very variable, it is not a good choice for the base parameter for a trend as the trend will be unstable too.

If the Excel Solver is used to estimate parameters (see below), there is no need for the model to have a particular structure, such as log linearity in the parameters. This gives additional flexibility in choosing parameters.

Figure 7 Fitted development pattern from the chain ladder model on a log scale, Dataset 1.



Parameter estimation

Next, it is necessary to decide how to fit the parameters. Statistical theory shows that you get the best predictions using an appropriate probability distribution with appropriate variance. In practice it is not easy with small triangles to decide what is appropriate. Some of the possibilities are:

- Over-dispersed Poisson. This is preferred in Gluck and Venter (2009);
- Gamma. This requires estimates of the shape parameter;
- Lognormal. This is used in Zehnwirth (1994). It requires a bias correction and often also adjustment of variances; and
- Normal. Triangle residuals are usually found to be skewed, which usually leads to this distribution being rejected.

In most cases, a simpler method like least squares (equivalent to assuming a normal distribution) will give estimates for the parameters that are sufficiently accurate for understanding how the development pattern is changing. If the payments are particularly variable in the early or late development periods, estimates may be improved by down-weighting those periods. A simple method of choosing the weights should be adequate.

Note that it is not necessary to calculate a complicated design matrix to perform the fitting. Putting the formulae for the levels and trends directly into a fitted triangle is usually the quickest and most transparent way to fit the model:

1. Set up a triangle of the actual incremental values;

2. Set up a table that will contain the fitted parameters for each accident period. Fill the first row with approximate starting values from the actual incremental values in the first accident period. Fill the remaining rows with formulae that set each row equal to the preceding one;

3. Set up a triangle of fitted values calculated from the levels and trends (or whatever alternative parameterisation you have used) in the parameter table;

4. Set up the deviation triangle as the difference between the actual and fitted triangles. Sum the squares of the deviations in each accident period (the deviation total).

5. For each accident period, run the solver, with the objective cell being the deviation total for that accident period, and the objective minimisation by changing the parameter cells for that accident period. An illustrative macro **RunSolver** to perform this is given in Appendix A, but it can be done manually (if tediously).

The parameters and fitted values can then be plotted to see how the development pattern changes with time.

This procedure makes very simple assumptions about the behaviour of the process error – that it is normally distributed and has uniform variance. It is straightforward to change these assumptions by using different formulae in the deviation triangle (step 4), if these assumptions are thought to be inappropriate. This might be the case, for example, if the variability is much higher in some development periods than others, if there are many zeroes or if the positive deviations tend to be much larger than the negative deviations.

Example 4 – Choosing a development pattern – parameter selection

Using Dataset 1, the first step is to estimate the development pattern based on the chain ladder model. This is calculated from the chain ladder ratios using Equations (3) and (4). The estimated development pattern is shown by the blue diamonds in Figure 6 and, on a log scale, in Figure 7.

The next step is to describe this pattern using a small number of parameters. A simple way of doing this is to use one parameter for the level in development period one, then exponential trends thereafter. On a log scale, the exponential trends become linear, so Figure 7 can be used to decide where the trends change. There are clear changes in trend at development periods 2, 4 and 7. After that, the choice becomes less clear.

After some experimentation, a 6-parameter model was chosen, with one level in development period 1, and five exponential trends (ending at development periods 2, 4, 5, 7 and 16 respectively). The trend from 4 to 7 was broken at 5 to give the peak some flexibility to move as it appears to do in Figure 2.

The fitted pattern from these 6 parameters is shown in Figure 6 by the red line with no symbols. It is not used directly in the following discussion, but it can be useful in its own right as a method of smoothing the chain ladder ratios in the tail. The line is fitted as follows:

1. Choose 6 empty cells to contain the parameter values. Put approximate starting values in these cells, for example six 1's.

2. Calculate the fitted levels from these parameter values, for each development period, as follows:

    a. the first level is parameter 1;

    b. the second level is the previous level times parameter 2;

    c. continue multiplying by parameter 2 until you reach the first trend change (in this case at development period 2 – so parameter 2 is only used once), then switch to multiplying by parameter 3;

    d. repeat the previous step with parameter 3 until you reach the second trend change (in this case at development period 4, so parameter 3 is used twice), then switch to multiplying by parameter 4;

    e. repeat with parameters 4, 5 and 6. Parameter 6 is used to the last development period.

    Note that this development pattern has been scaled to a value of 1 in development year 1, so parameter 1 will be 1. The actual fitted values in any accident year are given by multiplying this pattern by the fitted value for development year 1.

    If the 6 parameters are labelled p_1 to p_6, the first 7 fitted values are shown in Table 1.

Table 1 Formulae for fitted values for Dataset 1

Development period						
1	***2***	***3***	***4***	***5***	***6***	***7***
p_1	$p_1 p_2$	$p_1 p_2 p_3$	$p_1 p_2 p_3 p_3$	$p_1 p_2 p_3 p_3 p_4$	$p_1 p_2 p_3 p_3 p_4 p_5$	$p_1 p_2 p_3 p_3 p_4 p_5 p_5$

The remaining values are multiplied successively by p_6.

3. Calculate the deviation for each development period as the difference between the chain ladder levels and the fitted levels.

4. Calculate the total deviation as the sum of the squares of the deviations.

5. Set up and run the Solver to minimise the total deviation, using the six parameter cells as variable cells.

6. If fitted ratios are required, invert Equations (3) and (4): the $F(1, j)$ are the cumulative sum of the fitted levels, and the fitted ratios are the ratios of successive $F(1, j)$.

Example 5 – Fitting the development pattern to each accident period – parameter estimation

This example is based on the parameter selection in Example 4 using Dataset 1. The process of fitting a development pattern to each accident year is very similar to the above process, repeated for each accident year:

1. Start with a spreadsheet containing the incremental triangle.

2. Choose a 16 by 6 rectangle of empty cells to contain the parameter values: the **parameter table**. Put approximate starting values in the first row of these cells, for example 6004 (the first incremental value) and five 1's. Put formulae in all other rows setting them equal to the previous row.

3. Set up the **fitted triangle** from these parameter values, for each accident period, as follows:

    a. the first development period is parameter 1;

    b. the second development period is the previous development period times parameter 2;

    c. continue multiplying by parameter 2 until you reach the first trend change (in this case at development period 2 – so parameter 2 is only used once), then switch to multiplying by parameter 3;

    d. repeat the previous step with parameter 3 until you reach the second trend change (in this case at development period 4, so parameter 3 is used twice), then switch to multiplying by parameter 4;

    e. repeat with parameters 4, 5 and 6. Parameter 6 is used to the last development period.

    If the 6 parameters are labelled p_1 to p_6, the first 7 fitted values are shown in Table 1. The remaining values are multiplied successively by p_6.

4. Set up the deviation triangle as the difference between the actual and the fitted triangles.

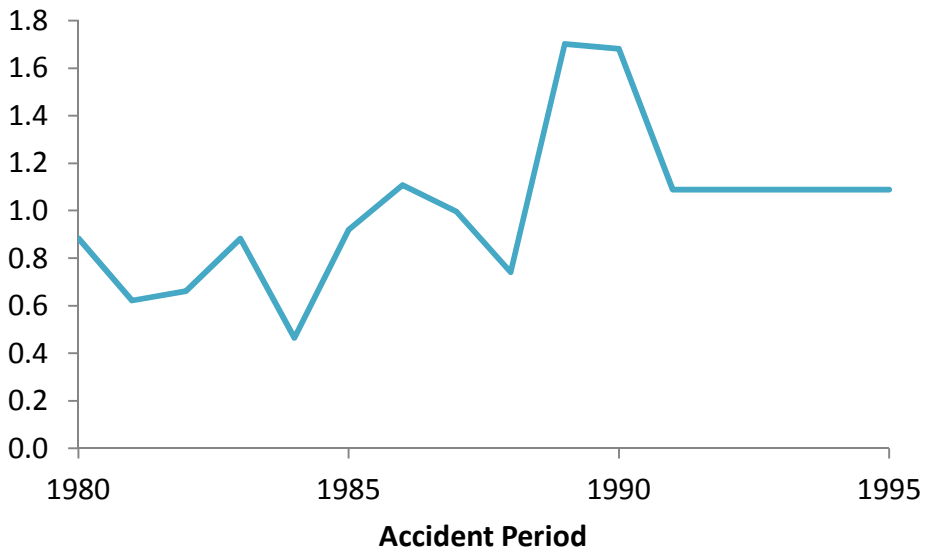
5. Calculate the deviation total for each accident period as the sum of the squares of the deviations in that accident period.

6. Set up and run the Solver to minimise the deviation total for the first accident period, using the six parameter cells at the top of the parameter table as variable cells. Repeat for each accident period, shifting the objective and variable cells down one row each time. Macro **RunSolver** can be used to do this, if the six ranges at the start of the macro are set to suitable values for your spreadsheet.

7. Graph each column of the parameter table to see how each parameter changes over the accident periods.

   The fitted values of the fourth parameter – the trend from development period 4 to 5 – are shown in
   Figure 8. In the earlier accident years, it is less than 1, so it represents a downward trend. In the later years, it is generally greater than 1, although it is quite variable. The estimates are constant from 1991 onwards, because there is no data to estimate this parameter from 1992.

Figure 8 Fitted values of parameter 4 for individual accident years, Dataset 1.



8. Graph selected accident periods in the fitted triangle to see how the development pattern changes with time.

   The way the development pattern is moving is shown by a selection of accident years in Figure 9. The peak is moving to the right. It has moved down at first but then increases.

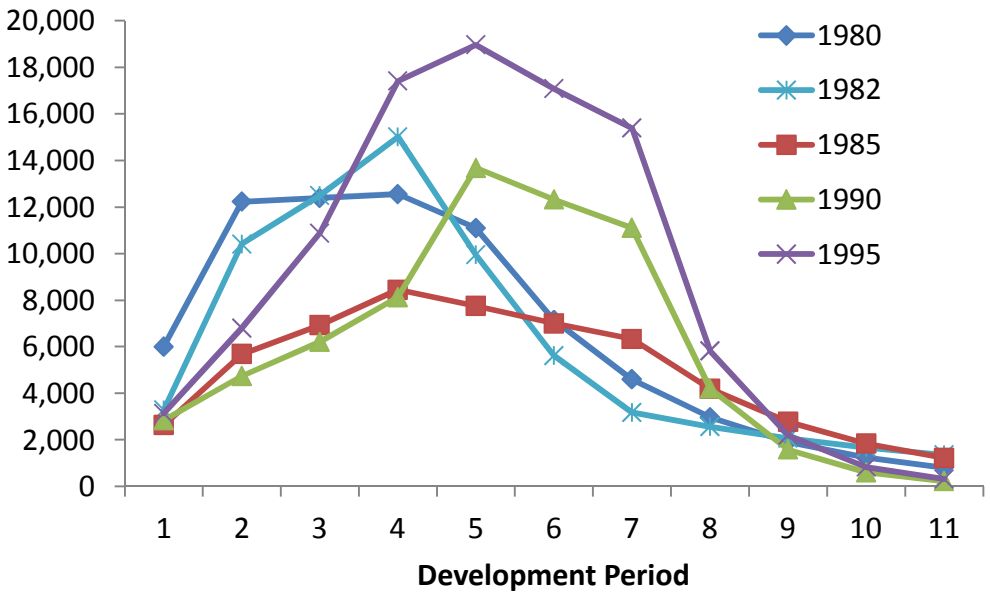
   The vertical scale is actual payments, so the area under the graph represents the ultimate payments.

What does this achieve?

The plots of how parameters change over accident years are typically quite noisy, but the eye is very good at smoothing the information to pick out trends. Some care must be taken to mentally down weight the most recent

accident periods, where there are fewer development periods (or none at all) to fit for that particular parameter.

Figure 9 Fitted development pattern for selected individual accident years, Dataset 1.



Some of the typical observations that might be made are:

- A relatively flat section, prior to the last accident period fitted to that parameter, will suggest an upper limit to the number of periods to be averaged over.

- A trend, prior to the last accident period fitted to that parameter, will suggest that an average may understate the future. It may be necessary to extrapolate the last few periods.

- Exceptionally high or low values should be questioned. Has there been one unusual payment? Has there been a change to processes or legislative requirements?

What next?

Smoothing by eye is probably adequate if you are intending only to use the information in selecting parameters for a standard model. If you would like to consider some alternatives to a standard model, a number of different methods of smoothing out the noise to produce a usable model are discussed in the next section.

All of the methods can be implemented relatively quickly in Excel, but they are ordered from the most simple to the more complex. Which is most appropriate will depend on the characteristics of your data and the time you have available.

6. Smoothing out the noise

Often the variability in the triangle makes it difficult to see what is happening from the individual accident period fits. There are two possible approaches:

- Group accident periods, or
- Apply some smoothing method. Three different methods of doing this are discussed below.

The general objective is to find a model that is an adequate fit to the data with as few parameters as possible. Having a smaller number of parameters can have a number of advantages:

- Better predictions, provided that you have enough parameters for a good fit;
- Fewer numbers to decide on future values for;
- Less chance of anchoring bias, that is, retaining last year's number because changes cannot be detected against a background of noisy data.

Zucchini (2000) has an excellent explanation of the principles behind model selection.

There are simple methods (for example, AIC, BIC and cross validation) to guide model selection for standard models that do not use smoothing. For their smoothed models, Gluck and Venter (2009) use the effective number of parameters from the approach of Ye (1998). They do not explain how they implement this method, but it is likely to be too complex for easy implementation in a spreadsheet. I am currently testing the performance of several methods for model selection that can be easily included in a spreadsheet: a simple approximation to the effective number of parameters to be used in the AIC or BIC and a cross validation method.

McGuire (2007) relies on judgement for the selection of the amount of smoothing, although this judgement is constrained by the limitations of the approximations being used. The methods described below can use either judgement or a version of the 1-step ahead forecast error.

Grouping accident periods

The first approach, grouping accident periods, can be done very simply:

1. Add together all the deviation totals to get a grand total, and use this grand total as the cell to be minimised by the Solver;
2. Set all but the first row of parameters equal to the previous row, as before;
3. Set the Solver variable cells to be the parameters in the first row of each group of accident periods.

Example 6 – Two or four groups of accident periods

This example is based on the parameter selection in Example 4 using Dataset 1. Based on Figure 2, a possible accident year grouping to try would be 1980-1984 and 1985-1995. The process is very similar to the previous example, except that the Solver is only run once, using a different objective cell and variable cells:

1. Set up the incremental triangle, 16 by 6 rectangle for starting parameter values and the fitted triangle from these parameter values, for each accident period, as in Example 5.

2. Set up the deviation triangle as the difference between the actual and the fitted triangles.

3. Calculate the total deviation as the sum of the squares of the whole deviation triangle.

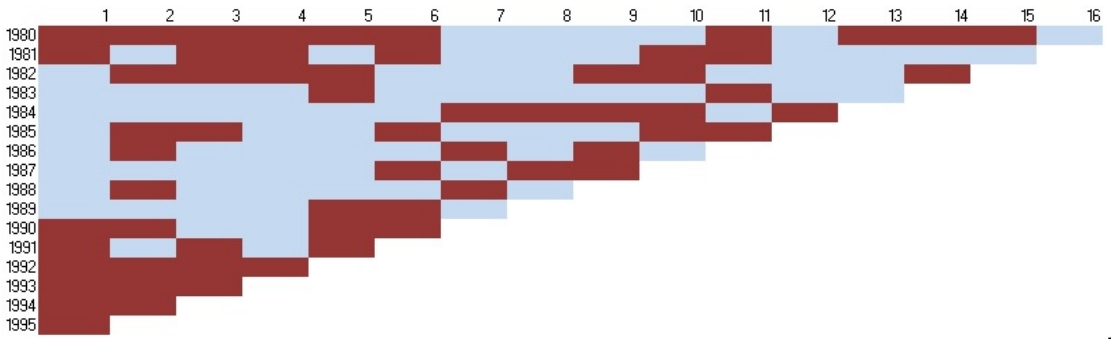
4. Set up and run the Solver to minimise the total deviation, using the twelve parameter cells in accident years 1980 and 1985 as variable cells.

5. Graph each column of the parameter table to see how each parameter changes over the accident periods.

6. Graph selected accident periods in the fitted triangle to see how the development pattern changes with time.

However, when this grouping is used, the plot of deviations analogous to Figure 2 still appears to be non-random: 1980 to 1981 and 1989 to 1995 are predominantly red (dark) in the first six development periods. This suggests breaking the two groups into four.

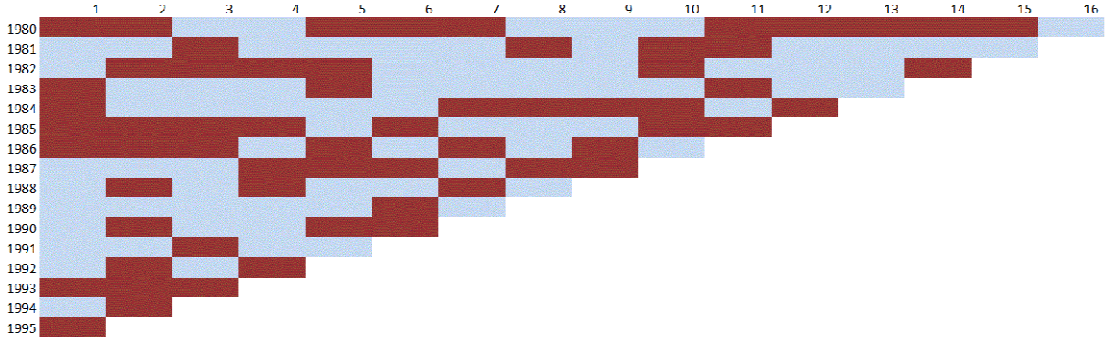
Figure 10 Deviations from the 2 group 6 parameter model, Dataset 1.



To use the grouping of 1980-1981, 1982-1984, 1985-1988, 1989-1995, the only change to the above process is to add another 12 variable cells to the Solver: the parameter cells for 1982 and 1989.

This grouping gives a more satisfactory pattern of deviations (see Figure 11), but other diagnostics should be checked before concluding that this model is adequate (see Shapland and Leong, 2010, for a discussion of diagnostics).

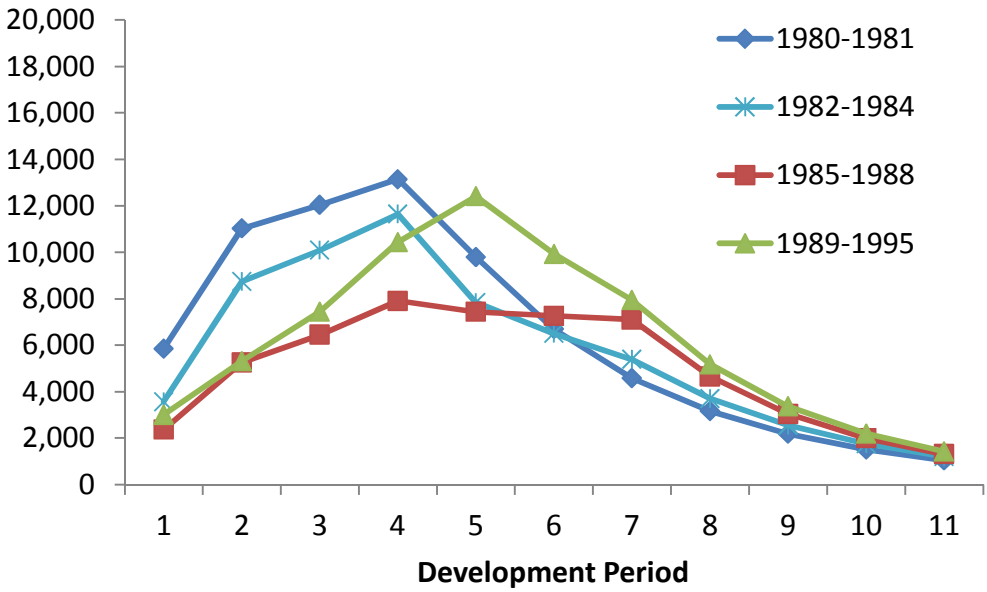
Figure 11 Deviations from the 4 group 6 parameter model, Dataset 1.



The development patterns for each of the 4 groups are shown in Figure 12.

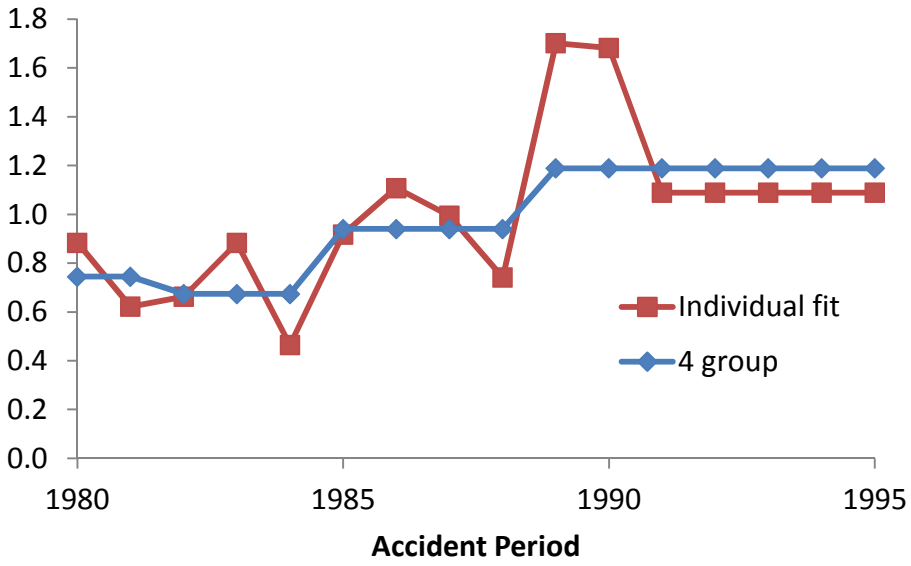
Note that the sixth parameter, the rate of decline in the tail, is estimated based essentially on a single number for the 1989-1995 group. It happens to have an estimate similar to the other groups, but in general it would be better to estimate it from more years' data. This could be done by setting it equal to the corresponding parameter for 1985-1988 and removing it from the list of Solver variable cells.

Figure 12 Fitted development patterns for the 4 accident year groups, for the 4 group 6 parameter model, Dataset 1.



As an example of how the noise is smoothed out of the parameters by grouping accident periods, Figure 13 shows the effect on parameter 4. By eye, it appears to be a reasonable representation of how that parameter is changing. It is likely that years 1980-1984 could be grouped for this parameter.

Figure 13 Fitted values of parameter 4 for individual accident years and the 4 group 6 parameter model, Dataset 1.



Smoothing method 1 – an exponential smoother

This method uses exponential smoothing to smooth the parameters fitted to the individual accident periods. The smoothed parameter value in a particular accident period is a weighted combination of the parameter estimate from the individual accident period fits and the smoothed parameter value from the previous accident period. Different smoothing weights are used for the different parameters.

The smoothed value for parameter r in accident period i, $s_{i,r}$, is given by

$$s_{i,r} = \phi_r p_{i,r} + (1-\phi_r) s_{i-1,r} \qquad (5)$$

where $p_{i,r}$ is the parameter value fitted by the individual accident period fits and ϕ_r is the smoothing weight for parameter r. The starting value $s_{0,r}$ can be set to $p_{1,r}$ or chosen by judgement.

When the parameters are levels and trends, the weights can be made to depend approximately on the number of data points available to fit the parameter. This means, for example, that a trend in an accident period with only two points to estimate it will get little weight compared to trends in previous accident periods.

In this case, the adjusted smoothing weight $\omega_{i,r}$ also depends on the accident period, for example:

$$s_{i,r} = \omega_{i,r} p_{i,r} + (1-\omega_{i,r}) s_{i-1,r}$$
$$\omega_{i,r} = \phi_r \left(\frac{n_{i,r}}{m_r} \right)^q \qquad (6)$$

where $n_{i,r}$ is the number of data points available to fit parameter r in accident period i and m_r is the maximum over i of $n_{i,r}$, and q is some positive number. A value of 1 for q seems to work well, but there may be a theoretically better value (further research is required).

The smoothing weights can be chosen to minimise some criterion such as the sum of squares of the 1-step ahead forecast error, where this is interpreted as the difference between the actual value in a cell and the fitted value in the same development period but previous accident period.

Smoothing weights should be constrained to be between 0 and 1. An illustrative macro **RunSolverOnce** to run the Solver with suitable constraints is given in Appendix A. However, the Solver can be set up manually to do the same minimisation.

Example 7 – Exponential smoothing of parameters

This example is based on the parameter selection in Example 4 and parameter estimation in Example 5, using Dataset 1. This method requires the following additional steps:

1. Set up a table that shows which development periods use each parameter:

Table 2 Development periods available to estimate each parameter

	Parameter					
	1	**2**	**3**	**4**	**5**	**6**
From	1	2	3	5	6	8
To	1	2	4	5	7	16

2. Set up a 16 by 6 table, the **count table**, containing formulae that count, for each accident year and parameter, how many of those development periods actually contain data, i.e. are within the triangle. These values are the $n_{i,r}$ of Equation (6).

3. Calculate the maximum value in each column of the count table. These values are the m_r of Equation (6).

4. Set up a 1 by 6 table to contain the smoothing weights, ϕ_r of Equation (6). Put in starting values of 0.5 – they will be set by the Solver later.

5. Set up a 16 by 6 table, the **weight table**, that contains the adjusted smoothing weights $\omega_{i,r}$ calculated using Equation (6). A value of 1 was used for q.

6. Set up a 16 by 6 table, the **parameter table**, that contains the $s_{i,r}$ calculated using Equation (6). The $p_{i,r}$ in this formula are the parameters calculated in Example 5. The first value $s_{1,r}$ is set to $p_{1,r}$.

7. Set up the fitted triangle from these parameter values, for each accident period, as before.

8. Set up the 1-step ahead error triangle as the difference between the actual incremental value and the fitted value in the same development year but the **previous** accident year.

9. Calculate the mean square error (**MSE**) as the square root of the average squared 1-step ahead error.

10. Set up and run the Solver to minimise the MSE, using the six smoothing weights ϕ_r as variable cells. Also set up constraints to limit the smoothing weights to be between 0 and 1. The macro **RunSolverOnce** can be used to do this.

11. Graph each column of the parameter table to see how each parameter changes over the accident periods.

12. Graph selected accident periods in the fitted triangle to see how the development pattern changes with time.

This gives the development pattern for selected accident years shown in Figure 14. The first four of the accident years shown are similar to the 4 group model in Figure 12. The final year shown is much higher, mainly due to the very large value in the last calendar year of accident year 1992. In the absence of sufficient data, it becomes a matter of judgement as to whether this large value is likely to be repeated in future accident years.

As an example of how the noise is smoothed out of the parameters by this method, Figure 15 shows the effect on parameter 4. By eye, it appears to be somewhat under smoothed. The smoothing weights ϕ_r could be manually adjusted to give what is judged to be a better representation of how the parameter would be expected to change gradually.

This method works best if the parameters are changing slowly and there is relatively little noise.

If any parameters are changing rapidly or there is seasonality, an approximate fixed parameter model could be used to estimate the major trends and shifts in the accident direction. The smoothing model would then be fitted to the "scaled" triangle. The smoothed parameters from that model would account for any remaining changes.

Figure 14 Fitted development patterns for selected accident years, for smoothing model 1, Dataset 1.

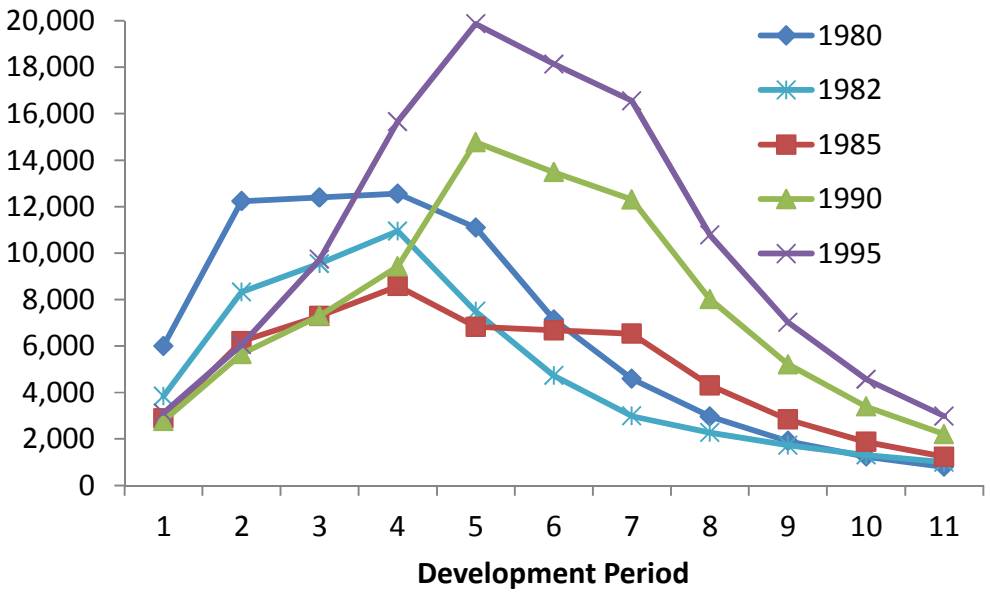
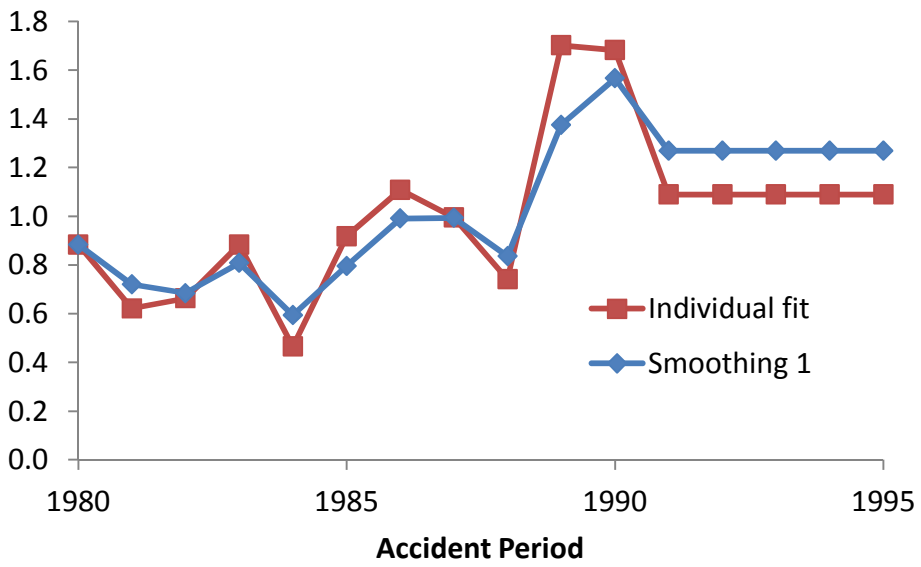


Figure 15 Fitted values of parameter 4 for individual accident years and smoothing model 1, Dataset 1.



If the estimates are very noisy, it might be necessary to apply some other strategies to reduce the noise:

- Put limits on the maximum change allowed in smoothed parameters between accident periods;
- Put limits on the minimum and maximum values of fitted parameters before smoothing them;
- Exclude outliers by ignoring their 1-step ahead forecast error;

- Include the starting values $s_{0,r}$ for each parameter in the variable cells in the optimisation.

Smoothing method 2 – a 2-way exponential smoother

This method smooths both forwards and backwards. This reduces the bias that occurs with exponential smoothing when there is a trend in the values and strong smoothing is needed. It is likely to work better than smoothing method 1 when the data is very noisy. However, it is a little more complicated to implement and the optimising criterion of 1-step ahead forecast errors is no longer independent of the fitted values. Nevertheless, it seems to give a reasonable fit.

The smoothed value for parameter r in accident period i, $s_{i,r}$, is given by

$$s_{i,r} = \frac{\sum_k \phi_r^{|i-k|} \lambda_{k,r} p_{k,r}}{\sum_k \phi_r^{|i-k|} \lambda_{k,r}} \qquad (7)$$

$$\lambda_{k,r} = \left(\frac{n_{k,r}}{m_r} \right)^q$$

where ϕ_r is the smoothing weight for parameter r, $p_{k,r}$ is the rth parameter value fitted by the individual accident period fits to accident period k, $\lambda_{k,r}$ is the weight adjustment for parameter r in accident period k, and the sum is over all accident periods k. As before, $n_{k,r}$ is the number of data points available to fit parameter r in accident period k and m_r is the maximum over k of $n_{k,r}$, and q is some positive number.

As with smoothing model 1, the smoothing weights ϕ_r can be chosen by minimising the sum of squares of the 1-step ahead forecast errors or by eye.

Example 8 – 2-way exponential smoothing of parameters

This example is based on the parameter selection in Example 4 and parameter estimation in Example 5, using Dataset 1. This method requires the following additional steps:

1. As in Example 7, set up:
    a. a table that shows which development periods use each parameter (Table 2);
    b. the count table;
    c. the maximum counts;
    d. the smoothing weights.

2. Set up a 16 by 6 table, the **weight adjustment table**, that contains the $\lambda_{k,r}$ calculated using Equation (7). A value of 1 was used for q.

3. Set up a 31 by 6 table, the **weight table**, that contains ϕ_r raised to the powers 16, 15, ..., 1, 0, 1,..., 15, 16, for r from 1 to 6.

4. Set up a 16 by 6 table, the **parameter table**, that contains the $s_{i,r}$ calculated using Equation (7). The $p_{i,r}$ in this formula are the parameters calculated in Example 5.

5. As in Example 7, set up:
   a. the fitted triangle;
   b. the 1-step ahead error triangle;
   c. the mean square error;
   d. the Solver (and run it);
   e. graphs of parameters;
   f. graphs of the development pattern.

This method, applied to Dataset 1, gives the development pattern for selected accident years shown in Figure 16. The first four of the accident years shown are similar to the smoothing model 1 in Figure 14. The final year shown is slightly higher.

The change in parameters is rather more gradual for this model than for smoothing model 1. Figure 17 shows the difference on parameter 4. Smoothing model 2 is much closer to what I would choose by eye.

Figure 16 Fitted development patterns for selected accident years, for smoothing model 2, Dataset 1.

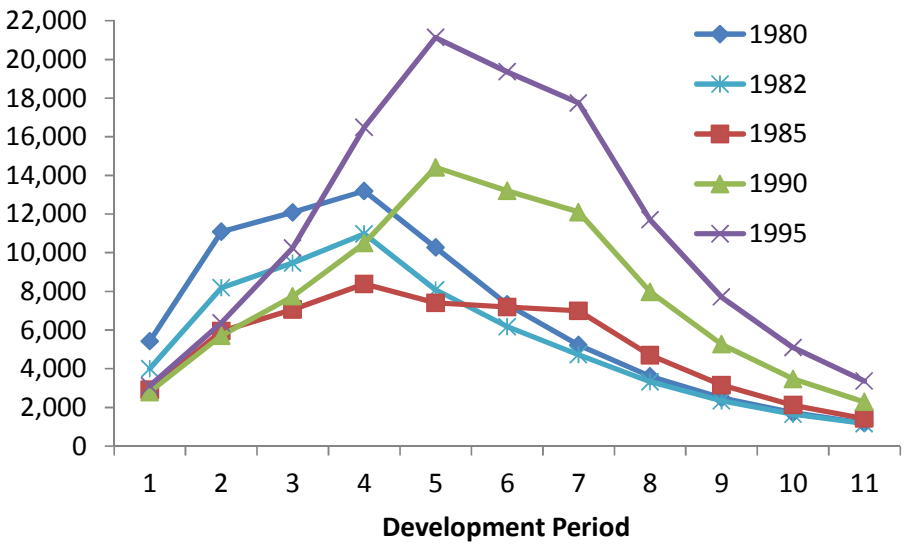
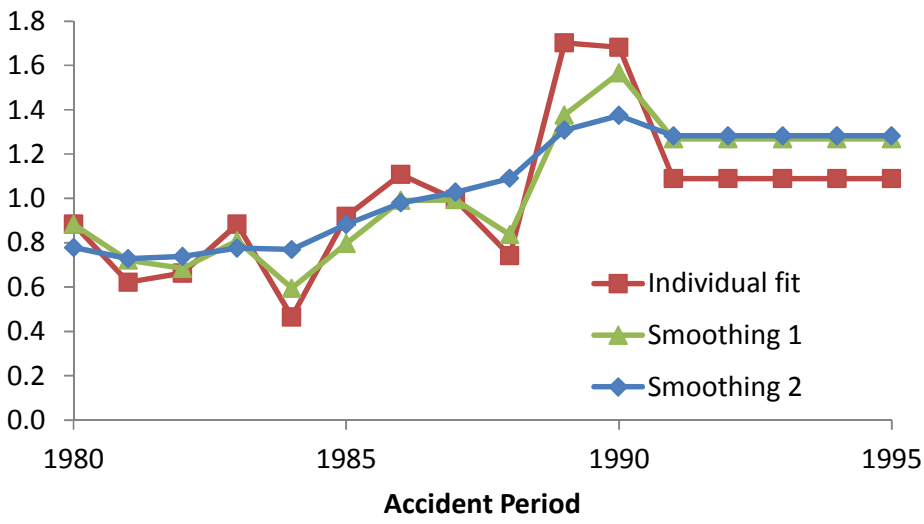


Figure 17 Fitted values of parameter 4 for individual accident years and smoothing models 1 and 2, Dataset 1.



Smoothing method 3 – a stochastic model

This method uses a statistical model that is similar to the basis of adaptive reserving in Taylor (2008) and McGuire (2007), and the stochastic trend model in Gluck and Venter (2009). It potentially provides a better fit to the data – as long as diagnostics are checked to make sure the model assumptions are satisfied. However, it is much more time-consuming to find an "optimal" set of smoothing weights, and, as with method 2, the optimising criterion of 1-step ahead forecast errors is no longer independent of the fitted values.

The model structure is similar to the individual models, except that the parameters are assumed to follow a random walk, that is, the parameters can change between accident periods by an amount that is assumed to be log normally distributed. Other distributions can be used if they are thought appropriate.

The parameter values are found by maximising the log likelihood, using the Solver. The log likelihood has one term for each cell of the triangle, and one term for each combination of parameter and accident period, except for the first accident period.

The contribution of each cell of the triangle to the log likelihood depends on the assumption made about the probability distribution that the process error follows. The simplest assumption is that the process error is normal, with constant variance. In this case, the contribution is:

$$-\left(y_{i,j} - \hat{y}_{i,j}\right)^2 \qquad (8)$$

where $y_{i,j}$ is the actual incremental value and $\hat{y}_{i,j}$ is the fitted incremental value for accident period i and development period j, and constant terms are omitted. This is the same (apart from a negative sign) as the cell's contribution to the total deviation in the previous models.

If the process error is assumed to be Poisson, the contribution becomes:

$$y_{i,j}\log(\hat{y}_{i,j}) - \hat{y}_{i,j} \tag{9}$$

If the process error is assumed to be Gamma distributed, the contribution becomes:

$$-\alpha_j \log(\hat{y}_{i,j}) - \frac{\alpha_j y_{i,j}}{\hat{y}_{i,j}} \tag{10}$$

where α_j is the shape parameter of the Gamma distribution for development period j.

The value for parameter r in accident period i, $s_{i,r}$, contributes the following term to the log likelihood, for $i \geq 2$:

$$\log\left(\phi\left(\frac{\log(s_{i,r}) - \log(s_{i-1,r})}{\sqrt{V_r}}\right)\right) \tag{11}$$

where ϕ is the standard normal density function and V_r is the variance for parameter r.

The variance of this lognormal distribution V_r controls the amount of smoothing. It can have a different value for each parameter. In the examples considered here, it is assumed to be constant over accident periods, but there is no reason why it cannot be varied. For example, if a level parameter appears to be generally constant with one or more sudden changes, the variance could be small over the constant periods and larger at the change point.

For specified values of the parameter variances, the fitted parameters can be found by maximising the log likelihood using the parameters $s_{i,r}$ for all accident periods i and parameters r as variable cells. However, the Excel Solver has a limit of 200 on the number of variable cells. If there are fewer than 200 parameter values (the number of accident periods times the number of parameters per accident period), the Solver can be used directly.

If there are too many parameters, the Solver can be applied iteratively to groups of parameters until the values stabilise. However, this might take a long time. Alternatively, some accident period parameters could be set equal to the previous period's parameters. For example, only every second period's parameters might be selected for the variable cells.

The parameter variances can be chosen to minimise some criterion such as the sum of squares of the 1-step ahead forecast error. Note, however, that

the log likelihood has to be maximised for each set of values of the parameter variances, so the minimisation cannot be done directly with the Solver. This can make the process of finding the minimum very time consuming as it has to be done by trial and error.

Example 9 – Stochastic smoothing of parameters

This example is based on the parameter selection in Example 4, using Dataset 1. This method requires the following steps:

1. Set up the incremental triangle, 16 by 6 rectangle for starting parameter values and fitted triangle from these parameter values, for each accident period, as in Example 5.

2. Set up the 1-step ahead error triangle and MSE, as in Example 7.

3. Set up a 1 by 6 table of variances V_r. Set them to 1 initially. Later they will be reduced in a search for the "best" values.

4. Set up a 1 by 16 table of shape parameters α_j. These are given the same values as in Figure 4.1 of McGuire (2007). Different values of the parameter alpha are assumed in each development period and they are treated as fixed in the optimisation. The intention is to set the variance of the Gamma distributed process error to approximately match the variance observed in the incremental values. Note that this means this example is using a different model to the previous examples.

5. Set up the log likelihood triangle using Equation (10). This means that, in contrast to the previous examples, a Gamma distribution is assumed, following the example in McGuire (2007).

6. Set up a 15 by 6 table for the parameter log likelihoods using Equation (11).

7. Calculate the total log likelihood as the sum of the log likelihood triangle and the parameter log likelihoods.

8. Set up and run the Solver to minimise the negative of the total log likelihood, using the 96 parameter cells as variable cells. This can take some time. Record the MSE when done.

9. Repeat steps 3 to 8 with variances reduced by a factor of 10 till a minimum MSE is found. Further fine tuning of individual variances may be done, either to further reduce the MSE, or to give a better smoothness as judged by eye.

10. Graph each column of the parameter table to see how each parameter changes over the accident periods.

11. Graph selected accident periods in the fitted triangle to see how the development pattern changes with time.

The resulting development pattern for selected accident years is shown in Figure 18. The first four of the accident years shown are broadly similar to the smoothing model 1 in Figure 14 and smoothing model 2 in Figure 16. The final year shown is significantly lower than the previous two smoothing models.

The change in parameters is more gradual than for smoothing Model 2. Figure 19 shows parameter 4 for each of the smoothing models. Smoothing model 3 is similar to what I would choose by eye.

Figure 18 Fitted development patterns for selected accident years, for smoothing model 3, Dataset 1.

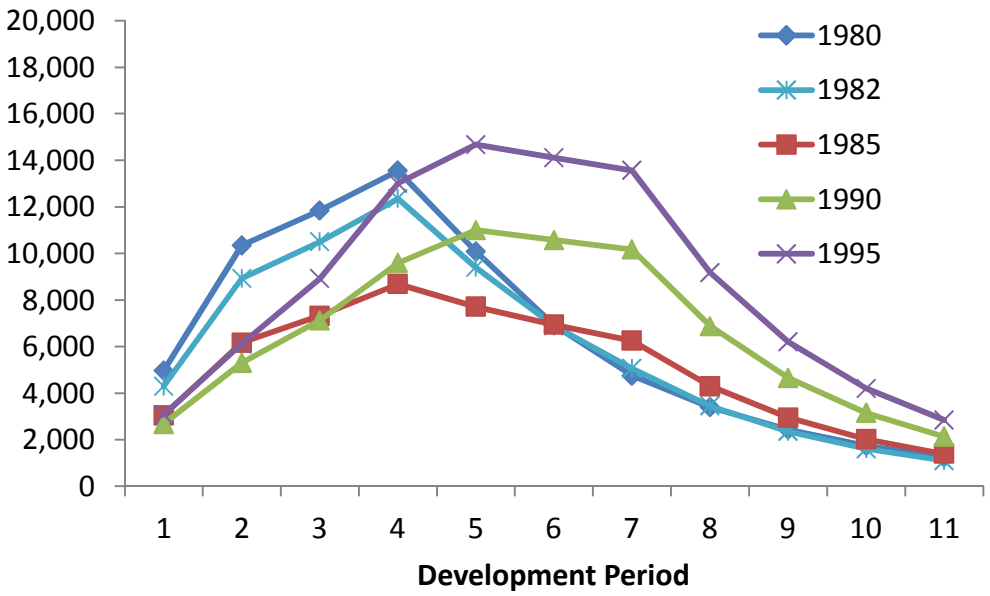
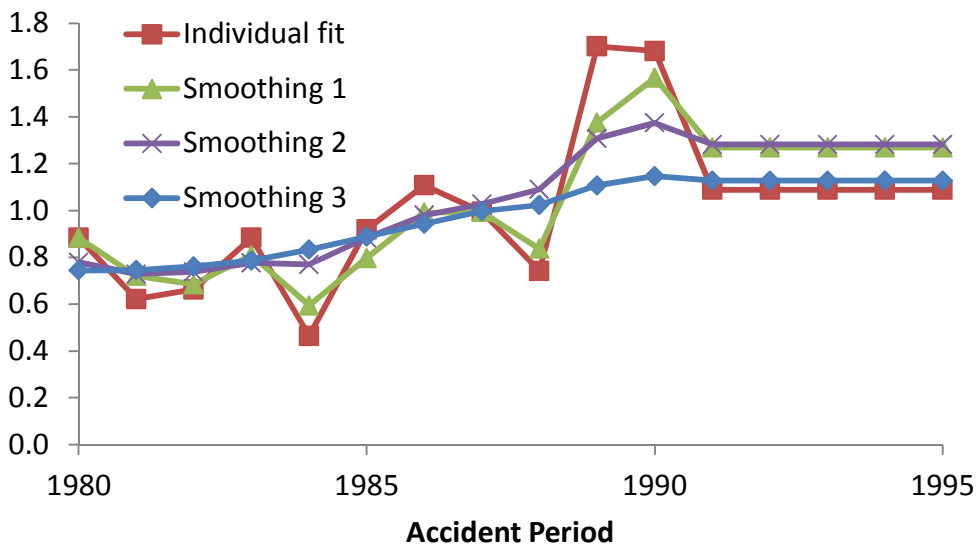


Figure 19 Fitted values of parameter 4 for individual accident years and smoothing models 1, 2 and 3, Dataset 1.



Comparison of models

Table 3 shows that the 1-step ahead forecast error is lowest for smoothing model 2. Thus, under this criterion, smoothing model 2 is "best". The next best models are the 4 group model and smoothing model 3. However, there are many other considerations that should be taken into account before deciding on a model. For example, smoothing models 1 and 2 are sensitive to the very high value in the most recent calendar year for accident year 1992. A judgement needs to be made on how much influence this value should have on expectations for the future.

Table 3 Comparison of the fitting criterion for all models

Model	1-step ahead forecast error
Chain ladder	2013
Individual fits	1918
2 groups	1801
4 groups	1548
Smoothing 1	1689
Smoothing 2	1377
Smoothing 3	1548

It is straightforward to implement smoothing model 3 in the free software R, using general purpose optimisation functions. This is much faster than using Excel and also allows the variability of the reserves to be assessed. It would be interesting to use an R implementation to investigate whether the approach of Gluck and Venter (2009) to model selection gives a similar result to the 1-step ahead forecast error criterion.

What does this achieve?

The models with grouped accident periods are easy to fit once you have decided on a set of parameters for the development pattern and a grouping of accident periods. The development parameters and accident grouping need to provide an adequate fit. These models are likely to provide more stable predictions than the chain ladder model. As more data is added, the models will need to be reviewed to see if the groupings need to be changed.

The first two smoothing models, one-way and two-way exponential smoothing, are easy to fit once you have decided on a set of parameters for the development pattern and suitable smoothing weights. There is a quick automatic method for choosing the smoothing weights, but the choice can also be made by eye if preferred. Two-way exponential smoothing performs better than one-way when there are strong trends in any of the parameters. Both methods can be sensitive to outliers, so it may be necessary to exclude some points from the fit, which can be done very easily.

The third smoothing model, adaptive or evolutionary reserving, is more time-consuming to fit and the smoothing weights have to be found by trial and error. It is less sensitive to outliers. The corresponding stochastic model can be used for a variability assessment in suitable software such as R or SAS.

Once the smoothing models have been set up with satisfactory smoothing weights, they could be used for robotic reserving. This means that the

predictions of the model would be relied upon at specified intervals, such as quarterly. They would be reviewed for significant deviations at less frequent intervals, such as yearly, or when the actual and fitted values differed by more than a specified amount.

7. What can be learnt from the smoothed models?

If the fit is judged to be adequate, the smoothed models can be used directly to forecast the outstanding reserves. They can also be used to make judgements about parameter estimates in a standard model, such as:

- The number of periods to average over in the standard model. The parameter estimates and fitted values from the smoothed model will indicate how long particular features of the data have been stable.

- Large deviations (differences between actual and fitted values from the smoothed model) should be examined to see whether they should be excluded from the standard model estimates.

- If the data is quarterly or monthly, the deviations in the smoothed model could be averaged over corresponding periods to see if there is a seasonal pattern that should be allowed for in the standard model.

- Robust selections for the decline in the tail of the standard model can be obtained from the fitted values in the tail of the smoothed model.

- Superimposed inflation may be evident from a trend in the level parameter in the smoothed model.

- Changing calendar trends may appear as patterns in the deviations of the smoothed model, especially in the plot of actual and fitted calendar totals versus calendar period.

See Appendix B for examples of how the smoothed models give useful information for selecting assumptions.

8. Implications for variability assessment

The methods described in this paper do not give an assessment of the variability of reserves when implemented in Excel. However, they can:

- Indicate deficiencies in models that assume a fixed development pattern;

- Show how often there have been changes in the past in levels and trends – large or frequent changes suggest a significant allowance should be made for model error;

- Indicate superimposed inflation, and changes in calendar trends that suggest that there may be changes in the future.

9. References

CAS Working Party on Quantifying Variability in Reserve Estimates, 2005, The Analysis and Estimation of Loss & ALAE Variability: A Summary Report, *Casualty Actuarial Society Forum Fall 2005*, pp. 29-146.

England P & Verrall R, 2001, A flexible framework for stochastic claims reserving, *Proceedings of the Casualty Actuarial Society*, 88, pp. 1-38.

Gluck S M & Venter G G, 2009, *Stochastic Trend Models in Casualty and Life Insurance*, 2009 Enterprise Risk Management Symposium, Society of Actuaries.

McGuire G, 2007, *Building a Reserving Robot*, Biennial Convention 2007, The Institute of Actuaries of Australia.

Shapland M R & Leong J, *2010, Bootstrap Modeling: Beyond the Basics*, 17th General Insurance Seminar, Institute of Actuaries of Australia.

Sims J, 2011, *Evolutionary Reserving Models – Are Particle Filters the Way to Go?*, GIRO conference and exhibition 2011, The Institute and Faculty of Actuaries.

Taylor G, 2000, *Loss Reserving: an Actuarial Perspective*, Kluwer Academic Publishers, London.

Taylor G, 2008, Second-order Bayesian Revision of a Generalised Linear Model, *Scandinavian Actuarial Journal*, 2008, 4, pp. 202-242.

Ye J, 1998, On measuring and correcting the effects of data mining and model selection, *Journal of the American Statistical Association*, 93, 441, pp. 120-131.

Zehnwirth B, 1994, Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals and Risk Based Capital, Casualty Actuarial Society Forum Spring 1994, 2, pp. 447-605.

Zucchini W, 2000, An introduction to model selection, *Journal of Mathematical Psychology*, 44, pp. 41-61.

10. Appendix A - Macros

The following macros were set up in Microsoft Excel 2010.

Macro to highlight the top two cells in a row

```
Sub HighlightTopTwo()
' Highlights top two cells in each row
' Make sure you select the first row of the triangle before running the macro

   Selection.FormatConditions.AddTop10

Selection.FormatConditions(Selection.FormatConditions.Count).SetFirstPriority
   With Selection.FormatConditions(1)
      .TopBottom = xlTop10Top
      .Rank = 2
      .Percent = False
   End With
   With Selection.FormatConditions(1).Font
      .Color = -16383844
      .TintAndShade = 0
   End With
   With Selection.FormatConditions(1).Interior
      .PatternColorIndex = xlAutomatic
      .Color = 13551615
      .TintAndShade = 0
   End With
   Selection.FormatConditions(1).StopIfTrue = False
   Selection.Copy
   n = Selection.Columns.Count
   For i = 1 To n-3
   Selection.Offset(1, 0).Select
   Selection.PasteSpecial Paste:=xlPasteFormats, Operation:=xlNone, _
      SkipBlanks:=False, Transpose:=False
   Next i
End Sub
```

Macro to colour cells greater and less than 0 red and blue

```
Sub AboveBelowZero()
'
' Colour cells red if greater than 0, blue if less than 0
' Text is coloured the same as the fill
'
   Selection.FormatConditions.Add Type:=xlCellValue, Operator:=xlLess, _
      Formula1:="=0"

Selection.FormatConditions(Selection.FormatConditions.Count).SetFirstPriority
   With Selection.FormatConditions(1).Font
      .ThemeColor = xlThemeColorLight2
      .TintAndShade = 0.799981688894314
   End With
   With Selection.FormatConditions(1).Interior
```

```
        .PatternColorIndex = xlAutomatic
        .ThemeColor = xlThemeColorLight2
        .TintAndShade = 0.799981688894314
    End With
    Selection.FormatConditions(1).StopIfTrue = False
    Selection.FormatConditions.Add Type:=xlCellValue, Operator:=xlGreater, _
        Formula1:="=0"

Selection.FormatConditions(Selection.FormatConditions.Count).SetFirstPriority
    With Selection.FormatConditions(1).Font
        .ThemeColor = xlThemeColorAccent2
        .TintAndShade = -0.249946592608417
    End With
    With Selection.FormatConditions(1).Interior
        .PatternColorIndex = xlAutomatic
        .ThemeColor = xlThemeColorAccent2
        .TintAndShade = -0.249946592608417
    End With
    Selection.FormatConditions(1).StopIfTrue = False
End Sub
```

Macro to run the Solver repeatedly

You may need to add the Solver to your References in Visual Basic (Tools->References) to run this and the next macro.

```
Sub RunSolver()
'
' Macro to run the Solver repeatedly
' for c1 to c2 accident periods
' with parameters in cells below r1
' to minimise deviations in cells below r3
' putting the current accident period being solved in c3
' and the solver result diagnostic number in cells below c4.
' 0-2 are acceptable diagnostics. See SolverSolve help for more details.
' May want to set AssumeNonNeg to True if all parameters are expected to be
positive
' Can also add constraints to stop unreasonable values in the tail
'
    Set c1 = Range("bo134") ' First accident period to run
    Set c2 = Range("bo135") ' Last accident period to run
    Set c3 = Range("bo133") ' Current accident period being run
    Set r1 = Range("bl68:bq68") ' Cells in the row above the first set of
parameters
    Set r3 = Range("bl132") ' Cell in the row above the deviation totals
    Set c4 = Range("br68") ' Cell in the row above the solver results

    For j = c1.Value To c2.Value
        Application.ScreenUpdating = True
        c3.Value = j
        Application.ScreenUpdating = False
        Set r2 = r1.Offset(j, 0)
        Set r4 = r3.Offset(j, 0)
```

```
      SolverReset
      SolverOptions MaxTime:=0, Iterations:=0, Precision:=0.0000000001,
Convergence:= _
      0.0000000001, StepThru:=False, Scaling:=False, AssumeNonNeg:=False,
Derivatives _
      :=2
      SolverOk SetCell:=r4, MaxMinVal:=2, ValueOf:=0, ByChange:=r2, _
         Engine:=1, EngineDesc:="GRG Nonlinear"
      n = SolverSolve(UserFinish:=True)
      c4.Offset(j, 0) = n
   Next j
   Application.ScreenUpdating = True
End Sub
```

Macro to set up the Solver with constraints

```
Sub RunSolverOnce()
'
' Macro to run the Solver once
' changing the cells in r1 within the range 0-1
' to minimise the deviation in cell r3
' putting the solver result diagnostic number in cells below c4.
' 0-2 are acceptable diagnostics. See SolverSolve help for more details.
'
   Set r1 = Range("bl66:bq66") ' Cells containing the smoothing constants
   Set r3 = Range("bo63") ' Cell containing the total of the sum of squares of
the 1-step ahead errors
   Set c4 = Range("bo65") ' Cell to contain the solver result

   Application.ScreenUpdating = False
   SolverReset
   Solveradd CellRef:=r1, Relation:=3, FormulaText:=0
   Solveradd CellRef:=r1, Relation:=1, FormulaText:="0.9999999" ' 1 seems to be
rejected
   SolverOptions MaxTime:=0, Iterations:=0, Precision:=0.0000000001,
Convergence:= _
   0.0000000001, StepThru:=False, Scaling:=False, AssumeNonNeg:=False,
Derivatives _
   :=2
   SolverOk SetCell:=r3, MaxMinVal:=2, ValueOf:=0, ByChange:=r1, _
      Engine:=1, EngineDesc:="GRG Nonlinear"
   n = SolverSolve(UserFinish:=True)
   c4.Value = n
   Application.ScreenUpdating = True
End Sub
```

11. Appendix B - Further examples

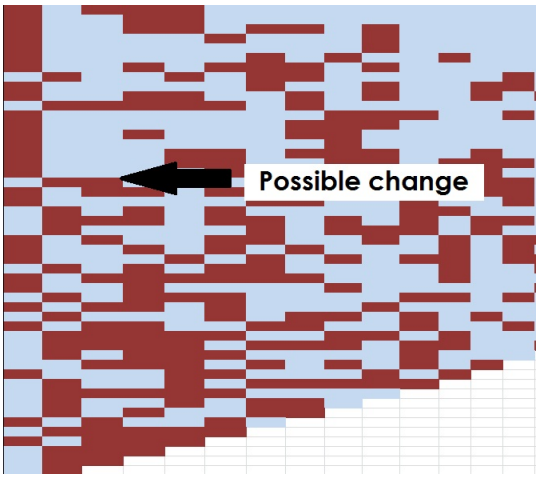
Example 10 – Domestic motor portfolio claim numbers

This example demonstrates how information can be obtained on:

- The number of periods to average over;
- The pattern of seasonality;
- Large deviations;
- Selections for the tail.

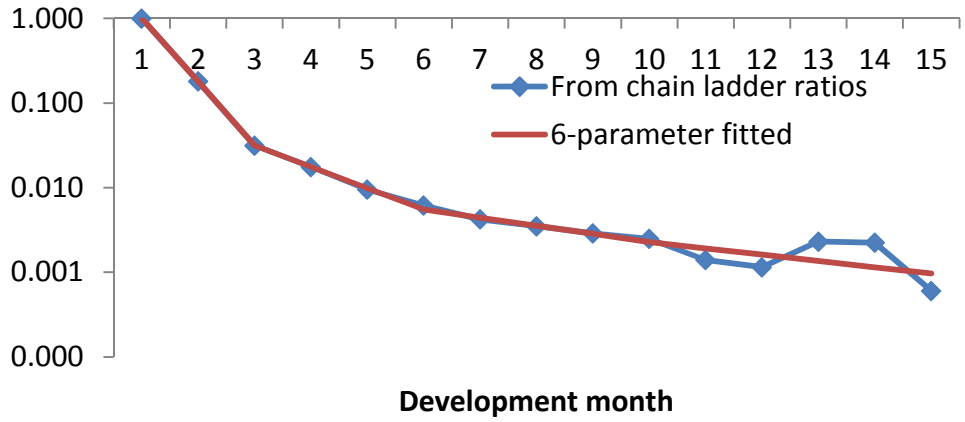
The triangle, Dataset 3, consists of 50 months of reported claim numbers for a domestic motor portfolio. The highest number of claims always occurs in the first month, so the first diagnostic test in Section 3 does not give any information on whether there is a changing development pattern. However, the second diagnostic test from Section 4 suggests there may have been changes after about 18 months.

Figure 20 Deviations from the chain ladder model, Dataset 3.



The chain ladder development pattern was used to decide on a parameter structure: a level parameter in the first development month and five trends ending in development months 2, 3, 6, 10 and 50 (Figure 21, note that a log scale has been used for the vertical axis to show the approximately piecewise linear structure of the development pattern).

Figure 21 Incremental claim count development pattern derived from the chain ladder model on a log scale, Dataset 3.

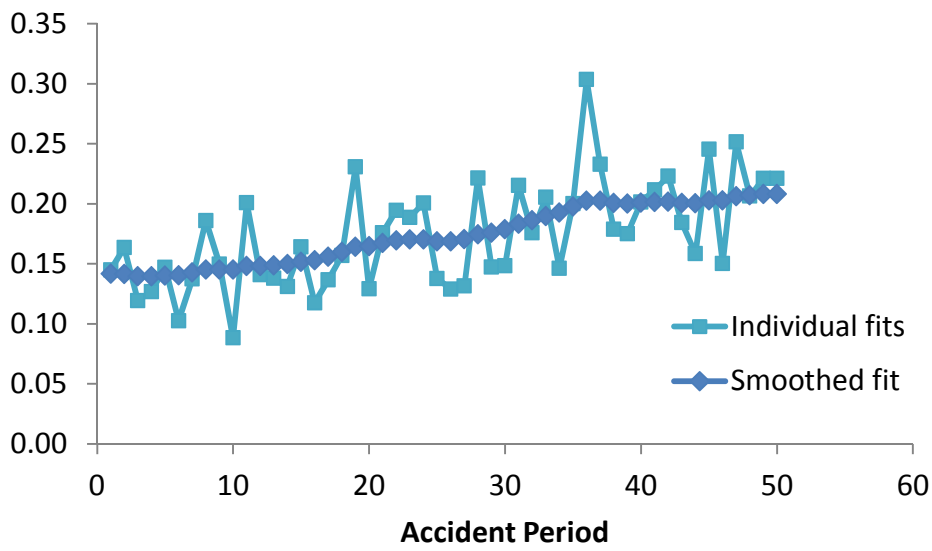


There are a large number of zeroes in the tail of this count triangle. The Poisson distribution is likely to give a better fit than a Normal distribution in this case, so instead of minimising the deviations to fit parameters in the following models, the negative of the Poisson log likelihood (Equation (9)) is minimised.

Individual accident period models were fitted as in Section 5, then smoothing was done using method 2 in Section 6, with the smoothing weights selected by eye.

The second parameter, the trend from development months 1 to 2, was the least stable of the trend parameters. It has been relatively stable for the last 15 months (see Figure 22). An averaging period of about 15 months should give the best result with this data.

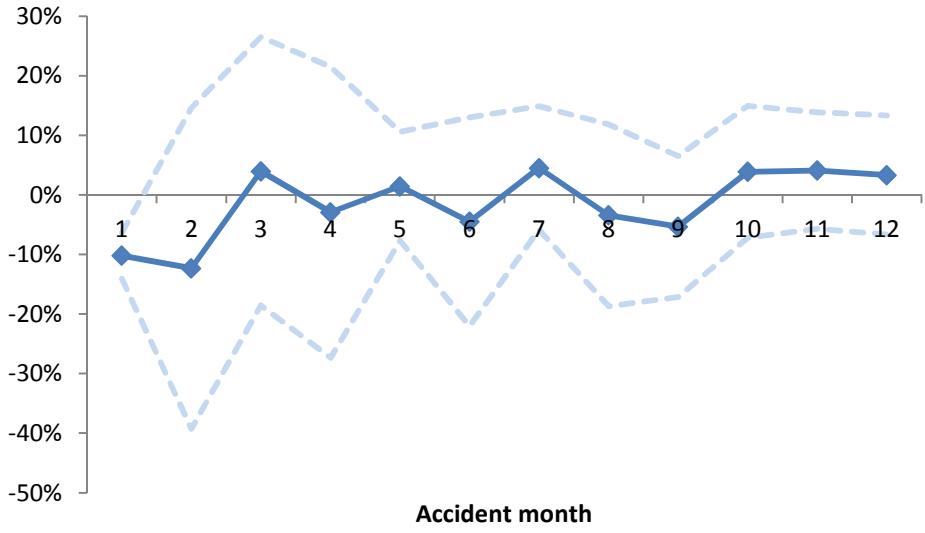
Figure 22 Fitted values of parameter 2 for individual accident years and smoothing model 2, Dataset 3.



The deviations for development months 1 and 2 were checked for seasonality. The calculated (actual – fitted)/fitted were grouped by accident month, then the mean and standard error were calculated for each

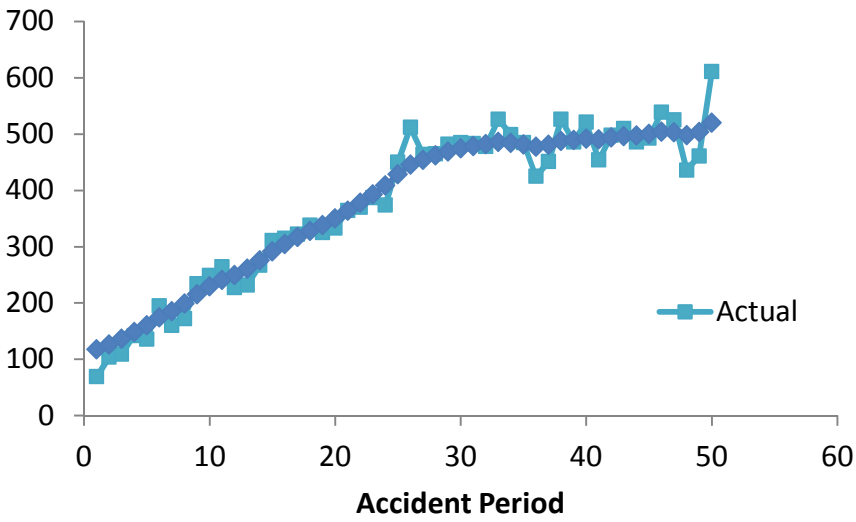
accident month. With only four years of data, the uncertainty is high, but it appears that there are significantly fewer claims than expected (about 10% less) in January for development month 1 (see Figure 23). This may tend to make the first ratio higher than normal for January. If a standard model is used, it should be checked that this is not significantly biasing the estimate of the first ratio.

Figure 23 Average percentage difference between actual and fitted reported claims in development month 1, for each accident month for smoothing model 2, Dataset 3.



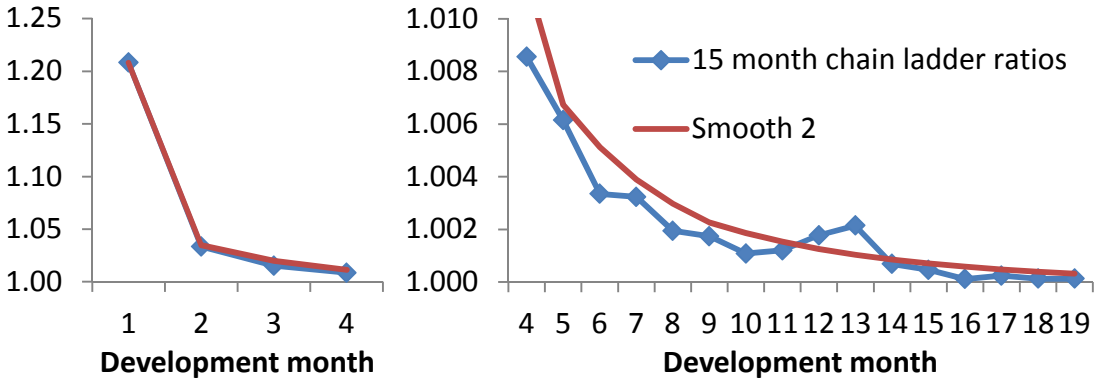
Over 40% of the unreported claims are in the last development month, so the chain ladder model prediction is quite sensitive to the number of claims reported in the most recent accident month. The actual value is unusually high compared to the fitted value from smoothing model 2 (the last accident period in Figure 24). This suggests that a better projection from a chain ladder model might be obtained by replacing the actual value of 611 with the fitted value of 520.

Figure 24 Actual and fitted values in development period 1 for smoothing model 2, Dataset 3.



As the trend parameters 3 to 6 are fairly stable, the development pattern in the last accident period could be used to calculate chain ladder ratios by inverting Equations (3) and (4). This will remove the need for guesswork in setting the ratios in the tail. The chain ladder ratios calculated from the recommended 15 months are compared with those from smoothing model 2 in Figure 25.

Figure 25 15 month chain ladder ratios and fitted ratios for the last accident period for smoothing model 2, Dataset 3.



Example 11 – Volatile incurred data – commercial property

This example demonstrates how to extract information from very volatile incurred loss data:

- The number of periods to average over;
- The development pattern.

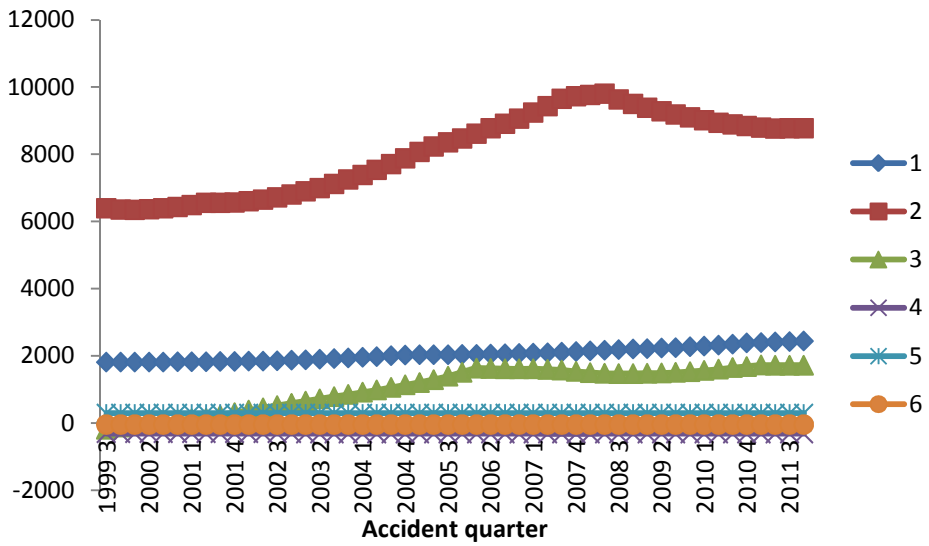
The triangle, Dataset 4, consists of 50 quarters of incurred losses for a commercial property portfolio. The highest increase in incurred losses generally occurs in the second development quarter, so the first diagnostic test in Section 3 does not give any information on whether there is a changing development pattern. The second diagnostic test from Section 4 does not have any apparently non-random patterns.

The chain ladder development pattern was used to decide on a parameter structure: because all development quarters are very variable, six level parameters were used, ending at development quarters 1, 2, 3, 5, 9 and 17. The tail is set to be zero from development period 18 onwards.

Individual accident period models were fitted as in Section 5. Smoothing was done using method 2 in Section 6, with the smoothing weights selected by eye.

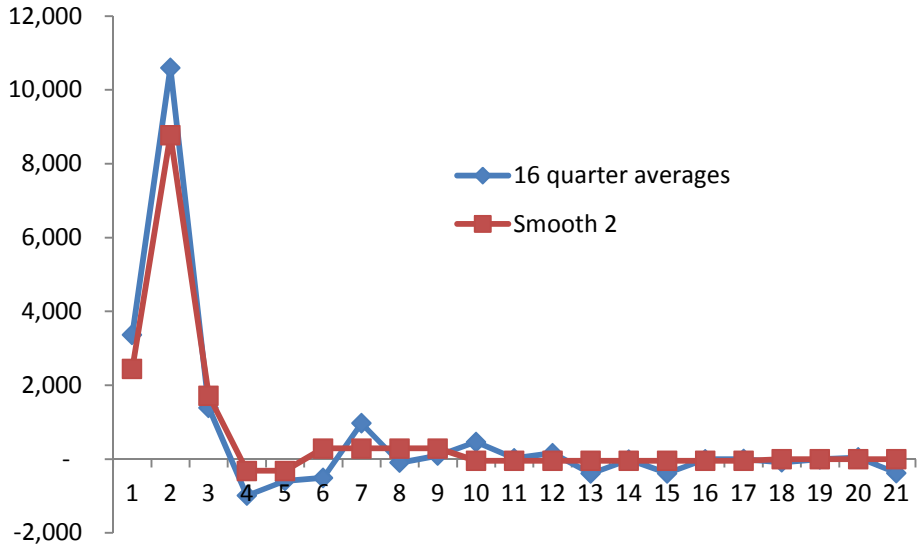
The second parameter, the level in development quarter 2, is the least stable of the parameters in recent quarters (see Figure 26). It has been relatively stable for the last 16 quarters but appears to have declined somewhat over that period. An averaging period of about 16 quarters should give a slightly conservative estimate of this parameter.

Figure 26 Fitted values of level parameters 1-6 for smoothing model 2, Dataset 4.



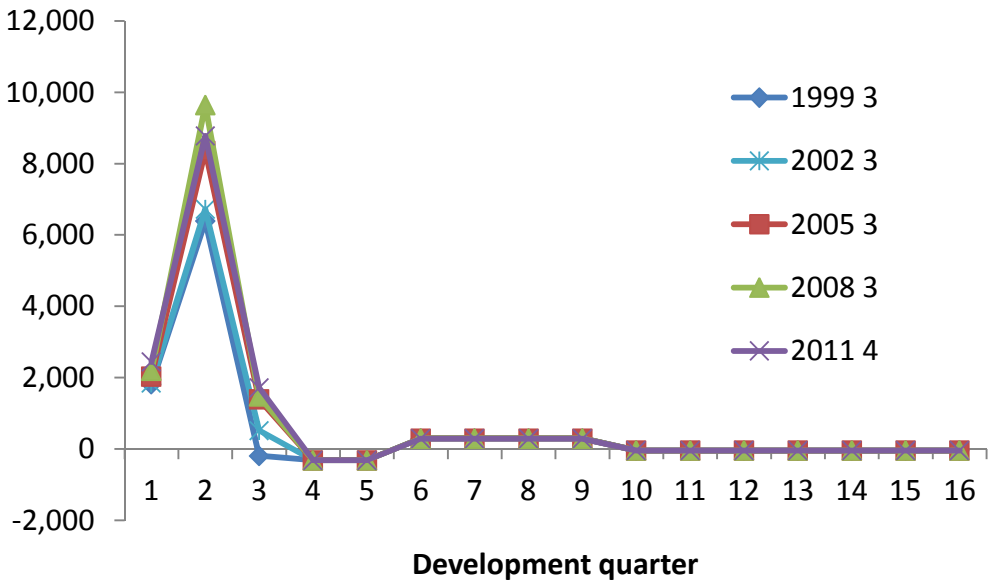
The 16 quarter average is compared with the smoothed development pattern for the latest accident quarter in Figure 27. As expected, the 16 quarter average for development quarter 2 is higher than the smoothing estimate for the latest accident quarter. The difference is not likely to be material given the large amount of volatility in the data.

Figure 27 16 quarter averages and fitted values for the last accident quarter for smoothing model 2, Dataset 4.



The changes in the smoothing model 2 development pattern over selected accident quarters is shown in Figure 28.

Figure 28 Fitted development patterns for selected accident quarters, for smoothing model 2, Dataset 4.



12. Appendix C – Dataset 1

Table 4 Payment per claim incurred for motor bodily injury data, Dataset 1

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1980	6004	12823	11228	13129	11004	7428	4949	1856	2069	803	1504	757	562	607	573	0
1981	5686	8698	13805	12720	8344	6648	4428	3243	1908	1919	2844	286	41	48	171	
1982	3271	11039	11457	15446	9826	6050	2712	1786	2403	3687	1026	331	478	401		
1983	3735	7794	8555	9675	8904	4404	4249	3336	2089	614	1233	717	381			
1984	3708	7530	10018	9912	6494	4973	12138	4265	4574	1792	961	938				
1985	2625	5366	7442	8233	7145	8352	5948	3861	1386	3271	1461					
1986	2369	6227	6452	7368	8466	6624	7291	4594	3246	944						
1987	2349	4328	6196	8053	7598	8831	6848	6726	4528							
1988	2128	6043	4177	8614	5900	6598	7851	2944								
1989	2670	4028	6082	6948	11626	10613	6027									
1990	2836	5498	5040	8578	13680	12325										
1991	2996	4543	9458	9086	10698											
1992	2883	6004	7309	17273												
1993	3561	5615	8985													
1994	2961	6430														
1995	3131															