Stochastic Reserving Methods

Advantages, Disadvantages, Key Issues and Lessons Learnt

by

Andrew Smith & Mitchell Prevett
Agenda

1. Introduction

2. Case study

3. Summary
Intro - Stochastic reserving?

- A Stochastic Model is any model that includes a random error term.
- Stochastic reserving is used “to estimate the full distribution of possible outcomes…” *
  not just the mean.

Intro - The Reserving Continuum

- **Deterministic**
  - 1970s +
    - Still used widely and is the basis of our education system

- **Sensitivity Testing**
  - Early 1990s
    - Introduction of PS300 in May 1994

- **Scenario/ Stress Testing**
  - Mid 1990s
    - Scenario testing to enhance understanding of uncertainty

- **Stochastic Overlay**
  - Mid 1990s/ 2000s +
    - Risk Margin discussions/ bootstrapping/ supplementary stochastic analysis

- **Stochastic**
  - 2001 +
    - Limited use of stochastic techniques but is on the increase
Intro - What are the Techniques?*

Techniques in common use…
- Generalised linear models
- Bootstrapping of deterministic models
- Top down risk quantification

Others…
- Bayesian or credibility models
- Kalman filter
- Wrights model
- Hoerl curve

Case study – Weekly Comp PPAC GLM

What?
Weekly compensation for a large accident compensation scheme

Why?
1. Superimposed inflation? Is it evident? What is the level?
2. Better understand the uncertainties in the reserving?
3. What insights can we derive to build into the main valuation?
4. Estimation of risk margins

How?

\[ E(Pmts_{i,j}) = Actives_{i,j-1} \times E(ContRate_j) \times E(PPAC_j), \text{ where } i = \text{devqtr}, j = \text{accqtr} \]

GLM1 – \[ ContRate_{i,j} = \frac{Actives_{i,j}}{Actives_{i,j-1}} = f(i, j, i+j) + \text{error} \]

GLM2 – \[ PPAC_{i,j} = \frac{Pmts_{i,j}}{Actives_{i,j}} = f(i, j, i+j) + \text{error} \]
Fitted GLM

- Short, medium and long term rates are modelled separately
- Over-dispersed Poisson / Log GLM on Active\(_{ij}\)
- Offset - Log(Active\(_{ij-1}\))
- Binary and linear segment transforms are used

---

**Model Parameter Estimates - Short Term**

<table>
<thead>
<tr>
<th>Model Predictor</th>
<th>Formula</th>
<th>Estimate</th>
<th>Multiple</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Development Quarter = 1</td>
<td>(devqtr = 1)</td>
<td>-0.1154</td>
<td>89%</td>
</tr>
<tr>
<td>Development Quarter = 2</td>
<td>(devqtr = 2)</td>
<td>-1.1367</td>
<td>32%</td>
</tr>
<tr>
<td>Development Quarter = 3</td>
<td>(devqtr = 3)</td>
<td>-0.4618</td>
<td>63%</td>
</tr>
<tr>
<td>Development Quarter (linear 4 to 7)</td>
<td>min(max(devqtr-4,0),7-4)</td>
<td>0.0498</td>
<td>105%</td>
</tr>
<tr>
<td>Development Quarter (linear 11 to 15)</td>
<td>min(max(devqtr-11,0),15-11)</td>
<td>0.0089</td>
<td>101%</td>
</tr>
<tr>
<td>Development Quarter (linear from 15)</td>
<td>max(devqtr-15,0)</td>
<td>0.0012</td>
<td>100%</td>
</tr>
<tr>
<td>Development Quarter (greater than 3)</td>
<td>(devqtr ge 4)</td>
<td>-0.2958</td>
<td>74%</td>
</tr>
<tr>
<td>Experience Quarter (step to 30 June XX)</td>
<td>(expqtrdate lt '30junXX'd)</td>
<td>0.0233</td>
<td>102%</td>
</tr>
<tr>
<td>(Development Quarter = 1) and Exp Qtr (prior to 30 June XX)</td>
<td></td>
<td>0.2132</td>
<td>124%</td>
</tr>
<tr>
<td>(Development Quarter = 2) and Exp Qtr (prior to 30 June XX)</td>
<td></td>
<td>-0.0708</td>
<td>93%</td>
</tr>
</tbody>
</table>

---

**Model Parameter Estimates - Medium Term**

<table>
<thead>
<tr>
<th>Model Predictor</th>
<th>Formula</th>
<th>Estimate</th>
<th>Multiple</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>-0.0264</td>
<td></td>
</tr>
<tr>
<td>Development Quarter (after 44)</td>
<td>(devqtr ge 44)</td>
<td>0.0128</td>
<td>101%</td>
</tr>
</tbody>
</table>
Comparison to deterministic model

- GLM fitted rates compare well with the deterministic fitted rates
- Confidence limits can be calculated for the GLM fitted function (this assumes no correlations in estimates)
Fitted GLM

- Over-dispersed Poisson / Log GLM on \( Pmts_{ij} \)
- Offset - Log(Active\(_{ij} \))
- Binary and linear segment transforms are used

<table>
<thead>
<tr>
<th>Model Predictor</th>
<th>Formula</th>
<th>Estimate</th>
<th>Multiple</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>10.4478</td>
<td></td>
</tr>
<tr>
<td>Development Quarter = 0</td>
<td>( (devqtr = 0) )</td>
<td>-0.9972</td>
<td>37%</td>
</tr>
<tr>
<td>Development Quarter = 1</td>
<td>( (devqtr = 1) )</td>
<td>-0.5187</td>
<td>60%</td>
</tr>
<tr>
<td>Development Quarter = 2</td>
<td>( (devqtr = 2) )</td>
<td>-0.2301</td>
<td>79%</td>
</tr>
<tr>
<td>Development Quarter = 3</td>
<td>( (devqtr = 3) )</td>
<td>-0.0436</td>
<td>96%</td>
</tr>
<tr>
<td>Development Quarter (linear 3 to 7)</td>
<td>( \min(\max(devqtr-3,0),7-3) )</td>
<td>0.0624</td>
<td>106%</td>
</tr>
<tr>
<td>Development Quarter (linear 7 to 11)</td>
<td>( \min(\max(devqtr-7,0),11-7) )</td>
<td>-0.0090</td>
<td>99%</td>
</tr>
<tr>
<td>Development Quarter (linear 11 to 15)</td>
<td>( \min(\max(devqtr-11,0),15-11) )</td>
<td>0.0166</td>
<td>102%</td>
</tr>
<tr>
<td>Development Quarter (linear 15 to 27)</td>
<td>( \min(\max(devqtr-15,0),27-15) )</td>
<td>0.0067</td>
<td>101%</td>
</tr>
<tr>
<td>Development Quarter (linear 27 to 39)</td>
<td>( \min(\max(devqtr-27,0),39-27) )</td>
<td>0.0006</td>
<td>100%</td>
</tr>
<tr>
<td>Experience Quarter (linear 30 Jun XX to 31 Dec XX)</td>
<td>( \min(\max(expqtr-93,0),30) )</td>
<td>0.0036</td>
<td>100%</td>
</tr>
</tbody>
</table>
Comparison to deterministic model

- Some differences in the GLM for early development qtrs
- Compare well after qtr 12
- Should we investigate why there is a difference?
What is SI?
• SI - increase in claims costs across time (payment quarter), after controlling for all known cost drivers (Devqtr and Accqtr)

Deterministic estimation
• Choose matching diagonals (or single Devqtrs)
• Estimate the increases in PPACs for the matched diagonals
Deterministic - Superimposed inflation in PPACs

• Appears to be a flat period followed by an upward trend
• Log regression is used to estimate the underlying rates
• What $R^2$ is significant?

Underlying superimposed inflation

$y = \text{XXX}e^{-0.0017x}$
$R^2 = 0.0837$
$\text{SI} = -0.7\%\text{pa}$

$y = \text{XXX}e^{0.0046x}$
$R^2 = 0.5877$
$\text{SI} = 1.9\%\text{pa}$
GLM Superimposed inflation in PPACs

- A linear segment is added from the beginning of the trend
- Slope of the estimate indicates an underlying SI of 2.8%pa
- Why might this be different? – Uses full triangle and correct error distribution
- But... what’s the cause of the SI?

Underlying Superimposed Inflation from PPAC GLM

- GLM hypothesis test show no significant SI (0%pa)
- GLM estimated SI = 2.8%pa

Experience Quarter

- Predicted PPAC
- Actual PPAC
Process - Simulate from the GLM Parameters*

Step 1 - Generate a series of random sets of parameters for each model (continuance rates and PPAC). The distribution of the GLM parameters is assumed to be Multivariate Normal

Step 2 - Calculate the model predictions using each of the sets of parameters

Step 3 - Generate a random observation from the process distribution (i.e. Poisson, Gamma)

Repeat for all future projection quarters, for each accident quarter.

Simulation of Active Claim Numbers

- A selected accident quarter’s projections are shown
- 5 simulations are displayed as well as the mean and 5th and 95th percentiles
- GLM simulated mean compares well with the valuation for this accident period
Simulation of Active Claim Numbers and PPACs

- Simulated cash flows are projected for the liability as at the valuation date
- The mean and 5th and 95th percentiles are calculated to show the implied the uncertainty
Central estimate

- Will not equal non-stochastic projection (must consider Jensen’s inequality*)

Distribution of reserves

- Close to Lognormal but with a heavier tail (>99%)
- CoV and P75 can be determined directly

* Source: Taylor and Mulquiny (2005)
Advantages of Stochastic Reserving

1. **Superimposed inflation and trends** – statistical significance can be used to determine if trends are “real”.

2. **Hypothesis testing** – impact of legislative changes, benefit changes, or any other claims administration changes

3. **Downside risk** - through estimation of full distribution the downside risk and hence risk margins can be estimated

4. **Individual claims modelling** – stochastic models are suitable for individual claims modelling

5. **Model update (control cycle)** – testing recent experience against the model we can determine if it needs to be updated
Disadvantages of Stochastic Reserving

1. **Less transparent** – can be harder to explain to management
2. **Reliable data** – generally more reliable data is required to construct a more sophisticated model
3. **Incorporating judgement** – more complex to overlay judgment based adjustments
4. **Costly and time consuming** – can be difficult to articulate the value
5. **Seen as the fix all** - If traditional methods fail then SR is unlikely to succeed…it’s not an automatic fix
6. **Model variability NOT risk** – the past may not be a guide to the future
Risk Framework

Empirical estimation of the claims distribution….

- Workshop the key risk drivers
- Use qualitative and quantitative methods to assess the range of scenarios / outcomes due to a given risk driver.
- Risk assessment phase likely to include stochastic analysis
- Determine a distribution for each risk driver
- Use stochastic simulation to combine the major sources of uncertainty
Summary

• Stochastic techniques are continuing to gain favour – part of tool suite

• They are not the magic answer but an important supplement / extension to current reserving practices

• Quantitative assessments should be combined with qualitative assessments to capture all aspects of risk
References


Stochastic Reserving Methods

Advantages, Disadvantages, Key Issues and Lessons Learnt

by

Andrew Smith & Mitchell Prevett