Making the most of your granular claims data

Matthew Pearlman, LCP
Richard Holloway, LCP
What were we trying to achieve?

• What is the distribution of outcomes of large open claims?
• More robust, better understood reserves
• Insights into claims development
• Capital modelling
• Reinsurance optimisation
Why is it difficult?

- Large claims data is sparse
- But significant proportion of total reserves
- Development patterns different from smaller claims
- Reserving practices change over time
Current approach

- Development triangles
- Graphical / Chain ladder
- Not enough data if just use large losses
- No indication of volatility
# Murphy McLennan model

<table>
<thead>
<tr>
<th></th>
<th>Yr 1</th>
<th>Yr 2</th>
<th>Yr 3</th>
<th>Yr 4…</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>400 Open</td>
<td>800 Closed</td>
<td>800 Closed</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>500 Open</td>
<td>1,500 Open</td>
<td>750 Closed</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1,000 Open</td>
<td>1,000 Open</td>
<td>1,500 Closed</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>200 Open</td>
<td>500 Open</td>
<td>250 Closed</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>750 Closed</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>300 Open</td>
<td>150 Closed</td>
<td>150 Closed</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>150 Open</td>
<td>300 Closed</td>
<td>300 Closed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>450 Open</td>
<td>225 Closed</td>
<td>675 Closed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150 Open</td>
<td>75 Closed</td>
<td>225 Closed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>375 Open</td>
<td>188 Closed</td>
<td>563 Closed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75 Closed</td>
<td>75 Closed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Key assumptions?

**Assumption of Murphy McLennan model**

<table>
<thead>
<tr>
<th>Assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each development period is independent of the next</td>
</tr>
<tr>
<td>Development does not depend on claim size</td>
</tr>
<tr>
<td>Claim closure is linked to the final period of development</td>
</tr>
</tbody>
</table>

26 September 2016
Our approach

• Throw data at a data scientist
  – Long-tail injury data
  – Various data sets with different characteristics

• Aerospace background
  – Similar to development of fatigue cracks on aircraft wings

• No baggage on prior expectations
  – Feed in “expert” knowledge at each iteration
What did the data show?

<table>
<thead>
<tr>
<th>Assumption of Murphy McLennan model</th>
<th>Data findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each development period is independent of the next</td>
<td>Claims behaviour has a variety of structural dependence</td>
</tr>
<tr>
<td>Development does not depend on claim size</td>
<td>Large claims have lower and less volatile development factors</td>
</tr>
<tr>
<td>Claim closure is linked to the final period of development</td>
<td>There is a longer closing phase where claims behave differently</td>
</tr>
</tbody>
</table>
Reserve movements

In previous period - claim x3
In next period - claim x1.6
Reserve movements – zoomed
Inflation allowance

Aggregated triangles
- All comes out in the wash
- Aggregated average inflation
- Project forward implicit past inflations
- Or adjust past data explicitly

Individual claims data
- Incurred estimate “sticky”
- Paid amounts need to be put on consistent money terms
- More research needed
Reserve movements – zoomed

![Graph showing reserve movements](image-url)
Translating movements to parameters

Transition matrix

<table>
<thead>
<tr>
<th>Previous development</th>
<th>This development</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;-3%</td>
</tr>
<tr>
<td>&lt;-3%</td>
<td>0.20</td>
</tr>
<tr>
<td>&lt;-2%</td>
<td>0.18</td>
</tr>
<tr>
<td>&lt;-0.01%</td>
<td>0.13</td>
</tr>
<tr>
<td>&lt;0.01%</td>
<td>0.02</td>
</tr>
<tr>
<td>&gt;0.01%</td>
<td>0.17</td>
</tr>
</tbody>
</table>
Diagnostic output

![Diagnostic output graph](image)

- **Probability** vs. **Sterling cost (000's)**

Legend:
- Blue: Both developments fell
- Cyan: One development fell
- Green: One rose; one fell
- Grey: Little movement
- Orange: One development rose
- Red: Both rose
Movements depend on size of claim

Insight
• Claims developments are both smaller on average, and less volatile for larger claims

Model
• Allows for the shift in development factors due to claim size by adjusting the sampled development path.
# What did we do differently?

<table>
<thead>
<tr>
<th>Assumption of Murphy McLennan model</th>
<th>Data findings</th>
<th>Model feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each development period is independent of the next</td>
<td>Claims behaviour has a variety of structural dependence</td>
<td>Use remainder of the development path</td>
</tr>
<tr>
<td>Development does not depend on claim size</td>
<td>Large claims have lower and less volatile development factors</td>
<td>Adjust development path based on claim size</td>
</tr>
<tr>
<td>Claim closure is linked to the final period of development</td>
<td>There is a longer closing phase where claims behave differently</td>
<td>Modelled closure separately</td>
</tr>
</tbody>
</table>

**26 September 2016**
Legs

Dotted lines = Historical claims

Solid lines = Simulated claim development scenarios

Case reserve (£m)

Developments from incident

26 September 2016
Overview of model

Initial claim
- development period
- status
- size

Population of development paths

Select an appropriate development path at random (based on development period size and status)

Claim closure probability (based on development period)

Does claim close?

Yes

Closure development factor and output results

No

More development periods left from selected path?

Develop claim

No

Yes
Complex problem

- How representative is the sample?
- How to lower noise floor without distorting?
- Spikes represent what?
- What rate at higher developments?
- Bayes badly behaved in tail region?
- How to increase information using adjacent data?
Single £1m claim – distribution of modelled outcomes
Single £1m claim – distribution of modelled outcomes

Ultimate claim (£m)

LCP
Murphy
McLennan

26 September 2016
Portfolio of claims – distribution of modelled outcomes
What did it tell us?

Ultimate aggregate claims as proportion of initial claims

Incident year

26 September 2016
What else have we learned?

• Stochastic development models are tricky
  – Small data effects can distort results
  – Many intricate sampling correction techniques needed
  – Add “sensible” limits on outcomes

• Data scientists are great
  – But need to inject a heavy dose of reality
  – Don’t know your data until someone pulls it apart

• Inflation is tricky
What next?

<table>
<thead>
<tr>
<th>Improvements to model / data capture</th>
<th>Business uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Simplify model back down to core components</td>
<td>• Large loss capital modelling including one-year recognition</td>
</tr>
<tr>
<td>• Project paid / outstandings separately</td>
<td>• Reinsurance optimisation</td>
</tr>
<tr>
<td>• Calendar year effects</td>
<td>• Early warning indicators to changes in claims practices</td>
</tr>
</tbody>
</table>
Expressions of individual views by members of the Institute and Faculty of Actuaries and its staff are encouraged.

The views expressed in this presentation are those of the presenter.