Performance Testing: Are you using the best reserving methods?

GIRO 2009, Edinburgh

Matthew Ball
John Charles
Today’s agenda

- Benefits of performance testing
- Defining the problem
- Performance testing — in general and in the context of reserves
- Embedding the reserving control cycle
- Case studies
- Conclusion

- This presentation is based on the paper “Loss Reserving: Performance Testing and the Control Cycle”, authored by Yi Jing, Joseph Lebens, and Stephen Lowe, that has recently been accepted for publication in Variance. It will be available at www.variancejournal.org
What do we mean by the best reserving methods?

Whatever method gets you closer to the actual outcome, on average, over time
Benefits of Performance Testing

- More accurate risk quantification
- More accurate capital utilisation
- Increased return on capital
- Economic value

Helping you to manage your risk, capital and return
Questions for the reserving actuary

- How do you know that the methods you are currently using are the “best”? 
  - What evidence supports your selection of methods? 
  - What are the optimal weights for combining the results of the methods? 
  - How do you decide when to change methods? 
  - What is the confidence range around estimates? 
  - Cost/benefit of developing new data sources or implementing more complex methods? 
  - How do you manage over-confidence?
The results of our research illustrate the prevalence of overconfidence

**Tillinghast Confidence Quiz**

### The Quiz

- **Objective:** To test respondents understanding of the *limits* of their knowledge
- Respondents were asked to answer ten questions related to their general knowledge of the global property/casualty industry
- For each answer, respondents were asked to provide a range that offered a 90% confidence interval that they would answer correctly
- Ideally (i.e., if “well calibrated”), respondents should have gotten nine out of ten questions correct

### Raw Scores of Respondents

<table>
<thead>
<tr>
<th>Number of Respondents</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td></td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td></td>
<td></td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: based on 374 respondents as of 4/5/04.
Profile of respondents: 86% work in P/C industry; 73% are actuaries.
Using Performance Testing to improve results

An actuarial method consists of:
• An algorithm
• A data set
• A set of intervention points

Without Performance Testing
Choose a combination of methods using ‘actuarial judgment’

Result:
Subjective best estimate

With Performance Testing
Choose a combination of methods that optimises the formal measure of ‘skill’ in a rigorous manner

Result:
More accurate best estimate; validated
The Approach

Hindsight
Review over historical time period

Compare ‘what if’ predictions with actual run-off

Estimate skill level by method or component of method

Estimate optimal combination of methods

Recommend Method given constraints

Constraints and considerations
• IT
• Data
• Tools etc.
Background to company data used in paper

- Commercial Auto BI Liability with heavy environmental influences
- Estimates of claim liabilities from 1979 to 1998
- Environmental influences during the period add difficulty to estimation
  - Economic and social inflation
  - Operational changes in claim department
  - Changes in underwriting posture
Performance testing is a formal analysis of prediction errors.

- Test a particular method by looking at historical performance – comparing estimates from the method with actual run-off.
- Giving us insights into the most accurate method to use.
Performance testing yields a formal measure of skill

- The skill of a method is measured by:
  \[ Skill_m = 1 - \frac{mse_m}{msa} \]
  - \( mse = \) mean squared error
  - \( msa = \) mean squared anomaly
- Skill is the proportion of variance “explained” by the method

**Actual Versus Projected Unpaid Claim Ratio Anomaly -- Accident Year @ 42 Months**

![Graph showing actual versus projected unpaid claim ratio anomaly from 1979 to 1998.](image-url)
Actuarial methods subjected to performance testing

<table>
<thead>
<tr>
<th>Actuarial Projection Method</th>
<th>Skill for Accident Year @ 42 Months</th>
<th>Overall Skill – for Latest Ten Accident Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paid Chain-Ladder</td>
<td>23%</td>
<td>13%</td>
</tr>
<tr>
<td>Incurred Chain-Ladder</td>
<td>52%</td>
<td>32%</td>
</tr>
<tr>
<td>Case Reserve Development</td>
<td>60%</td>
<td>22%</td>
</tr>
<tr>
<td>Reported Count Chain-Ladder</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>Case Adequacy Adjusted Incurred Chain-Ladder</td>
<td>52%</td>
<td>52%</td>
</tr>
</tbody>
</table>

- Note that absolute level of skill results are low due to changing case reserve adequacy and claim settlement patterns
Case study 1- Installing performance testing and a control cycle

- Corporate Actuary responsible for reserves set by decentralized organization of actuaries within each business unit
- Standard templates and database used to capture quarterly projections on an ongoing basis
- Actuaries review performance test results prior to each quarterly reserve-setting exercise; perform more detailed analysis annually
Embedding Performance Testing into Business Operations

The Actuarial Control Cycle for the Reserving Process - Embedding Reserve Risk Management

1. Define/Refine Process
2. Implement Process
3. Measure Performance

Formal Performance Testing
- Are the current methods appropriate? Would changes to methods improve estimation skill?
- Are the data and other input accurate and sufficient? Would improvements or expansion of data improve estimation skill?
- Are there opportunities to improve process flow?
- Are emerging estimation errors within tolerances?

Reserving Process Elements
- Data used
- Actuarial methods employed
- Operational input
- Judgments and intervention points
- Process flow and timeline
- Quality assurance process

Ensuring the reserving approach is continually monitored and adapted as required
Case study 2 - Empirical hindsight performance test data indicates that Mack may understate reserve risk

- Sample of 20 lines of business, “more difficult” US casualty lines
- Experience over a 15-20 year period
  - Historical best estimate reserve errors
- Mack based on most recent development triangle
  - includes parameter risk and tail factor volatility

![Mack Reserve Risk Performance Versus Hindsight](image)
Case Study 3 - UK Private Motor:
Median Skill scores by method and maturity

Median Skill by Maturity

Average Skill

-100% -80% -60% -40% -20% 0% 20% 40% 60% 80% 100%

Maturity (Months)

12 24 36 48 60 72 84 96 108

Paid CL
Inc CL
Case OS
Case Study 3 - Distribution of Skill Scores by company
Case Study 3 - Company D Skill scores by method and maturity
The minimum-variance weighting of methods depends on their variances and their correlation.

- For a given correlation, the optimal weights are those with the minimum combined variance.
  - Minimum starts at the very right, when correlation is 100%.
  - Minimum gradually shifts leftward as correlation decreases.
The best place to start is with a pilot project

- Test a few lines of business to gain some initial learnings
  - Lines where there is a ready data history
  - Cross-section of lines with varying degrees of difficulty
- Test current methods and new methods
  - Stochastic methods versus traditional
  - Man versus machine
- Use learnings to educate staff and demonstrate value
- Develop plan for further implementation
Benefits of Performance Testing

- Supports Solvency II
  - Formal validation of best estimates and ranges
- Embeds reserving control cycle
  - Improve accuracy of estimates
    - Inflation risk
    - Reserving cycle
  - Manage over-confidence
- Cost / benefit of enhancements to data and systems
Discussion