Effectiveness of Reserving Methods
working party
34th Annual GIRO Convention
2-5 October 2007
Celtic Manor, Wales
Richard Winter & Steven Fisher

What are we going to talk about today?
- Why hasn't this been done before?
- Some early conclusions
- Measuring effectiveness quantitatively
- Obtaining (or creating) sufficient data
- Designing the testing process

Why hasn't this been done before?
- Because it is hard!
Conclusion 1 – need lots of actuaries

- Testing the unadjusted operation of a calculation method is informative
- The obvious next question is whether, with the application of judgement, an experienced actuary can use information to get to a “better” answer
- To control the range of answers introduced by individual judgement will need a lot of actuaries to be involved

Conclusion 2 – need to know what the answer is

- If very recent claims information is used, then there is no way of telling what the final answer actually is – so cannot assess the effectiveness of the method
- A more effective method is to take claims data that is fully run off, and to present actuaries with the year-end information they would have received over several years
- Using data this old means that limited knowledge of distortions caused by, say, changes/delays in claims handling will now be available

Conclusion 3 – need to use lots of different datasets

- One aim of the working party is to show…
  …how effective a method is…
  …for a class of business…
  …at each stage of development
- To prove this, it is going to be necessary to demonstrate that effectiveness over many similar datasets, rather than just one or two
Conclusion 4 – need an objective measure

- It is possible to illustrate a method’s effectiveness simply (e.g. graphical representation).
- We believe that it will also be helpful to develop a quantitative measure of effectiveness to compare different methods.
- This will need to measure the effectiveness at different stages of claims development and under different circumstances.

A philosophical question

- What do we mean by an “effective” method?
- Which is more “effective”?
  - A method that frequently differs widely from the eventual outcome but, on average over many trials, comes very close to the eventual outcome; or
  - A method that has less variability from the eventual outcome, but on average over many trials is not as close to the answer; or
  - A method that gives a good answer at an early stage of development, but the accuracy of that answer doesn’t improve over time.

- Different methods may be more effective in different circumstances.

Objective measure of effectiveness (1)

\[
(\text{Estimated Ultimate at time } t) - (\text{True Ultimate})
\]

True Ultimate
Incurred chain ladder

Incurred BF to 70% development; Incurred CL thereafter

Objective measure of effectiveness (2)

(Estimated Ultimate at time t) – (True Ultimate)
True Ultimate – Paid to date
Incurred BF to 70% development; Incurred CL thereafter

A testing masterplan
- We need to test many different methods…
- Based on many different datasets…
- Covering many different classes…
- Run by many different actuaries…
- At many different year-ends!

Begin at the beginning: Data
- Limited availability of real company data
- Need complete datasets to cover full spectrum of methods (paid, incurred, premium/exposure, rating indices, number of reported/settled/nil claims, etc)
- Also need extensive data history – in order to know actual outcome
- Donations of real data gratefully received – full anonymity guaranteed!
Generating “pseudo-data”

- We believe that it is necessary to use “pseudo-data” as well as real data in order to handle some of the problems identified.
- The method proposed to generate this models:
  - The probability of claims
  - Reporting delays
  - Settlement delays
  - Changes in case estimates
  - Payment accounts compared to case estimates
  - Trying to reflect the way the real world operates

More than just methods

- Effectiveness of the “pure” method
- Testing of mechanical operation of methods
- Value added by actuarial judgement
- Testing by individual actuaries with limited background info
- Value added by understanding the business
- Testing by individual actuaries with detailed background info

Separate testing streams

- Mechanical testing
  - Macro-based
  - Use of pseudo-data
  - Multiple year-ends
  - Many methods and variations
  - Test effectiveness of each method in isolation
- Manual testing
  - Individual actuaries
  - Use of real data
  - Multiple year-ends
  - Core methods
  - Test individual methods and overall selected results
  - 2 subgroups: limited/detailed background info
Core methods for testing

- Paid and incurred chain-ladder
- Paid and incurred Bornhuetter-Ferguson
- ACPC-type methods: payments per claim incurred; payments per claim finalised
- Probabilistic trend family methods (eg ICRFS)
- Stochastic methods?
- Operational time method?
- Other methods?

Variations on standard methods

<table>
<thead>
<tr>
<th>Development factor selection basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>volume-weighted</td>
</tr>
<tr>
<td>time-weighted</td>
</tr>
<tr>
<td>unweighted</td>
</tr>
<tr>
<td>ex high/low</td>
</tr>
<tr>
<td>last few years</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BF exposure measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>premium</td>
</tr>
<tr>
<td>vehicle-years (motor)</td>
</tr>
<tr>
<td>ultimate claim count</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BF IEULR selection basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>premium</td>
</tr>
<tr>
<td>vehicle-years (motor)</td>
</tr>
<tr>
<td>wage-roll (EL)</td>
</tr>
<tr>
<td>ultimate claim count</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calling all volunteers…</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-20 distinct classes, with mix of issues</td>
</tr>
<tr>
<td>At least 5 successive year-ends</td>
</tr>
<tr>
<td>10-20 testers for each class</td>
</tr>
<tr>
<td>Need lots of volunteers to test 1-2 classes each</td>
</tr>
<tr>
<td>100 volunteers from outside working party (half a day each)</td>
</tr>
<tr>
<td>Working party members each to test all classes</td>
</tr>
<tr>
<td>All volunteers to receive confidential individualised feedback on results</td>
</tr>
</tbody>
</table>
Six simple steps

- **STEP 1**: Ready-to-use projection template will be provided, fully populated with data and diagnostic exhibits
- **STEP 2**: Select development factors and other parameters; initial selections will be automatically recorded
- **STEP 3**: Compare results of different methods; revisit selected assumptions if desired; updated selections will be automatically recorded
- **STEP 4**: Select your best estimate
- **STEP 5**: Repeat steps 1-4 at next four year-ends
- **STEP 6**: Do not adjust older projections in the light of the subsequent emerging experience; that would be cheating!

Next steps

- Working party open to new joiners
- Search for data
- “Dry run” of testing methodology prior to roll-out
- Recruit legions of volunteers, and complete testing of methods
- Analyse the results and draw conclusions
- Other aspects of working party terms of reference:
  - Investigation of existing literature
  - Documentation of known strengths and weaknesses of main methods
  - Documentation of useful diagnostics
- Report back in Sorrento

Questions

- To what extent would you trust results from the working party based on pseudo-data?
- Are there any critically important methods that we’ve missed?
- Would you be prepared to volunteer to participate in the main testing exercise? How much time?
- Are you able to provide real company data for use in testing? (We will help you to anonymise the data)