HARBORPOINT

HARBOR POINT Re LIMITED

Exposure Data Quality's Impact on Reinsurance Pricing

Jed Rhoads

CARe Meeting – May 18, 2009

Quotes on Data

- Experts often possess more data than judgment.
  Colin Powell

- Errors using inadequate data are much less than those using no data at all.
  Charles Babbage

Quotes on Data - continued

- Data is not information, information is not knowledge, knowledge is not understanding, understanding is not wisdom.
  Clifford Stoll

Exposure Data Quality's Impact on Reinsurance Pricing

- 'Garbage in, Perfection out'
- Personal Lines vs. Commercial Lines
- International vs. US
- Ike and other Anecdotal Evidence
- Severity Adjustment Factors

HarborPoint

‘Garbage in, Perfection out’

- Data Quality vs. Data Integrity
- The six “C’s: completeness, current, consistent, correct, controlled, credible
- Data Quality = completeness, current, consistent
- Data Integrity = correct, controlled, credible

“Garbage in, Perfection out’ - continued

- Priorities: why don’t we exert similar effort and expense on data input that we do on modeling and data systems?

‘Chinese Whispers’

- Lost in translation: who inputs the data?
- Third parties, agents, MGA’s, brokers, customers themselves, clerical staff, underwriters (rarely), inspectors, engineers, etc.
- Do they have an “agenda” and/or do they care?
- The data odyssey to a reinsurance underwriter: original inputs to insurer to modeling company to intermediary to reinsurer.

After being handled so many times, what are we actually receiving?

- How many times has it been manipulated?
- Did each step of the process actually add value, or did it take value away in the final analysis?
- What level of employee actually inputs the data? Is there an understanding as to the value we ascribe to the function?

Personal Lines – exposure data quality

- Historically, personal lines cat losses have outperformed commercial lines in actual to modeled losses.
- Why?
  - More homogenous (less complex risks),
  - Smaller value = single errors matter less,
  - Fewer data fields
  - Typically wind or EQ driven, not both.

Commercial Lines

- EQ and Wind are both issues, but EQ is a larger issue than wind
- Single risk multiple locations: all coded to one location
- Single risk multiple locations: all with identical characteristics: age of construction, etc.
- Producers are more sophisticated = know what drives pricing
- Risks are more complex = difficult to input properly

International vs. US/Canada

- Non-US/Canada exposure data is clearly inferior
- Compounding the issue: more unmodelled perils (e.g. recent China losses)
- US/Canada data is not great, but it is superior to non-US/Canada data

Hurricane Ike

- “Due to the fact that most of our claims were at least 30 miles inland, we now believe that a combination of insufficient building codes, and possibly unreliable enforcement of those codes, was a major contributor to the increased incurred figure. In addition, tree damage was much higher than anticipated and not handled well in the commercial models, older roof age due to the lack of an event in the area for many years and an inability of our insureds to fix the damage quickly due to demand surge also have contributed significantly.”
- Duh….. Nothing new here. Why do we act like this is new?

Ike - continued

- Personal lines performed best relative to model error (~25% off)
- Light – Medium commercial (~50% off)
- Heavy commercial (~80%+ off)
- D&F, E&S, Binders (MGA) performed worst (~300%+).
- No surprise: the data quality/integrity mirrors the results relative to modeled output

Ike - continued

- Ike was not unlike other recent cat losses.
- ITV emerged as major issue with some companies.
- Regional differences? Indeed.
- Agents/producers play games. For example, most masonry stucco construction were coded “masonry”. It should have been coded “masonry veneer” or “frame”.
- Is inland data poorer than coastal data?

Other Anecdotal Evidence

- Found in exposure data:
  - 14 story frame building
  - Every residence has a HIP roof (vs. Gabled)
  - Appurtenant Structure: mobile home
  - Roof age: 15 years, on 3 year old home
  - Age of Construction: all coded 1998 (code change year)
- Wind Mitigation: can your grandmother in Florida actually hang the shutters when a storm is coming?

Adjustment Factors – ‘Rumsfeld’ style

- Pricing the ‘known knowns’
- Pricing the ‘known unknowns’
- Pricing the ‘unknown unknowns’
- True story (post Northridge): CEO asks modeling company why Northridge was not in the model. Response: we don’t know all the fault lines. CEO: how many fault lines are not in the model you don’t know about? Response: 113.

Severity Adjustment Factors

- Can you name three examples in the past 15 years where the cat models were “heavy” post-event?
- There is either something wrong with the data or the model, or both.
- The models, while not perfect, probably contribute less to the problem.
- Main culprit: poor exposure data.
- The problem relates to EQ as well as Wind.

Severity Adjustment Factors – continued

- How do we compensate?: apply adjustment factors = pricing higher than models.
- There are more data points with Wind than EQ. Wind data has improved faster and therefore is probably of higher quality and more predictive in a model than EQ.
- Ironically, EQ loads are probably less used than Wind loads.
- With the new 2009 model changes in EQ, will adjustment factors be more commonplace?

Severity Adjustment Factors - continued

- If we don’t know or question the data quality, we apply severity adjustment factors:
  - More for commercial than personal
  - More for EQ than Wind (atypical for industry)
  - More for major peril zones with fewer loss data points, than those with data points.
  - Ground them in experience
  - Actual vs. modeled historical performance
  - Perform a gates analysis

Final Quotes

- Life is made up of a series of judgments on insufficient data, and if we waited to run down all our doubts, it would flow past us.
  Learned Hand
- Data is a precious thing and will last longer than the systems themselves.
  Tim Berners-Lee

Conclusion

- Cat models are developmentally well ahead of data quality and integrity. The next major advancement in understanding cat exposures will be in data capture. **Jed Rhoads**