Benchmarking Residential Mortgage Loss Forecasts

Third Annual Stress Test Modeling Symposium
Federal Reserve Bank of Boston

Michael Szwejbka
VP Credit Risk Analytics
June 26, 2014
Hurricane CCAR

The Stress Testing Hurricane

Is predicting the path of a hurricane similar to predicting the path of a financial crisis?
Benchmark: Standard, or comparison, for assessing performance

- **Examples**
  - Performance of a security against a bond index
  - Return on USB stock compared to the S&P 500
  - Individual bank loss rates vs. industry-wide loss rates

- **Benefits of benchmarking**
  - Gives broader view of possible outcomes
  - Assess the effect of business assumptions, which may vary across models (ex. treatment of new volume, off-balance sheet exposures)
  - Assess impact of inherent statistical assumptions
  - Begin to address model risk

- **Evolution of benchmarks**
  - As modeling progresses, previous production models become benchmarks
  - Other R&D models
  - Models with external data or from vendors
### Model Spectrum

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model Level</strong></td>
<td>Loan-level</td>
<td>Hybrid of loan-level and portfolio-level</td>
<td>Portfolio-level (time series regression)</td>
<td>Portfolio-level (seemingly unrelated regressions)</td>
</tr>
<tr>
<td><strong>Default Model</strong></td>
<td>Multinomial logistic regressions and Markov Chains</td>
<td>Logistic regressions and portfolio-level Markov Chains</td>
<td>Built in to NCO</td>
<td>Built in to NCO</td>
</tr>
<tr>
<td><strong>Loss Given Default</strong></td>
<td>Account-level regressions</td>
<td>Account-level regressions</td>
<td>Built in to NCO</td>
<td>Built in to NCO</td>
</tr>
<tr>
<td><strong>Auto-regressive Terms</strong></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Explicit Portfolio Effects</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Macroeconomic Variables</strong></td>
<td>State-level variables</td>
<td>State-level variables</td>
<td>National-level variables</td>
<td>National-level variables</td>
</tr>
<tr>
<td><strong>Complexity</strong></td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

*Note: Models 1, 2, and 3 are validated; Model 4 is not validated*
Forecasts

Different Models, Different Forecasts

- Each model uses different statistical techniques
- Two account-level and two portfolio-level models
- Treatment of macroeconomic variables results in differing sensitivities and timing of peak losses

Residential Mortgage Benchmark Loss Forecasts

Note: Graph for illustrative purposes only; uses dummy data and a hypothetical stress scenario.
Variation Summary

How can output from several models be summarized?
- Graph below shows the maximum, minimum, and average across the four models
- Averaging may be more useful as the number of benchmarks increase
- Minimum and maximum should not be confused with confidence intervals

Residential Mortgage Benchmark Loss Forecasts

Note: Graph for illustrative purposes only; uses dummy data and a hypothetical stress scenario.
Sources of Variation

What causes variation in the forecasts between models?

- Top-down models “paint with a broad brush”
  - Components of losses are rolled into the NCO metric
  - Provides less insight into what’s driving losses – defaults, LGDs, portfolio composition

- Borrower Characteristics
  - Models 1 and 2 control for portfolio composition (ex. credit score, LTV, age)
  - Impacts the ‘starting’ point of the forecast

- Macroeconomic Variables
  - Top-down models use national macroeconomic variables
  - Bottom-up models use more granular levels of macroeconomic variables
    - State or MSA-level macroeconomic variables tailor models to the portfolio’s footprint
    - Ability to stress regional concentrations
Use of Benchmarks

How to use the results of benchmarking?

- Identify models weaknesses for management overlays
  - Provide quantitative support for overlays
  - Does a particular model type reflect the current portfolio better than the others?
  - Are there known business changes that make one model preferable over the others?
  - How does backtesting compare across models?
  - Is data limited for a certain model type?

- Promote discussion and R&D
  - Consensus across benchmark models not necessary
  - Concentration risk identification
  - Differences promote discussion amongst modelers, management, and between groups