



Performance Testing

Sugata Chakravarty
SVP - Wholesale Model Development

Fifth Annual Stress Test Modeling Symposium
Federal Reserve Bank of Boston
October 6th, 2016

Disclaimer: Views expressed are those of the speaker and do not necessarily represent the views of PNC

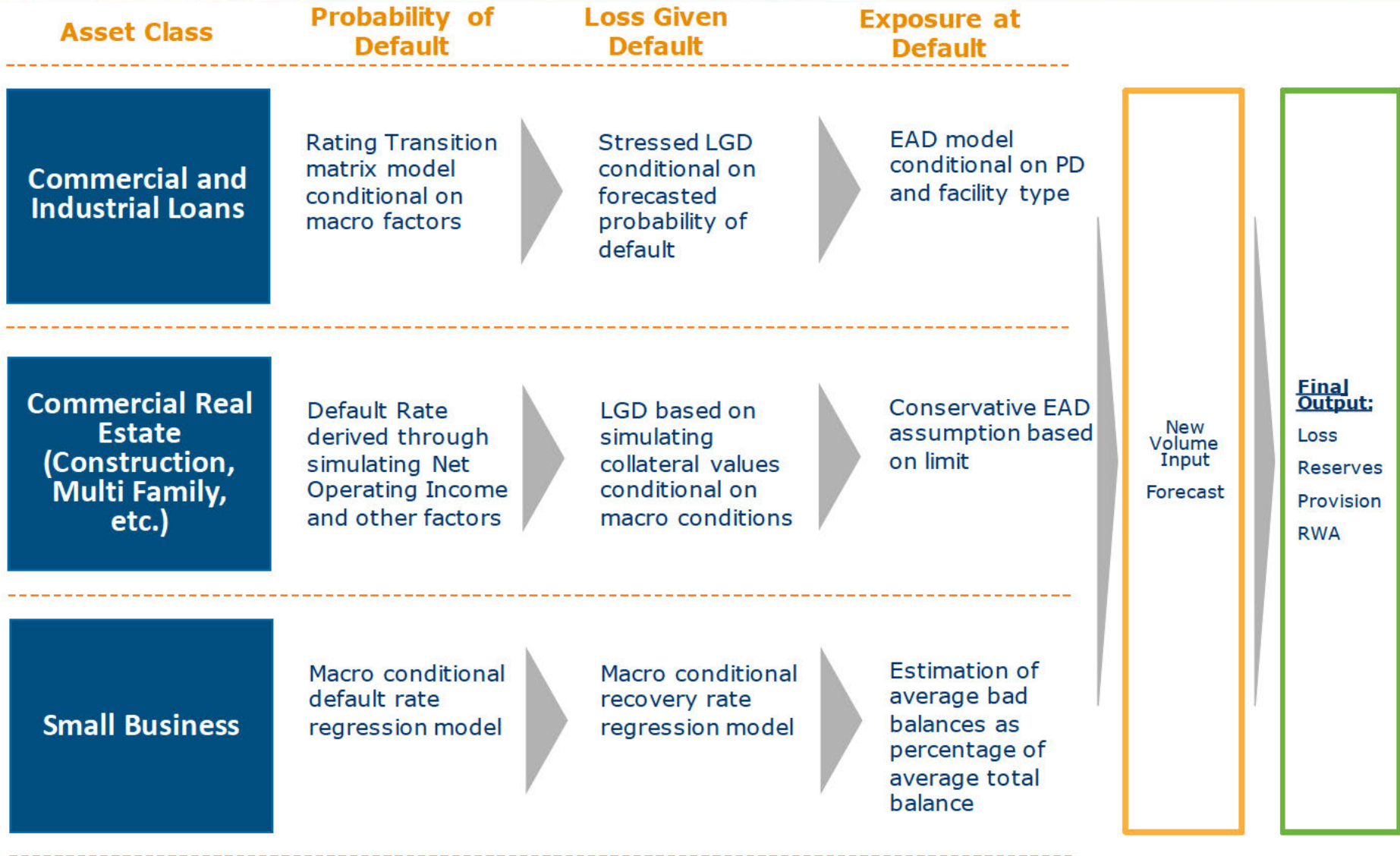
Contents

- Overview of Wholesale Loss Forecasting Models
- Performance Testing metrics used for model monitoring
- Benchmarking and Sensitivity Testing
- Challenges/Issues in monitoring models and mitigation processes
- Q&A

Glossary: Champion, Challenger and Benchmark

	Champion	Challenger	Benchmark
Use	Primary model used for all decision making, loss reporting	Can be used to inform overlays and to validate champion model results	Informative and used for directional guidance
Data	Based on internal data when available otherwise may rely on external sources	Based on internal data when available otherwise may rely on external sources	Based on external data or aggregate portfolio information
Methodology	Generally moderate to complex methodologies to estimate individual components such as PD, LGD and EAD	Methods may be more complex than Champion models and typically based on different modeling techniques	Simple modeling techniques like OLS or heuristics
Model Risk Validation	Subject to rigorous validation and backtesting standards. Requires quarterly or half yearly monitoring plans	Subject to slightly less rigorous validation and backtesting standards. Requires quarterly or half yearly monitoring plans	Subject to validation and monitoring. Intensity is commensurate with importance
Example	Simulation, generalized linear models	Survival, Regression Trees	Simple regression of Net Chargeoff data on macro variables

Wholesale Loss Forecasting Models



Common performance metrics aim to measure overall goodness of fit and underestimation error

Performance Measures

Objective

Adjusted R²

- Allows broad comparison across models and asset classes (independent of dependent variable definition)
- Adjusted for number of macro variables to avoid overly optimistic view of accuracy
- Measure of in-sample performance

Root Mean Squared Error (RMSE)

- Measures typical error; "how far off is the model on average?"
- Results in the same unit as model output (e.g. PD%)

Under-estimation frequency

- Analyzes whether underestimation is recurring (systematic model optimism or lag) or due to a few isolated "misses"
- Suggested measure by quarterly time steps

Average under-estimation

- Measures underestimation error

Sensitivity Ratio

- Measures relative sensitivity of model forecasts to portfolio shocks

Typical Performance Testing and Monitoring

Model Back-testing and Monitoring

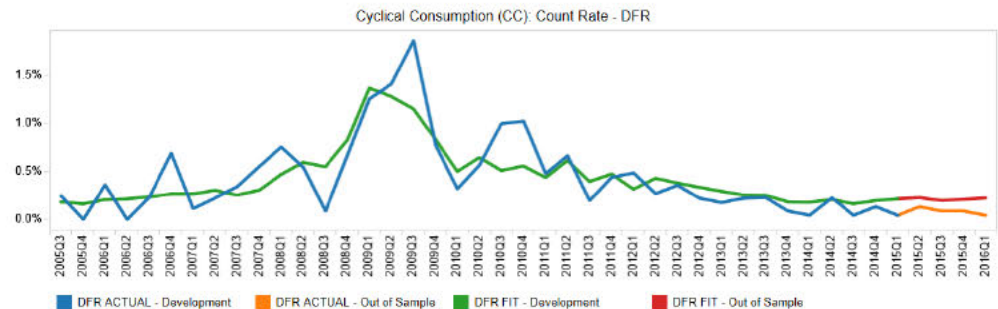
- Standard practice of separating Training Data and Out of Sample for performance testing
- Frequent monitoring comparing Actual vs. Predicted values with Red/ Yellow/ Green Thresholds
- Continued evaluation of Stress Test Model applicability to the portfolio
- Additional monitoring from Model Risk Group (2nd line) review help identify model limitations and key assumptions

Example: Monitoring Report

Summary		Threshold			CC
Performance	Metrics	Green	Amber	Red	Q1 2016
Accuracy	KS Test	0 <= 0.1	0.1 <= 0.2	0.2 <= ∞	●
	RMSE (APD)	0 ·		: <= ∞	●
	RMSE (DFR)	0 ·		: <= ∞	●
Stability	PSI (By DHE)	0 ·			●
	PSI (By PD Bands)	0 ·			●
	PSI (By Sectors)	0 <= 0.1	0.1 <= 0.25	0.25 <= ∞	●

Illustrative thresholds

Example: Default Rate series out of sample testing

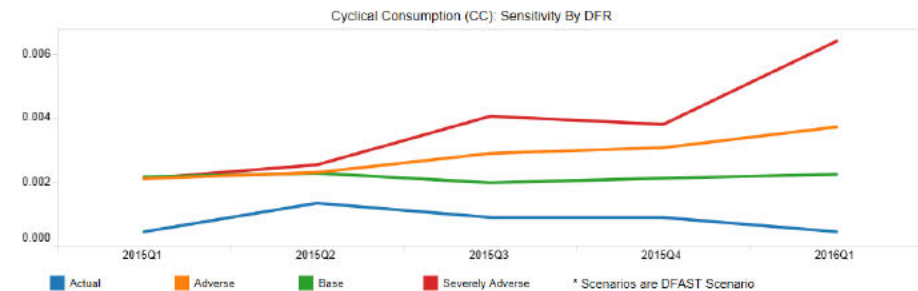


Champion/challenger framework and Model Sensitivity

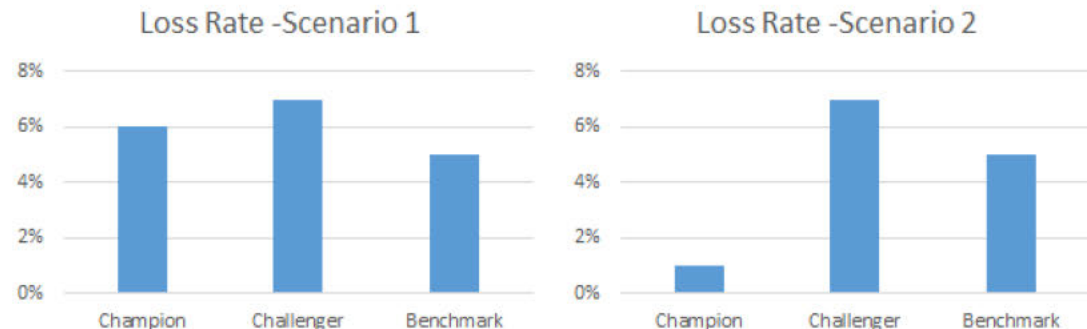
Benchmarking and Sensitivity Analyses

- Champion – Challenger Framework compares Champion results against Challenger model that employs a completely different modelling concept/ construct
- Independent challenge and review process involving senior execs, and Line of Business to consider any overlays to Champion model results
- Sensitivity Analyses to inform changes in results from changes in key model assumptions

Example: Sensitivity measured by separation in Default Rates



Results may vary across models and require further investigation



Most challenges to performance testing stem from issues common to the model development process

1. **Limited History** with at best data covering just one credit cycle

- Testing impacted as expanding time period of 'hold-out test data' limits the time period on training
- Use random sampling of all historical data to provide various training / test data splits

2. **Model Segmentation** Schema that balances homogeneity of model segments and ensures sufficient default count across model segments

- Limited segmentation may captures the key differences in portfolio attributes
- Usage of borrower attributes as variables rather than segmentation parameters

3. **Idiosyncratic events** are difficult to predict and can manifest as over/ under prediction errors, especially in *Low Default portfolios*

- Model overlays informed through Champion or Benchmark approaches

4. **Definitional Differences between Historical loss metrics and Forecast metrics**

- Historical loss metrics (e.g. net charge-offs) are Accounting constructs impacted by Credit policies whereas Forecast losses are often Economic constructs
- Historical Reserves are often impacted by management decisions

Management judgment plays a critical role in interpreting / using results from Challenger / Benchmark models

1. Applicability

- Challenger models built on external data may not be representative of the Bank's current portfolio

2. Challenger Model Methodology

- Is the Challenger model methodology robust enough to account for changes in Portfolio Credit quality

3. Justification

- Is there a well identified Champion model limitation that is specifically addressed by the Challenger model? (e.g. Low Default portfolio results)
- Is the Champion model limitation better addressed via a separate analytical exercise quantifying the impact of the limitation and providing results under alternative assumptions. (e.g. Individually Impaired Loans)

Questions?