



Life Cycle of a Model: Benchmarking

FRB Modeling Symposium

Michael Szwejbka
SVP, Quantitative Model Development
U.S. Bank


Disclaimer



Views expressed in this presentation are mine and do not necessarily reflect those of U.S. Bank.

Agenda



- Expectations
 - SR Letter 15-18
 - SR Letter 11-7
 - Benchmarks for Ongoing Monitoring
 - Examples
- 

SR Letter 15-18



“A firm should use benchmark or challenger models to assess the performance of its primary models for all material portfolios or to supplement, where appropriate, the primary models.”

“Benchmark models that are developed and run independently of primary models can be used to more effectively calibrate the firm’s final estimates.”

“Benchmark models used to arrive at the firm’s final estimates should be subject to model risk management.”

*Appendix C, SR Letter 15-18
Board of Governors of the Federal Reserve System*

SR Letter 11-7



“Benchmarking is the comparison of a given model's inputs and outputs to estimates from alternative internal or external data or models. It can be incorporated in model development as well as in ongoing monitoring. For credit risk models, examples of benchmarks include models from vendor firms or industry consortia and data from retail credit bureaus. Pricing models for securities and derivatives often can be compared with alternative models that are more accurate or comprehensive but also too time consuming to run on a daily basis.”

“Discrepancies between the model output and benchmarks should trigger investigation into the sources and degree of the differences, and examination of whether they are within an expected or appropriate range given the nature of the comparison. The results of that analysis may suggest revisions to the model. However, differences do not necessarily indicate that the model is in error.”

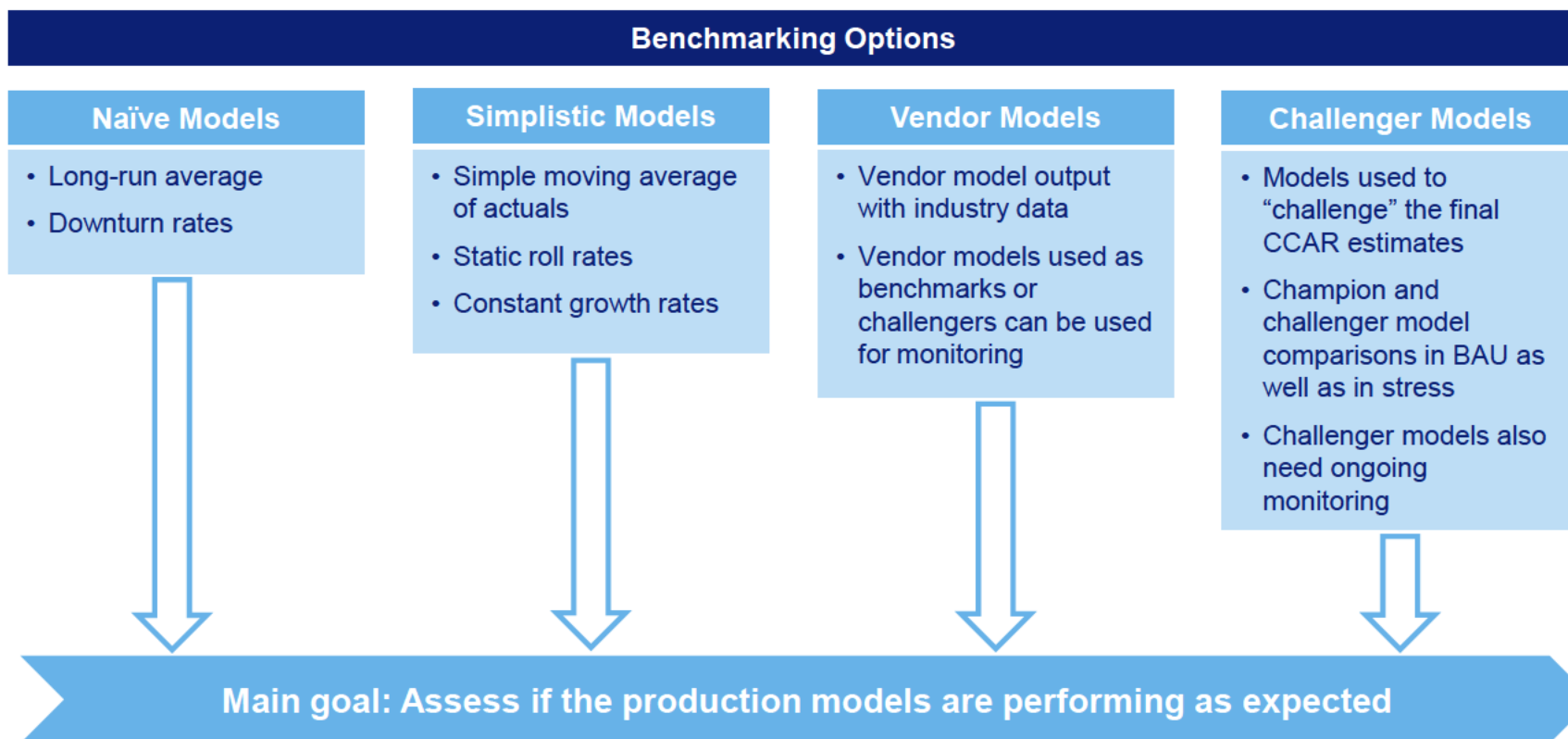
*SR Letter 11-7
Board of Governors of the Federal Reserve System
Office of the Comptroller of the Currency*

Benchmarks for Ongoing Monitoring



- Benchmarks are diagnostic tools to assess the production model
- Similar to outcomes analysis/backtesting
 - Outcomes analysis typically compares actuals vs predicted (\hat{p}_{prod})
 - Adding benchmarking there are multiple comparisons: actuals vs \hat{p}_{prod} vs \hat{p}_{b1} vs \hat{p}_{b2}
 - Most interested in actuals vs \hat{p}_{prod} but can also consider \hat{p}_{b1} vs \hat{p}_{b2}
- Benchmarks for ongoing monitoring may be the same benchmarks used during model development
- A more complex benchmark is not necessarily better; need to balance complexity and practicality
- Do not need to be production models

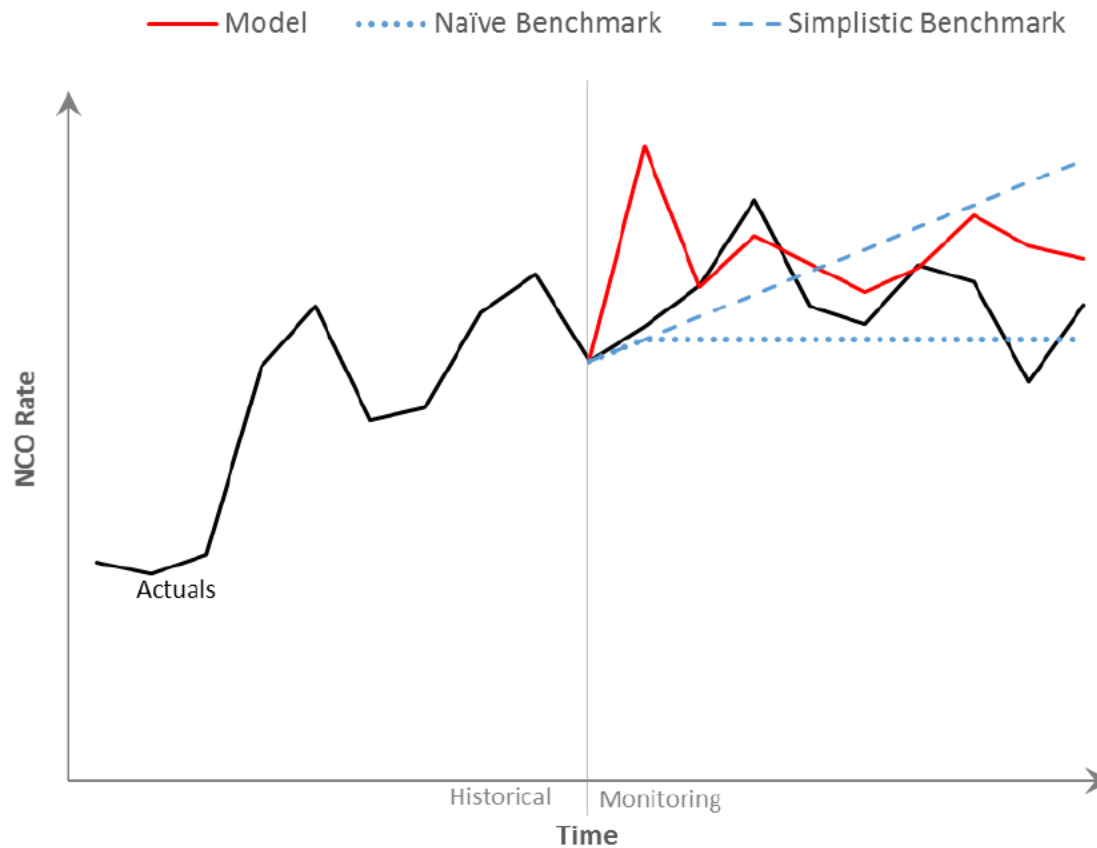
Benchmarks for Ongoing Monitoring



Example

Example of when simple methods are appropriate

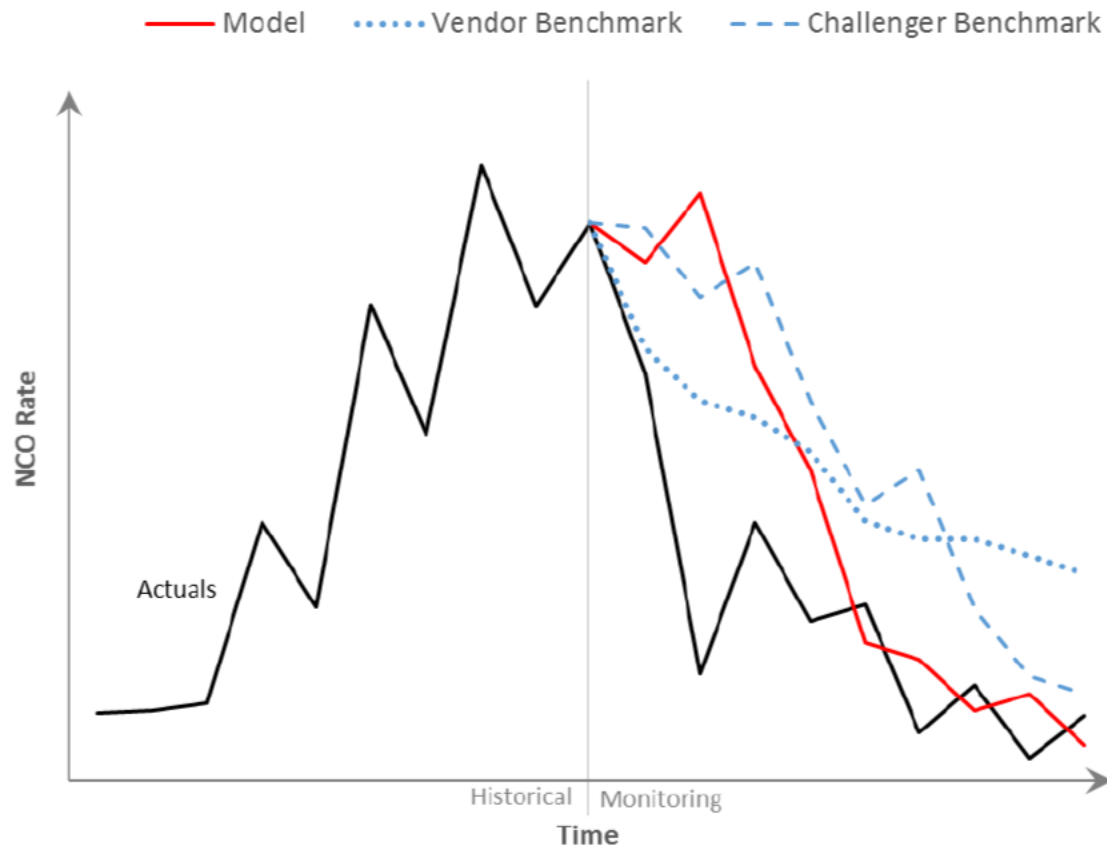
Stylized Example of Naïve and Simplistic Benchmarks



Example

Example of when more complex methods may be needed

Stylized Example of Vendor and Challenger Benchmarks

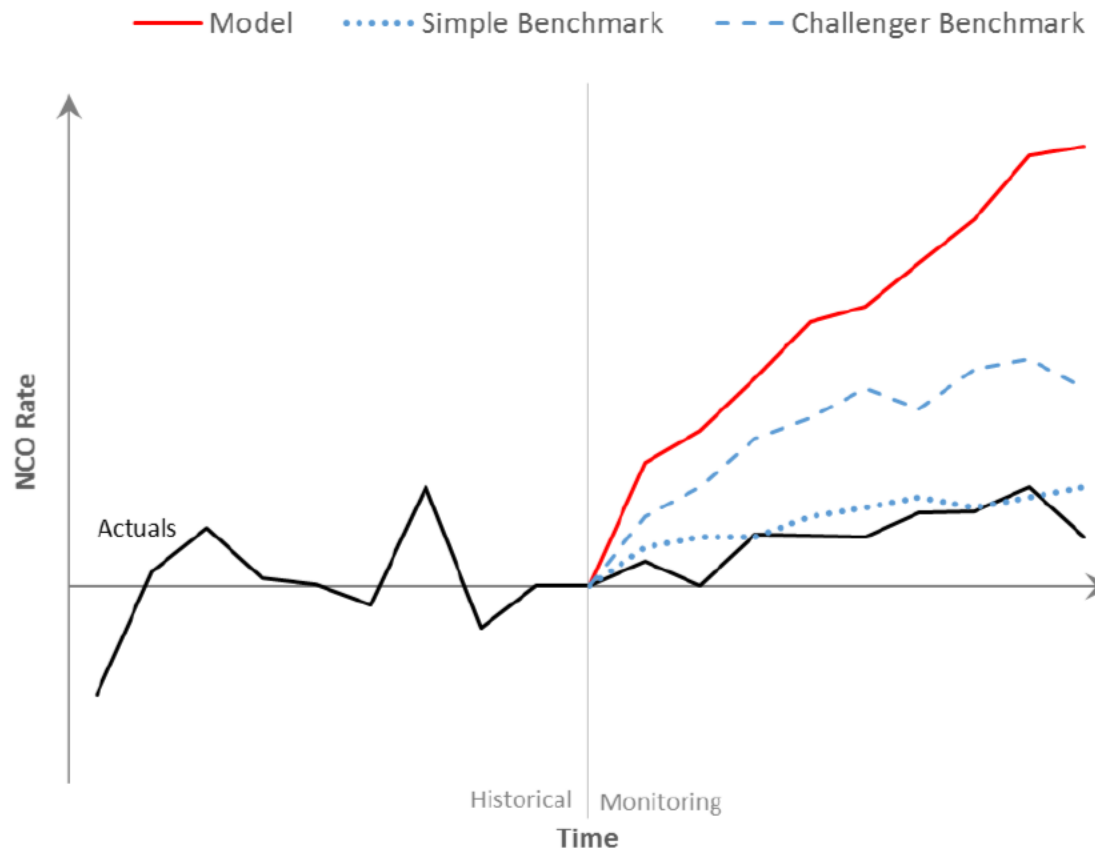


Note: Simulated data;
for illustration only.

Example

Example where stress testing models may not perform well – very low NCO rates

Stylized Example of Simple and Challenger Benchmarks



Note: Simulated data;
for illustration only.

Conclusion



- Ongoing monitoring for stress testing models is unique
- There are times in an economic cycle when a simple roll rate model may be better than a complicated transition matrix
- Benchmarking can give a measure of relative performance of the production model
- Benchmarks can reveal weaknesses and/or limitations of the production model
- Benchmarks may help establish performance thresholds
- The same type of benchmark may not be suitable for every situation



Michael Szwejbka
michael.szwejbka@usbank.com