

BACK TESTING INTEREST RATE – RISK MODELS

Prudent risk managers should not make any decisions about their rate-risk exposure unless they have confidence that their models are at least reasonably accurate. Of course, most risk managers are quite confident. The question is, do they have sufficient reason to be so confident?

While model error is unavoidable, input and output tests can make a difference.

All too often, confidence in one's model results seems to correlate with the cost of the model. Managers who have spent hundreds of thousands on models like to believe that they have Ferraris. Managers who have spent tens of thousands on models may feel envious. Experience tells us that both impressions can be right or wrong, depending on how well the models are used.

Confidence in one's model forecasts should be based on well-structured testing. In its 2004 guidelines for interest-rate risk (IRR), the Bank for International Settlements (BIS) notes that:

Reviews of the interest rate risk measurement system should include assessments of the assumptions, parameters, and methodologies used. Such reviews should seek to understand, test, and document the current measurement process, evaluate the system's accuracy, and recommend solutions to any identified weaknesses. If the measurement system incorporates one or more subsidiary systems or processes, the review should include testing aimed at ensuring that the subsidiary systems are well-integrated and consistent with each other in all critical respects.¹

Collectively, those assessments are called "back testing." Back testing is an activity that gets lots of favorable press but not enough practical use.

THE SCOPE OF THE PROBLEM

In 2007, a U.S. bank with total assets in the range of \$8 billion to \$12 billion had its interest rate risk measured, as of the

same date, using three different models. The actual results² are shown in Table 1.

Table 1: Interest Rate–Risk Measures: Three Different Models

	Model A	Model B	Model C
Change in EVE if interest rates rise 200bp	-21%	-18%	-24%
Change in EVE if interest rates fall 200bp	-21%	+3%	+1%

The dispersion in the measurements for a 200 basis point (bp) rise in rates is six percent of the economic value of equity (EVE). The dispersion in the measurements for a 200 bp fall in rates is 22 percent of EVE.

Consider limits. Most banks would likely be in violation of EVE limits for sensitivities greater than 20 percent. If the bank only used Model A, it would report a limit violation. But if the bank used either model B or C, management would think that the bank was well within its limits for the 200 bp falling rate shock. Which is correct?

Consider risk management costs. Many banks in this size range use rate swaps to hedge their IRR exposures. If the bank used only Model B, it might underhedge and lose an unexpectedly large amount of value in the event of a major increase in rates. If the bank used only Model C, it might spend too much on hedges and suffer unexpected losses in the other direction.

WHAT CAN BE DONE?

Earnings at risk (EAR) back testing is always easier than EVE back testing. Forecasts of either net-interest income (NII) or net income from EAR models can be compared to the actual results subsequently observed. Common output back tests include the following:

- Compare the forecasted NII for a subsequent period to the "normalized NII" actually observed for that period. The normalized NII is the NII adjusted for nonrecurring income or expense.

- Use rate/volume/mix variance analysis to isolate the variances observed between forecasted and actual earnings that result only from the actual rate changes. Then see how closely the observed variance resulting from the actual rate changes compares to the forecasted changes.
- Save prior model runs. At a later date, rerun the model with the actual observed market rates instead of the forecasted rates used in the original model run. Keep the data, the assumptions and everything else the same as in the original model run. Then see how closely the income reported from rerun with actual rates compares to actual income for that period.

Output testing for EVE models is far more difficult. Obviously, EVE cannot be compared to any independent value. EVE is not the same as book equity, regulatory capital, economic capital or market cap.

EVE model users can, and should, compare their models' calculated market values to observed market values for actively traded securities owned by the bank.

BEYOND OUTPUT TESTING

Notice that the BIS quotation cited above does not explicitly refer to output testing. Instead, it lists assessments of the assumptions, parameters and methodologies used. This boils down to data testing and assumption testing.

Data inputs into both EAR and EVE models should be tested. This does not mean that all inputs from bank records must be reconciled to the penny. It does mean that inputs from other bank records should be reconciled, with some allowance

for error, to general ledger values to ensure that all material asset or liability volumes are included in the model.

Data inputs to both EAR and EVE models also require scrubbing. Much of the data important to rate-risk managers, such as loan rate caps and securities calls, cannot be reconciled. Errors tend to accumulate. Checking data inputs to find and fix such errors (data scrubbing) is a full-time process for most medium-sized and all large banks.

Assumptions used in both EAR and EVE models should also be back tested. Observed changes in variables, such as changes in loan prepayment speeds following a change in market rates, should be regularly compared to the assumed change employed in the modeling process.

SUMMARY

Model error is unavoidable. After all, models are, by definition, simplifications of reality. Both EAR and EVE models have huge quantities of data inputs and heavily depend on rate and volume assumptions.

It is often noted that the skill of the craftsman, not the tools, makes the difference between a job done and a job well done. For every rate-risk manager, input and output tests are essential tools. Wise rate-risk managers go beyond mere policy or regulatory compliance and use these tools to carefully assess the size of their model error before making risk management decisions.

This article was contributed by Leonard Matz, Director, Liquidity and Interest Rate Risk Consulting, SunGard.

For more information contact ambitinfo@sungard.com

Footnotes:

¹ Principle 10, Paragraph 66.

² The numbers shown are slightly modified to protect the anonymity of the source. The relationships between the figures remain almost unchanged.

First published in
BANK ACCOUNTING & FINANCE